

Universidade Federal de Alfenas

Thaís Brenda Martins

**Análise de máximos de risco Brasil e resseguros: uma
abordagem Bayesiana**

Alfenas/MG

2018

Thaís Brenda Martins

**Análise de máximos de risco Brasil e resseguros: uma
abordagem Bayesiana**

Dissertação apresentada ao Programa de Pós-Graduação 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. Davi Butturi-Gomes.

Colaborador: Prof. Me. Reinaldo Antônio Gomes Marques.

Alfenas/MG

2018

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

Martins, Thaís Brenda.

Análise de máximos de risco Brasil e resseguros: uma abordagem
Bayesiana / Thaís Brenda Martins. -- Alfenas/MG, 2018.

110 f.: il. --

Orientador: Luiz Alberto Beijo.

Dissertação (Mestrado em Estatística Aplicada e Biometria) -
Universidade Federal de Alfenas, 2018.

Bibliografia.

1. Teoria bayesiana de decisão estatística. 2. Previsão estatística. 3. Testes
de hipóteses estatísticas. I. Beijo, Luiz Alberto. II. Título.

CDD-519.54

Ficha Catalográfica elaborada por Marlom Cesar da Silva
Bibliotecário-Documentalista CRB6/2735



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/>



THAÍS BRENDA MARTINS

“Análise de máximos de risco Brasil e resseguros: uma abordagem Bayesiana”

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: 27 de abril de 2018.

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

Assinatura: 

Profa. Dra. Patrícia Siqueira Ramos
Instituição: UNIFAL-MG

Assinatura: 

Prof. Dr. Leandro Ferreira
Instituição: UNIFAL-MG

Assinatura: 

AGRADECIMENTOS

Início meus agradecimentos por *Deus*, sem Ele nada seria possível.

Ao meu orientador, *Prof. Dr. Luiz Alberto Beijo*, por todo conhecimento compartilhado, orientação, incentivo e apoio.

Aos *Profs. Reinaldo e Davi*, respectivamente, colaborador e coorientador do meu trabalho, pela contribuição, incentivo, apoio e disponibilidade.

Aos *Profs. Leandro e Patrícia*, por participarem das minhas bancas de qualificação e defesa, e pelas valiosas sugestões que colaboraram para a melhoria do meu trabalho.

Aos meus pais, *Marli e Valdir*, meu infinito agradecimento por sempre acreditarem em minha capacidade e me incentivarem a buscar e fazer o melhor continuamente. Obrigada pelo amor incondicional.

Aos meus amados irmãos, *Carol e Matheus*, por estarem sempre comigo e serem o maior presente que meus pais poderiam me dar, e a *Paula* por trazer para as nossas vidas a pessoa que alegre e transborda de amor nossos dias, minha pequena, *Manu*.

As minhas velhas e melhores amigas, *Francielly e Camila*, por se fazerem sempre presente e me apoiarem em todas as escolhas que fiz, vocês são muito mais que amigas, são minhas irmãs de alma, são presentes de Deus em minha vida.

À todos os amigos/colegas que fiz durante o mestrado, Deus me permitiu conhecer pessoas maravilhosas, que tornaram essa jornada mais prazerosa, em especial a Gisele, Larissa e Isabela.

À todos os professores e funcionários do Departamento de Estatística, obrigada pelo conhecimento e experiência compartilhada. À *Martha* pela amizade e apoio. E agradeço à Deus a oportunidade de ter conhecido e convivido com a dona *Kátia* (em memória).

Agradeço, também, à CAPES pelo apoio financeiro.

E, por fim, gostaria de agradecer à Universidade Federal de Alfenas e a todos os professores com quem tive contato por me proporcionarem mais que o conhecimento profissional e científico adquirido durante a graduação e o mestrado, mas sim pela experiência de vida incrível que me fez mudar o modo de pensar e agir muitas vezes, me fez amadurecer e crescer pessoalmente.

Ninguém vence sozinho, muito obrigada a todos!

Ainda que a minha mente e o meu corpo enfraqueçam, Deus é a minha força, Ele é tudo o que eu sempre preciso (Salmos 73:26).

Dedico este trabalho à Deus, aos meus pais e aos meus irmãos.

ANÁLISE DE MÁXIMOS DE RISCO BRASIL E RESSEGUROS: UMA ABORDAGEM BAYESIANA

Thaís Brenda Martins; Luiz Alberto Beijo; Davi Butturi-Gomes; Reinaldo Antônio Gomes Marques

RESUMO: A teoria de valores extremos é frequentemente utilizada para análise e previsão de dados meteorológicos, mas sua aplicação tem se expandido para diversas áreas em que a ocorrência de eventos extremos causam grandes perdas financeiras, como na área de finanças, mercado de ações e seguros. Um método de estimação de parâmetros que vem sendo cada vez mais estudado e tem obtido bons resultados é a inferência Bayesiana, tais resultados deve-se, em grande parte, a incorporação de uma informação *a priori* que tende a melhorar o processo de inferência, reduzindo as incertezas a respeito das estimativas dos parâmetros. O objetivo deste trabalho foi avaliar a aplicação da inferência Bayesiana, considerando diferentes estruturas de distribuições *a priori*, no ajuste das distribuições de valores extremos, bem como comparar os resultados com os obtidos via estimação por máxima verossimilhança. Para isso, analisaram-se duas situações práticas: na primeira, analisou-se dados de risco Brasil máximo por semestre, investigando a possível presença de tendência linear nos dados, devido ao fato de serem extraídos de uma séria histórica, por meio da distribuição generalizada de valores extremos e distribuição Gumbel; e na segunda aplicação, utilizou-se a distribuição generalizada de Pareto, com estimação por máxima verossimilhança e inferência Bayesiana, para modelar as perdas referentes a reclamações automobilísticas de companhias seguradoras europeias. Conclui-se, com a realização deste estudo, que a inferência Bayesiana proporciona melhores resultados que os estimadores de máxima verossimilhança quando utilizada para estimação de parâmetros das distribuições generalizada de valores extremos e Gumbel para os dados de risco Brasil e para estimação de parâmetros da distribuição generalizada de Pareto aplicada a dados de resseguros.

Palavras-chave: Inferência Bayesiana, distribuição generalizada de valores extremos, distribuição Gumbel, distribuição generalizada de Pareto.

ANALYSIS OF MAXIMUM BRAZIL RISK AND REINSURANCE: A BAYESIAN APPROACH

Thaís Brenda Martins; Luiz Alberto Beijo; Davi Butturi-Gomes

ABSTRACT: Extreme value theory is often used for the analysis and forecasting of meteorological data, but its application has expanded to several areas where the occurrence of extreme events causes major financial losses, such as finance, stock market and insurance. A method of parameter estimation that has been increasingly studied and has obtained good results is the Bayesian inference. Such results are due in large part to the incorporation of a priori information that tends to improve the process of inference, reducing the uncertainties regarding parameter estimates. The objective of this work is to evaluate the application of Bayesian inference, considering different structures of a priori distributions, in the adjustment of the distributions of extreme values. We also compare the results with those obtained through estimation by maximum likelihood. For this, two practical situations were analyzed: in the first one, Brazil maximum risk data per semester was analyzed, investigating the possible presence of linear trend in the data, due to the fact of being extracted from a historical series, through the generalized distribution of extreme values and Gumbel distribution; and in the second application, the generalized Pareto distribution was used, with estimation by maximum likelihood and Bayesian inference, to model the losses related to auto insurance claims of European insurance companies. We concluded that Bayesian inference provides better results than the maximum likelihood estimators when used for estimation of parameters of generalized distributions of extreme values and Gumbel for Brazil risk data and for estimation of parameters of the distribution generalized Pareto applied to reinsurance data.

Keywords: *Bayesian inference, generalized extreme value distributions, Gumbel distribution, generalized Pareto distribution.*

LISTA DE TABELAS

Tabela 1 –	Estruturas de <i>prioris</i> informativas usadas para os parâmetros da distribuição GEV e Gumbel.	52
Tabela 2 –	Estimativas dos parâmetros, calculados pelo método da máxima verossimilhança (MV) e inferência Bayesiana com <i>priori</i> não informativa (PNI), das distribuições GEV e Gumbel e seus respectivos intervalos de confiança e HPD de 95%.	58
Tabela 3 –	Estimativas, intervalos de confiança (IC) ou HPD de 95%, amplitude dos intervalos (<i>Amp</i>), amplitude média (<i>Amp_m</i>), número de acertos de acurácia (Nº A), DIC (para os modelos Bayesianos) e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com estimação por máxima verossimilhança (EMV) e inferência Bayesiana com <i>priori</i> não informativa (PNI).	60
Tabela 4 –	Estimativas dos parâmetros da distribuição GEV e Gumbel calculadas via método de máxima verossimilhança, para os dados de risco país da Argentina, México e a média dos dois países (MED).	60
Tabela 5 –	Variâncias das estimativas dos parâmetros da GEV e Gumbel calculadas via Método de Máxima Verossimilhança, para os dados de risco Brasil máximo da Argentina, México e da média dos dois países (MED).	61
Tabela 6 –	Estimativas, intervalos HPD de 95%, amplitude dos intervalos (<i>Amp</i>), amplitude média (<i>Amp_m</i>), número de acertos de acurácia (Nº A), DIC e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com estimação por inferência Bayesiana e informação a <i>priori</i> da Argentina (ARG) para as estruturas I e II.	62

Tabela 7 –	Estimativas, intervalos HPD de 95%, amplitude dos intervalos (Amp), amplitude média (Amp_m), número de acertos de acurácia ($N^{\circ} A$), DIC e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com estimação por inferência Bayesiana e informação <i>a priori</i> do México (MEX) para as estruturas I e II.	63
Tabela 8 –	Estimativas, intervalos HPD de 95%, amplitude dos intervalos (Amp), amplitude média (Amp_m), número de acertos de acurácia ($N^{\circ} A$), DIC e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com inferência Bayesiana com informação <i>a priori</i> da média entre o risco país da Argentina e do México (MED) para as estruturas I e II.	64
Tabela 9 –	Modelos selecionados a partir dos critérios para verificação de desempenho dos ajustes.	66
Tabela 10 –	Resultados dos ajustes dos três modelos selecionados à série completa de risco Brasil máximo semestral (2003-2 a 2017-2), sendo as estimativas dos parâmetros, os intervalos HPD de 95% de credibilidade, o DIC de cada modelo, amplitude (Amp) e amplitude média (Amp_m)	67
Tabela 11 –	Estimativa dos parâmetros e os resultados dos critérios de Geweke (CG), de Raftery e Lewis (R-L) e de Heidelberger e Welch (H-W) para os parâmetros do modelo selecionado, Gumbel com informação <i>a priori</i> do risco país do México, estrutura II e variância multiplicada por 4.	68
Tabela 12 –	Níveis de retorno para os tempos de 2 a 10 semestres estimados através do modelo selecionado.	69
Tabela 13 –	Probabilidades do risco Brasil (RB) ser maior que os quantis (Q) definidos, intervalos $HPD_{95\%}$ e critérios de convergência, calculadas a partir do melhor modelo.	69
Tabela 1 –	Estimativas (Est. MV), intervalos de confiança de 95% e amplitude (Amp) dos intervalos para os parâmetros da distribuição generalizada de Pareto com estimação por máxima verossimilhança para os três limites escolhidos.	92

Tabela 2 –	Estimativas (Est. Bayes), intervalo HPD _{95%} e amplitude (<i>Amp</i>) do intervalo para os parâmetros (P) da distribuição generalizada de Pareto com estimação por inferência Bayesiana com <i>priori</i> não informativa para os três limiares com respectivos resultados dos critérios de convergência.	92
Tabela 3 –	Probabilidades de ocorrência (em %) de sinistros acima de <i>x</i> milhões calculadas pelos ajustes da distribuição generalizada de Pareto com estimativas pelo método de máxima verossimilhança (MV) e as probabilidades (em %) e intervalos HPD _{95%} (em %) calculados a partir dos modelos com inferência Bayesiana (Bayes) para os três limiares, <i>u</i> . . .	95
Tabela 1 –	Resultados dos critérios de convergência para os parâmetros do ajuste das distribuições Gumbel e GEV com estimação por inferência Bayesiana: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberg e Welch (H-W).	102
Tabela 2 –	Estimativas dos parâmetros μ e σ , da distribuição Gumbel via inferência Bayesiana com informação <i>a priori</i> do risco país da Argentina para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberg e Welch (H-W).	102
Tabela 3 –	Estimativas dos parâmetros μ e σ , da distribuição Gumbel via inferência Bayesiana com informação <i>a priori</i> do risco país do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberg e Welch (H-W).	103
Tabela 4 –	Estimativas dos parâmetros μ e σ , da distribuição Gumbel via inferência Bayesiana com informação <i>a priori</i> da média do risco país da Argentina e do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberg e Welch (H-W).	103

Tabela 5 –	Estimativas dos parâmetros μ , σ e ξ , da distribuição GEV via inferência Bayesiana com informação <i>a priori</i> do risco país da Argentina para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).	104
Tabela 6 –	Estimativas dos parâmetros μ , σ e ξ , da distribuição GEV via inferência Bayesiana com informação <i>a priori</i> do risco país do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).	105
Tabela 7 –	Estimativas dos parâmetros μ , σ e ξ , da distribuição GEV via inferência Bayesiana com informação <i>a priori</i> da média entre o risco país da Argentina e do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).	106
Tabela 8 –	Resultados dos critérios de convergência para os níveis de retorno para os tempos de retorno de 2 e 3 semestres do ajuste das distribuições Gumbel e GEV com estimação por inferência Bayesiana: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).	107
Tabela 9 –	Estimativas dos níveis de retorno para os tempos de retorno (TR) de 2 e 3 semestres, para os modelos de distribuição Gumbel e GEV via inferência Bayesiana com informação <i>a priori</i> do risco país da Argentina para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).	108

Tabela 10 –	Estimativas dos níveis de retorno para os tempos de retorno (TR) de 2 e 3 semestres, para os modelos de distribuição Gumbel e GEV via inferência Bayesiana com informação <i>a priori</i> do risco país do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).	109
Tabela 11 –	Estimativas dos níveis de retorno para os tempos de retorno (TR) de 2 e 3 semestres, para os modelos de distribuição Gumbel e GEV via inferência Bayesiana com informação <i>a priori</i> da média do risco país da Argentina e do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).	110

LISTA DE FIGURAS

Figura 1 –	Função densidade de probabilidade da distribuição GEV para $\mu = 12$, $\sigma = 3$ e $\xi = -0,45$ (Weibull), $\xi = 0$ (Gumbel) e $\xi = 0,45$ (Fréchet).	24
Figura 2 –	Exemplo de gráfico de vida média residual.	25
Figura 1 –	Fluxograma da metodologia do Capítulo 2.	46
Figura 2 –	Série de máximos semestrais de risco Brasil do segundo semestre de 2003 ao primeiro semestre de 2016.	57
Figura 3 –	Gráficos P-P para as distribuições GEV e Gumbel considerando o método de máxima verossimilhança (a) e inferência Bayesiana (b).	58
Figura 4 –	Gráfico P-P para o modelo selecionado para descrever o risco Brasil máximo semestral.	68
Figura 1 –	Exemplo gráfico de contrato XL por risco.	82
Figura 2 –	Série de sinistros automobilísticos reclamados em companhias europeias de 1988 a 2001.	91
Figura 3 –	Gráfico da média dos excessos e quantis de 70%, 75% e 80%.	91
Figura 4 –	Gráficos P-P para os modelos ajustados com a distribuição GP com máxima verossimilhança (a) e inferência Bayesiana (b) para os três limiares escolhidos.	93
Figura 5 –	Probabilidades de ocorrência de sinistros com severidade maior que x milhões, calculadas pelos modelos com máxima verossimilhança e inferência Bayesiana considerando os três limiares, e os intervalos HPD de 95% para os modelos Bayesianos.	96

SUMÁRIO

	RESUMO	7
	CAPÍTULO 1 - INTRODUÇÃO	17
1	OBJETIVO GERAL	20
1.1	OBJETIVOS ESPECÍFICOS	20
2	REVISÃO DE LITERATURA	21
2.1	TEORIA DE VALORES EXTREMOS	21
2.1.1	Máximos em blocos	22
2.1.2	Excessos sobre um limiar	24
2.2	ESTIMAÇÃO DE PARÂMETROS	26
2.2.1	Método de máxima verossimilhança	27
2.2.2	Inferência Bayesiana	28
2.2.2.1	Informação <i>a priori</i>	29
2.2.2.2	Distribuição <i>a posteriori</i>	30
2.2.2.3	Métodos MCMC e análise de convergência das cadeias	31
3	REFERÊNCIAS	34
	CAPÍTULO 2 - ANÁLISE DO RISCO BRASIL MÁXIMO SEMESTRAL VIA DISTRIBUIÇÕES GUMBEL E GEV: ABORDAGENS CLÁSSICA E BAYESIANA	
	RESUMO	38
1	INTRODUÇÃO	40
2	REVISÃO DE LITERATURA	42
2.1	RISCO BRASIL	42
2.2	DISTRIBUIÇÕES GEV E GUMBEL	43
2.3	ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA	44
2.4	ESTIMAÇÃO POR INFERÊNCIA BAYESIANA	45
3	MATERIAL E MÉTODOS	46
3.1	FLUXOGRAMA DA METODOLOGIA	46
3.2	DADOS	47
3.3	RECURSOS COMPUTACIONAIS	47
3.4	TESTES DE HIPÓTESES PRELIMINARES	48
3.5	ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA	48
3.6	ESTIMAÇÃO POR INFERÊNCIA BAYESIANA	49
3.6.1	Prioris não informativas	50
3.6.2	Prioris informativas	51
3.7	NÍVEIS DE RETORNO	52
3.8	ANÁLISE DE CONVERGÊNCIA DAS CADEIAS GERADAS PELO MÉTODO MCMC	53
3.9	ANÁLISE DE DESEMPENHO DOS MODELOS	54
3.10	CÁLCULO DE PROBABILIDADES E NÍVEIS DE RETORNO PARA TEMPOS FUTUROS	55
4	RESULTADOS E DISCUSSÃO	57
4.1	SELEÇÃO DE MODELOS	57
4.2	ANÁLISE DA SÉRIE COMPLETA DE RISCO BRASIL MÁXIMO SEMESTRAL	67
5	CONCLUSÃO	72

6	REFERÊNCIAS	73
	CAPÍTULO 3 - MODELAGEM DE RESSEGURO VIA DISTRIBUIÇÃO GENERALIZADA DE PARETO UTILIZANDO AS ABORDAGENS CLÁSSICA E BAYESIANA	76
	RESUMO	76
1	INTRODUÇÃO	78
2	REVISÃO DE LITERATURA	80
2.1	RESSEGURO	80
2.2	DISTRIBUIÇÃO GENERALIZADA DE PARETO	83
2.3	MÉTODO DE MÁXIMA VEROSSIMILHANÇA	84
2.4	INFERÊNCIA BAYESIANA	85
3	MATERIAL E MÉTODOS	87
3.1	DADOS	87
3.2	RECURSOS COMPUTACIONAIS	87
3.3	ANÁLISE DE INDEPENDÊNCIA	87
3.4	SELEÇÃO DE LIMAR u DA DISTRIBUIÇÃO GENERALIZADA DE PARETO	88
3.5	ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA	88
3.6	ESTIMAÇÃO POR INFERÊNCIA BAYESIANA	89
3.7	CÁLCULO DE PROBABILIDADE DE OCORRÊNCIA DE VALORES EXTREMOS	89
3.8	ANÁLISE DE CONVERGÊNCIA DAS CADEIAS GERADAS PELO MÉTODO MCMC	90
4	RESULTADOS E DISCUSSÃO	91
5	CONCLUSÃO	97
6	REFERÊNCIAS	98
	CAPÍTULO 4 - CONSIDERAÇÕES FINAIS	101
	APÊNDICES	102

CAPÍTULO 1 - INTRODUÇÃO

Eventos extremos ou raros são aqueles que assumem valores afastados da média, podendo ser máximos ou mínimos, e que possuem baixa probabilidade de ocorrência. Valores fora do comum ou que excedem um determinado valor podem ser considerados eventos extremos.

Esses eventos podem ocorrer em ambientes naturais, tecnológicos, sociais, econômicos e financeiros. Geralmente são associados a desastres naturais, como terremotos e furacões, a eventos meteorológicos, como velocidade do vento e precipitações, mas também estão presentes em áreas como finanças, seguros e teoria do risco. Eventos raros ou extremos estão frequentemente relacionados à perdas de vida e/ou financeiras (BORGES, 2013). Logo, é de vital importância a predição probabilística desses eventos para o planejamento das atividades sujeitas a seus efeitos adversos.

O estudo da ocorrência destes eventos é feito pela teoria de valores extremos (TVE), que é um ramo da probabilidade que estuda os eventos extremos (BAUTISTA; ZOCCHI; ANGELOCCI, 2004). Valores extremos podem ser os máximos ou os mínimos de um conjunto de variáveis aleatórias em blocos ou excessos acima (ou abaixo) de um determinado limiar alto (ou baixo). A distribuição de probabilidade dos extremos em blocos é dada por um dos três tipos de distribuições de valores extremos: Gumbel ou tipo I, Fréchet ou tipo II e Weibull ou tipo III, ou pela distribuição generalizada de Valores Extremos (*generalized extreme values* - GEV), que é a junção dos três tipos em uma única forma paramétrica. Quando os extremos são obtidos através dos excessos sobre um limiar, os dados seguem uma distribuição generalizada de Pareto (*generalized Pareto* - GP). As características da distribuição desses extremos são determinadas pelas caudas extremas (inferior ou superior) da distribuição dos dados.

As distribuições de extremos têm sido aplicadas em diversos campos, por exemplo, Sangiolo (2008) usou distribuições de extremos para estudar a ocorrência de precipitação máxima, temperatura máxima e mínima e velocidade do vento em Piracicaba (SP); Bautista, Zocchi e Angelocci (2004) também estudaram velocidade máxima do vento em Piracicaba (SP) por meio da distribuição GEV; Beijo, Muniz e Castro Neto (2005) utilizaram a distribuição Gumbel para calcular níveis de retorno para precipitação máxima em Lavras (MG); Silva (2006) aplicou a TVE em gerenciamento de risco operacional; Arraes e Rocha (2006) estimaram as perdas máximas esperadas em séries financeiras utilizando TVE na metodologia e Borges (2013) aplicou

TVE a redes de dados complexas.

O comportamento de uma distribuição de probabilidade é determinado pelas suas propriedades que podemos chamar de parâmetros. Com isso, é preciso estimar os valores desses parâmetros, sendo esse processo denominado estimação. De acordo com Coles (2001), existem vários métodos para estimação de parâmetros, sendo que um desses métodos, que é amplamente utilizado, é o método da máxima verossimilhança. Este método utiliza a abordagem clássica da Estatística. Outra abordagem que vem sendo muito utilizada para estimação de parâmetros é a inferência Bayesiana. Essa abordagem possui vantagens, como o fato de permitir a incorporação de informações passadas à distribuição (*a priori*) que melhoram o processo de inferência (PAULINO; TURKMAN; MURTEIRA, 2003; BEIJO; VIVANCO; MUNIZ, 2009).

O uso da inferência Bayesiana para a estimação de parâmetros das distribuições de extremos e a comparação com métodos clássicos têm sido abordados em alguns trabalhos como: Smith e Naylor (1987) que desenvolveram e compararam estimadores de máxima verossimilhança e Bayesianos para a distribuição Weibull; Cooley, Nychka e Naveau (2007) que usaram a modelagem Bayesiana de distribuições de extremos para produzir mapas de níveis de retorno de precipitação para uma região do Colorado/EUA; Beijo, Vivanco e Muniz (2009) que estudaram o tempo de retorno das precipitações pluviométricas máximas em Jaboticabal (SP) e Portes (2017) que modelou os níveis máximos de IPCA.

Um dos objetivos deste trabalho foi comparar o uso da inferência Bayesiana com o método de máxima verossimilhança quando utilizadas distribuições de extremos, por isso foram realizadas duas aplicações: a primeira, em dados de risco Brasil e a segunda, em dados de sinistros de resseguros.

O risco país é uma medida utilizada para determinar o risco de investimento em um determinado país. É calculado para países emergentes, sendo que através dele investidores conseguem analisar o risco de se investir em um determinado país. Quanto maior o risco país, menos investimentos estrangeiros o país pode receber. O Brasil passou por períodos com grandes elevações no seu risco país, como por exemplo, de acordo com Leichsenring (2004), em setembro de 1998, com as eleições presidenciais brasileiras e a falência do LTCM (*Long Term Capital Management*), o risco Brasil passou do patamar de 380 para 1.700 pontos-base. Como altos níveis de risco Brasil ocasionam retração de investimentos, que influenciam diretamente no desenvolvimento econômico do país e assim na população, é necessário realizar previsões probabilísticas a respeito do nível de tal risco.

O resseguro é, informalmente, o seguro das seguradoras. É um contrato, geralmente entre pessoas jurídicas, em que a seguradora transfere o risco de uma apólice para uma outra seguradora (resseguradora) em troca de um prêmio. Então, o ressegurador assume o compromisso de indenizar a seguradora cedente de parte das consequências financeiras de certas exposições à prejuízos (AYRES, 2009). O resseguro é usado contra consequências financeiras devido à ocorrência de eventos catastróficos que ocasionam em muitos sinistros, como incêndios, fenômenos da natureza, desastres de avião, entre outros. Um montante alto de sinistros pode levar à diminuição dos resultados da seguradora e ameaçar a solvência da empresa.

1 OBJETIVO GERAL

O objetivo geral deste trabalho foi avaliar os resultados obtidos no uso da inferência Bayesiana comparando com método de máxima verossimilhança na análise dos dados de risco Brasil por meio das distribuições Gumbel e generalizada de valores extremos e modelagem dos dados de resseguro por meio da distribuição generalizada de Pareto.

1.1 OBJETIVOS ESPECÍFICOS

Considerando-se os dados de risco Brasil, especificamente, objetivou-se:

- Analisar dados de risco Brasil via distribuições generalizada de valores extremos e Gumbel;
- Averiguar a presença de tendência linear na série de risco Brasil máximo semestral;
- Comparar os resultados obtidos pelos métodos de estimação de máxima verossimilhança e inferência Bayesiana;
- Avaliar, em inferência Bayesiana, diferentes estruturas e informações *a priori*;
- Determinar por meio de critérios para seleção de modelos qual o melhor para descrever probabilisticamente os dados de risco Brasil máximo semestral;
- Realizar previsões de níveis de retorno de risco Brasil e calcular probabilidades de valores extremos de risco Brasil.

Considerando-se os dados de resseguro, especificamente, objetivou-se:

- Modelar a série de reivindicações de resseguro utilizando a distribuição generalizada de Pareto;
- Comparar os resultados obtidos pelos métodos de estimação de máxima verossimilhança e inferência Bayesiana;
- Calcular probabilidades de valores extremos de sinistros, ou seja, sinistros com alta severidade.

2 REVISÃO DE LITERATURA

Nesta seção são apresentados conceitos e métodos que foram utilizados para o desenvolvimento deste trabalho.

Eventos extremos (ou raros) são aqueles que possuem baixa probabilidade de ocorrência, mas que causam grandes danos econômicos, sociais e financeiros. Geralmente, são associados a desastres naturais ou fenômenos meteorológicos como terremotos e furacões, chuvas intensas e velocidade do vento. Também podem estar presentes no cenário financeiro como, por exemplo, as crises econômicas, que levam a altas taxas de juros e desequilíbrio no mercado econômico-financeiro. Portanto, a ocorrência de eventos extremos em qualquer área pode causar danos e prejuízos. Com isso, surge a importância e necessidade de modelar e realizar previsões probabilísticas desses eventos. Uma metodologia utilizada para descrever o comportamento de eventos extremos é a teoria de valores extremos (BEIJO; AVELAR, 2010).

2.1 TEORIA DE VALORES EXTREMOS

Fisher e Tippett (1928) desenvolveram os fundamentos da teoria de valores extremos ao definirem os três tipos de distribuições assintóticas de valores extremos: tipo I (ou Gumbel), tipo II (ou Fréchet) e tipo III (ou Weibull). Jenkinson (1955) propôs uma expressão que representa em uma única forma paramétrica as três distribuições, denominada de distribuição generalizada de valores extremos (GEV).

A distribuição GEV tem sido utilizada em diversos trabalhos para estudo de ocorrências extremas em fenômenos meteorológicos. Bautista, Zocchi e Angelocci (2004) usaram a distribuição GEV para modelar velocidade máxima de ventos em Piracicaba (SP). Beijo, Vivanco e Muniz (2009) usaram a distribuição GEV e inferência Bayesiana para prever precipitações pluviométricas máximas em Jaboticabal (SP). Há também trabalhos nas áreas de finanças e ciências atuariais. Arraes e Rocha (2006) estudaram perdas extremas em mercados de risco utilizando distribuição GEV e Campos (2003) analisou eventos extremos no mercado de derivativos brasileiro usando distribuição GEV.

De acordo com Coles (2001), a série de máximos a ser estudada pode ser extraída da série de dados de duas formas distintas: selecionando os máximos de cada período de uma série

de dados, que é chamado de máximos em blocos, ou por meio dos valores que ultrapassam um determinado limiar alto, que é chamado de excessos sobre um limiar. O modo como os máximos são selecionados na amostra influencia na distribuição de probabilidade da série de máximos.

2.1.1 Máximos em blocos

De acordo com Coles (2001), seja uma amostra X_1, \dots, X_n de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) com função de distribuição $F(x)$. A variável aleatória M_n é definida como o máximo das n variáveis aleatórias X_i , isto é, $M_n = \max(X_1, \dots, X_n)$.

A distribuição de M_n é dada por:

$$\begin{aligned} F_{M_n} &= P(M_n \leq z) \\ &= P(X_1 \leq z, X_2 \leq z, \dots, X_n \leq z) \\ &= P(X_1 \leq z) \times P(X_2 \leq z) \times \dots \times P(X_n \leq z) \\ &= [F(z)]^n. \end{aligned}$$

No entanto, quando $n \rightarrow \infty$, a função F_{M_n} pode ser degenerada. Contudo, o teorema de Fisher e Tippett (1928) garante que através de uma transformação linear de M_n a partir de constantes de normalização, $a_n > 0$ e $b_n \in \mathbb{R}$, tem-se uma distribuição assintótica não degenerada, G , tal que

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} G,$$

e segundo os autores essa função de distribuição não degenerada, G , segue um dos três tipos de distribuições de valores extremos, que são:

Distribuição Gumbel (Tipo I)

$$G_I(x) = \exp(-\exp^{-x}) \quad \text{para } -\infty < x < \infty. \quad (2.1)$$

Distribuição Fréchet (Tipo II)

$$G_{II}(x) = \begin{cases} \exp(-x^{-\alpha}), & \text{para } \alpha > 0 \quad \text{e para } x > 0 \\ 0, & \text{para } x \leq 0. \end{cases} \quad (2.2)$$

Distribuição Weibull (Tipo III)

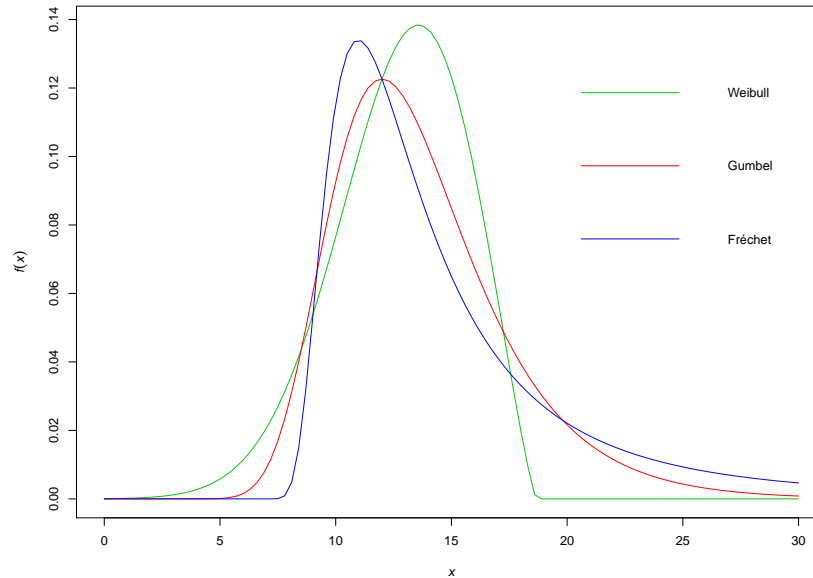
$$G_{III}(x) = \begin{cases} \exp(-(-x)^\alpha), & \text{para } \alpha > 0 \quad \text{e para } x \leq 0 \\ 1, & \text{para } x > 0. \end{cases} \quad (2.3)$$

Jenkinson (1955), a partir dos estudos feitos por Gnedenko (1943), propôs a distribuição generalizada de valores extremos (GEV) que consiste na representação dos três tipos de distribuições de valores extremos em uma única forma paramétrica:

$$G_\xi(x) = \begin{cases} \exp \left\{ - \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} & \xi \neq 0 \\ \exp \left\{ -\exp \left[- \left(\frac{x-\mu}{\sigma} \right) \right] \right\} & \xi = 0, \end{cases} \quad (2.4)$$

definida em $-\infty < x < \mu - \frac{\sigma}{\xi}$, para $\xi < 0$ e $\mu - \frac{\sigma}{\xi} < x < \infty$, para $\xi > 0$. De acordo com Bautista, Zocchi e Angelocci (2004), as distribuições de extremos Fréchet e Weibull, correspondem aos casos em que $\xi < 0$ e $\xi > 0$, respectivamente. E quando ξ tende a zero, $G_\xi(x)$ corresponde à função de distribuição Gumbel, que é definida em $-\infty < x < \infty$. Logo, a partir da distribuição GEV é possível obter as três distribuições de extremos, conforme são apresentadas na Figura 1.

Figura 1 – Função densidade de probabilidade da distribuição GEV para $\mu = 12$, $\sigma = 3$ e $\xi = -0,45$ (Weibull), $\xi = 0$ (Gumbel) e $\xi = 0,45$ (Fréchet).



Fonte: Elaborada pela autora

Usando a função densidade de probabilidade da distribuição GEV e fixando os parâmetros de posição $\mu = 12$ e escala $\sigma = 3$ e variando o parâmetro de forma em $\xi = -0,45$, $\xi = 0$ e $\xi = 0,45$, obtemos as funções densidade de probabilidade das distribuições Weibull, Gumbel e Fréchet, respectivamente, conforme apresenta a Figura 1.

2.1.2 Excessos sobre um limiar

Seja uma amostra X_1, \dots, X_n de variáveis aleatórias i.i.d. com função de distribuição $F(x)$, a diferença entre os valores X_i e um limiar u , $Y = \{X_i - u | X_i > u\}$, de acordo com Gomes e Guillou (2014), é denominada de excesso sobre um limiar e segue uma distribuição generalizada de Pareto (GP), dada por:

$$F(y, \sigma, \xi) = \begin{cases} 1 - \left[1 + \xi \left(\frac{y}{\sigma}\right)\right]^{-1/\xi}, & \sigma > 0 \quad 1 + \xi \left(\frac{y}{\sigma}\right) \geq 0, \quad \text{se } \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\sigma}\right), & y \in \mathbb{R}, \quad \text{se } \xi = 0, \end{cases} \quad (2.5)$$

em que σ e ξ são os parâmetros de escala e forma, respectivamente.

Uma dificuldade citada por Coles (2001), Mendes (2004) e Silva (2008) quando se usa a GPD é quanto à seleção do limiar u , devido ao fato de não ser trivial encontrar um valor u ótimo. Segundo os autores, um limiar muito alto pode resultar em um número pequeno de observações na cauda, levando a uma maior variabilidade dos estimadores. Por outro lado, um limiar não suficientemente alto não irá satisfazer as suposições teóricas e pode resultar em estimativas distorcidas.

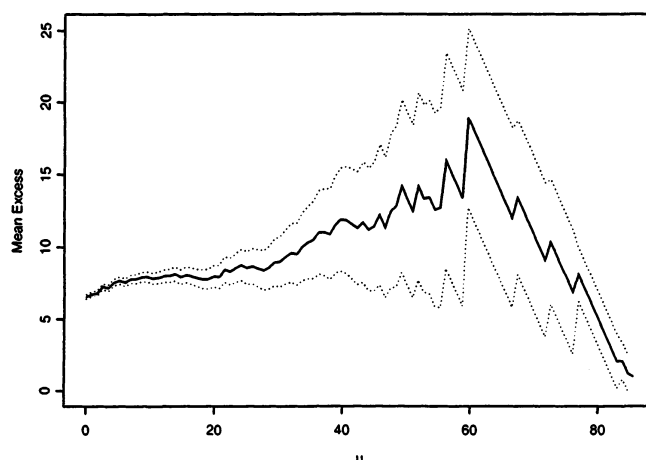
Os mesmos autores citam métodos utilizados na escolha do limiar. Mendes (2004) cita o uso das estatísticas de ordem para escolha do limiar, as funções média e mediana de excessos empíricas junto com a análise gráfica exploratória. Coles (2001) destaca que a prática padrão é adotar um limite mais baixo possível que forneça uma aproximação razoável, apresentando dois métodos: a técnica exploratória antes da estimação do modelo, por meio do gráfico de vida média residual, e a avaliação da estabilidade das estimativas de parâmetros com base em modelos ajustados com diferentes limiares. Silva (2008) também utiliza em seu estudo a análise do gráfico de vida média residual ou média de excessos.

O gráfico de vida média residual, de acordo com Coles (2001), é dado por:

$$\left\{ \left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_i - u) : u < x_{max} \right) \right\},$$

em que x_1, x_2, \dots, x_{n_u} consistem nas observações que excedem u e x_{max} é o valor mais elevado das observações. Na Figura 2 é apresentado um exemplo de gráfico de vida média residual.

Figura 2 – Exemplo de gráfico de vida média residual.



Fonte: Coles (2001)

De acordo com Coles (2001), a curva de vida média residual deve ser aproximadamente linear em u , o limiar é definido sobre o valor de u que diferencia a curva em linear e não linear

(ou aproximadamente linear). Segundo o mesmo, no exemplo da Figura 2, pode-se observar uma estabilidade entre $u = 30$ e $u = 60$ e após isso decai bruscamente, mas o número de observações acima de 60 seria muito baixo para realizar inferências significativas. Por isso, é melhor concluir que há alguma evidência de linearidade acima de $u = 30$ e trabalhar inicialmente com um limite definido em $u = 30$.

2.2 ESTIMAÇÃO DE PARÂMETROS

Uma variável aleatória pode ser descrita por uma distribuição de probabilidade, que por sua vez é descrita por seus parâmetros populacionais. As distribuições são usadas para modelar diversas variáveis, sendo que as distribuições de valores extremos modelam a ocorrência de eventos extremos. Geralmente, os parâmetros populacionais da distribuição não são conhecidos, por isso é necessário estimá-los (CASELLA; BERGER, 2010). A distribuição generalizada de valores extremos apresenta três parâmetros desconhecidos: os parâmetros μ e σ que correspondem, respectivamente, aos parâmetros de posição e de escala e o parâmetro ξ que é o parâmetro de forma que determina a natureza das caudas da distribuição (BEIJO; AVELAR, 2010).

Segundo Coles (2001), existem muitos métodos para a estimação dos parâmetros para modelos de valores extremos, entre eles as técnicas gráficas, estimadores baseados em momentos, método dos momentos ponderados pela probabilidade, técnicas baseadas em verossimilhança e métodos Bayesianos.

O método da máxima verossimilhança é a técnica mais popular para derivar estimadores (CASELLA; BERGER, 2010). De acordo com Naguettini e Pinto (2007), também é considerado o método mais eficiente de estimação, pois produz estimadores de menor variância. Mas, em alguns casos, essa eficiência é apenas assintótica, fazendo com que a aplicação para amostras pequenas forneça estimadores de qualidade igual ou inferior à de outros métodos (SMITH, 1985). Por outro lado, com a inferência Bayesiana, tem-se a possibilidade de introduzir uma informação *a priori*, que melhora o processo de inferência, tornando possível o uso de pequenas amostras (PAULINO; TURKMAN; MURTEIRA, 2003).

2.2.1 Método de máxima verossimilhança

No método de máxima verossimilhança, busca-se maximizar a probabilidade de um parâmetro representar uma população por meio da densidade conjunta dos elementos amostrais que é uma função dos parâmetros, chamada de função de verossimilhança (BOLFARINE; SANDOVAL, 2001).

Seja uma amostra aleatória $\mathbf{x} = (x_1, x_2, \dots, x_n)$ independente e identicamente distribuída (i.i.d.), com função densidade de probabilidade $f(x_i, \theta)$, então a função de verossimilhança é dada por:

$$L(\theta|\mathbf{x}) = \prod_{i=1}^n f(x_i, \theta). \quad (2.6)$$

Para encontrar os pontos de máximo (estimativas para os parâmetros que maximizam a função de verossimilhança) é preciso derivar essa função em relação a cada parâmetro e igualar cada derivada a zero. A fim de facilitar esse processo, é aplicado o logaritmo natural nessa função antes de derivá-la:

$$\begin{aligned} l(\theta|\mathbf{x}) &= \ln[L(\theta|\mathbf{x})] \\ &= \ln[f(x_1, \theta) \times f(x_2, \theta) \times \dots \times f(x_n, \theta)] \\ &= \ln[f(x_1, \theta)] + \ln[f(x_2, \theta)] + \dots + \ln[f(x_n, \theta)] \\ &= \sum_{i=1}^n \ln[f(x_i, \theta)]. \end{aligned} \quad (2.7)$$

As equações resultantes das derivadas podem não possuir solução analítica, por isso utilizam-se métodos iterativos. No pacote *optim* do R (R CORE TEAM, 2017) existem métodos iterativos implementados, como os algoritmos Nelder-Mead (NM) e Gradiente Conjugado (CG), que são baseados nos métodos propostos por Nelder e Mead (1965) e Fletcher e Reeves (1964), respectivamente (MARTINS; RIBEIRO JUNIOR; BONAT, 2009).

Segundo Bolfarine e Sandoval (2001), os estimadores de máxima verossimilhança apresentam as propriedades de invariância e eficiência. A propriedade de invariância permite obter estimadores para funções dos parâmetros. De acordo com Mood, Graybill e Boes (1950), seja uma variável aleatória X com função densidade (ou de probabilidade) $f(x_i, \theta)$ e uma amostra

aleatória dessa variável. Se $\hat{\theta}$ é um estimador de máxima verossimilhança de θ e se $g(\cdot)$ é uma função real inversível definida em Θ , então $g(\hat{\theta})$ é um estimador de máxima verossimilhança de $g(\theta)$.

A propriedade de eficiência é válida quando o tamanho da amostra é grande. Então tem-se que:

$$\sqrt{n}(\hat{\theta} - \theta) \stackrel{a}{\sim} N(0, I_F(\theta)^{-1})$$

e

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \stackrel{a}{\sim} N(0, (g'(\theta))^2 I_F(\theta)^{-1}),$$

em que $\stackrel{a}{\sim}$ significa distribuição assintótica e $I_F(\theta)$ é a matriz de informação de Fisher. Para grandes amostras, de acordo com Bolfarine e Sandoval (2001), os estimadores de θ e $g(\theta)$, pelo método da máxima verossimilhança, são aproximadamente não viciados, cujas variâncias coincidem com os limites inferiores das variâncias dos estimadores não viciados de θ e $g(\theta)$. Com isso, em grandes amostras, o estimador de máxima verossimilhança é eficiente.

De acordo com Coles e Dixon (1999), os estimadores de máxima verossimilhança não mantêm as boas propriedades em pequenas amostras. Quando não há disponibilidade de amostras grandes é necessário usar métodos que possibilitem o uso desses dados. Beijo, Vivanco e Muniz (2009) citam como uma alternativa ao método de máxima verossimilhança na análise de valores extremos, a inferência Bayesiana, uma vez que permite incorporar informações passadas, diminuindo assim as incertezas quanto aos parâmetros a serem estudados.

2.2.2 Inferência Bayesiana

A abordagem Bayesiana surgiu na década de 30 com base em estudos teóricos que exigiam resoluções de integrais complexas. Por isso, a utilização de métodos Bayesianos ficou limitada até a década de 90, quando foi solucionado o problema dos cálculos integrais (BEIJO, 2006).

O uso da inferência Bayesiana na modelagem de valores extremos é bem recente. Havia poucos trabalhos até o ano de 1995 que ligavam a teoria de valores extremos e inferência Bayesiana, sendo quase sempre aplicados à distribuição Weibull (COLES; POWELL, 1996).

Segundo Beijo (2006), um dos primeiros estudos a aplicar inferência Bayesiana junto à distribuição GEV foi realizado por Coles e Tawn (1996) que se basearam em conhecimentos de especialistas da área para formular informações *a priori* como base para uma análise Bayesiana de precipitações extremas em uma estação do sudoeste da Inglaterra.

O processo de inferência na abordagem Bayesiana é todo baseado na distribuição *a posteriori* do parâmetro θ (BROEMELING, 1985). Essa distribuição é resultado da combinação entre a informação obtida dos dados amostrais por meio da função de verossimilhança e a informação *a priori* sobre o parâmetro θ , conhecida como distribuição *a priori*. Essa combinação de informações é feita por meio do teorema de Bayes (CASELLA; BERGER, 2010).

De acordo com Nogueira et al. (2003), a diferença formal entre a inferência frequentista e Bayesiana é que, na abordagem Bayesiana, a quantidade de interesse θ é variável aleatória e, por isso, é descrita por uma distribuição de probabilidade e na inferência frequentista os parâmetros são valores fixos. A distribuição *a priori* trata-se de uma distribuição subjetiva, baseada na crença do experimentador e definida antes da observação dos dados (PAULINO; TURKMAN; MURTEIRA, 2003).

De acordo com o teorema de Bayes, seja $p(\theta)$ a distribuição *a priori* e $L(\theta|\mathbf{x})$ a função de verossimilhança, então a função de distribuição *a posteriori* é dada por:

$$p(\theta|\mathbf{x}) = \frac{L(\theta|\mathbf{x})p(\theta)}{\int L(\theta|\mathbf{x})p(\theta)d\theta} \quad (2.8)$$

em que $\mathbf{x} = (x_1, x_2, \dots, x_n)$. No denominador, resolvendo a integral, teremos a marginal de \mathbf{X} , sendo que tal denominador não depende de θ e funciona como uma constante normalizadora. Dessa forma, podemos reescrever o teorema de Bayes da seguinte maneira:

$$p(\theta|\mathbf{x}) \propto L(\theta|\mathbf{x})p(\theta). \quad (2.9)$$

em que \propto significa proporcionalidade, ou seja a distribuição *a posteriori* é proporcional ao produto da função de verossimilhança e da distribuição *a priori*.

2.2.2.1 Informação *a priori*

A informação *a priori* sobre uma quantidade de interesse θ é a característica essencial da inferência Bayesiana, de acordo com Gelman et al. (2014). Essa informação é a representação probabilística do conhecimento que se tem a respeito de θ antes da observação dos dados, isto é, uma distribuição de probabilidade para o parâmetro, θ , que é considerado uma variável aleatória na abordagem Bayesiana.

Essa informação *a priori* que se pretende incorporar na análise é possuída por um especialista do problema que se estuda, podendo ser um investigador, estatístico ou outrem, e contém elementos subjetivos. Esses elementos podem ser radicados em fontes objetivas através de dados históricos do problema ou de problemas análogos (PAULINO; TURKMAN; MURTEIRA, 2003).

Pode-se classificar a distribuição *a priori* quanto ao nível de informação como *priori* informativa e *priori* não informativa. O uso da *priori* informativa é feito quando o pesquisador possui um conhecimento ou informação prévia do que está sendo estudado. Quando o pesquisador possui pouca ou nenhuma informação para incorporar à análise é usada uma *priori* não informativa.

De acordo com Box e Tiao (1992), a *priori* não informativa não representa, necessariamente, o desconhecimento do pesquisador sobre o experimento. Esse tipo de *priori* pode ser usada de forma a viabilizar o uso da inferência Bayesiana, ou quando se espera que a informação dos dados seja dominante.

Para a informação *a priori*, deve-se escolher a sua forma funcional, ou seja, sua distribuição de probabilidade e o valor dos seus parâmetros, que são chamados de hiperparâmetros.

2.2.2.2 Distribuição *a posteriori*

A distribuição *a posteriori* é obtida por meio do teorema de Bayes e, conforme a Equação 2.9, é proporcional ao produto entre a distribuição *a priori* e a função de verossimilhança dos dados. Sendo assim, a *posteriori* contém toda a informação probabilística a respeito dos parâmetros, sendo que toda a inferência sobre eles é baseada nessa distribuição (MARTINS FILHO et al., 2008).

O processo de inferência sobre um parâmetro necessita da sua marginal *a posteriori*, que é obtida pela integral da distribuição conjunta *a posteriori* em relação aos demais parâmetros. A partir da distribuição marginal, é possível calcular as estimativas pontuais que são os valores centrais da distribuição (PAULINO; TURKMAN; MURTEIRA, 2003).

Além das estimativas pontuais, é possível calcular estimativas intervalares que são chamadas de regiões ou intervalos de credibilidade. De acordo com Paulino, Turkman e Murteira (2003), entre os infinitos intervalos de credibilidade possíveis com o mesmo grau de credibilidade, aquele com probabilidade *a posteriori* máxima é chamado de intervalo HPD (*Highest Posterior Density*). E segundo Brunello e Nakano (2015), o intervalo HPD é o que apresenta a menor amplitude dentre todos os possíveis intervalos de credibilidade.

A solução analítica das integrais para encontrar as marginais *a posteriori*, de acordo com Gamerman e Lopes (2006), geralmente é complexa, o que torna necessário a utilização de técnicas iterativas para solucioná-las, como os algoritmos Amostrador de Gibbs e Metropolis-Hastings da técnica Monte Carlo via Cadeias de Markov (*Monte Carlo Markov Chain - MCMC*).

2.2.2.3 Métodos MCMC e análise de convergência das cadeias

A inferência Bayesiana para modelos complexos ficou algumas décadas limitada devido à dificuldade de se obterem analiticamente as integrais complexas da distribuição *a posteriori* para se obterem as distribuições marginais. A partir dos avanços na área de computação, foram surgindo métodos para obter soluções aproximadas com menores erros. As técnicas Monte Carlo via Cadeias de Markov (MCMC) têm sido usadas para obter as distribuições marginais de forma iterativa, quando a forma analítica é inviável (NOGUEIRA; SÁFADI; FERREIRA, 2004).

Segundo Martins Filho et al. (2008), os algoritmos MCMC mais usados são o Amostrador de Gibbs (GEMAN; GEMAN, 1984; GELFAND; SMITH, 1990; CASELLA; GEORGE, 1992) e o Metropolis-Hastings (METROPOLIS et al., 1953; HASTINGS, 1970). O algoritmo usado neste trabalho é o Metropolis-Hastings. Segundo Gazzinelli (2013), o algoritmo Metropolis-Hastings é de grande importância para inferência Bayesiana, uma vez que garantem a convergência da cadeia para a distribuição de equilíbrio: *a posteriori*.

De acordo com Albert (2009), supondo que se deseja simular valores a partir de uma

densidade *a posteriori* $g(\theta|y)$ (ou simplesmente $g(\theta)$ para simplificar a notação). O algoritmo Metropolis-Hastings começa com um valor inicial θ^0 e especifica uma regra para simular o t -ésimo valor, θ^t , dado o valor $(t-1)$ -ésimo, θ^{t-1} . Essa regra consiste em uma densidade proposta que simula um valor candidato θ^* e o cálculo de uma probabilidade de aceitação P que indica a probabilidade do valor candidato ser aceito como o próximo valor da sequência. O algoritmo pode ser descrito conforme os passos a seguir:

- Um valor candidato θ^* é simulado a partir de uma densidade proposta $p(\theta^*|\theta^{t-1})$;
- Calcula-se a razão $R = \frac{g(\theta^*)p(\theta^{t-1}|\theta^*)}{g(\theta^{t-1})p(\theta^*|\theta^{t-1})}$;
- Calcula-se a probabilidade de aceitação $P = \min\{R, 1\}$;
- Amostra-se um valor θ^t tal que $\theta^t = \theta^*$ com probabilidade P . Caso contrário, $\theta^t = \theta^{t-1}$.

Sob algumas condições de regularidade, facilmente satisfeitas na densidade proposta $p(\theta^*|\theta^{t-1})$, a sequência de valores simulados $\theta^1, \theta^2, \dots$, conhecida como cadeia de Markov, irá convergir para uma variável aleatória que é distribuída de acordo com a distribuição *a posteriori* $g(\theta)$ (ALBERT, 2009).

As cadeias de Markov geradas devem convergir para uma distribuição estacionária. Na teoria, isso acontece apenas quando o número de iterações tende ao infinito. Na prática, admite-se convergência quando a série se torna estacionária, o que indica que as condicionais completas estão suficientemente próximas das distribuições marginais (NOGUEIRA; SÁFADI; FERREIRA, 2004). Problemas como determinar o número necessário de iterações para garantir a convergência da cadeia para o estado de equilíbrio podem surgir durante o uso de técnicas MCMC, assim como a possibilidade das iterações iniciais da amostra serem influenciadas pelos valores iniciais dos parâmetros e o fato das sequências de valores serem autocorrelacionadas (COWLES; CARLIN, 1996).

Alguns critérios de diagnóstico são usados para monitorar a convergência das cadeias. Entre os mais usados estão o critério de Geweke (1992), o critério de Raftery e Lewis (1992) e o critério de Heidelberger e Welch (1983), que serão usados neste trabalho para verificar a convergência das cadeias geradas para uma distribuição estacionária.

De acordo com Gazzinelli (2013), o critério de Geweke consiste em dividir a cadeia de Markov em duas partes e testar se os valores da primeira parte são iguais em média aos valores da segunda parte, pois, se a distribuição é estacionária, espera-se que a média da primeira parte

seja igual à da segunda. O critério de Heidelberger e Welch avalia a estacionariedade da cadeia e se o tamanho da amostra é adequado para estimar a média com precisão. O critério de Raftery e Lewis avalia a autocorrelação amostral por meio do fator de dependência calculado, determina o período de *burn-in* (número de iterações iniciais a serem descartadas) e a distância mínima de uma interação a outra para diminuir a autocorrelação amostral. Autores como Gazzinelli (2013) e Nogueira, Sáfadi e Ferreira (2004) afirmam que com um fator menor que 5 a convergência da série é obtida.

3 REFERÊNCIAS

- ALBERT, J. **Bayesian computation with R**. New York: Springer Science & Business Media, 2009.
- ARRAES, R. A.; ROCHA, A. S. Perdas extremas em mercados de risco. **Revista Contabilidade e Finanças**, p. 22-34, 2006.
- AYRES, T. R. **A necessidade de capital de uma seguradora em função do resseguro**. 2009. 163f. Dissertação (Mestrado em Gestão de Riscos Financeiros e Atuariais) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.
- BAUTISTA, E. A. L.; ZOCCHI, S. S.; ANGELOCCI, L. R. A distribuição generalizada de valores extremos aplicada ao ajuste dos dados de velocidade máxima do vento em Piracicaba, São Paulo, Brasil. **Revista de Matemática e Estatística**, v. 22, p. 95–111, 2004.
- BEIJO, L. A. **Construção de uma priori para os parâmetros do modelo de valores extremos generalizados baseada em quantis com distribuição Gumbel**. 2006. 98f. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2006.
- 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**, v. 1, 2010.
- BEIJO, L. A., MUNIZ, J. A., CASTRO NETO, P. Tempo de retorno das precipitações máximas em Lavras (MG) pela distribuição de valores extremos do tipo I. **Ciência e agrotecnologia**, v. 29, n. 3, p. 657-667, 2005.
- BEIJO, L. A.; VIVANCO, M. J. F.; MUNIZ, J. A. Análise Bayesiana no estudo do tempo de retorno das precipitações pluviais máximas em Jaboticabal (SP). **Ciência e Agrotecnologia**, v. 33, p. 261–270, 2009.
- BOLFARINE, H.; SANDOVAL, M. C. **Introdução à inferência estatística**. São Paulo: SBM, 2001.
- BORGES, R. R. **Teoria de Valores Extremos Aplicada a Redes Complexas**. Ponta Grossa: UEPG, 2013. 116f. Dissertação (Mestrado em Ciências - Área de Concentração: Física) - Programa de Pós-Graduação em Ciência, Universidade Estadual de Ponta Grossa, Ponta Grossa, 2013.
- BOX, G. E. P.; TIAO, G. C. **Bayesian inference in statistical analysis**. New York: Wiley, 1992.
- BROEMELING, L. D. **Bayesian analysis of linear models**. New York: M. Dekker, 1985.
- BRUNELLO, G. H. V.; NAKANO, E. Y. Inferência bayesiana no modelo weibull discreto em dados com presença de censura. **TEMA (São Carlos)**, SciELO Brasil, v. 16, n. 2, p. 97–110, 2015.

- CAMPOS, A. L. S. **Eventos extremos no mercado de derivativos acionários brasileiro**. 2003. 110f. Monografia (Bacharel em Engenharia de Produção) - EPUSP, São Paulo, 2003.
- CASELLA, G.; BERGER, R. L. **Inferência Estatística**. São Paulo: Cengage Learning, 2010.
- CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. **The American Statistician**, Taylor & Francis, v. 46, n. 3, p. 167–174, 1992.
- COLES, S. G. **An introduction to statistical modeling of extreme values**. London: Springer, 2001.
- COLES, S. G.; DIXON, M. J. Likelihood-based inference for extreme value models. **Extremes**, Springer, 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 Statistical Review**, v. 64, n. 1, p. 119–136, 1996.
- COLES, S. G.; TAWN, J. A. A bayesian analysis of extreme rainfall data. **Applied statistics**, Oxford, v. 45, n. 4, p. 463–478, 1996.
- COOLEY, D.; NYCHKA, D.; NAVEAU, P. Bayesian spatial modeling of extreme precipitation return levels. **Journal of the American Statistical Association**, Taylor & Francis, v. 102, n. 479, p. 824–840, 2007.
- COWLES, M. K.; CARLIN, B. P. Markov chain monte carlo convergence diagnostics: a comparative review. **Journal of the American Statistical Association**, Taylor & Francis, v. 91, n. 434, p. 883–904, 1996.
- FLETCHER, R; REEVES, C. M. Function minimization by conjugate gradients. **The computer journal**, v. 7, n. 2, p. 149-154, 1964.
- FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In: **Mathematical Proceedings of the Cambridge Philosophical Society**. Cambridge University Press, v. 24, n. 02, p. 180–190, 1928.
- GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. Boca Raton, FL: CRC press, 2006.
- GAZZINELLI, R. M. **Inferência bayesiana na análise de dados de experimentos planejados**. 2013. 90 f., il. Monografia (Bacharelado em Estatística) - Universidade de Brasília, Brasília, 2013.
- GELFAND, A. E.; SMITH, A. F. Sampling-based approaches to calculating marginal densities. **Journal of the American statistical association**, Taylor & Francis Group, v. 85, n. 410, p. 398–409, 1990.
- GELMAN, A. et al. **Bayesian data analysis**. 3. ed. Boca Raton, FL: CRC press, 2014.

GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, n. 6, p. 721–741, 1984.

GEWEKE, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: BERNARDO, J. M. et al. (Ed.). **Bayesian Statistics 4**. New York: Oxford University Press, 1992. p. 625–631.

GNEDENKO, B. Sur la distribution limite du terme maximum d'une serie aleatoire. **Annals of mathematics**, JSTOR, p. 423–453, 1943.

GOMES, M. I.; GUILLOU, A. Extreme value theory and statistics of univariate extremes: A review. **International Statistical Review**, 2014.

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. **Biometrika**, v. 57, n. 1, p. 97–109, 1970.

HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. **Operations Research**, v. 31, n. 6, p. 1109–1144, 1983.

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

LEICHSENRING, D. R. **Endogeneidade e mecanismos de transmissão entre a taxa de juros doméstica e o risco soberano: uma revisita aos determinantes do risco-Brasil**. 2004. Tese (Doutorado) — Universidade de São Paulo, São Paulo, 2004.

MARTINS, A. B. T.; RIBEIRO JUNIOR, P. J.; BONAT, W. H. Um modelo geoestatístico bivariado para dados composicionais. **Revista Brasileira de Biometria**, v. 27, n. 3, p. 456–477, 2009.

MARTINS FILHO, S. et al. Abordagem bayesiana das curvas de crescimento de duas cultivares de feijoeiro. **Ciência Rural**, SciELO Brasil, v. 38, n. 6, p. 1516–1521, 2008.

MENDES, B. V. M. **Introdução à análise de eventos extremos**. Rio de Janeiro: E-papers Serviços Editoriais Ltda, 2004.

METROPOLIS, N. et al. Equation of state calculations by fast computing machines. **The journal of chemical physics**, AIP, v. 21, n. 6, p. 1087–1092, 1953.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**. USA: McGraw-hill, 1950.

NAGUETTINI, M.; PINTO, É. J. d. A. **Hidrologia Estatística**. Belo Horizonte: CPRM, 2007.

NELDER, J. A.; MEAD, R. A. A simplex method for function minimization. **The computer**

journal, v. 7, p. 308–313, 1965.

NOGUEIRA, D. A.; SÁFADI, T.; FERREIRA, D. F. Avaliação de critérios de convergência para o método de monte carlo via cadeias de markov. **Revista Brasileira de Estatística**, IBGE, v. 65, n. 224, 2004.

NOGUEIRA, D. A. et al. Análises clássica e bayesiana de um modelo misto aplicado ao melhoramento animal: uma ilustração. **Ciência e Agrotecnologia**, Lavras, v. 27, p. 1614–1624, 2003.

PAULINO, C. D. M.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: FCG, 2003.

PORTES, P. C. **Modelagem Bayesiana dos níveis máximos do Índice de Preços ao Consumidor**. 2017. 73 f. Dissertação (Mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Alfenas, Alfenas, 2017.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>. Acesso em: 21 de mar. 2018.

RAFTERY, A. E.; LEWIS, S. M. [Practical Markov Chain Monte Carlo]: Comment: one long run with diagnostics: implementation strategies for Markov Chain Monte Carlo. **Statistical science**, JSTOR, v. 7, n. 4, p. 493–497, 1992.

SANSIGOLO, C. A. Distribuições de extremos de precipitação diária, temperatura máxima e mínima e velocidade do vento em Piracicaba, SP (1917-2006). **Revista Brasileira de Meteorologia**, SciELO Brasil, v. 23, n. 3, p. 341–346, 2008.

SILVA, J. V. d. M. **Modelagem estocástica em risco operacional aplicando teoria dos valores extremos**. 2006. 73 f. Dissertação (Mestrado em Economia)-Universidade de Brasília, Brasília, 2006.

SILVA, R. R. **A distribuição Generalizada de Pareto e mistura de distribuições de Gumbel no estudo da vazão e da velocidade máxima do vento em Piracicaba, SP**. 2008. 101 f. Dissertação (Mestrado em Agronomia) - Universidade de São Paulo, Piracicaba, 2008.

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

SMITH, R. L.; NAYLOR, J. A comparison of maximum likelihood and bayesian estimators for the three-parameter weibull distribution. **Applied Statistics**, JSTOR, p. 358–369, 1987.

CAPÍTULO 2 - ANÁLISE DO RISCO BRASIL MÁXIMO SEMESTRAL VIA DISTRIBUIÇÕES GUMBEL E GEV: ABORDAGENS CLÁSSICA E BAYESIANA

Thaís Brenda Martins; Luiz Alberto Beijo; Davi Butturi-Gomes; Reinaldo Antônio Gomes Marques

RESUMO: O risco Brasil é um *spread* que expressa o grau de risco de investimento no país. Portanto, a predição de risco Brasil máximo para determinados períodos de tempo é uma informação relevante para investidores e gestores. As distribuições Gumbel e generalizada de valores extremos (GEV) descrevem o comportamento probabilístico de eventos extremos máximos. A série de dados pode apresentar tendência ao longo do tempo, por ser medida diariamente, e por isso deve ser considerada no processo de modelagem. O objetivo deste capítulo é modelar a série de risco Brasil máximo semestral via distribuições GEV e Gumbel e verificar qual dessas distribuições melhor se ajusta à série de dados do período de 2003-2 a 2017-2. Para estimar os parâmetros, usou-se o método de máxima verossimilhança e inferência Bayesiana considerando diferentes distribuições *a priori*. Na abordagem Bayesiana, utilizaram-se *prioris* não informativas e informativas, sendo que as informações foram extraídas de dados de risco país da Argentina e México. Essas informações foram combinadas dentro dos hiperparâmetros de duas formas diferentes, as quais chamamos de estruturas. A tendência não foi significativa e a distribuição GEV com inferência Bayesiana apresentou melhores resultados que os modelos com máxima verossimilhança. O modelo Bayesiano com distribuição Gumbel e informação *a priori* do risco México com estrutura de *priori* II e a variância multiplicada por 4, foi considerado o modelo mais adequado para descrever o comportamento do risco Brasil máximo semestral. O modelo apresentou acerto nos 2 níveis de acurácia, menor amplitude média e erro médio de predição, o que indica que o modelo é mais preciso, e também apresentou erro médio de predição baixo e, por isso, foi usado para predição do risco Brasil e cálculo de probabilidades de ocorrência de valores extremos de risco Brasil. Portanto, a inferência Bayesiana apresentou melhores resultados que a máxima verossimilhança, sendo considerada como melhor método de estimação, quando usada com as distribuições Gumbel e GEV aplicadas a dados de risco Brasil máximo semestral.

Palavras-chave: Inferência Bayesiana, distribuição generalizada de valores extremos, distribuição Gumbel, taxa de retorno.

ANALYSIS OF MAXIMUM COUNTRY RISK BRAZIL PER SEMESTER VIA GUMBEL AND GEV DISTRIBUTIONS: CLASSIC AND BAYESIAN APPROACHES

Thaís Brenda Martins; Luiz Alberto Beijo; Davi Butturi-Gomes

ABSTRACT: The country risk Brazil is a spread that expresses the degree of investment risk in the country. Therefore, the prediction of maximum Brazil risk for certain periods of time is relevant information for investors and managers. The Gumbel and generalized extreme value distributions (GEV) describe the probabilistic behavior of extreme maximum events. The data series may present a trend over time, as it is measured daily, and therefore should be considered in the modeling process. The objective of this chapter is to model the series of Brazil maximum semester risk via GEV and Gumbel distributions and to verify which of these distributions best fits the data series from the 2003-2 to 2017-2 period. To estimate the parameters, the maximum likelihood and Bayesian inference method was used considering different a priori distributions. In the Bayesian approach, non-informative and informative prioris were used, and the information was extracted from country risk data from Argentina and Mexico. This information was combined within the hyperparameters in two different ways, which we call structures. The trend was not significant and the GEV distribution with Bayesian inference presented better results than the models with maximum likelihood. The Bayesian model with Gumbel distribution and a priori information of the risk Mexico with priori structure II and the variance multiplied by 4, was considered the most adequate model to describe the behavior of the Brazil maximum semester risk. The model presented accuracy in the two levels of accuracy, lower mean amplitude and mean prediction error, which indicates that the model is more accurate, and also presented a low prediction error, and was therefore used to predict Brazil and calculation of probability of occurrence of extreme risk values Brazil. Therefore, Bayesian inference presented better results than the maximum likelihood, being considered as the best estimation method, when used with the Gumbel and GEV distributions applied to data of maximum country risk Brazil per semester.

Keywords: *Bayesian inference, Generalized extreme value distributions, Gumbel distribution, return rate.*

1 INTRODUÇÃO

O desenvolvimento e funcionamento da economia de todo país estão significativamente ligados às suas taxas de juros. Os juros cumprem papel fundamental na determinação do nível de atividade, do emprego, da taxa de câmbio e de outras variáveis econômicas (GARCIA; DIDIER, 2003).

O risco país é uma medida que reflete a situação econômica e financeira de um país, assim como o desempenho histórico no cumprimento de suas obrigações financeiras e a estabilidade política do país (GARCIA; DIDIER, 2003). É válido ressaltar que o risco país não faz nenhum juízo de valor sobre os países, e sim expressa o risco de crédito a que investidores estrangeiros são submetidos quando investem no país. Por isso, a interpretação dos investidores é de que quanto maior o risco país, mais perigoso será investir nesse país. Com isso, é importante ter conhecimento sobre as previsões de risco Brasil máximo.

Autores como Garcia e Didier (2003), Moreira e Rocha (2003) e Leichsenring (2004) buscaram em seus trabalhos determinar variáveis que influenciam na variação do risco Brasil. No histórico feito por Leichsenring (2004), são notáveis os efeitos negativos na economia quando ocorre uma elevação do risco Brasil, como por exemplo, a retração de investimentos internacionais no país, uma vez que as taxas de juros elevadas indicam que há algo arriscado no país. Com isso, torna-se necessário estudar as observações máximas de risco Brasil.

Um ramo da Estatística que estuda a ocorrência de eventos extremos, sejam eles máximos ou mínimos, é a teoria de valores extremos (TVE). Essa teoria teve início com Fisher e Tippett (1928) que desenvolveram os três tipos de distribuições de valores extremos: Weibull, Gumbel e Fréchet. Estudos posteriores deram base para Jenkinson (1955) unir essas três distribuições em uma única forma paramétrica, conhecida como distribuição generalizada de valores extremos (*Generalized Extreme Values* - GEV). Por meio da TVE, é possível fazer previsões para o risco Brasil máximo para um determinado período de tempo, a partir de uma série de máximos desse risco.

O objetivo desse capítulo é modelar a série de máximos semestrais de risco Brasil. Para isso pretende-se identificar entre as distribuições GEV e Gumbel, a que melhor descreve probabilisticamente os dados e paralelamente avaliar as abordagens clássica e Bayesiana para estimação de parâmetros dessas distribuições. E, a partir do modelo mais adequado, realizar previsões de risco Brasil para tempos futuros (níveis de retorno) e calcular probabilidades de ocorrência

de valores máximos de Risco Brasil.

2 REVISÃO DE LITERATURA

2.1 RISCO BRASIL

A restrição aos fluxos internacionais de capitais dos países emergentes começou a ser retirada no início da década de 1990. O fluxo de capitais equaliza as taxas de juros entre os países quando as taxas são denominadas em uma mesma moeda. Isso acontece quando há paridade coberta das taxas de juros. O diferencial dessa paridade é conhecido como risco país. O termo é risco país pois afeta os rendimentos de todos os ativos financeiros (GARCIA; DIDIER, 2003).

O risco país é calculado para países emergentes. No mercado brasileiro, um indicador que expressa o risco Brasil é o EMBI+Br (*Emerging Markets Bond Index Plus*), calculado pelo Banco J. P. Morgan Chase. De acordo com o Banco Central do Brasil (2016), é um *spread* que corresponde à média ponderada dos prêmios pagos pelos títulos da dívida externa brasileira em relação a papéis de prazo equivalente do Tesouro dos Estados Unidos que, segundo Toledo (2002), é tido como o país mais estável, de risco praticamente nulo.

O risco país é uma medida que busca expressar de forma objetiva o risco de crédito a que investidores estrangeiros estão submetidos quando investem no país, ou seja, expressa a capacidade de o país honrar com seus compromissos financeiros com o exterior (BANCO CENTRAL DO BRASIL, 2016). Logo, quanto maior a pontuação do risco Brasil, maior é o risco de crédito do país.

Um país com *spread* elevado, para atrair capital estrangeiro em montante suficiente para o financiamento de sua dívida externa, necessita oferecer altas taxa de juros em seus papéis (BANCO CENTRAL DO BRASIL, 2016).

Em seus trabalhos, Leichsenring (2004) e Couto e Meurer (2017) encontraram uma relação positiva entre o risco Brasil e a taxa de câmbio entre o real e o dólar norte-americano. Essa relação indica que uma queda constante no risco Brasil repercute melhor percepção dos investidores em relação às perspectivas macroeconômicas do país, tendo melhor acesso a capitais de longo prazo, na forma de investimento direto, e investimento em carteira. Isto representa maior oferta de capitais para o país, possibilitando um deslocamento da taxa de câmbio no sentido da valorização do real. De forma análoga, o aumento contínuo do risco Brasil piora as condições

de acesso dos agentes econômicos domésticos ao financiamento externo, deslocando a taxa de câmbio no sentido de desvalorização do real (BANCO CENTRAL DO BRASIL, 2016).

Nunes et al. (2003) investigaram em seu trabalho a relação entre variáveis macroeconômicas e risco Brasil como os retornos do Ibovespa. Seus resultados apontaram para uma cointegração entre risco Brasil e retornos do Ibovespa. De acordo com os autores, a presença de cointegração entre as variáveis fornece evidência forte de que o risco Brasil seja um fator significativo para explicar os retornos esperados do Ibovespa no longo prazo.

Oliveira (2002) propôs em seu trabalho analisar se a taxa Selic contribui para a formação do risco Brasil. Por meio de uma metodologia para estabelecer a relação entre as variáveis, chegou à conclusão de que há uma relação positiva entre as variações da taxa Selic e as variações do risco Brasil.

As distribuições de extremos têm sido utilizadas para modelar variáveis econômicas. Arraes e Rocha (2006) estimaram perdas máximas em séries financeiras utilizando a distribuição GEV, Oliveira (2012) utilizou a distribuição Gumbel para modelagem de dados de perdas operacionais de instituições financeiras e Portes (2017) modelou níveis máximos de IPCA (Índice de Preços ao Consumidor Amplo) através da distribuição GEV. As distribuições de extremos Gumbel e GEV são apresentadas a seguir.

2.2 DISTRIBUIÇÕES GEV E GUMBEL

A função densidade da distribuição GEV não-estacionária, com tendência linear no parâmetro posição (μ), é obtida acrescentando-se um valor para a inclinação no parâmetro posição. Com algumas adaptações na expressão apresentada por Moreira, Naghettini e Eleutério (2016), obtém-se a função densidade de probabilidade para a distribuição GEV com ou sem a presença de tendência no parâmetro posição:

$$f(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu_T}{\sigma} \right)^{-\left(\frac{1+\xi}{\xi}\right)} \right] \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu_T}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}, \quad (2.1)$$

definida em, $-\infty < x < \mu_T - \frac{\sigma}{\xi}$, para $\xi < 0$ e $\mu_T - \frac{\sigma}{\xi} < x < \infty$, para $\xi > 0$. Quando a distribuição GEV é estacionária (não apresenta tendência) $\mu_T = \mu$ e quando a distribuição é não-estacionária

o parâmetro posição passa a ser uma função linear em relação ao tempo, $\mu_T = \mu + \mu_1 T$. A distribuição GEV sem tendência será chamada de distribuição GEV estacionária (GEV) e a distribuição GEV com tendência linear no parâmetro de posição, de GEV não-estacionária (GEV_T).

A partir da expressão apresenta por Clarke (2002), é possível obter uma função densidade de probabilidade para a distribuição Gumbel com e sem tendência no parâmetro posição:

$$f(x) = \frac{1}{\sigma} \exp \left\{ - \left(\frac{x - \mu_T}{\sigma} \right) - \exp \left[- \left(\frac{x - \mu_T}{\sigma} \right) \right] \right\}, \quad (2.2)$$

definida em $-\infty < x < \infty$, $-\infty < \mu_T < \infty$ e $\sigma > 0$. Quando a distribuição Gumbel é estacionária (não apresenta tendência) o parâmetro posição é $\mu_T = \mu$ e quando a distribuição é não-estacionária (apresenta tendência) esse parâmetro passa a ser uma função linear em relação ao tempo, $\mu_T = \mu + \mu_1 T$.

2.3 ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA

Para a distribuição generalizada de valores extremos (GEV) tem-se a seguinte função de verossimilhança:

$$L(\mu, \sigma, \xi | \mathbf{x}) = \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[1 + \xi \left(\frac{x_i - \mu_T}{\sigma} \right) \right]^{-\left(\frac{1+\xi}{\xi}\right)} \right\} \exp \left\{ \sum_{i=1}^n \left\{ - \left[1 + \xi \left(\frac{x_i - \mu_T}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \right\}. \quad (2.3)$$

Aplicando o logaritmo natural na equação 2.3, tem-se que:

$$l(\mu, \sigma, \xi, \mathbf{x}) = \sum_{i=1}^n \left\{ -\ln(\sigma) - \left(\frac{1+\xi}{\xi} \right) \ln \left[1 + \xi \left(\frac{x_i - \mu_T}{\sigma} \right) \right] - \left[1 + \xi \left(\frac{x_i - \mu_T}{\sigma} \right) \right]^{\frac{1}{\xi}} \right\}. \quad (2.4)$$

A função de verossimilhança para a distribuição Gumbel é dada por:

$$L(\mu, \sigma | \mathbf{x}) = \frac{1}{\sigma^n} \exp \left\{ \sum_{i=1}^n \left[- \left(\frac{x_i - \mu_T}{\sigma} \right) \right] - \sum_{i=1}^n \exp \left[- \left(\frac{x_i - \mu_T}{\sigma} \right) \right] \right\}. \quad (2.5)$$

Para a distribuição Gumbel, o logaritmo natural da função de verossimilhança é dado por:

$$l(\mu, \sigma, \mathbf{x}) = -n \ln(\sigma) + \sum_{i=1}^n \left[- \left(\frac{x_i - \mu_T}{\sigma} \right) \right] - \sum_{i=1}^n \exp \left\{ - \frac{x_i - \mu_T}{\sigma} \right\}. \quad (2.6)$$

Derivando o logaritmo natural da função de verossimilhança da distribuição GEV e da distribuição Gumbel em relação a cada parâmetro e igualando-as a zero, obtém-se o sistema de equações com uma equação para cada parâmetro da distribuição.

Os sistemas de equações obtidos das derivadas das Equações 2.4 e 2.6, geralmente, não têm solução analítica, por isso utilizam-se métodos numéricos como o de Nelder-Mead (NELDER; MEAD, 1965).

2.4 ESTIMAÇÃO POR INFERÊNCIA BAYESIANA

Na inferência Bayesiana, além da informação obtida dos dados amostrais representada pela função de verossimilhança é também utilizada a informação *a priori*. A combinação dessas duas informações resulta na distribuição *a posteriori* que contém toda informação probabilística a respeito do parâmetro de interesse (conforme revisão apresentada na Seção 2 do Capítulo 1).

A incorporação da informação *a priori* ao processo de estimação tende a melhorar o processo de inferência, reduzindo as incertezas sobre as estimativas dos parâmetros, de acordo com Coles (2001), essa vantagem é um dos motivos que têm influenciado autores a utilizarem a inferência Bayesiana para análise de valores extremos. Alguns trabalhos que utilizaram a abordagem Bayesiana e a teoria de valores extremos são: Coles e Tawn (1996) que analisaram precipitações extremas, Beijo, Vivanco e Muniz (2009) que estudaram o tempo de retorno das precipitações pluviais máximas em Jaboticabal (SP) e Bernardini (2010) que modelou os máximos trimestrais das taxas de desemprego nos EUA e os máximos semestrais dos níveis de maré em Newlyn.

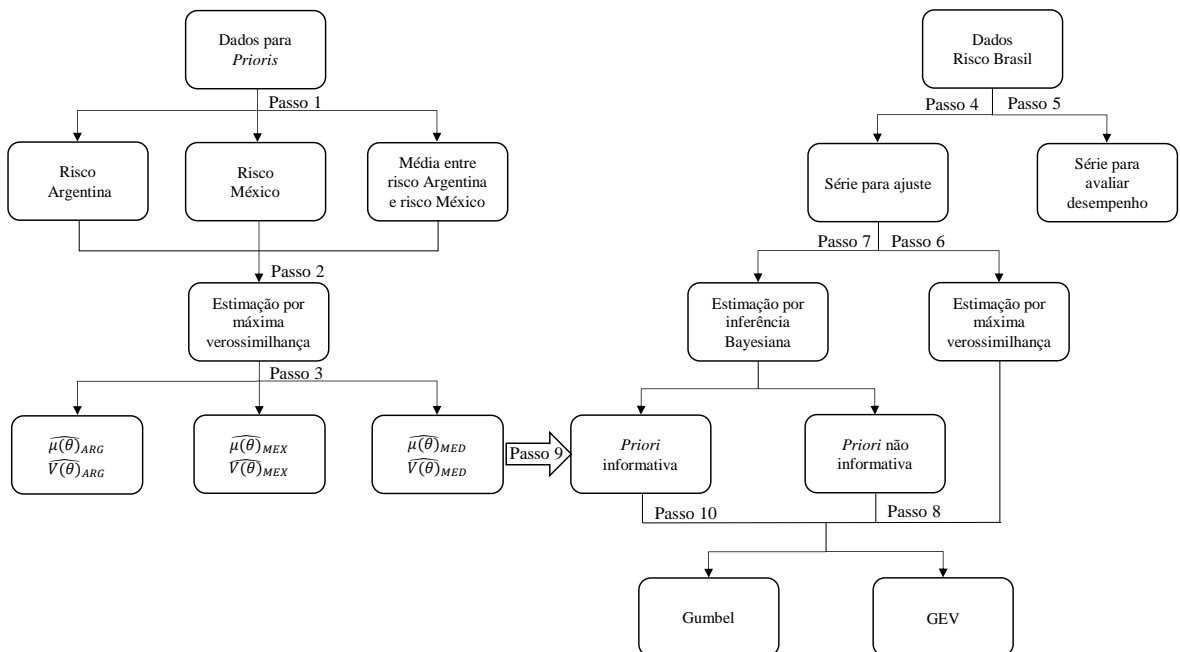
3 MATERIAL E MÉTODOS

Nesta seção serão apresentados em subseções os dados, recursos computacionais e métodos utilizados para modelagem do risco Brasil.

3.1 FLUXOGRAMA DA METODOLOGIA

A Figura 1 apresenta o fluxograma que descreve a metodologia usada neste capítulo. Esse fluxograma foi elaborado para ilustrar e elucidar o entendimento da análise do risco Brasil máximo semestral por meio das distribuições Gumbel e GEV, ambas estacionárias e não-estacionárias, com estimação dos parâmetros pelos métodos de máxima verossimilhança e inferência Bayesiana.

Figura 1 – Fluxograma da metodologia do Capítulo 2.



Fonte: Elaborada pela autora

O fluxograma apresentado na Figura 1 descreve a metodologia utilizada neste capítulo, sendo que detalhes da execução de cada passo do fluxograma será descrito nas subseções a seguir.

3.2 DADOS

Os dados utilizados nesse estudo podem ser divididos em dois grupos: os dados de risco Brasil, que são os dados analisados nesse capítulo, e os dados para *prioris* usados para elicitación das informações *a priori*.

Os dados do risco Brasil foram obtidos na base de dados do Instituto de Pesquisa Econômica Aplicada - IPEA (IPEA-DATA, 2017). As observações de risco Brasil são diárias do período de 01/07/2003 a 31/12/2017. A partir dessa série, foi obtida a série de máximos semestrais, obtendo um máximo para cada semestre o que totalizou uma série de 29 observações. A série de 29 observações foi dividida em duas, sendo que de 2003-2 a 2016-1 foi usada para ajustar as distribuições (Passo 4, na Figura 1) e de 2016-2 a 2017-2 foi usada para avaliar o desempenho desses ajustes (Passo 5, na Figura 1).

Os dados para as informações *a priori* foram obtidos na base de dados Knoema (KNOEMA, 2017) e correspondem às observações mensais do risco país da Argentina, do período de janeiro de 1994 a abril de 2014, e do risco país do México, do período de janeiro de 1994 a setembro de 2013. Dessas séries mensais, foram obtidas séries de máximos anuais para os anos de 1994 a 2013 para o risco Argentina e para o risco México. A partir da série de risco país máximo anual da Argentina e do México, foi composta uma nova série com a média dos dois riscos. As séries de risco Argentina, risco México e a média dos dois riscos foram usadas para elicitación das informações *a priori* (Passo 1, na Figura 1). Vale mencionar que a unidade de medida de qualquer risco país é o ponto-base, dez pontos-base equivalem a um décimo de 1%.

3.3 RECURSOS COMPUTACIONAIS

Os testes e análises foram realizadas utilizando a linguagem R (R CORE TEAM, 2017) e o *freeware* OpenBugs (THOMAS; O'HARA, 2004), que é um programa para análise de modelos usando o método Monte Carlo via Cadeias de Markov. Os pacotes utilizados no R foram o *evd* (STEPHENSON, 2002), o *evdbayes* (STEPHENSON; RIBATET, 2014), o *R2OpenBUGS* (STURTZ; LIGGES; GELMAN, 2005), o *tseries* (TRAPLETTI; HORNIK, 2017) e o *coda* (PLUMMER et al., 2006).

3.4 TESTES DE HIPÓTESES PRELIMINARES

A pressuposição de independência da série de risco Brasil foi verificada por meio do teste de hipótese Durbin-Watson (DURBIN; WATSON, 1951), ao nível de 5% de significância. As hipóteses do teste foram:

$$\begin{cases} H_0 : \text{a série é independente,} \\ H_1 : \text{a série é dependente.} \end{cases} \quad (3.1)$$

Para verificar a não-estacionariedade (tendência) da série de risco Brasil, foi utilizado o teste de hipótese Mann-Kendall (MANN, 1945; KENDALL, 1975), ao nível de significância de 5%. As hipóteses do teste Mann-Kendall são:

$$\begin{cases} H_0 : \text{a série é estacionária,} \\ H_1 : \text{a série é não-estacionária.} \end{cases} \quad (3.2)$$

O teste para tendência não indicou a presença dessa na série de risco Brasil. Tal resultado é apresentado nesta Seção para que as próximas subseções sejam coerentes com a análise que foi realizada, uma vez que a tendência foi desconsiderada a partir do teste Mann-Kendall (valor- $p = 0,3580$). Portanto, os ajustes apresentados a seguir foram para as distribuições Gumbel e GEV nas formas estacionárias (sem tendência).

3.5 ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA

Os primeiros ajustes das distribuições GEV e Gumbel, estacionárias, à série de dados de ajuste (2003-2 a 2016-1), foram feitos usando o método de máxima verossimilhança para estimação dos parâmetros (Passo 6, na Figura 1). Portanto, obtiveram-se dois ajustes com máxima verossimilhança. Esses ajustes foram comparados com os ajustes Bayesianos com *prioris* não informativas e informativas.

3.6 ESTIMAÇÃO POR INFERÊNCIA BAYESIANA

A estimação por inferência Bayesiana (Passo 7, na Figura 1) foi dividida em duas etapas: a estimação com *priori* não informativa e com *priori* informativa.

Para possibilitar o uso da inferência Bayesiana, é preciso especificar uma distribuição *a priori* e os valores de seus hiperparâmetros. Para os parâmetros da distribuição GEV foi usada como distribuição *a priori* conjunta a distribuição normal trivariada sugerida por Coles e Powell (1996):

$$p(\boldsymbol{\theta}) = (2\pi)^{-3/2} |\boldsymbol{\Sigma}_0|^{-1/2} \exp \left\{ -\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\Phi}_0)^t \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\theta} - \boldsymbol{\Phi}_0) \right\}, \quad (3.3)$$

em que $\boldsymbol{\theta} = (\mu, \ln(\sigma), \xi)$ é o vetor de parâmetros da distribuição GEV. Os hiperparâmetros $\boldsymbol{\Phi}_0$ e $\boldsymbol{\Sigma}_0$ são, respectivamente, o vetor de médias dos parâmetros e a matriz de variâncias e covariâncias dos parâmetros, sendo que na diagonal estão as variâncias e no restante da matriz as covariâncias.

Logo, com a função de verossimilhança da distribuição GEV e a distribuição *a priori*, a partir do teorema de Bayes, é possível obter a distribuição *a posteriori* conjunta para a distribuição GEV, que é dada por:

$$P(\boldsymbol{\theta}|\mathbf{x}) \propto \frac{1}{\sigma^n} \prod_{i=1}^n \left\{ \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{-\left(\frac{1+\xi}{\xi}\right)} \right\} \exp \left\{ \sum_{i=1}^n \left\{ - \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \right\} \times \\ \times (2\pi)^{-3/2} |\boldsymbol{\Sigma}_0|^{-1/2} \exp \left\{ -\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\Phi}_0)^t \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\theta} - \boldsymbol{\Phi}_0) \right\},$$

em que $\boldsymbol{\theta} = (\mu, \ln(\sigma), \xi)$ é o vetor de parâmetros da distribuição GEV.

Para os parâmetros da distribuição Gumbel, foram adotadas como distribuições *a priori*, distribuições normais para cada parâmetro, sendo que, de acordo com Mood, Graybill e Boes (1950), a função densidade da distribuição normal é dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\}.$$

sendo que para o estudo, assumimos que μ é a média da distribuição que chamaremos de m e σ^2 que é a variância substituiremos pela precisão que é $\tau = 1/\sigma^2$. Logo, a distribuição *a priori*

para os parâmetros de posição (μ) e escala (σ) da Gumbel é dada por:

$$\begin{aligned}\mu &\sim N(m_\mu, \tau_\mu) \\ \sigma &\sim N(m_\sigma, \tau_\sigma).\end{aligned}\tag{3.4}$$

Portanto, a distribuição *a posteriori* conjunta para a distribuição Gumbel estacionária é dada por:

$$\begin{aligned}P(\mu, \sigma | \mathbf{x}) &\propto \frac{1}{\sigma^n} \exp \left\{ \sum_{i=1}^n \left[- \left(\frac{x_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^n \exp \left[- \left(\frac{x_i - \mu}{\sigma} \right) \right] \right\} \times \\ &\times \exp \left\{ - \frac{\tau_\mu}{2} (\mu - m_\mu)^2 \right\} \times \exp \left\{ - \frac{\tau_\sigma}{2} (\sigma - m_\sigma)^2 \right\},\end{aligned}$$

sendo m_μ e m_σ os hiperparâmetros média e τ_μ e τ_σ os hiperparâmetros de precisão.

Como vimos anteriormente, o processo de inferência Bayesiana é realizado sobre a distribuição *a posteriori* e para obter essa distribuição, usou-se o algoritmo Metropolis-Hastings do método MCMC.

Para os ajustes Bayesianos com distribuição GEV, foi utilizado o pacote *evdbayes* e para os ajustes Bayesianos com distribuição Gumbel, foi utilizado o pacote *R2OpenBUGS*.

3.6.1 *Prioris* não informativas

Os primeiros ajustes Bayesianos foram feitos com *prioris* não informativas (Passo 8, na Figura 1). Para os parâmetros da distribuição GEV, foi adotada como *priori* não informativa a normal trivariada com vetor de médias Φ_0 e matriz de variâncias e covariâncias Σ_0 sendo:

$$\Phi_0 = (0, 0, 0) \quad \text{e} \quad \Sigma_0 = \begin{bmatrix} 1.000.000 & 0 & 0 \\ 0 & 1.000.000 & 0 \\ 0 & 0 & 10.000 \end{bmatrix}.$$

A *priori* não informativa foi correspondente para as duas distribuições, logo para os parâmetros da distribuição Gumbel, foram adotadas distribuições normais com hiperparâmetros dados por:

$$\mu \sim N\left(0, \frac{1}{1.000.000}\right) \quad \text{e} \quad \sigma \sim N\left(0, \frac{1}{1.000.000}\right),$$

ou seja, médias iguais a zero e variâncias iguais a 1.000.000 para os parâmetros de posição e escala. Logo, foram obtidos dois modelos Bayesianos com *priori* não informativa.

Antes de começar os ajustes Bayesianos com *prioris* informativas, foi verificado se o ajuste das duas distribuições eram adequados aos dados. Para isso, usou-se a análise dos gráficos probabilidade-probabilidade (ou gráficos P-P). Esses gráficos comparam a função de distribuição acumulada empírica de uma série de dados plotada no eixo das abscissas com a função de distribuição acumulada teórica plotada no eixo das ordenadas. A interpretação para o gráfico P-P, de acordo com Silva (2008), é que se o modelo está bem ajustado aos dados, o gráfico deve apresentar tendência linear e a amostra observada deve estar dentro dos limites de confiança do envelope simulado.

3.6.2 *Prioris* informativas

As distribuições Gumbel e GEV foram ajustada via inferência Bayesiana com *prioris* informativas.

Primeiramente, foram ajustados as distribuições Gumbel e GEV com estimação dos parâmetros pelo método de máxima verossimilhança (Passo 2, na Figura 1) aos dados de risco Argentina, risco México e a média entre os dois riscos.

A partir dos ajustes foram obtidas as estimativas para cada parâmetro e variância dessa estimativa (Passo 3, na Figura 1), que foram usadas como valores dos hiperparâmetros para as *prioris* informativas (Passo 10, na Figura 1). Para o vetor de médias da normal trivariada (Φ_0) e para a média das normais (m), foram usados os valores obtidos das estimativas e o logaritmo natural desses valores. Para a matriz de variâncias e covariâncias (Σ_0) e o hiperparâmetro de precisão (τ), foram usadas as variâncias obtidas nos ajustes. Essas variâncias foram multiplicadas por 1, 2, 4 e 6 para flexibilização (elevação) da variância, sendo as 4 formas de variância: 1var, 2var, 4var e 6var.

A Tabela 1 apresenta as duas estruturas de *priori* que foram usadas a partir das estimativas e variâncias obtidas de cada informação *a priori* (Argentina, México e a média dos dois risco país).

Tabela 1 – Estruturas de *prioris* informativas usadas para os parâmetros da distribuição GEV e Gumbel.

Estrutura	Distribuição GEV	Distribuição Gumbel
Priori I	$(\mu, \ln(\sigma), \xi) \sim \text{MVN}(\Phi_0, \Sigma_0)$ $\Phi_0 = (m_\mu, m_\sigma, m_\xi)$	$\mu \sim \text{N}(m_\mu, \tau_\mu)$ $\ln \sigma \sim \text{N}(m_\sigma, \tau_\sigma)$
Priori II	$(\mu, \ln(\sigma), \xi) \sim \text{MVN}(\Phi_0, \Sigma_0)$ $\Phi_0 = (m_\mu, \ln(m_\sigma), m_\xi)$	$\mu \sim \text{N}(m_\mu, \tau_\mu)$ $\ln \sigma \sim \text{N}(\ln(m_\sigma), \tau_\sigma)$

Fonte: Elaborada pela autora

A diferença entre as duas *prioris* está no uso do logaritmo natural na média do parâmetro de escala na *priori* II. O logaritmo natural foi usado para que o parâmetro de escala assumisse apenas valores positivos.

As distribuições Gumbel e GEV foram ajustadas com informações obtidas do risco país da Argentina, México e da média dos dois países. Foram usadas as duas estruturas de *priori* e as 4 formas da variância. Então, para cada distribuição foram ajustados 24 modelos Bayesianos com *prioris* informativas. Assim, foram ajustados 48 modelos Bayesianos informativos, 2 modelos Bayesianos não informativos e 2 modelos com máxima verossimilhança.

Para verificar o desempenho desses 52 modelos e então determinar qual melhor método de estimação para esses dados a partir das distribuições GEV e Gumbel, foram calculados níveis de retorno de risco Brasil máximo.

3.7 NÍVEIS DE RETORNO

Para cada ajuste foram calculados os níveis de retorno (ou predições), $x(T)$, que são os níveis de risco Brasil máximo esperado para determinados tempos de retorno (T). Os níveis de retorno correspondem aos quantis da distribuição GEV e Gumbel.

De acordo com Coles (2001), para a distribuição GEV, os níveis de retorno são dados por:

$$x(T) = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left\{ 1 - \left[-\ln \left(1 - \frac{1}{T} \right) \right]^{-\hat{\xi}} \right\},$$

em que $\hat{\mu}$, $\hat{\sigma}$ e $\hat{\xi}$ são as estimativas dos parâmetros da distribuição GEV e T é o tempo de retorno dado em semestres.

E para a distribuição Gumbel, a equação para os níveis de retorno é dada por:

$$x(T) = \hat{\mu} - \hat{\sigma} \ln \left[-\ln \left(1 - \frac{1}{T} \right) \right],$$

em que $\hat{\mu}$ e $\hat{\sigma}$ são as estimativas dos parâmetros da distribuição Gumbel e T o tempo de retorno em semestres.

Foram calculados níveis de retorno para 2 e 3 semestres para avaliar o desempenho dos ajustes.

3.8 ANÁLISE DE CONVERGÊNCIA DAS CADEIAS GERADAS PELO MÉTODO MCMC

As distribuições marginais para cada parâmetro são obtidas por meio da integração da distribuição *a posteriori* conjunta e, como não podem ser solucionadas analiticamente, foram utilizadas simulações MCMC.

Os ajustes com distribuição GEV foram feitos utilizando-se o pacote *evdbayes*. Primeiramente, foram geradas cadeias com 6.000 valores, sendo excluídos os 1.000 primeiros e realizado salto a cada 1 valor. Neste caso não obteve-se indícios de convergência das cadeias nos critérios. Então, aumentou-se o número de valores gerados, a queima e o salto, obtendo cadeias eficientes com 110.000 valores gerados, excluindo os 10.000 primeiros e saltos realizados a cada 26 valores, obtendo cadeias finais de aproximadamente 3.846 valores.

Os ajustes com a distribuição Gumbel foram feitos utilizando o pacote *R2OpenBUGS* e para obter simulações eficientes foi preciso gerar cadeias com 22.000 valores, sendo que os primeiros 4.000 foram excluídos e foram realizados saltos a cada 4 valores, o que resultou em cadeias finais com 4.000 valores.

A verificação de convergência das cadeias, dos parâmetros e dos níveis de retorno para os modelos Bayesianos foi realizada com a utilização do pacote *coda* e, conforme descrito na Seção 2, foram usados três critérios. No critério de Raftery e Lewis, foi verificado se o fator de dependência era próximo de 1,0. No critério de Geweke, avaliou-se o valor normalizado, onde usou-se 95% de confiança para o critério e o módulo do resultado não poderia ser maior que 1,96, ou seja, deve estar entre -1,96 e 1,96. E no critério de Heidelberger e Welch, verificou-se se o valor-p do teste era maior que 0,05 (NOGUEIRA; SÁFADI; FERREIRA, 2004).

3.9 ANÁLISE DE DESEMPENHO DOS MODELOS

Para avaliar o desempenho dos modelos foi verificada a sua acurácia, o erro médio de predição, o DIC para modelos Bayesianos e a amplitude média das estimativas intervalares de 95%.

Os níveis de retorno foram calculados, com base nas equações apresentados na subseção 3.7, para os tempos de 2 e 3 semestres. Os valores observados, foram extraídos da série reservada para análise de desempenho, que são as observações do segundo semestre de 2016 ao segundo semestre de 2017, sendo os valores observados o máximo em 2 e 3 semestres.

Uma predição tem acurácia quando o intervalo de confiança ou HPD de 95% estimado para um nível de retorno engloba o valor observado para o respectivo tempo de retorno. Como estimativas pontuais foram adotadas as médias *a posteriori*.

O erro médio de predição foi calculado com base nos dois valores máximos observados na série de 2016-2 a 2017-2 em 2 e 3 semestres e os valores preditos para os mesmos tempos. A equação do erro médio de predição (EMP) é dada por:

$$EMP = \frac{\sum_{i=1}^2 |VO_i - VP_i| / VO_i}{2} * 100,$$

em que VO_i é o valor observado e VP_i é o valor predito, ambos para o i -ésimo tempo de retorno com $i = 1, 2$.

A amplitude média foi calculada com base nas amplitudes dos intervalos de confiança ou HPD de 95%. Sendo a amplitude dada por:

$$Amp = LS - LI,$$

sendo que, LS e LI são, respectivamente, o limite superior e o limite inferior da estimativa intervalar de 95% dos níveis de retorno. Assim, a amplitude média é dada por:

$$Amp_m = \frac{\sum_{i=1}^2 Amp_i}{2},$$

em que $i = 1, 2$, pois tem-se níveis de retorno para 2 e 3 semestres.

De acordo com Spiegelhalter et al. (2002), existem alguns métodos para comparar modelos ajustados a um mesmo conjunto de dados e escolher o que melhor se ajusta, entre eles o

critério de Akaike (AIC), critério de informação Bayesiano (BIC) e o critério de informação do desvio (DIC). Esses critérios são usados para modelos Bayesianos uma vez que são baseados na média *a posteriori* do desvio. Para comparação dos modelos Bayesianos neste trabalho será usado o DIC.

As melhores previsões de curto prazo serão feitas pelo modelo que apresentar o menor DIC, porém, Spiegelhalter et al. (2002) sugerem utilizar o seguinte critério para o módulo da diferença entre os valores de DIC de dois modelos, A e B, analisados. Esse critério é apresentado pela equação:

$$D = |DIC_A - DIC_B|.$$

Sendo assim,

- se $D < 5$, a diferença entre os modelos não é substancial;
- se $5 \leq D \leq 10$, a diferença entre os modelos é substancial;
- se $D > 10$, pode-se definitivamente excluir o modelo com maior DIC.

Para selecionar o melhor modelo seguiu-se os seguintes passos: primeiramente, selecionou-se todos os modelos que apresentaram acurácia para os dois níveis de retorno, ou seja, os modelos que tiveram 2 acertos. Posteriormente, ordenaram-se os modelos em relação aos valores do DIC, sendo escolhidos os três modelos que apresentaram menor amplitude média, entre os que apresentam diferença de DIC menor que 10, em relação ao de menor DIC. Esses três modelos foram ajustados à série completa de dados (2003-2 a 2017-2). Então, foi selecionado um modelo entre esses três ajustes, o que apresentou menor amplitude média dos intervalos HPD de 95% para as estimativas dos parâmetros.

A partir do modelo selecionado foram estimados os níveis de retorno para tempos futuros e calculadas as probabilidades de ocorrência de valores extremos máximos de risco Brasil.

3.10 CÁLCULO DE PROBABILIDADES E NÍVEIS DE RETORNO PARA TEMPOS FUTUROS

Após selecionar o modelo mais adequado para a série de risco Brasil máximo semestral foram estimados os níveis de retorno futuros para os tempos de 2 a 10 semestres. E também

foram calculadas as probabilidades de ocorrência de valores extremos máximos de risco Brasil. Os valores dos quantis para calcular as probabilidades de ocorrência foram obtidos a partir do maior valor observado na série de risco Brasil de 2003-2 a 2017-2, sendo esse valor igual a 898 pontos. Então, multiplicou-se esse valor por 0,80; 1,00; 1,20; 1,40; 1,60; 1,80 e 2,00, obtendo sete quantis. De acordo com Coles (2001), a probabilidade é calculada a partir da função de distribuição acumulada (f.d.a), sendo:

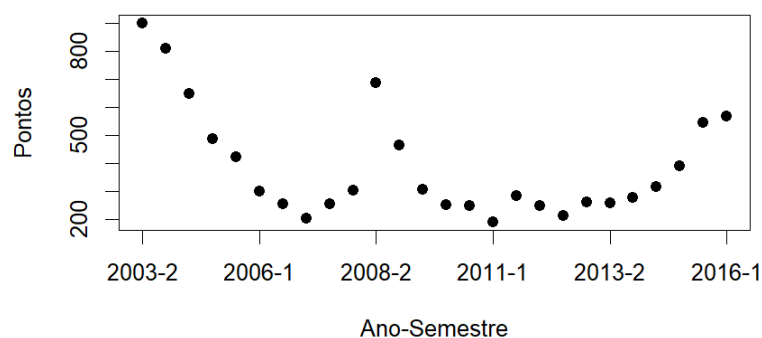
$$P(X > x) = 1 - F_G(x),$$

em que F_G é a função de distribuição acumulada do modelo escolhido, Gumbel (Equação 2.1) ou GEV (Equação 2.4).

4 RESULTADOS E DISCUSSÃO

O gráfico das observações de risco Brasil máximo do segundo semestre de 2003 ao primeiro semestre de 2016 é apresentado na Figura 2.

Figura 2 – Série de máximos semestrais de risco Brasil do segundo semestre de 2003 ao primeiro semestre de 2016.



Fonte: Elaborada pela autora

4.1 SELEÇÃO DE MODELOS

Primeiramente, nos testes preliminares foi verificado que a série de risco Brasil é independente pelo teste de Durbin-Watson (valor-p = 0,0745). O teste Mann-Kendall (valor-p = 0,2171) indicou que a série é estacionária, ou seja, não apresenta tendência. Por isso, a tendência não foi considerada nos ajustes das distribuições Gumbel e GEV.

As distribuições Gumbel e GEV foram ajustadas utilizando o método da máxima verossimilhança e inferência Bayesiana com *priori* não informativa para estimar os parâmetros e seus respectivos intervalos de confiança ou intervalos HPD de 95%. A Tabela 2 apresenta o resultado desses ajustes.

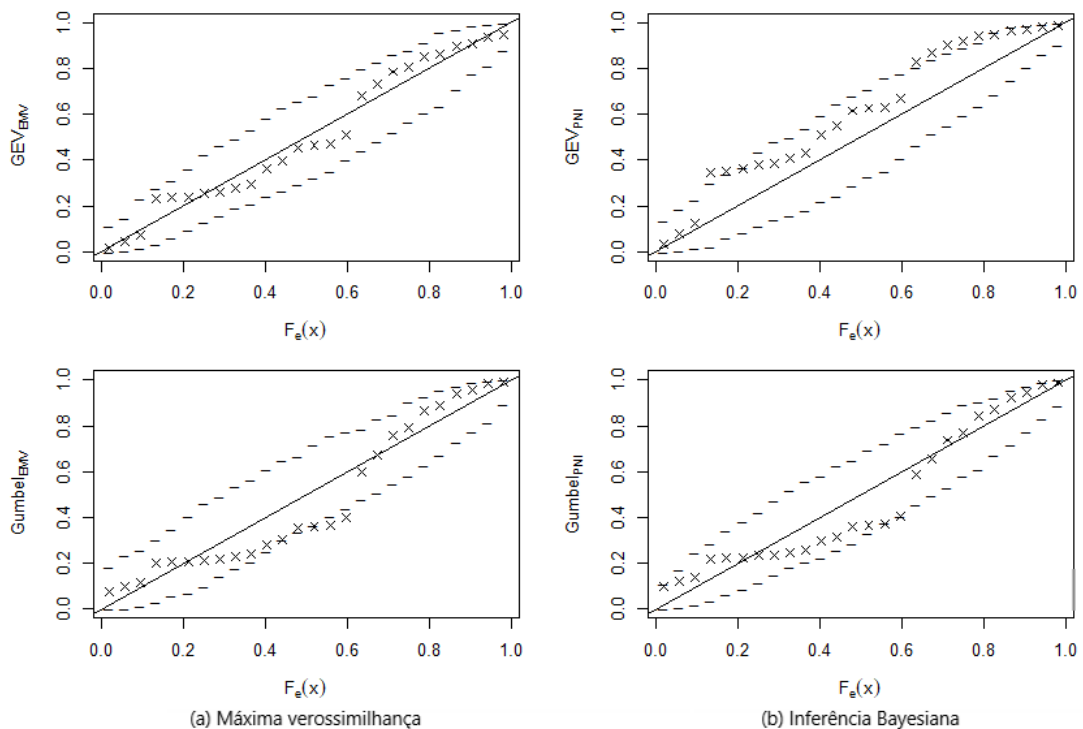
Tabela 2 – Estimativas dos parâmetros, calculados pelo método da máxima verossimilhança (MV) e inferência Bayesiana com *priori* não informativa (PNI), das distribuições GEV e Gumbel e seus respectivos intervalos de confiança e HPD de 95%.

Método de Estimação	Distribuição	Parâmetro	Estimativa	IC _{95%} ou HPD _{95%}
MV	GEV	μ	279,57	[240,01; 319,13]
		σ	87,92	[49,38; 126,46]
		ξ	0,52	[0,09; 0,94]
	Gumbel	μ	307,20	[257,99; 356,50]
		σ	123,00	[81,82; 164,19]
PNI	GEV	μ	252,33	[238,58; 326,23]
		σ	60,41	[56,86; 148,30]
		ξ	0,35	[0,09; 1,00]
	Gumbel	μ	305,12	[246,50; 361,00]
		σ	134,35	[91,03; 184,10]

Fonte: Elaborada pela autora

Foram plotados os gráficos P-P para os quatro ajustes apresentados da Tabela 2, a fim de verificar se as duas distribuições se ajustam aos dados. Os gráficos podem ser observados na Figura 3.

Figura 3 – Gráficos P-P para as distribuições GEV e Gumbel considerando o método de máxima verossimilhança (a) e inferência Bayesiana (b).



Fonte: Elaborada pela autora

Pode-se perceber, pela análise da Figura 3, que a distribuição GEV com *priori* não informativa apresenta alguns pontos fora do envelope simulado, mas como uma grande parte dos dados encontra-se dentro do envelope simulado, foi considerado que ambas distribuições se ajustaram aos dados. Logo, foram ajustados os modelos para as duas distribuições com inferência Bayesiana e *prioris* informativas.

Os resultados das estimativas dos parâmetros, seus respectivos intervalos HPD_{95%} e os resultados dos critérios de convergência para os parâmetros e para os níveis de retorno dos modelos Bayesianos, com *prioris* informativas e não informativa, são apresentados nas tabelas dispostas nos APÊNDICES deste trabalho (Tabela 1 a Tabela 11). Sendo que, no APÊNDICE A são apresentados os resultados para os parâmetros e no APÊNDICE B, os resultados para os níveis de retorno.

Para todos os parâmetros dos modelos com distribuição Gumbel e GEV (Tabela 1 a Tabela 7 do APÊNDICE A), assim como para os respectivos níveis de retorno de cada modelo (Tabela 8 a Tabela 11 do APÊNDICE B), os critérios de convergência indicaram que há convergência das cadeias para a distribuição de equilíbrio, pois, o módulo das estatísticas de Geweke foram menores que 1,96, os fatores de dependência de Raftery e Lewis foram próximos de 1, e valores-p do teste de Heidelberger e Welch foram maiores que 0,05.

A seguir serão apresentados e discutidos os resultados dos níveis de retorno, seus respectivos intervalos de confiança e HPD de 95%, amplitude dos intervalos (*Amp*), número de acertos de acurácia, DIC para os modelos Bayesianos, amplitude média (*Amp_m*) dos intervalos e EMP. Os valores observados no período de 2016-2 a 2017-2 para os tempos de 2 e 3 semestres foram, respectivamente, 357 e 357 pontos-base, ou seja, o mesmo valor para os dois tempos de retorno, pois é o máximo em dois e três semestres.

Na Tabela 3 são apresentados os resultados para os modelos ajustados com máxima verossimilhança e inferência Bayesiana com *priori* não informativa.

Tabela 3 – Estimativas, intervalos de confiança (IC) ou HPD de 95%, amplitude dos intervalos (Amp), amplitude média (Amp_m), número de acertos de acurácia ($N^o A$), DIC (para os modelos Bayesianos) e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com estimação por máxima verossimilhança (EMV) e inferência Bayesiana com *priori* não informativa (PNI).

Modelo	TR	NR	IC ou HPD _{95%}	Amp	$N^o A$	DIC	Amp_m	EMP
Gumbel	EMV	2	352,33	[297,23; 408,73]	111,50	-	125,66	9,24%
		3	418,28	[348,84; 488,67]	139,83			
	PNI	2	354,36	[291,70; 418,50]	126,80	340,90	141,20	10,09%
		3	426,40	[350,10; 505,70]	155,60			
GEV	EMV	2	315,05	[262,81; 365,42]	102,61	-	129,43	9,19%
		3	380,69	[301,25; 457,50]	156,25			
	PNI	2	322,60	[267,46; 383,99]	116,53	335,40	152,37	10,71%
		3	399,10	[313,59; 501,81]	188,22			

Fonte: Elaborada pela autora

Os modelos ajustados pelo método de máxima verossimilhança e inferência Bayesiana com *priori* não informativa apresentaram acurácia nos dois níveis de retorno ($N^o A = 2$), pois os valores observados em 2 e 3 semestres (357 pontos-base) pertencem aos respectivos intervalos de confiança ou HPD de 95%. A partir dos resultados da Tabela 3, pode-se fazer a interpretação prática para o nível de retorno para o modelo Gumbel com máxima verossimilhança, em que para um tempo médio de 3 semestres, espera-se que pelo menos um dia ocorra um valor de risco Brasil que seja maior ou igual a 418,28 pontos-base, e ainda, espera-se com 95% de confiança que esse valor seja maior ou igual a um valor que esteja entre 348,84 e 488,67 pontos-base.

Antes de apresentar os resultados obtidos dos ajustes Bayesianos da GEV e Gumbel com *prioris* informativas, serão apresentadas as informações extraídas das séries de máximos anuais do risco país da Argentina, México e da média dos dois países que foram usadas para compor as informações *a priori*. A Tabela 4 apresenta as estimativas obtidas para os parâmetros com os ajustes das distribuições GEV e Gumbel estacionárias que foram usadas para os hiperparâmetros de média das distribuições *a priori* adotadas para Gumbel (Equação 3.4) e GEV (Equação 3.3).

Tabela 4 – Estimativas dos parâmetros da distribuição GEV e Gumbel calculadas via método de máxima verossimilhança, para os dados de risco país da Argentina, México e a média dos dois países (MED).

	GEV			Gumbel	
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	$\hat{\mu}$	$\hat{\sigma}$
Argentina	114,07	34,33	-0,31	108,73	31,39
México	212,46	125,15	0,07	218,00	130,00
MED	170,67	68,05	-0,01	170,59	67,92

Fonte: Elaborada pela autora

A Tabela 5 apresenta as variâncias encontradas nos ajustes com os dados da Argentina, México e a média entre dois países.

Tabela 5 – Variâncias das estimativas dos parâmetros da GEV e Gumbel calculadas via Método de Máxima Verossimilhança, para os dados de risco Brasil máximo da Argentina, México e da média dos dois países (MED).

	GEV			Gumbel	
	Var($\hat{\mu}$)	Var($\hat{\sigma}$)	Var($\hat{\xi}$)	Var($\hat{\mu}$)	Var($\hat{\sigma}$)
Argentina	73,11	41,12	0,04	52,64	27,56
México	1410,53	966,31	0,13	892,71	538,56
MED	322,07	187,68	0,06	243,94	141,19

Fonte: Elaborada pela autora

As estimativas apresentadas na Tabela 4 e as variâncias expostas na Tabela 5 foram usadas como valores dos hiperparâmetros das distribuições adotadas como *prioris* da Gumbel (Equação 3.4) e da GEV (Equação 3.3), de acordo com as estruturas da Tabela 1. Para exemplificar, as estruturas I e II das distribuições *a priori* para os parâmetros (μ e σ) da Gumbel usando informação da Argentina com variância multiplicada por 2 (poderia ser 1, 2, 4 ou 6 que foram as flexibilizações usadas) são:

Priori I

$$\begin{aligned}\mu &\sim N(m_\mu, \tau_\mu) \sim N\left(108,73, \frac{1}{2 * 52,64}\right) \\ \sigma &\sim N(m_\sigma, \tau_\sigma) \sim N\left(31,39, \frac{1}{2 * 27,56}\right); \end{aligned}$$

Priori II

$$\begin{aligned}\mu &\sim N(m_\mu, \tau_\mu) \sim N\left(108,73, \frac{1}{2 * 52,64}\right) \\ \sigma &\sim N(\ln(m_\sigma), \tau_\sigma) \sim N\left(\ln(31,39), \frac{1}{2 * 27,56}\right). \end{aligned}$$

Então, a seguir são apresentados os resultados obtidos pelos modelos Bayesianos com *prioris* informativas. Primeiramente, serão mostrados os modelos que foram ajustados com elicitación das *prioris* com base nas informações do risco país da Argentina. Na Tabela 6 são apresentadas as estimativas para os níveis de retorno, HPD de 95%, amplitude dos intervalos e amplitude média, número de acertos, DIC e o erro médio de predição.

Tabela 6 – Estimativas, intervalos HPD de 95%, amplitude dos intervalos (Amp), amplitude média (Amp_m), número de acertos de acurácia ($N^o A$), DIC e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com estimação por inferência Bayesiana e informação *a priori* da Argentina (ARG) para as estruturas I e II.

Modelo	TR	NR	HPD _{95%}	Amp	$N^o A$	DIC	Amp_m	EMP	
Gumbel	I _{ARG} (1var)	2 3	151,49 185,89	[137,80; 165,60] [171,80; 200,70]	27,80 28,90	0	433,50	28,35	52,75%
	II _{ARG} (1var)	2 3	152,03 180,60	[138,30; 166,00] [166,50; 194,60]	27,70 28,10	0	461,50	27,90	53,41%
	I _{ARG} (2var)	2 3	167,45 207,08	[147,40; 186,20] [186,80; 226,20]	38,80 39,40	0	406,70	39,10	47,54%
	II _{ARG} (2var)	2 3	169,62 203,72	[149,40; 188,30] [183,90; 222,70]	38,90 38,80	0	421,60	38,85	47,71%
	I _{ARG} (4var)	2 3	190,51 235,50	[163,60; 217,20] [208,80; 261,90]	53,60 53,10	0	384,20	53,35	40,33%
	II _{ARG} (4var)	2 3	194,44 234,09	[167,80; 221,20] [207,90; 259,90]	53,40 52,00	0	391,10	52,70	39,98%
	I _{ARG} (6var)	2 3	208,01 255,83	[176,50; 239,90] [224,90; 286,40]	63,40 61,50	0	373,00	62,45	35,04%
	II _{ARG} (6var)	2 3	212,64 255,25	[180,80; 243,20] [225,00; 284,70]	62,40 59,70	0	377,00	61,05	34,47%
	I _{ARG} (1var)	2 3	218,50 350,80	[180,32; 256,72] [274,07; 425,28]	76,40 151,21	1	365,64	113,81	20,26%
	II _{ARG} (1var)	2 3	215,80 344,90	[176,84; 256,02] [269,66; 424,35]	79,18 154,69	1	365,69	116,94	21,47%
	I _{ARG} (2var)	2 3	216,90 343,20	[177,04; 257,67] [272,34; 431,50]	80,63 159,16	1	365,23	119,90	21,56%
	II _{ARG} (2var)	2 3	215,20 339,10	[177,72; 258,97] [261,85; 418,31]	81,25 156,46	1	365,17	118,86	22,36%
	I _{ARG} (4var)	2 3	220,70 338,90	[179,77; 264,62] [266,86; 418,39]	84,85 151,53	1	363,59	118,19	21,63%
	II _{ARG} (4var)	2 3	219,80 336,90	[177,83; 262,04] [267,39; 423,69]	84,21 156,30	1	363,64	120,26	22,02%
	I _{ARG} (6var)	2 3	227,00 335,80	[181,85; 267,20] [264,23; 411,54]	85,35 147,31	1	361,03	116,33	21,17%
	II _{ARG} (6var)	2 3	226,40 334,00	[181,98; 268,47] [266,72; 415,28]	86,49 148,56	1	360,91	117,53	21,51%

Fonte: Elaborada pela autora

Todos os modelos para a distribuição Gumbel apresentados na Tabela 6 não apresentaram acurácia para nenhum tempo de retorno ($N^o A = 0$) e os EMP foram elevados, sendo o menor EMP de 34,47% para a estrutura de *priori* II e variância 6var. Entre os modelos GEV, todos apresentaram acurácia em 1 nível de retorno, sendo que o modelo com estrutura de *priori* I e 1var apresentou menor EMP (20,26%). A amplitude média dos modelos Gumbel foram menores, enquanto os valores de DIC foram os maiores em relação aos modelos GEV.

Os níveis de retorno estimados para os modelos das distribuições Gumbel e GEV com informação *a priori* do risco país do México com estruturas I e II e variâncias multiplicadas por 1, 2, 4 e 6 são apresentados na Tabela 7.

Tabela 7 – Estimativas, intervalos HPD de 95%, amplitude dos intervalos (*Amp*), amplitude média (*Amp_m*), número de acertos de acurácia (Nº A), DIC e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com estimação por inferência Bayesiana e informação *a priori* do México (MEX) para as estruturas I e II.

		TR	NR	HPD _{95%}	<i>Amp</i>	Nº A	DIC	<i>Amp_m</i>	EMP	
Gumbel	I _{MEX} (1var)	2	312,77	[269,20; 352,90]	83,70	1	341,40	89,40	9,73%	
		3	382,26	[335,80; 430,90]	95,10					
	II _{MEX} (1var)	2	303,08	[268,50; 335,80]	67,30	1	345,20	70,85	8,38%	
		3	351,12	[312,50; 386,90]	74,40					
	I _{MEX} (2var)	2	327,24	[277,90; 373,70]	95,80	2	340,70	103,60	9,72%	
		3	396,62	[341,30; 452,70]	111,40					
	II _{MEX} (2var)	2	316,91	[276,10; 355,90]	79,80	1	342,20	85,15	7,58%	
		3	371,05	[326,10; 416,60]	90,50					
	I _{MEX} (4var)	2	338,21	[285,90; 391,90]	106,00	2	340,50	115,95	9,79%	
		3	408,11	[347,10; 473,00]	125,90					
	II _{MEX} (4var)	2	328,84	[282,30; 374,30]	92,00	2	340,90	99,45	8,35%	
		3	388,48	[335,70; 442,60]	106,90					
	I _{MEX} (6var)	2	342,89	[285,90; 396,90]	111,00	2	340,50	122,15	9,85%	
		3	413,24	[347,20; 480,50]	133,30					
	II _{MEX} (6var)	2	334,60	[284,60; 383,50]	98,90	2	340,70	107,40	8,74%	
		3	396,99	[340,60; 456,50]	115,90					
	GEV	I _{MEX} (1var)	2	303,60	[263,19; 346,56]	83,37	1	335,15	104,21	8,81%
			3	366,50	[305,58; 430,63]	125,05				
II _{MEX} (1var)		2	303,10	[259,61; 344,64]	85,03	1	335,31	106,53	8,79%	
		3	365,80	[301,57; 429,63]	128,06					
I _{MEX} (2var)		2	311,30	[263,72; 361,36]	97,64	2	335,10	122,47	9,38%	
		3	378,30	[307,27; 454,56]	147,29					
II _{MEX} (2var)		2	311,10	[266,73; 361,72]	94,99	2	335,01	119,31	9,39%	
		3	378,10	[310,95; 454,58]	143,63					
I _{MEX} (4var)		2	315,90	[267,25; 368,82]	101,57	2	335,16	128,59	9,88%	
		3	386,40	[314,38; 469,98]	155,60					
II _{MEX} (4var)		2	315,80	[264,73; 368,02]	103,29	2	335,25	131,30	9,88%	
		3	386,30	[313,28; 472,59]	159,31					
I _{MEX} (6var)		2	318,60	[266,80; 374,06]	107,26	2	335,13	137,41	10,23%	
		3	391,60	[312,07; 479,62]	167,55					
II _{MEX} (6var)		2	317,90	[266,06; 373,86]	107,80	2	335,35	137,04	10,14%	
		3	390,30	[313,80; 480,07]	166,27					

Fonte: Elaborada pela autora

Os modelos para a distribuição Gumbel e GEV com informação *a priori* do risco país do México (TABELA 7) apresentaram menores EMP que os modelos com informação da Argentina (TABELA 6). Os modelos GEV apresentaram menores DIC, mas as amplitudes médias

dos modelos foram maiores que os respectivos modelos Gumbel. O menor EMP foi do modelo Gumbel com *priori* II e variância multiplicada por 2 (7,58%). Todos os modelos apresentaram pelo menos 1 acerto de acurácia.

Na Tabela 8 são apresentados os níveis de retorno estimados para os modelos das distribuições Gumbel e GEV com informação *a priori* elicitada através da média entre o risco país da Argentina e do México, com as variâncias multiplicadas por 1, 2, 4 e 6.

Tabela 8 – Estimativas, intervalos HPD de 95%, amplitude dos intervalos (*Amp*), amplitude média (*Amp_m*), número de acertos de acurácia (Nº A), DIC e EMP para os níveis de retorno (NR) dos tempos de retorno (TR) de 2 e 3 semestres de acordo com os ajustes da distribuição Gumbel e GEV com inferência Bayesiana com informação *a priori* da média entre o risco país da Argentina e do México (MED) para as estruturas I e II.

		TR	NR	HPD _{95%}	<i>Amp</i>	Nº A	DIC	<i>Amp_m</i>	EMP	
Gumbel	I _{MED} (1var)	2	244,40	[216,60; 270,90]	54,30	0	356,80	55,35	24,37%	
		3	295,58	[267,10; 323,50]	56,40					
	II _{MED} (1var)	2	247,97	[221,90; 271,90]	50,00	0	367,40	50,15	25,24%	
		3	285,79	[260,10; 310,40]	50,30					
	I _{MED} (2var)	2	266,92	[231,40; 299,90]	68,50	1	349,60	69,80	17,59%	
		3	321,46	[286,60; 357,70]	71,10					
	II _{MED} (2var)	2	269,29	[236,80; 297,70]	60,90	0	354,50	61,50	18,51%	
		3	312,54	[281,10; 343,20]	62,10					
	I _{MED} (4var)	2	290,66	[250,90; 331,30]	80,40	1	344,60	82,70	10,51%	
		3	348,33	[305,70; 390,70]	85,00					
	II _{MED} (4var)	2	290,74	[255,30; 325,70]	70,40	1	346,80	72,75	11,71%	
		3	339,64	[301,80; 376,90]	75,10					
	I _{MED} (6var)	2	304,10	[259,10; 344,20]	85,10	1	342,70	88,90	8,35%	
		3	363,70	[317,30; 410,00]	92,70					
	II _{MED} (6var)	2	302,06	[262,80; 339,40]	76,60	1	344,20	72,75	8,07%	
		3	354,33	[311,80; 395,30]	83,50					
	GEV	I _{MED} (1var)	2	256,90	[225,80; 292,34]	66,54	1	348,25	80,05	17,70%
			3	330,70	[275,41; 391,70]	116,29				
II _{MED} (1var)		2	255,90	[222,07; 287,57]	65,50	1	347,81	89,26	18,29%	
		3	327,50	[275,83; 388,85]	113,02					
I _{MED} (2var)		2	275,80	[244,30; 312,90]	68,60	1	339,05	85,16	14,54%	
		3	334,40	[285,93; 387,64]	101,71					
II _{MED} (2var)		2	275,40	[243,89; 312,23]	68,34	1	339,08	86,48	14,68%	
		3	333,90	[284,54; 389,16]	104,62					
I _{MED} (4var)		2	291,40	[249,75; 329,38]	79,63	1	336,20	98,84	9,79%	
		3	352,70	[294,14; 412,18]	118,04					
II _{MED} (4var)		2	291,10	[252,00; 331,14]	79,14	1	336,22	99,00	9,89%	
		3	352,30	[297,14; 416,00]	118,86					
I _{MED} (6var)		2	298,60	[259,60; 344,25]	84,65	1	335,62	108,09	8,99%	
		3	362,80	[299,63; 431,15]	131,52					
II _{MED} (6var)		2	298,30	[257,07; 343,43]	86,36	1	335,66	108,55	9,00%	
		3	362,50	[298,93; 429,67]	130,74					

Fonte: Elaborada pela autora

Os modelos Gumbel e GEV com informação *a priori* da média entre o risco país da Argentina e do México, apresentados na Tabela 8, apresentaram EMP menores que os modelos com informações *a priori* da Argentina (TABELA 6) e maiores que os modelos com informações *a priori* do México (TABELA 7). Entre os modelos Gumbel e GEV, os modelos GEV apresentaram menor DIC e o EMP menor na maioria das estruturas de *priori*, sendo o EMP do modelo Gumbel menor apenas para as *prioris* I e II com 6var. Quanto à amplitude média, os modelos Gumbel tiveram menor amplitude média dos intervalos HPD de 95% de credibilidade.

Em geral, os modelos Gumbel apresentaram menores valores de amplitude média dos intervalos que os modelos GEV, porém os modelos GEV apresentaram menores valores de DIC e de erro médio de predição.

Os modelos Gumbel com informação *a priori* do risco país da Argentina não apresentaram nenhum acerto em acurácia em todas as *prioris*.

A estratégia de flexibilizar a variância, multiplicando-a por 1, 2, 4 e 6, não apresentou um padrão nos resultados, isto é, em alguns casos aumentou o EMP e em outros diminuiu o EMP à medida em que se aumentava o fator multiplicador. Quanto à amplitude média dos intervalos HPD de 95% de credibilidade, a flexibilização da variância fez com que a amplitude média aumentasse. De acordo com Paulino, Turkman e Murteira (2003) o uso de informação *a priori* possibilita a obtenção de intervalos de credibilidade mais precisos (estritos), logo, quando aumenta-se a variabilidade da informação *a priori* perde-se em precisão, que refere-se a amplitude dos intervalos de credibilidade.

Todos os modelos que apresentaram acurácia nos dois níveis de retorno foram organizados em ordem crescente de amplitude e podem ser observados na Tabela 9, assim como o EMP e o DIC de cada um desses modelos.

Tabela 9 – Modelos selecionados a partir dos critérios para verificação de desempenho dos ajustes.

Modelo	DIC	Amp_m	EMP
Gumbel II_{MEX} (4var)	340,90	99,45	8,35%
Gumbel I_{MEX} (2var)	340,70	103,60	9,72%
Gumbel II_{MEX} (6var)	340,70	107,40	8,74%
Gumbel I_{MEX} (4var)	340,50	115,95	9,79%
GEV II_{MEX} (2var)	335,01	119,31	9,39%
Gumbel I_{MEX} (6var)	340,50	122,15	9,85%
GEV I_{MEX} (2var)	335,10	122,47	9,38%
Gumbel _{EMV}	-	125,66	9,24%
GEV I_{MEX} (4var)	335,16	128,59	9,88%
GEV _{EMV}	-	129,43	9,19%
GEV II_{MEX} (4var)	335,25	131,30	9,88%
GEV II_{MEX} (6var)	335,35	137,04	10,14%
GEV I_{MEX} (6var)	335,13	137,41	10,23%
Gumbel _{PNI}	340,90	141,20	10,09%
GEV _{PNI}	335,40	152,37	10,71%

Fonte: Elaborada pela autora

Modelos com menores amplitudes médias entre os modelos com acurácia nos dois níveis de retorno.

Pode-se observar na Tabela 9 que os modelos que apresentaram acurácia nos dois níveis de retorno foram modelos Bayesianos ajustados com distribuições GEV e Gumbel com informação *a priori* obtida do risco país do México e os modelo Gumbel e GEV com *priori* não informativa e ajustados por máxima verossimilhança (EMV). A maior diferença entre os critérios de informação do desvio (DIC) dos modelos Bayesianos foi de 5,89, considerando o menor (335,01) e o maior DIC (340,90). Sendo assim, não houve diferença substancial para descartar modelos. Por isso, não foi possível selecionar modelos considerando apenas o DIC. Então, analisou-se a amplitude e foram selecionados os três modelos com menor amplitude. Sendo eles, o modelo Gumbel com *priori* II com 4 vezes a variância ($Amp_m = 99,45$), o modelo Gumbel com *priori* I com 2 vezes a variância ($Amp_m = 103,60$) e o modelo Gumbel com *priori* II com 6 vezes a variância ($Amp_m = 107,40$), lembrando que na *priori* I usa-se os próprios valores das estimativas nas médias dos parâmetros e na *priori* II usa-se o logaritmo natural da estimativa para a média do parâmetro de escala.

4.2 ANÁLISE DA SÉRIE COMPLETA DE RISCO BRASIL MÁXIMO SEMESTRAL

Os três modelos selecionados foram ajustados para a série completa de dados de 2003-2 a 2017-2. Os resultados desses ajustes podem ser observados na Tabela 10.

Tabela 10 – Resultados dos ajustes dos três modelos selecionados à série completa de risco Brasil máximo semestral (2003-2 a 2017-2), sendo as estimativas dos parâmetros, os intervalos HPD de 95% de credibilidade, o DIC de cada modelo, amplitude (Amp) e amplitude média (Amp_m)

Modelo	Parâmetro	Estimativa	HPD _{95%}	DIC	Amp	Amp_m
I _{MEX} (2var)	Gumbel μ	283,75	[238,90; 324,50]	375,60	85,60	76,93
	σ	122,49	[89,84; 158,10]		68,26	
II _{MEX} (4var)	Gumbel μ	290,69	[250,60; 330,20]	375,70	79,70	68,35
	σ	105,96	[78,89; 135,90]		57,01	
II _{MEX} (6var)	Gumbel μ	294,05	[250,90; 334,90]	375,50	84,00	72,97
	σ	110,16	[80,35; 142,30]		61,95	

Fonte: Elaborada pela autora

Conforme pode ser observado na Tabela 10, os três modelos ajustados apresentaram critério de desvio de informação (DIC) bem próximos, com diferença menor que cinco, por isso não foi possível selecionar o melhor ajuste por esse critério. Então, o melhor modelo foi selecionado com base na amplitude média dos intervalos HPD dos parâmetros.

Com isso, o modelo selecionado, como melhor modelo para descrever probabilisticamente os dados de risco Brasil, foi o ajuste com distribuição Gumbel e informação *a priori* do risco México com estrutura de *priori* II e a variância multiplicada por 4. A *priori* utilizada para este modelo selecionado foi a estrutura II com informações do risco país do México e é dada por:

$$\mu \sim N \left(218, \frac{1}{4 * 892,71} \right)$$

$$\sigma \sim N \left(\ln(130), \frac{1}{4 * 538,56} \right).$$

As estimativas dos parâmetros, seus respectivos intervalos HPD de 95% e os resultados dos critérios de convergência para esse modelo são apresentados na Tabela 11.

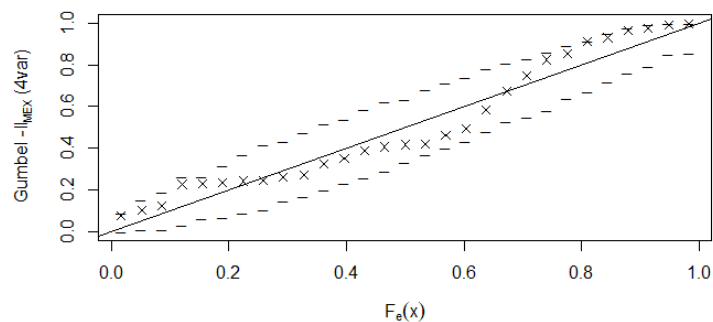
Tabela 11 – Estimativa dos parâmetros e os resultados dos critérios de Geweke (CG), de Raftery e Lewis (R-L) e de Heidelberger e Welch (H-W) para os parâmetros do modelo selecionado, Gumbel com informação *a priori* do risco país do México, estrutura II e variância multiplicada por 4.

Modelo	Parâmetro	Estimativa	HPD _{95%}	Critérios de Convergência		
				CG	R-L	H-W
Gumbel	μ	290,69	[250,60; 330,20]	0,027	0,987	0,248
Π_{MEX} (4var)	σ	105,96	[78,89; 135,90]	0,347	1,010	0,485

Fonte: Elaborada pela autora

Para verificar a adequabilidade do modelo selecionado aos dados, foi realizada a análise do gráfico P-P apresentado na Figura 4.

Figura 4 – Gráfico P-P para o modelo selecionado para descrever o risco Brasil máximo semestral.



Fonte: Elaborada pela autora

Pela análise do gráfico probabilidade-probabilidade, que tem no eixo das abscissas (x) a função densidade de probabilidade empírica e no eixo das ordenadas (y) a função densidade de probabilidade da distribuição Gumbel, o modelo está bem ajustado aos dados, uma vez que apresenta tendência linear e a amostra observada está dentro dos limites de confiança do envelope simulado. Então, esse modelo Gumbel Bayesiano com *priori* do México, estrutura II e variância multiplicada por 4, foi utilizado para realizar previsões de níveis de retorno de risco Brasil máximo semestral e para calcular probabilidades de ocorrência de valores altos de risco Brasil. Na Tabela 12, são apresentados os níveis de retorno para os tempos de 2 a 10 semestres estimados através do modelo selecionado.

Tabela 12 – Níveis de retorno para os tempos de 2 a 10 semestres estimados através do modelo selecionado.

Tempos de retorno	Níveis de retorno	HPD _{95%}	Critérios de Convergência		
			CG	R-L	H-W
2	329,53	[285,80; 369,90]	0,490	1,020	0,236
3	386,35	[336,40; 435,00]	0,543	1,020	0,321
4	422,71	[370,30; 480,90]	0,554	0,998	0,394
5	449,63	[391,00; 511,70]	0,552	0,987	0,450
6	471,04	[408,40; 537,80]	0,550	0,993	0,492
7	488,82	[422,30; 559,50]	0,548	1,010	0,525
8	504,04	[434,50; 578,40]	0,545	0,998	0,553
9	517,34	[444,60; 594,70]	0,542	1,000	0,577
10	529,15	[453,60; 608,80]	0,538	1,020	0,597

Fonte: Elaborada pela autora

A interpretação prática, considerando a Tabela 12, para o tempo de retorno de 10 semestres é que se espera, em um tempo médio de 10 semestres, que ocorra pelo menos um dia em que o valor do risco Brasil seja maior ou igual a 529,15 pontos, ou seja, espera-se com 95% de credibilidade que o risco Brasil seja um valor maior ou igual a um valor que esteja entre 453,60 e 608,80 pontos.

E ainda, foram calculadas as probabilidades de ocorrência de valores máximos de risco Brasil. Os valores para os quais foram calculadas as probabilidades de ocorrência foram definidos com base no maior valor observado na série de risco Brasil máximo semestral (898 pontos-base no segundo semestre de 2003), sendo eles os quantis 718,40; 898,00; 1.077,60; 1.257,20; 1.436,80; 1.616,40 e 1.796,00. Os resultados podem ser observados na Tabela 13.

Tabela 13 – Probabilidades do risco Brasil (RB) ser maior que os quantis (Q) definidos, intervalos HPD_{95%} e critérios de convergência, calculadas a partir do melhor modelo.

Quantis (Q)	$P(RB > Q)$	HPD _{95%}	Critérios de Convergência		
			CG	R-L	H-W
718,40	1,9201%	[0,2615%; 4,2050%]	0,305	1,010	0,700
898,00	0,3992%	[0,0108%; 1,0650%]	0,022	1,010	0,790
1077,60	0,0859%	[0,0007%; 0,2726%]	0,317	1,010	0,825
1257,20	0,0192%	[0,0001%; 0,0699%]	0,574	1,000	0,844
1436,80	0,0045%	[0,0000%; 0,0179%]	0,792	0,995	0,842
1616,40	0,0011%	[0,0000%; 0,0046%]	0,975	0,995	0,820
1796,00	0,0003%	[0,0000%; 0,0012%]	1,123	1,000	0,780

Fonte: Elaborada pela autora

A interpretação para as probabilidades calculadas, por exemplo, para o quantil de 718,40 pontos, a probabilidade de ocorrência de um valor de risco Brasil maior ou igual a 718,40 pontos-base é de 1,9201%.

A percepção de risco mais elevada (medida do risco Brasil alta) pode causar uma retração nos investimentos por meio de saída de capitais ou interrupção na entrada, conforme ressalta Leichsenring (2004). E ainda, o aumento do risco percebido provoca movimentações correspondentes na taxa de câmbio, devido a correlação positiva entre as duas variáveis.

A taxa de câmbio é o indexador (taxa de reajuste utilizada para acompanhar a atividade econômica, corrigir preços e evitar volatilidade) de grande parte da dívida total brasileira (pública e privada). Sendo o risco Brasil positivamente correlacionado com a taxa de câmbio, segundo Leichsenring (2004), a sua alta refletirá no aumento da taxa de câmbio, elevando assim a proporção da dívida em relação ao PIB, ocasionando em uma retroalimentação na percepção de risco num ciclo vicioso.

De acordo com Moreira e Rocha (2003), há uma interação dinâmica entre risco Brasil, taxa de câmbio e taxa de juros domésticas. Segundo Oliveira (2002), existe uma relação entre a taxa de juros doméstica (taxa Selic) e o risco Brasil, isto é, que variações no risco Brasil são acompanhadas de variações na taxa Selic, e vice-versa. Conforme, salientado por Garcia e Didier (2003), a taxa de juros tem papel fundamental na determinação do nível de atividade, do emprego, da taxa de câmbio e outras variáveis econômicas. Sabendo dessa relação entre risco Brasil e taxa de juros, as previsões sobre risco Brasil, assim como as probabilidades de ocorrência de valores extremos, conforme as Tabelas 12 e 13, podem ser de grande relevância para o governo na elaboração de políticas monetárias para que as taxas de juros sejam calibradas, evitando perturbações na economia e a retração do fluxo de investimentos estrangeiros.

A partir dos resultados deste artigo, pode-se considerar que a distribuição Gumbel seja a melhor distribuição quando comparada à GEV para descrever o comportamento de risco Brasil máximo semestral, uma vez que os modelos ajustados com a distribuição Gumbel apresentaram menores erros médios de predição e maiores níveis de acurácia (acertos).

Com os resultados obtidos, pode-se dizer que o uso da inferência Bayesiana melhorou o processo de inferência em relação ao método da máxima verossimilhança, pois os modelos Bayesianos apresentaram maior precisão e menores erros médios de predição.

A informação *a priori* obtida do risco país do México apresentou melhores resultados que as informações da Argentina e da média do risco dos dois países, pois apresentou as menores amplitudes nos intervalos HPD de 95% indicando precisão dos modelos e acurácia nos dois níveis de retorno em 11 dos 16 modelos ajustados a partir da informação *a priori* do México. Ou seja, a *a priori* do risco México apresentou resultados mais acurados e precisos. Em contra-

partida, os modelos ajustados com a informação *a priori* elicitada a partir dos dados do risco país da Argentina e da média do risco dos dois países apresentaram apenas 0 ou 1 acerto em acurácia e erros médios de predição elevados, se comparados aos resultados obtidos com *priori* do México. Alguns modelos com informação *a priori* da Argentina tiveram erros médios de predição quase 5 vezes (52,75% e 53,41%) maiores que o EMP do modelo escolhido (8,35%).

A flexibilização da variância aumentou a amplitude média dos intervalos HPD de 95% de credibilidade e, em alguns casos, como no uso da informação *a priori* da Argentina, diminuiu o EMP, mas em outros casos aumentou o EMP, como no uso da informação do México para a distribuição GEV.

Alguns autores encontraram resultados que indicaram que a inferência Bayesiana é importante na modelagem de extremos, como Beijo, Vivanco e Muniz (2009), que observaram resultados melhores usando inferência Bayesiana em comparação com o resultado usando estimação por máxima verossimilhança em análise de precipitações máximas. Esses resultados também concordam com os resultados encontrados por Coles e Powell (1996) para velocidade do vento, que observaram melhores resultados para estimativas dos parâmetros da GEV e para a velocidade máxima quando usaram inferência Bayesiana em comparação com máxima verossimilhança. Existem também trabalhos que envolvem aplicações na área econômico-financeira que apresentaram melhores resultados ao usar inferência Bayesiana, como Otrók e Whiteman (1998), que fizeram predição das condições econômicas de um estado nos EUA; Giannone et al. (2014), que concluíram a partir de seus resultados que inferência Bayesiana pode melhorar os resultados de projeção da inflação. O ganho da inferência Bayesiana em relação ao método de máxima verossimilhança é proporcionado pela incorporação de informação *a priori* ao processo de inferência, pois melhora o processo de inferência ao reduzir as incertezas a respeito dos parâmetros.

5 CONCLUSÃO

Com os resultados obtidos, pode-se concluir que, a série de risco Brasil não apresentou tendência linear. O uso da inferência Bayesiana no processo de estimação apresentou melhores resultados que o método da máxima verossimilhança. Nos modelos Bayesianos, as estruturas de *priori* I, na qual usou-se os próprios valores das estimativas nas médias da distribuição *a priori* para os parâmetros de posição e escala, e *priori* II, na qual usou-se o logaritmo natural da estimativa para a média do parâmetro de escala, apresentaram melhores resultados com informação *a priori* elicitada a partir dos dados do risco país do México.

O modelo selecionado como mais adequado para descrever o risco Brasil semestral foi a distribuição Gumbel com informação *a priori* do risco país do México, com a estrutura de *priori* II, na qual foi utilizado o logaritmo natural na média do parâmetro de escala, e a variância multiplicada por 4.

Os níveis de retorno e probabilidades calculados com o modelo mais adequado podem ser usados pelo governo para planejamento de políticas públicas que a curto ou longo prazo evitem a retração de investimentos estrangeiros e aumentem o nível desses investimentos no país, evitando também o aumento da taxa de câmbio e taxa Selic.

Como estudos futuros, pretende-se obter dados sobre risco país de outros países emergentes e incorporar o uso dessas informações como *prioris* na análise do risco país do Brasil.

6 REFERÊNCIAS

BANCO CENTRAL DO BRASIL. Risco Brasil. **Série Perguntas Mais Frequentes**. Diretoria de Política Econômica, Departamento de Relacionamento com Investidores e Estudos Especiais, 2016.

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**, v. 33, p. 261–270, 2009.

BERNARDINI, D. F. **Inferência Bayesiana para valores extremos**. 2010. 69 f. Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica, Campinas, SP.

CLARKE, R. T. Estimating time trends in gumbel-distributed data by means of generalized linear models. **Water Resources Research**, Wiley Online Library, v. 38, n. 7, 2002.

COLES, S. G. **An introduction to statistical modeling of extreme values**. London: Springer, 2001.

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

COUTO, S. V. V.; MEURER, R. Os impactos da política monetária na taxa de câmbio no Brasil: identificação via heterocedasticidade. **Revista de Economia Contemporânea**, v. 21, n. 2, 2017.

DURBIN, J.; WATSON, G. S. Testing for serial correlation in least squares regression. II. **Biometrika**, JSTOR, v. 38, n. 1/2, p. 159–177, 1951.

FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In: CAMBRIDGE UNIV PRESS. **Mathematical Proceedings of the Cambridge Philosophical Society**, v. 24, n. 02, p. 180–190, 1928.

GARCIA, M. G. P.; DIDIER, T. Taxa de juros, risco cambial e risco Brasil. **Pesquisa e Planejamento Econômico**, v. 33, n. 2, p. 253–297, 2003.

GIANNONE, D. et al. Short-term inflation projections: a bayesian vector autoregressive approach. **International journal of forecasting**, Elsevier, v. 30, n. 3, p. 635–644, 2014.

IPEA-DATA. **Instituto de Pesquisa Econômica Aplicada**. 2017. Disponível em: <http://www.ipeadata.gov.br>. Acesso em: 27 mar. 2017.

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

KENDALL, M. G. **Rank Correlation Methods**. 4. ed. London: Charles Griffen, 1975.

KNOEMA. **World Bank Global Economic Monitor, April 2014**. 2017. Disponível em: <https://pt.knoema.com/atlas>. Acesso em: 18 mar. 2017.

LEICHSENDRING, D. R. **Endogeneidade e mecanismos de transmissão entre a taxa de juros doméstica e o risco soberano: uma revisita aos determinantes do risco Brasil**. 2004. Dissertação (Mestrado em Economia) — Universidade de São Paulo, São Paulo, 2004.

MANN, H. B. Nonparametric tests against trend. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 245–259, 1945.

MOREIRA, A. R.; ROCHA, K. Determinantes do risco brasil: fundamentos e expectativas: uma abordagem de modelos de risco de crédito. **Instituto de Pesquisa Econômica Aplicada (Ipea)**, 2003.

MOREIRA, J. G. d. V.; NAGHETTINI, M.; ELEUTÉRIO, J. C. Frequency and risk in nonstationary pluviometric records in the drainage basin of tarauacá river, acre. **RBRH**, SciELO Brasil, v. 21, n. 1, p. 232–241, 2016.

NELDER, J. A.; MEAD, R. A. A simplex method for function minimization. **The computer journal**, v. 7, p. 308–313, 1965.

NOGUEIRA, D. A.; SÁFADI, T.; FERREIRA, D. F. Avaliação de critérios de convergência para o método de monte Carlo via cadeias de Markov. **Revista Brasileira de Estatística**, IBGE, v. 65, n. 224, 2004.

NUNES, M. S. et al. Co-integração e causalidade entre variáveis macroeconômicas, "risco brasil" e retornos no mercado de ações brasileiro. **Revista de economia e administração**, v. 2, n. 3, 2003.

OLIVEIRA, A. X. d. Mudanças na taxa de juros e mudanças no risco brasil. **FGV EESP - Textos para Discussão / Working Paper Series**, 2002.

OLIVEIRA, D. R. **Valor em risco operacional com cópulas de valores extremos**. 2012. 59 f. Monografia (Bacharelado em Estatística) - Universidade de Brasília, Brasília, 2012.

OTROK, C.; WHITEMAN, C. H. Bayesian leading indicators: measuring and predicting economic conditions in iowa. **International Economic Review**, JSTOR, p. 997–1014, 1998.

PAULINO, C. D. M.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: FCG, 2003.

PLUMMER, M. et al. Coda: Convergence diagnosis and output analysis for mcmc. **R News**, v. 6, n. 1, p. 7–11, 2006. Disponível em: <<https://journal.r-project.org/archive/>>. Acesso em: 21 mar. 2018.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>. Acesso em: 21 de mar. 2018.

SILVA, R. R. **A distribuição Generalizada de Pareto e mistura de distribuições de Gumbel no estudo da vazão e da velocidade máxima do vento em Piracicaba, SP**. 2008. 101 f. Dissertação (Mestrado em Agronomia) - Universidade de São Paulo, Piracicaba, 2008.

SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 64, n. 4, p. 583-639, 2002.

STEPHENSON, A.; RIBATET., M. **evdbayes: Bayesian Analysis in Extreme Value Theory**. R package version 1.1-1, 2014. Disponível em: <<https://CRAN.R-project.org/package=evdbayes>>. Acesso em: 21 mar. 2018.

STEPHENSON, A. G. evd: Extreme value distributions. **R News**, v. 2, n. 2, 2002. Disponível em: <<http://CRAN.R-project.org/doc/Rnews/>>. Acesso em: 21 mar. 2018.

STURTZ, S.; LIGGES, U.; GELMAN, A. R2winbugs: A package for running winbugs from r. **Journal of Statistical Software**, v. 12, n. 3, p. 1–16, 2005. Disponível em: <<http://www.jstatsoft.org>>. Acesso em: 21 mar. 2018.

THOMAS, A.; O'HARA, R. B. **OpenBUGS**. 2004.

TOLEDO, J. E. C. Risco brasil: o efeito-lula e os efeitos-banco central. **Revista de Economia Política**, v. 22, n. 3, 2002.

TRAPLETTI, A.; HORNIK, K. **tseries: Time Series Analysis and Computational Finance**. 2017. R package version 0.10-42. Disponível em: <<https://CRAN.R-project.org/package=tseries>>. Acesso em: 21 mar. 2018.

CAPÍTULO 3 - MODELAGEM DE RESSEGURO VIA DISTRIBUIÇÃO GENERALIZADA DE PARETO UTILIZANDO AS ABORDAGENS CLÁSSICA E BAYESIANA

Thaís Brenda Martins; Luiz Alberto Beijo; Davi Butturi-Gomes

RESUMO: O resseguro é uma operação contratada por uma seguradora para transferência de parte ou totalidade de seus riscos, mediante o pagamento de um prêmio a uma resseguradora. Um dos tipos de contratos é o não proporcional em que o ressegurador se responsabiliza por valores que ultrapassem o limite de retenção, determinado pela seguradora, até um limite máximo de cobertura, ambos preestabelecidos no contrato. As perdas da resseguradora relativas a sinistros ocorridos que ultrapassem o limite de retenção determinado pela seguradora podem ser modelados pela distribuição generalizada de Pareto. Essa distribuição é usada para descrever o comportamento dos excessos de valores extremos que ultrapassem um determinado valor alto, chamado de limiar (que vem do inglês *threshold*). A inferência Bayesiana tem sido utilizada na teoria de valores extremos para estimação de parâmetros, um dos motivos é a vantagem que essa abordagem possui sobre outros métodos de estimação ao permitir incorporar informações passadas ao processo de inferência que reduzem as incertezas das estimativas dos parâmetros. Portanto, o objetivo deste capítulo é modelar uma série de perdas de resseguro por meio da distribuição generalizada de Pareto, comparando a estimação Bayesiana com *priori* não informativa com a estimação via máxima verossimilhança. Para modelagem, foram usados dados sobre reclamações automobilísticas de companhias seguradoras europeias que excederam 1,2 milhões de euros no período de 1988 a 2001 e estão disponíveis no pacote *ReIns* do *freeware* R. Para verificar a adequabilidade dos ajustes, foram utilizados gráficos probabilidade-probabilidade. Foram calculadas probabilidades de ocorrência de reivindicações com valores extremos, uma vez que as ocorrências de sinistros com valores extremos podem influenciar na solvência de uma resseguradora. A inferência Bayesiana possibilitou calcular intervalos HPD para essas probabilidades calculadas, o que foi uma vantagem sobre o método de máxima verossimilhança. O uso da inferência Bayesiana, mesmo com *priori* não informativa, mostrou-se melhor ao permitir a estimação dos intervalos de credibilidade HPD.

Palavras-chave: Distribuição generalizada de Pareto, resseguro, inferência Bayesiana.

REINSURANCE MODELING VIA GENERALIZED PARETO DISTRIBUTIONS USING CLASSIC AND BAYESIAN APPROACHES

Thaís Brenda Martins; Luiz Alberto Beijo; Davi Butturi-Gomes; Reinaldo Antônio Gomes Marques

ABSTRACT: Reinsurance is an operation contracted by an insurer to transfer part or all of its risks, upon payment of a premium to a reinsurer. One of the types of contracts is the non-proportional one in which the reinsurer is responsible for amounts that exceed the retention limit determined by the insurer up to a maximum coverage limit, both pre-established in the contract. The reinsurer's losses relating to claims incurred that exceed the retention limit determined by the insurer may be modeled by the generalized Pareto distribution. This distribution is used to describe the behavior of excesses of extreme values that exceed a certain high value, called a threshold. Bayesian inference has been used in the theory of extreme values for parameter estimation, one reason is the advantage that this approach has over other estimation methods by allowing to incorporate information passed to the process of inference that reduces the uncertainties of the parameter estimates. Therefore, the objective of this chapter is to model a series of reinsurance losses through the generalized Pareto distribution, comparing the Bayesian estimation with non-informative priori with the estimation via maximum likelihood. For modeling, data were used on automobile claims of European insurance companies that exceeded 1.2 million euros in the period from 1988 to 2001 and are available in the ReIns package of freeware R. To verify the suitability of the adjustments, probability-probability . Odds of claims with extreme values were calculated, since the occurrence of claims with extreme values can influence the solvency of a reinsurer. Bayesian inference made it possible to calculate HPD intervals for these calculated probabilities, which was an advantage over the maximum likelihood method. The use of Bayesian inference, even with prior non-informative, was shown to be better in allowing the estimation of HPD credibility intervals.

Keywords: *Generalized Pareto distribution, reinsurance, Bayesian inference.*

1 INTRODUÇÃO

O resseguro é a transferência de um risco técnico, parcial ou total, da seguradora para a resseguradora, ou seja, o resseguro é um seguro para o risco assumido pela seguradora (LUCAS FILHO, 2011).

As pessoas e seus bens estão expostos a inúmeros riscos no dia a dia. Por meio do pagamento de um valor previamente calculado, o segurado (pessoa física ou jurídica) transfere esses riscos a um segurador. O resseguro é um seguro feito pelas seguradoras e é usado para diluir o risco da seguradora, diminuir a necessidade de capital da seguradora, otimizar a sua saúde financeira e proteger a seguradora contra riscos catastróficos. Portanto, não existe um contrato de resseguro sem a existência do seguro (FUNDAÇÃO MAPFRE, 2011).

De acordo com Cabral (1988), o primeiro contrato de resseguro ocorreu na Itália, em Gênova, em julho de 1370, e se referia ao seguro de um carregamento que seria transportado pelo mar de Cádiz para Sluis nos Países Baixos. Porém, devido à periculosidade da travessia, o segurador transferiu a maior parte do risco para um segundo segurador. Isso resultou em um resseguro entre segurador e ressegurador, sem que o proprietário da carga tivesse relação contratual com o ressegurador.

Os contratos de resseguro podem ser classificados em: facultativos e obrigatórios (ou automáticos) e simultaneamente em contratos proporcionais e não proporcionais. Conforme Ferreira (2007), nos contratos proporcionais "os prêmios e sinistros são divididos proporcionalmente entre o segurador e o ressegurador numa proporção preestabelecida". Já nos não proporcionais, não há valores segurados cedidos, e sim limites pré-definidos da participação do ressegurador nos sinistros. Um exemplo de contrato não proporcional é o contrato de excesso de danos ou *excess of loss*, sendo que esse tipo de contrato será o foco deste capítulo.

Nos contratos de resseguros não proporcionais de excesso de danos, a resseguradora paga a diferença que excede um determinado valor, preestabelecido, denominado limite de retenção (R). Esse valor é o máximo que a seguradora suporta, sendo que acima dele é responsabilidade da resseguradora. Logo, intuitivamente, pressupõe-se que o limite de retenção seja um valor alto. Sendo assim, os sinistros podem ser vistos como ocorrências de valores extremos máximos e podem ser descritos pela distribuição generalizada de Pareto (GP), por se tratar de uma distribuição de extremos acima de um limiar, que pode ser considerado o limite de retenção da seguradora.

Determinar o valor desse limiar pode ser uma dificuldade ao se trabalhar com a distribuição GP, porém, autores como Coles (2001) e Silva (2008) descrevem e apontam métodos para seleção do limiar da distribuição generalizada de Pareto, como o método do gráfico da vida média residual, também conhecido como gráfico da média de excessos.

No que diz respeito a estimação de parâmetros, de acordo com Casella e Berger (2010), entre os vários métodos existentes, o método de máxima verossimilhança é o mais popular e tem sido amplamente utilizado. Segundo Naguettini e Pinto (2007), esse método é considerado o mais eficiente ao produzir estimadores de menor variância. Porém, suas boas propriedades são assintóticas, ou seja, para amostras pequenas pode fornecer estimadores de qualidade equivalente ou menor que os obtidos por outros métodos. A inferência Bayesiana é uma abordagem que tem sido utilizada para contornar essa limitação de se trabalhar com amostras pequenas, pois permite incorporar informações passadas ao processo de estimação, o que melhora a inferência sobre os parâmetros, reduzindo as incertezas das estimativas (PAULINO; TURKMAN; MURTEIRA, 2003).

O objetivo deste capítulo é ajustar a distribuição generalizada de Pareto aos dados de resseguro disponíveis no pacote *ReIns* (REYNKENS; VERBELEN, 2017) do *freeware* R (R CORE TEAM, 2017) e comparar os resultados obtidos usando estimação por inferência Bayesiana e o método de máxima verossimilhança. Para isso, selecionou-se três limiares para os dados para ajustar a distribuição generalizada de Pareto. Os dados referem-se a reivindicações de seguros automotivos do período de 1988 a 2001 ocorridos na Europa. Através dos modelos da distribuição generalizada de Pareto ajustados com estimação por inferência Bayesiana e por máxima verossimilhança, foi calculado a probabilidade de ocorrência de sinistros com valores altos, sendo que a inferência Bayesiana proporciona uma vantagem sobre o método de máxima verossimilhança ao possibilitar a estimação dos intervalos de credibilidade HPD para as probabilidades calculadas.

2 REVISÃO DE LITERATURA

2.1 RESSEGURO

Risco é um evento futuro e incerto, cuja ocorrência pode resultar em prejuízos de natureza econômica. Nesse sentido, é intrínseca a prevenção ao risco. A ideia de prevenção ao risco é o que norteia a operação de seguros (FRANCO, 1993).

As pessoas, sejam elas físicas ou jurídicas, estão expostas a inúmeros riscos que podem lhes atingir ou acometer seus bens. Logo, contratam um seguro, mediante pagamento de um prêmio, para se prevenir das perdas financeiras decorrentes de algum risco. De acordo com Grosman (1990), o resseguro é a cessão de uma parte dos riscos assumidos por uma seguradora perante os segurados, com base em contratos de seguros, a uma instituição, sem nenhuma relação direta com o segurado - o ressegurador. Essa definição é comumente simplificada ao definir o resseguro como o seguro da seguradora.

Segundo Rios (2005), "o resseguro é um mecanismo de pulverização de riscos, no qual a seguradora transfere a outrem, total ou parcialmente, o risco assumido". Além da pulverização de riscos, a operação de resseguro pode ter as funções mercadológica, gerencial, financeira e educacional.

A função mercadológica ou estrutural corresponde à geração de capacidade para a seguradora cedente visando alavancar seus negócios. Na função gerencial, o objetivo é a proteção direta da seguradora cedente e indireta de seus clientes primários, ou seja, os seus segurados. Na função financeira, o resseguro é usado para mitigar os efeitos das flutuações de resultados nos negócios das companhias seguradoras. E na função educacional, o resseguro é um elemento de aprimoramento na troca e disseminação de novas tecnologias securitárias, na assessoria e na capacitação das companhias seguradoras (HADDAD, 2003).

A companhia seguradora que transfere seus riscos é chamada de seguradora ou cedente e a companhia que aceita esses riscos é chamada de resseguradora ou cessionária. O contrato entre as duas companhias é denominado de contrato de resseguro e o valor pago por essa transferência de risco é chamado de prêmio de resseguro.

De acordo com a Fundação MAPFRE (2011), os contratos de resseguros podem ser divididos em facultativos e obrigatórios e, simultaneamente, em proporcionais e não proporcionais.

Quando facultativo, a cedente e o ressegurador têm liberdade de decisão, a cedente para ceder ou não um risco, e o ressegurador para aceitar ou não a operação oferecida. O resseguro facultativo é indicado para riscos isolados que possuem severidade alta, ou seja, perdas e danos elevados, como por exemplo, acidentes em plataformas de petróleo (FUNDAÇÃO MAPFRE, 2011).

No resseguro obrigatório, a cedente oferece uma carteira inteira e a resseguradora pode aceitar ou não esse contrato de resseguro, mas uma vez estabelecidas as condições do contrato, o ressegurador não pode retirar riscos individuais da carteira, sendo obrigado a aceitar todos os sinistros decorrentes de riscos daquela carteira (FUNDAÇÃO MAPFRE, 2011).

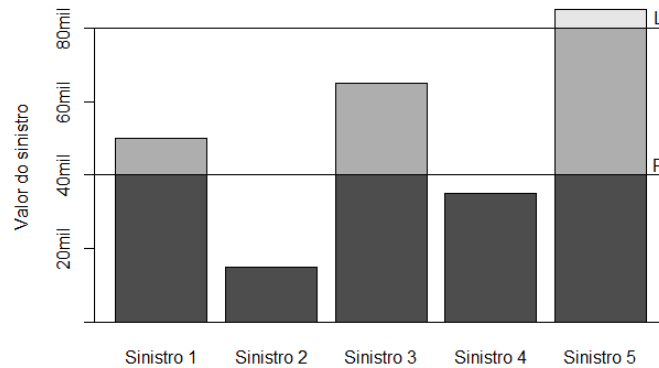
Nos contratos proporcionais, o ressegurador aceita a responsabilidade sobre uma parte fixa e proporcional de todos e cada um dos riscos incluídos. Assim, deve participar de uma quota determinada dos sinistros ocorridos e também receber uma participação dos prêmios na mesma proporção. As duas modalidades de resseguro proporcional são a quota-parte e excedente de responsabilidade (FERREIRA, 2007).

Já nos contratos não proporcionais, a divisão de responsabilidades é estabelecida sobre os sinistros e não sobre a importância segurada. O ressegurador se compromete a indenizar o ressegurado de todos os sinistros ou grupos de sinistros cujos valores ultrapassem uma quantia previamente fixada, denominada limite de retenção ou prioridade, até um determinado valor, chamado limite de cobertura. Os contratos não proporcionais têm sido denominados de resseguros de excesso de danos (*excess of loss reinsurance*), e também são conhecidos como contratos XL (FUNDAÇÃO MAPFRE, 2011).

Grande parte dos contratos de resseguro não proporcionais são baseados na estrutura de excesso de danos (PARODI, 2014), sendo eles os contratos de excesso de danos por risco (XL), de excesso de danos agregados, de excesso de danos por catástrofe e os contratos de excesso de sinistralidade ou *stop-loss*. Neste capítulo, serão apresentados os contratos de excesso de danos por risco.

Neste tipo de resseguro, a seguradora é indenizada por qualquer perda ocorrida que seja superior a um valor pré-determinado R, chamado de limite de retenção, até um limite de cobertura L (há casos em que o contrato é ilimitado, então não há um limite de cobertura) (PARODI, 2014). A Figura 1 apresenta um exemplo gráfico de contrato XL por risco.

Figura 1 – Exemplo gráfico de contrato XL por risco.



Fonte: Elaborada pela autora

No exemplo apresentado pela Figura 1, o limite de retenção definido pela seguradora é de 40.000 unidades monetárias e o limite máximo de cobertura da resseguradora é de 80.000 unidades monetárias, sendo assim diz se que o contrato é de 40.000 *xs* 40.000, isto é, a resseguradora irá pagar até 40.000 depois que a cedente desembolsar 40.000. No sinistro 1, a resseguradora terá de indenizar a cedente em 10 mil unidades monetárias que excederam os 40 mil do limite de retenção. No sinistro 2, a resseguradora não terá que indenizar a seguradora, pois a perda (15 mil) não atingiu o limite de retenção pré-determinado. No sinistro 5, a perda foi de 85.000 unidades monetárias, com isso a seguradora arca com os 40.000 de prioridade, a resseguradora é obrigada a indenizar 40.000 para a seguradora e os 5.000 que ultrapassaram o limite de cobertura fica a cargo da seguradora.

Matematicamente, de acordo com Lee (1988), a responsabilidade da resseguradora em um contrato de excesso de danos por risco pode ser expresso em função da perda P :

$$h(P; R, L) = \begin{cases} 0, & 0 < P \leq R \\ P - R, & R < P \leq L \\ L, & L < P, \end{cases}$$

em que R é o limite de retenção e L é o limite máximo de cobertura.

2.2 DISTRIBUIÇÃO GENERALIZADA DE PARETO

Seja X a quantidade de interesse com função de distribuição F , e $Y = X - u$ os excedentes sobre o limiar u . Balkema e de Hann (1974), Pickands (1975) e Li, Tang e Jiang (2016) mostraram que a distribuição de Y , dado que $X > u$, é a distribuição generalizada de Pareto (GP) e sua função densidade de probabilidade é dada por:

$$g(y, \sigma, \xi) = \begin{cases} \frac{1}{\sigma} \left(1 + \xi \left(\frac{y}{\sigma}\right)\right)^{-(1+\xi)/\xi}, & 1 + \xi \left(\frac{y}{\sigma}\right) \geq 0, \quad \text{se } \xi \neq 0 \\ \frac{1}{\sigma} \exp\left(-\frac{y}{\sigma}\right), & \text{se } \xi = 0, \end{cases}$$

em que ξ é o parâmetro de forma e σ o parâmetro de escala ($\sigma > 0$). A função de distribuição acumulada da generalizada de Pareto é dada por:

$$G(y, \sigma, \xi) = \begin{cases} 1 - \left[1 + \xi \left(\frac{y}{\sigma}\right)\right]^{-1/\xi}, & \sigma > 0 \quad 1 + \xi \left(\frac{y}{\sigma}\right) \geq 0, \quad \text{se } \xi \neq 0 \\ 1 - \exp\left(-\frac{y}{\sigma}\right), & y \in \mathbb{R}, \quad \text{se } \xi = 0, \end{cases}$$

Quando o parâmetro de forma é igual a zero, $\xi = 0$, a distribuição generalizada de Pareto assume sua forma mais simples, a distribuição exponencial (MENDES, 2004; SILVA, 2013).

De acordo com Coles (2001), a escolha do limiar u é uma etapa crucial na modelagem de excessos por uma distribuição generalizada de Pareto, pois quando o valor de u é muito pequeno pode não garantir a convergência dos excessos Y para a distribuição generalizada de Pareto, levando a um viés. No entanto, a medida que se aumenta o valor u , a amostra de excedentes Y diminui, levando ao aumento da variância das estimativas.

Um método usado por alguns autores como Coles(2001), Silva (2008), Silva (2013) e Braga (2015), é o gráfico da vida média residual ou média dos excessos. Conforme Coles (2001), se Y tem uma distribuição generalizada de Pareto com parâmetros σ e ξ , então:

$$E[Y] = \frac{\sigma}{1 - \xi},$$

sendo $\xi < 1$, pois quando $\xi \geq 1$ a média é infinita. Supondo que a distribuição generalizada de Pareto seja válida para um modelo de excessos de um limiar u_0 gerado por uma série X_1, \dots, X_n , da qual um termo arbitrário é denominado X , então:

$$E[X - u_0 | X > u_0] = \frac{\sigma_{u_0}}{1 - \xi},$$

sendo σ_{u_0} é o parâmetro de escala para os excessos do limiar u_0 . Se a distribuição GP é válida para os excessos do limiar u_0 , então deve ser igualmente válida para todos os limites $u > u_0$. Logo, para $u > u_0$,

$$\begin{aligned} E[X - u | X > u] &= \frac{\sigma_u}{1 - \xi} \\ &= \frac{\sigma_{u_0} + \xi_u}{1 - \xi}. \end{aligned}$$

$E(X - u | X > u)$ é uma função linear de u , é a média dos excessos do limiar u , para os quais a média amostral dos excessos fornece uma estimativa empírica. De acordo com Coles (2001), espera-se que essas estimativas se modifiquem linearmente com u , em níveis de u para os quais a distribuição GP é apropriada. O gráfico da média de excessos ou vida média residual é obtido por:

$$\left\{ \left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} (x_{(i)} - u) \right) : u < x_{max} \right\},$$

onde $x_{(1)}, \dots, x_{(n_u)}$ são as observações que excedem u e x_{max} é o máximo entre os valores observados.

A interpretação do gráfico da média de excessos não é um processo simples (COLES, 2001). Porém, Silva (2008) sugere que o valor do limiar u seja o menor valor de limiar a partir do qual a função das médias dos excessos amostral segue aproximadamente uma função afim.

2.3 MÉTODO DE MÁXIMA VEROSSIMILHANÇA

O processo de estimação pelo método de máxima verossimilhança parte da função de verossimilhança da amostra aleatória $\mathbf{y} = (y_1, \dots, y_n)$ que é independente e identicamente distribuída. De acordo com Bolfarine e Sandoval (2001), a função de verossimilhança é dada por:

$$L(\theta | \mathbf{y}) = \prod_{i=1}^n f(y_i, \theta).$$

Em seguida, aplica-se o logaritmo natural nessa função de verossimilhança para obter a função suporte, que é derivada em relação a cada parâmetro. Compõe-se um sistema de

equações com as derivadas igualadas a zero.

As estimativas dos parâmetros é obtida por meio da solução do sistema de equações, que geralmente não possui solução analítica, por isso faz-se necessária a utilização de métodos iterativos para solucionar o sistema, como o método de Nelder e Mead (1965).

2.4 INFERÊNCIA BAYESIANA

Conforme apresentado na revisão de literatura geral (Seção 2), o processo de inferência Bayesiana, de acordo com Paulino, Turkman e Murteira (2003), é baseado na distribuição *a posteriori* de θ (parâmetro ou vetor de parâmetros), que é obtida por meio do teorema de Bayes, sendo:

$$p(\theta|\mathbf{y}) \propto L(\theta|\mathbf{y})p(\theta)$$

em que $L(\theta|\mathbf{y})$ é a função de verossimilhança dos dados e $p(\theta)$ é a distribuição *a priori* dos parâmetros.

A distribuição *a posteriori*, geralmente, não pode ser obtida analiticamente, por isso são usados métodos como Monte Carlo via Cadeias de Markov (MCMC) para gerar amostras dessa distribuição, e as estimativas pontuais dos parâmetros podem ser obtidas através da média da distribuição *a posteriori* e as estimativas intervalares pelos intervalos HPD (*Highest Posterior Density*). Para um determinado nível de credibilidade, o intervalo HPD é o que apresenta menor amplitude dentre todos os possíveis intervalos de credibilidade (BRUNELLO; NAKANO, 2015).

A inferência Bayesiana tem sido utilizada em estudos com a distribuição generalizada de Pareto, como em Behrens, Lopes e Gamerman (2004), que utilizaram mistura de distribuições para modelar o índice financeiro, Nasdaq 100, para estimar o limiar da distribuição GP ajustada para a cauda dos modelos de mistura propostos. Li, Tang e Jiang (2016) propõem uma mistura de modelos com a GPD e abordagem Bayesiana para estimação dos parâmetros e do limiar dos modelos de mistura para dados das perdas ocorridas com o terremoto Yunnan.

O uso de *priori* não informativa possibilita a aplicação da abordagem Bayesiana para estimação de parâmetros quando não se tem informações passadas sobre o estudo (PAULINO; TURKMAN; MURTEIRA, 2003). A inferência Bayesiana apresenta como vantagem, mesmo com o uso de *priori* não informativa, a possibilidade de estimar intervalos de credibilidade

HPD (*Highest Posterior Density*) para as probabilidades calculadas, esses intervalos permitem quantificar a incerteza diretamente sobre as probabilidades, isso geralmente não é possível com o uso de abordagens clássicas na inferência de funções dos parâmetros.

3 MATERIAL E MÉTODOS

3.1 DADOS

Os dados foram obtidos no pacote *ReIns* (REYnkENS; VERBELEN, 2017) da linguagem de programação R (R CORE TEAM, 2017). Esse pacote contém funções do livro *Reinsurance: Actuarial and Statistical Aspects* de Albrecher de Beirlant e Teugels (2017). Os dados usados correspondem aos dados *Secura Re automobile reinsurance* e referem-se a reclamações automobilísticas de companhias seguradoras europeias do período de 1988 a 2001 que excederam 1.200.000 euros. Os dados são corrigidos pela inflação e totalizam 371 observações.

3.2 RECURSOS COMPUTACIONAIS

As análises e testes foram realizados utilizando a linguagem R e o *freeware* OpenBugs (THOMAS; O'HARA, 2004). Para a linguagem R foram utilizados os pacotes *ReIns*, *R2OpenBUGS* (STURTZ; LIGGES; GELMAN, 2005), *extRemes* (GILLELAND; KATZ, 2016), *POT* (RIBATET; DUTANG, 2016) e *coda* (PLUMMER et al., 2006).

3.3 ANÁLISE DE INDEPENDÊNCIA

O teste de hipótese de Durbin-Watson (DURBIN; WATSON, 1951), ao nível de significância de 5%, foi utilizado para verificar a independência da série de sinistros automobilísticos. As hipóteses do teste podem ser observadas nos testes de hipóteses preliminares na Seção 3.4 do Capítulo 2 (p. 48).

3.4 SELEÇÃO DE LIMIAR u DA DISTRIBUIÇÃO GENERALIZADA DE PARETO

A escolha do limiar foi realizada com base no gráfico da média de excessos. Para obtenção deste gráfico utilizou-se o comando "*mrlplot()*" no R que constrói o gráfico automaticamente a partir das observações.

A interpretação do gráfico é subjetiva, e o valor de u segundo Silva (2008) é o menor valor a partir do qual a curva da média de excessos é aproximadamente uma função afim, por isso, foram determinados três valores de limiar que correspondem aos quantis próximos ao início da linearidade da curva plotada pelo gráfico.

A distribuição generalizada de Pareto foi ajustada para os três limiares com estimação de parâmetros pelo método de máxima verossimilhança e por inferência Bayesiana.

3.5 ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA

A função de verossimilhança da distribuição generalizada de Pareto, de acordo com Silva (2008), é dada por:

$$L(\sigma, \xi | \mathbf{y}) = \prod_{i=1}^n \frac{1}{\sigma} \left(1 + \xi \frac{y_i}{\sigma}\right)^{-\left(\frac{1+\xi}{\xi}\right)}.$$

O logaritmo natural da função de verossimilhança é dado por:

$$l(\sigma, \xi, \mathbf{y}) = -n \cdot \ln \sigma - \frac{1+\xi}{\xi} \sum_{i=1}^n \ln \left(1 + \xi \frac{y_i}{\sigma}\right).$$

A partir da solução do sistema de equações, obtido das derivadas parciais em relação a σ e ξ , são obtidos os estimadores de máxima verossimilhança dos parâmetros da distribuição generalizada de Pareto. Foi utilizado o método iterativo de Nelder-Mead para obtenção das soluções numéricas.

3.6 ESTIMAÇÃO POR INFERÊNCIA BAYESIANA

As distribuições *a priori* adotadas para os parâmetros σ e ξ da distribuição generalizada de Pareto foram distribuições normais com média m e precisão τ :

$$\begin{aligned}\sigma &\sim N(m_\sigma, \tau_\sigma) \\ \xi &\sim N(m_\xi, \tau_\xi)\end{aligned}$$

Pelo teorema de Bayes, a distribuição *a posteriori* conjunta para a distribuição generalizada de Pareto é dada por:

$$P(\boldsymbol{\theta}|\mathbf{y}) \propto \prod_{i=1}^n \frac{1}{\sigma} \left(1 + \xi \frac{y_i}{\sigma}\right)^{-\left(\frac{1+\xi}{\sigma}\right)} \exp\left\{-\frac{\tau_\sigma}{2}(\sigma - m_\sigma)^2\right\} \exp\left\{-\frac{\tau_\xi}{2}(\xi - m_\xi)^2\right\}.$$

Foi utilizado *priori* não informativa, considerando a distribuição normal com média 0 e variância de 10.000, sendo a precisão o inverso da variância, tem-se que:

$$\begin{aligned}\sigma &\sim N\left(0, \frac{1}{10.000}\right) \\ \xi &\sim N\left(0, \frac{1}{10.000}\right).\end{aligned}$$

3.7 CÁLCULO DE PROBABILIDADE DE OCORRÊNCIA DE VALORES EXTREMOS

Através dos ajustes da distribuição GP por inferência Bayesiana e máxima verossimilhança, foi calculada a probabilidade de ocorrência de alguns sinistros com valores extremos. A probabilidade de ocorrência de um valor maior ou igual a x dado que esse valor seja maior que o limiar u , de acordo com Coles (2001), é dada por:

$$P(X > x | X > u) = \left[1 + \xi \left(\frac{x - u}{\sigma}\right)\right]^{-\frac{1}{\xi}}.$$

3.8 ANÁLISE DE CONVERGÊNCIA DAS CADEIAS GERADAS PELO MÉTODO MCMC

Do mesmo modo que no Capítulo 2, foram utilizadas simulações MCMC por meio do algoritmo Metropolis-Hastings para obter as soluções das integrais da distribuição *a posteriori* da distribuição generalizada de Pareto.

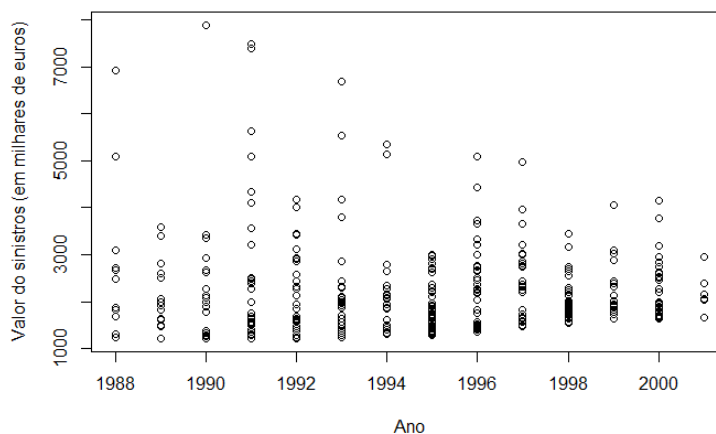
Os ajustes da distribuição generalizada de Pareto foram realizados utilizando-se o pacote *R2OpenBUGS* juntamente com o *freeware OpenBUGS*. Primeiramente, foram geradas cadeias com 6.000 valores, excluídos os 1.000 primeiros e saltos a cada 1 valor. Neste caso, não obtiveram-se indícios de convergência das cadeias nos critérios. Dessa maneira, aumentou-se o número de valores gerados, obtendo cadeias eficientes com 90.000 valores gerados, excluindo os 10.000 primeiros e saltos realizados a cada 26 valores, obtendo cadeias finais de, aproximadamente, 3.846 valores.

A verificação de convergência das cadeias foi realizada com a utilização do pacote *coda* e foram usados os critérios de Geweke, de Raftery e Lewis e de Heidelberger e Welch. De acordo com Nogueira, Sáfadi e Ferreira (2004), as cadeias convergem para distribuição de equilíbrio quando no critério de Geweke o módulo da estatística é menor que 1,96, no critério de Raftery e Lewis se o fator de dependência for próximo de 1,0 e no critério de Heidelberger e Welch se o valor-p do teste for maior que 0,05.

4 RESULTADOS E DISCUSSÃO

A série dos 371 sinistros automobilísticos reclamados em companhias europeias no período de 1988 a 2001 pode ser observada na Figura 2.

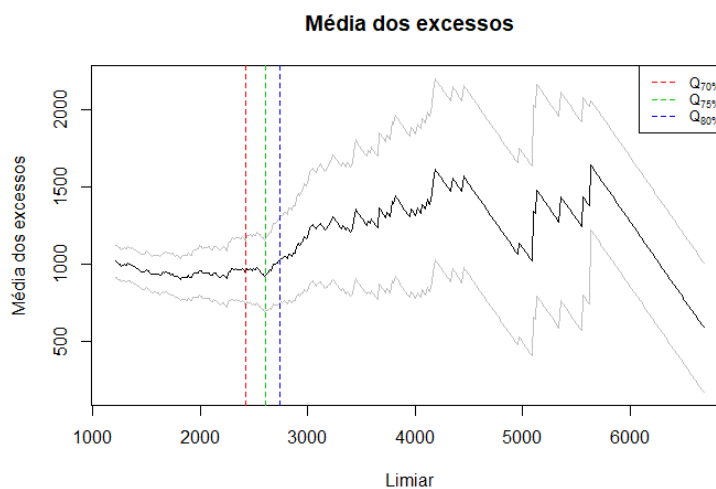
Figura 2 – Série de sinistros automobilísticos reclamados em companhias europeias de 1988 a 2001.



Fonte: Elaborada pela autora

Para escolher os limiares, foi analisado o gráfico da média dos excessos. Os limiares escolhidos foram os quantis de 70%, 75% e 80%, pois são valores próximos ao início da linearidade do gráfico. Os quantis foram $Q_{70\%} = 2.425,34$, $Q_{75\%} = 2.609,29$ e $Q_{80\%} = 2.736,90$. A Figura 3 apresenta o gráfico de vida média residual e os quantis escolhidos.

Figura 3 – Gráfico da média dos excessos e quantis de 70%, 75% e 80%.



Fonte: Elaborada pela autora

Pelo teste de Durbin-Watson, foi verificada a independência da série de sinistros (valor- $p=0,1102$). Então, primeiramente, foi ajustada a distribuição generalizada de Pareto para a série

de sinistros considerando os três valores de limiar e utilizando o método da máxima verossimilhança para a estimação dos parâmetros. Os resultados das estimativas dos parâmetros e seus respectivos intervalos de confiança são apresentados na Tabela 1.

Tabela 1 – Estimativas (Est. MV), intervalos de confiança de 95% e amplitude (*Amp*) dos intervalos para os parâmetros da distribuição generalizada de Pareto com estimação por máxima verossimilhança para os três limiares escolhidos.

Limiar	Parâmetro	Est. MV	IC _{95%}	<i>Amp</i>
$u_{70\%} = 2425,34$	σ	748,75	[522,67; 974,83]	452,16
	ξ	0,21	[-0,02; 0,46]	0,48
$u_{75\%} = 2609,29$	σ	642,55	[406,71; 878,39]	471,68
	ξ	0,34	[0,03; 0,65]	0,62
$u_{80\%} = 2736,90$	σ	785,97	[469,67; 1102,28]	632,61
	ξ	0,25	[-0,08; 0,59]	0,67

Fonte: Elaborada pela autora

A distribuição generalizada de Pareto também foi ajustada com estimação dos parâmetros via inferência Bayesiana com *priori* não informativa. Na Tabela 2 são apresentados os resultados desses ajustes para os três limiares.

Tabela 2 – Estimativas (Est. Bayes), intervalo HPD_{95%} e amplitude (*Amp*) do intervalo para os parâmetros (P) da distribuição generalizada de Pareto com estimação por inferência Bayesiana com *priori* não informativa para os três limiares com respectivos resultados dos critérios de convergência.

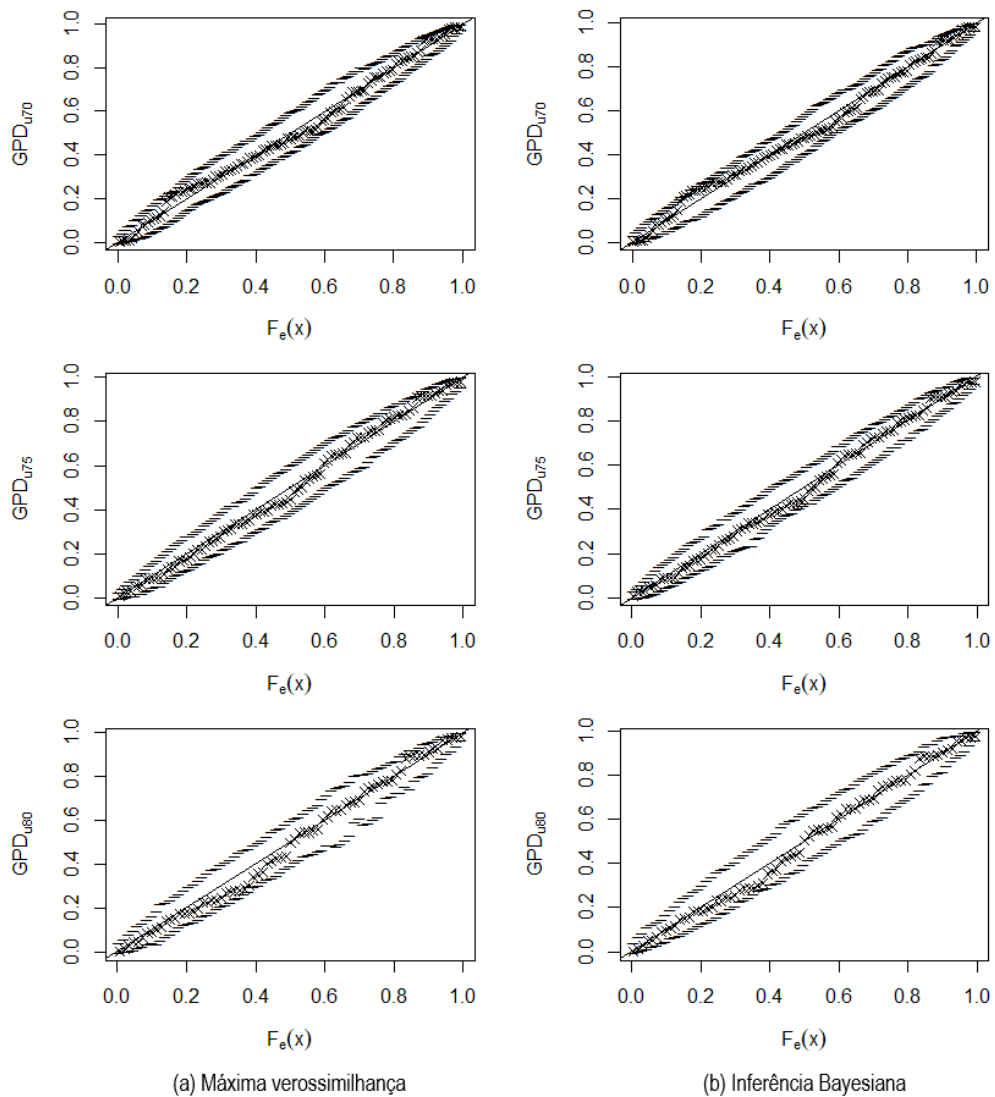
Limiar	P	Est. Bayes	HPD _{95%}	<i>Amp</i>	Critérios de Convergência		
					CG	R-L	H-W
$u_{70\%} = 2425,34$	σ	734,14	[708,60; 775,40]	66,80	0,578	1,020	0,373
	ξ	0,26	[0,08; 0,45]	0,37	0,130	1,000	0,353
$u_{75\%} = 2609,29$	σ	631,70	[596,40; 689,80]	93,40	0,079	1,020	0,815
	ξ	0,39	[0,15; 0,64]	0,49	0,158	0,995	0,870
$u_{80\%} = 2736,90$	σ	741,93	[709,00; 797,80]	88,80	0,325	1,010	0,972
	ξ	0,33	[0,09; 0,61]	0,52	1,119	0,998	0,359

Fonte: Elaborada pela autora

Para os três modelos com inferência Bayesiana, de acordo com os resultados dos critérios apresentados na Tabela 2, houve indícios de convergência das cadeias dos parâmetros, pois as estatísticas de Geweke, em módulo, são menores que 1,96, o fator de Raftery e Lewis é próximo de 1,0 e o valor-p do teste de Heidelberger e Welch é menor que 0,05.

Para verificar se os modelos se ajustaram bem aos dados foram plotados os gráficos P-P, apresentados na Figura 4.

Figura 4 – Gráficos P-P para os modelos ajustados com a distribuição GP com máxima verossimilhança (a) e inferência Bayesiana (b) para os três limiares escolhidos.



Fonte: Elaborada pela autora

De acordo com Silva (2008), um modelo está bem ajustado quando apresenta tendência linear e a amostra observada está contida dentro dos limites de confiança dos envelopes simulados. Analisando os gráficos P-P da Figura 4, nos quais o eixo das abscissas corresponde à função densidade de probabilidade empírica e o eixo das ordenadas à função densidade de probabilidade da distribuição ajustada em cada modelo, pode-se concluir que os seis modelos se ajustaram bem aos dados.

A distribuição generalizada de Pareto tem sido utilizada em estudos de valores extremos possibilitando bons resultados. Holmes e Moriarty (1999) analisaram a velocidade máxima de rajadas de ventos geradas pela passagem de *downbursts* em Moore (Austrália) e apontam como

vantagem no uso da distribuição GP, o uso de todos os dados relevantes e não só os máximos anuais, não sendo necessário ter um máximo para cada ano. Braga (2015) utilizou as distribuições GEV e generalizada de Pareto para estudo de precipitação máxima em Juiz de Fora e obteve bons resultados com ambas as distribuições. Melo (2006) comparou as distribuições Gamma, Log-Normal, GP e GP modificada na análise de sinistros de grandes riscos em contratos de resseguro e obteve como melhores resultados com o uso da distribuição generalizada de Pareto.

O uso da inferência Bayesiana aliado à distribuição generalizada de Pareto também tem sido alvo de estudos e tem alcançado bons resultados. Behrens, Lopes e Gamerman (2004) utilizaram inferência Bayesiana e distribuição GP para análise do índice de mercado financeiro Nasdaq 100. Li, Tang e Jiang (2016) utilizaram mistura de modelos, utilizando a distribuição GP como distribuição condicional limitante dos excessos e abordagem Bayesiana para estimação dos modelos ajustados à perdas causadas por um terremoto, os autores concluíram que os modelos propostos são adequados para ajustes de dados de perdas de terremotos.

A partir dos ajustes da distribuição generalizada de Pareto, utilizando inferência Bayesiana e máxima verossimilhança para os três limiares diferentes, foram calculadas as probabilidades de ocorrência de sinistros elevados. O maior valor observado na série foi de 7,9 milhões de euros, então calculou-se a probabilidade de ocorrer valores maiores que os limiares adotados e maiores que esse valor de 7,9 milhões. As probabilidades foram calculadas para 2,5; 2,7; 2,8; 5; 7,9; 8; 10; 15 e 20 milhões de Euros, os resultados são apresentados na Tabela 3.

Tabela 3 – Probabilidades de ocorrência (em %) de sinistros acima de x milhões calculadas pelos ajustes da distribuição generalizada de Pareto com estimativas pelo método de máxima verossimilhança (MV) e as probabilidades (em %) e intervalos HPD_{95%} (em %) calculados a partir dos modelos com inferência Bayesiana (Bayes) para os três limiares, u .

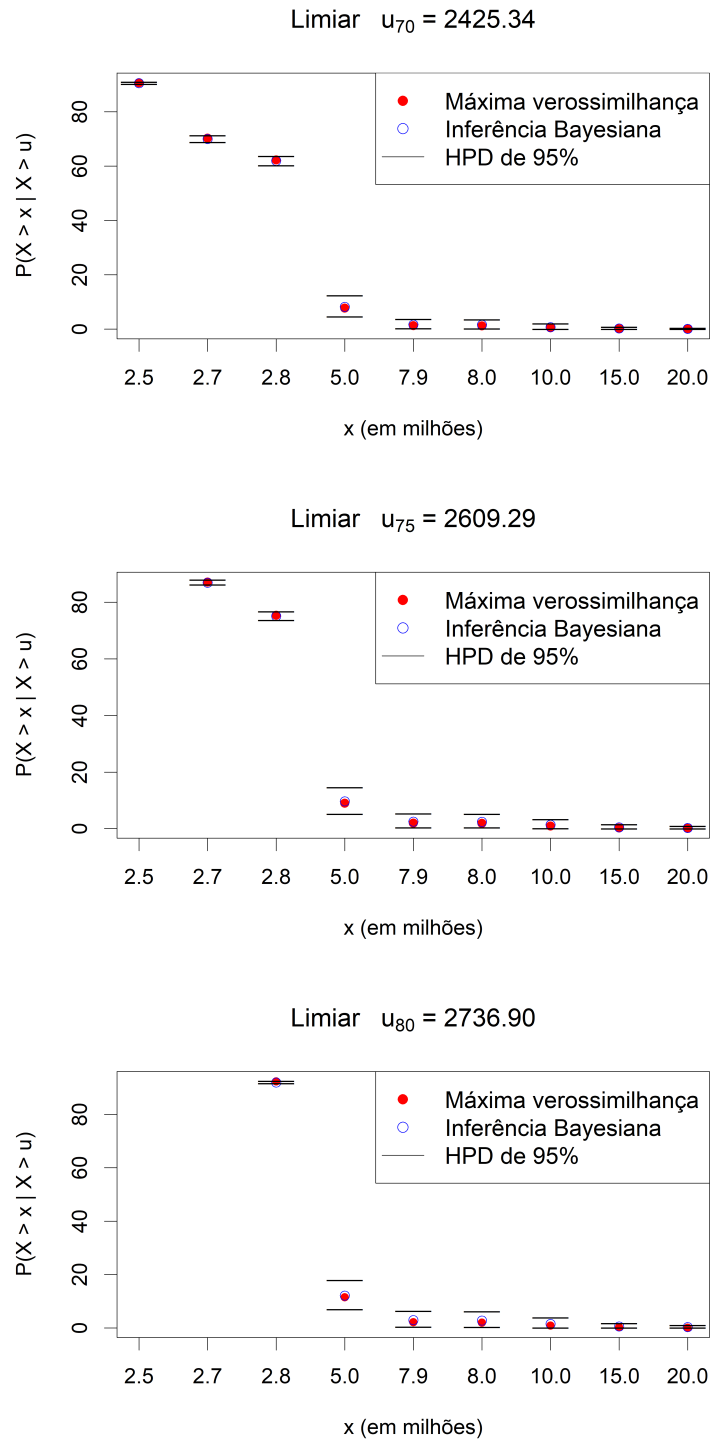
x	$u_{70\%} = 2425,34$		$u_{75\%} = 2609,29$		$u_{80\%} = 2736,90$	
	MV	Bayes	MV	Bayes	MV	Bayes
2,5	90,61	90,44	-	-	-	-
		[90,09; 90,89]	-	-	-	-
2,7	70,27	69,95	87,12	86,94	-	-
	-	[68,71; 71,24]	-	[86,19; 87,88]	-	-
2,8	62,20	61,89	75,37	75,13	92,36	91,95
	-	[60,20; 63,57]	-	[73,63; 76,71]	-	[91,58; 92,46]
5,0	7,70	8,21	9,02	9,69	11,48	12,12
	-	[4,58; 12,33]	-	[5,13; 14,54]	-	[6,92; 17,88]
7,9	1,27	1,67	1,97	2,52	2,08	2,88
	-	[0,18; 3,64]	-	[0,35; 5,28]	-	[0,30; 6,24]
8,0	1,21	1,60	1,89	2,43	1,98	2,78
	-	[0,16; 3,51]	-	[0,33; 5,15]	-	[0,28; 6,08]
10	0,48	0,76	0,92	1,35	0,85	1,47
	-	[0,01; 1,98]	-	[0,04; 3,26]	-	[0,01; 3,78]
15	0,09	0,22	0,26	0,51	0,18	0,52
	-	[0,00; 0,73]	-	[0,00; 1,51]	-	[0,00; 1,71]
20	0,03	0,09	0,11	0,27	0,06	0,26
	-	[0,00; 0,37]	-	[0,00; 0,91]	-	[0,00; 0,99]

Fonte: Elaborada pela autora

Na Tabela 3, os valores são interpretados da seguinte maneira: pelo modelo com limiar $u = 2.425,34$ e estimação de parâmetros pelo método de máxima verossimilhança, a probabilidade de ocorrer um sinistro maior que 2,5 milhões de unidades monetárias é de 90,61%, e para o modelo com inferência Bayesiana a probabilidade é de 90,44%. Pode-se observar que os valores das probabilidades calculados por máxima verossimilhança e inferência Bayesiana são próximos. Isso acontece porque os ajustes Bayesianos foram realizados com *priori* não informativa, onde a informação é vaga. De acordo com Box e Tiao (1992), o uso de *priori* não informativa deve ser considerado quando pretende-se que a informação dos dados seja dominante. Por isso, prevalece a informação da verossimilhança que está presente na distribuição *a posteriori* conforme o teorema de Bayes.

A Figura 5 apresenta as probabilidades calculadas pelos modelos Bayesianos e seus respectivos intervalos HPD de 95% de credibilidade e as probabilidades calculadas pelos modelos com estimação por máxima verossimilhança.

Figura 5 – Probabilidades de ocorrência de sinistros com severidade maior que x milhões, calculadas pelos modelos com máxima verossimilhança e inferência Bayesiana considerando os três limiares, e os intervalos HPD de 95% para os modelos Bayesianos.



Fonte: Elaborada pela autora

Por meio da análise gráfica é possível observar que os valores pontuais das probabilida-

des calculadas pelos modelos com inferência Bayesiana e máxima verossimilhança são semelhantes.

O fato da inferência Bayesiana possibilitar a estimação dos intervalos HPD de 95% de credibilidade pode ser visto como uma vantagem sobre os modelos com estimação por máxima verossimilhança, pois permitem quantificar a incerteza sobre as probabilidades calculadas.

5 CONCLUSÃO

A distribuição generalizada de Pareto se ajustou aos dados de perdas relacionadas à reivindicações automobilísticas em companhias securitárias europeias. O uso de *priori* não informativa para a inferência Bayesiana possibilitou a utilização dessa abordagem. A vantagem da inferência Bayesiana sobre o método de máxima verossimilhança foi notada no cálculo das probabilidades, em que foi possível obter os intervalos de credibilidade HPD.

Como trabalhos futuros, pretende-se obter dados atuais sobre contratos de resseguros e ocorrência de sinistros, para que possam ser analisados pela distribuição generalizada de Pareto e obter resultados que comprovem os bons resultados da inferência Bayesiana. Outra proposta, seria verificar se a seleção do limiar para a distribuição generalizada de Pareto, por meio do gráfico da média dos excessos, com a escolha do limite de retenção por meio da determinação com base no PLA (Patrimônio Líquido Ajustado) da empresa seguradora, resultam em valores semelhantes ou próximos.

6 REFERÊNCIAS

- ALBRECHER, H.; BEIRLANT, J.; TEUGELS, J. L. **Reinsurance: Actuarial and Statistical Aspects**. Chichester: Wiley, 2017.
- BALKEMA, A. A.; HAAN, L. D. Residual life time at great age. **The Annals of probability**, p. 792-804, 1974.
- BEHRENS, C. N.; LOPES, H. F.; GAMERMAN, D. Bayesian analysis of extreme events with threshold estimation. **Statistical Modelling**, Sage Publications Sage CA: Thousand Oaks, CA, v. 4, n. 3, p. 227–244, 2004.
- BOX, G. E. P.; TIAO, G. C. **Bayesian inference in statistical analysis**. New York: Wiley, 1992.
- BRAGA, D. d. O. M. **Aplicação da Teoria de Valores Extremos para índice pluviométrico da cidade de Juiz de Fora - MG**. 2015. 52 f. Monografia (Bacharel em Estatística) - Universidade Federal de Juiz de Fora, Juiz de Fora, 2015.
- BRUNELLO, G. H. V.; NAKANO, E. Y. Inferência bayesiana no modelo weibull discreto em dados com presença de censura. **TEMA (São Carlos)**, SciELO Brasil, v. 16, n. 2, p. 97–110, 2015.
- CABRAL, P. S. Seguro: sua evolução histórica. **Revista de Ciência Política**, v. 31, n. 1, p. 84-91, 1988.
- COLES, S. G. **An introduction to statistical modeling of extreme values**. London: Springer, 2001.
- DURBIN, J.; WATSON, G. S. Testing for serial correlation in least squares regression. II. **Biometrika**, JSTOR, v. 38, n. 1/2, p. 159–177, 1951.
- FERREIRA, P. P. **Precificação: Credibilidade, Risco no Resseguro e Aplicações Diversas**. Rio de Janeiro: Funenseg, 2007.
- FRANCO, V. H. M. **Lições de Direito Securitário: Seguros Terrestres privados**. São Paulo: Ed. Maltese, 1993.
- FUNDAÇÃO MAPFRE. **Introdução ao Resseguro**. Instituto de Ciências do Resseguro, Escola Nacional de Seguros. Madrid, 2011.
- GILLELAND, E.; KATZ, R. W. extRemes 2.0: An extreme value analysis package in R. **Journal of Statistical Software**, v. 72, n. 8, p. 1–39, 2016.
- GROSMANN, M. **Resseguro - Uma Introdução**. Instituto de Economia de Seguros da Escola Superior de St. Gallen, 1990, 3. ed, Suíça.

HADDAD, M. **O resseguro internacional**. Rio de Janeiro: Funenseg, 2003.

HOLMES, J.; MORIARTY, W. Application of the generalized pareto distribution to extreme value analysis in wind engineering. **Journal of Wind Engineering and Industrial Aerodynamics**, Elsevier, v. 83, p. 1–10, 1999.

KANG, S.; SONG, J. Parameter and quantile estimation for the generalized pareto distribution in peaks over threshold framework. **Journal of the Korean Statistical Society**, Elsevier, v. 46, n. 4, p. 487–501, 2017.

LEE, Y. S. The mathematics of excess of loss coverages and retrospective rating—a graphical approach. **PCAS LXXV**, Citeseer, v. 49, p. 49–77, 1988.

LI, Y.; TANG, N.; JIANG, X. Bayesian approaches for analyzing earthquake catastrophic risk. **Insurance: Mathematics and Economics**, Elsevier, v. 68, p. 110–119, 2016.

LUCCAS FILHO, O. **SEGUROS: fundamentos, formação de preço, provisões e funções biométricas**. São Paulo: Atlas, 2011.

MELO, E. F. L. d. **Uma Aplicação da Teoria de Valores Extremos para Avaliação do Risco de Contratos de Resseguro**. Rio de Janeiro: Revista Brasileira de Risco e Seguro, 2006.

MENDES, B. V. M. **Introdução à análise de eventos extremos**. Rio de Janeiro: E-papers Serviços Editoriais Ltda, 2004.

NELDER, J. A.; MEAD, R. A. A simplex method for function minimization. **The computer journal**, v. 7, p. 308–313, 1965.

PARK, M. H.; KIM, J. H. Estimating extreme tail risk measures with generalized pareto distribution. **Computational Statistics & Data Analysis**, Elsevier, v. 98, p. 91–104, 2016.

PARODI, P. **Pricing in general insurance**. Boca Raton: CRC Press, 2014.

PICKANDS, J. Statistical inference using extreme order statistics. **The Annals of Statistics**, p. 119-131, 1975.

PLUMMER, M. et al. Coda: Convergence diagnosis and output analysis for mcmc. **R News**, v. 6, n. 1, p. 7–11, 2006. Disponível em: <<https://journal.r-project.org/archive/>>. Acesso em: 21 mar. 2018.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2017. Disponível em: <<https://www.R-project.org/>>. Acesso em: 21 mar. 2018.

REYNKENS, T.; VERBELEN, R. **ReIns: Functions from "Reinsurance: Actuarial and Statistical Aspects"**. 2017. R package version 1.0.6. Disponível em: <<https://CRAN.R-project.org/package=ReIns>>. Acesso em: 21 mar. 2018.

RIBATET, M.; DUTANG, C. **POT: Generalized Pareto Distribution and Peaks Over Th-**

reshold. 2016. R package version 1.1-6. Disponível em: <<https://CRAN.R-project.org/package=POT>>. Acesso em: 21 mar. 2018.

RIOS, V. T. **O Resseguro no Brasil**. 2005. 94 f. Monografia (Bacharel em Direito) - Pontifícia Católica do Rio de Janeiro, Rio de Janeiro, 2005.

SILVA, N. P. da. **Extremos de vento sobre o Oeste do Oceano Atlântico Sul: análise direcional das ocorrências**. 2013. 74f. Dissertação de Mestrado - Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo, 2013.

SILVA, R. R. **A distribuição Generalizada de Pareto e mistura de distribuições de Gumbel no estudo da vazão e da velocidade máxima do vento em Piracicaba, SP**. 2008. 101 f. Dissertação (Mestrado em Agronomia) - Universidade de São Paulo, Piracicaba, 2008.

STURTZ, S.; LIGGES, U.; GELMAN, A. R2winbugs: A package for running winbugs from r. **Journal of Statistical Software**, v. 12, n. 3, p. 1–16, 2005. Disponível em:<<http://www.jstatsoft.org>>. Acesso em: 21 mar. 2018.

THOMAS, A.; O'HARA, R. B. **OpenBUGS**. 2004.

CAPÍTULO 4 - CONSIDERAÇÕES FINAIS

Analisando os resultados obtidos no Capítulo 3 e Capítulo 4, nas aplicações de TVE e inferência Bayesiana para análise e modelagem de risco Brasil e resseguros, respectivamente, pode-se considerar que a inferência Bayesiana apresentou melhores resultados nas duas análises.

Para a análise de risco Brasil, o uso da inferência Bayesiana possibilitou ajuste de modelos com melhor capacidade preditiva, considerando acurácia e erro médio de predição.

Para a modelagem de resseguros, destaca-se o ganho do uso da abordagem Bayesiana em relação a máxima verossimilhança, ao obter intervalos HPD de 95% de credibilidade para as probabilidades calculadas que permitem quantificar a incerteza sobre a probabilidade.

Por fim, conclui-se que a inferência Bayesiana aliada a teoria de valores extremos apresentou melhores resultados.

Para trabalhos futuros, pretende-se obter novas fontes para extração de informações *a priori* com séries de dados maiores que melhorem o processo de estimação, visando a produção de resultados mais satisfatórios.

APÊNDICES

APÊNDICE A - Estimativas dos parâmetros para os ajustes da distribuição Gumbel e GEV para *priori* não informativa e *prioris* informativas com estruturas I e II com informação do risco país da Argentina, do México e da média entre os dois países e resultados dos critérios de convergência dos parâmetros dos modelos Bayesianos.

Tabela 1 – Resultados dos critérios de convergência para os parâmetros do ajuste das distribuições Gumbel e GEV com estimação por inferência Bayesiana: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

Modelo	Parâmetro	Estimativa	Critérios de Convergência		
			CG	R-L	H-W
Gumbel	μ	305,12	0,738	0,997	0,480
	σ	134,35	0,675	0,999	0,932
GEV	μ	282,15	0,754	1,070	0,350
	σ	99,86	0,847	1,110	0,932
	ξ	0,53	0,800	1,020	0,140

Fonte: Elaborada pela autora

Tabela 2 – Estimativas dos parâmetros μ e σ , da distribuição Gumbel via inferência Bayesiana com informação *a priori* do risco país da Argentina para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	Parâmetro	Estimativa	HPD _{95%}	Critérios de Convergência		
					CG	R-L	H-W
Gumbel	I _{ARG} (1var)	μ	127,98	[114,00; 142,10]	0,377	1,030	0,987
		σ	64,15	[57,27; 71,21]	0,509	1,010	0,479
	II _{ARG} (1var)	μ	132,49	[118,80; 147,00]	0,689	1,000	0,249
		σ	53,29	[47,34; 59,01]	0,056	0,998	0,969
	I _{ARG} (2var)	μ	140,35	[119,70; 159,60]	1,032	0,985	0,768
		σ	73,92	[64,39; 83,30]	0,155	1,010	0,850
	II _{ARG} (2var)	μ	146,31	[125,90; 165,90]	1,273	0,993	0,284
		σ	63,59	[55,58; 72,08]	1,525	0,993	0,827
	I _{ARG} (4var)	μ	159,76	[132,00; 187,70]	1,239	1,020	0,595
		σ	83,90	[71,19; 96,82]	0,189	0,980	0,685
	II _{ARG} (4var)	μ	167,34	[139,60; 195,20]	0,456	1,030	0,888
		σ	73,94	[62,64; 85,69]	0,176	0,990	0,826
	I _{ARG} (6var)	μ	175,32	[143,40; 210,30]	0,723	1,000	0,907
		σ	89,18	[73,96; 105,00]	0,202	0,987	0,982
	II _{ARG} (6var)	μ	183,51	[152,30; 218,00]	1,504	0,985	0,053
		σ	79,47	[65,76; 93,62]	1,978	1,010	0,161

Fonte: Elaborada pela autora

Tabela 3 – Estimativas dos parâmetros μ e σ , da distribuição Gumbel via inferência Bayesiana com informação *a priori* do risco país do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	Parâmetro	Estimativa	HPD _{95%}	Critérios de Convergência		
					CG	R-L	H-W
Gumbel	I _{MEX} (1var)	μ	265,27	[221,90; 305,60]	0,888	0,976	0,880
		σ	129,60	[97,41; 161,40]	0,109	0,981	0,799
	II _{MEX} (1var)	μ	270,24	[237,70; 303,80]	0,351	0,973	0,253
		σ	89,60	[70,05; 109,60]	0,395	1,010	0,405
	I _{MEX} (2var)	μ	279,82	[232,50; 325,70]	0,790	1,020	0,983
		σ	129,39	[94,07; 166,70]	0,402	0,986	0,907
	II _{MEX} (2var)	μ	279,91	[239,70; 317,00]	0,607	0,999	0,721
		σ	100,96	[75,70; 125,90]	1,217	0,985	0,262
	I _{MEX} (4var)	μ	290,42	[240,80; 340,80]	0,292	1,000	0,327
		σ	130,37	[91,21; 170,60]	0,809	1,020	0,138
	II _{MEX} (4var)	μ	288,08	[243,30; 331,00]	1,523	0,985	0,716
		σ	111,22	[83,22; 144,60]	0,084	1,000	0,302
	I _{MEX} (6var)	μ	294,80	[242,20; 345,40]	0,791	1,010	0,426
		σ	131,21	[90,42; 173,00]	0,151	1,020	0,805
	II _{MEX} (6var)	μ	291,96	[244,00; 337,20]	1,398	0,997	0,326
		σ	116,35	[84,55; 152,00]	0,489	1,030	0,761

Fonte: Elaborada pela autora

Tabela 4 – Estimativas dos parâmetros μ e σ , da distribuição Gumbel via inferência Bayesiana com informação *a priori* da média do risco país da Argentina e do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	Parâmetro	Estimativa	HPD _{95%}	Critérios de Convergência		
					CG	R-L	H-W
Gumbel	I _{MED} (1var)	μ	209,41	[180,80; 236,80]	0,174	1,020	0,300
		σ	95,45	[79,28; 112,30]	1,302	1,010	0,778
	II _{MED} (1var)	μ	222,12	[196,30; 247,90]	0,047	0,995	0,888
		σ	70,53	[58,17; 82,74]	1,755	1,000	0,543
	I _{MED} (2var)	μ	229,64	[194,20; 265,40]	0,096	1,030	0,633
		σ	101,71	[81,51; 124,30]	1,000	0,925	0,797
	II _{MED} (2var)	μ	239,73	[208,90; 271,50]	0,804	1,010	0,970
		σ	80,65	[64,29; 96,86]	1,648	0,993	0,646
	I _{MED} (4var)	μ	251,24	[208,80; 291,30]	0,149	1,010	0,703
		σ	107,55	[82,51; 134,70]	0,252	1,020	0,809
	II _{MED} (4var)	μ	257,32	[221,40; 292,40]	0,666	1,030	0,778
		σ	91,20	[70,65; 112,10]	0,572	0,999	0,630
	I _{MED} (6var)	μ	263,37	[219,40; 304,90]	0,232	0,999	0,840
		σ	111,14	[84,18; 141,60]	1,565	1,000	0,514
	II _{MED} (6var)	μ	266,33	[229,40; 305,90]	1,241	1,030	0,333
		σ	97,48	[74,73; 121,50]	1,151	1,010	0,147

Fonte: Elaborada pela autora

Tabela 5 – Estimativas dos parâmetros μ , σ e ξ , da distribuição GEV via inferência Bayesiana com informação *a priori* do risco país da Argentina para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	Parâmetro	Estimativa	HPD _{95%}	Critérios de Convergência		
					CG	R-L	H-W
GEV	I _{ARG} (1var)	μ	118,09	[101,22; 135,53]	0,730	0,970	0,548
		σ	286,26	[170,31; 401,53]	0,936	1,010	0,444
		ξ	-0,22	[-0,44; 0,04]	0,584	1,010	0,623
	II _{ARG} (1var)	μ	118,44	[101,84; 135,83]	0,730	0,970	0,548
		σ	276,76	[160,12; 396,57]	0,936	1,010	0,444
		ξ	-0,20	[-0,44; 0,05]	0,584	1,010	0,623
	I _{ARG} (2var)	μ	122,96	[96,98; 146,97]	0,485	1,000	0,349
		σ	265,53	[147,07; 400,64]	0,469	1,000	0,726
		ξ	-0,17	[-0,46; 0,13]	1,207	1,320	0,447
	II _{ARG} (2var)	μ	123,48	[99,46; 148,06]	0,682	0,951	0,946
		σ	258,69	[149,94; 390,99]	1,012	1,040	0,790
		ξ	-0,16	[-0,44; 0,14]	0,813	0,970	0,421
	I _{ARG} (4var)	μ	134,88	[99,26; 169,03]	0,690	1,020	0,296
		σ	240,17	[131,22; 195,17]	1,635	1,040	0,737
		ξ	-0,11	[-0,44; 0,21]	1,036	1,350	0,986
	II _{ARG} (4var)	μ	135,17	[99,04; 169,48]	0,464	0,991	0,881
		σ	236,75	[121,68; 369,67]	1,799	0,991	0,333
		ξ	-0,09	[-0,42; 0,27]	1,264	1,080	0,573
	I _{ARG} (6var)	μ	150,21	[105,27; 201,01]	0,659	1,020	0,101
		σ	212,75	[92,66; 347,31]	0,208	1,140	0,074
		ξ	-0,03	[-0,43; 0,36]	0,280	1,410	0,064
	II _{ARG} (6var)	μ	150,87	[102,84; 78,53]	0,456	1,080	0,263
		σ	208,86	[78,53; 342,99]	0,264	0,970	0,706
		ξ	-0,02	[-0,40; 0,41]	0,307	1,030	0,835

Fonte: Elaborada pela autora

Tabela 6 – Estimativas dos parâmetros μ , σ e ξ , da distribuição GEV via inferência Bayesiana com informação *a priori* do risco país do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	Parâmetro	Estimativa	HPD _{95%}	Critérios de Convergência		
					CG	R-L	H-W
GEV	I _{MEX} (1var)	μ	268,37	[236,13; 301,33]	0,215	0,979	0,447
		σ	88,67	[54,71; 130,02]	0,703	1,020	0,196
		ξ	0,43	[0,08; 0,76]	0,402	1,110	0,679
	II _{MEX} (1var)	μ	268,00	[233,89; 301,01]	0,780	0,970	0,549
		σ	88,43	[53,16; 128,51]	0,028	1,020	0,268
		ξ	0,43	[0,11; 0,78]	1,392	0,970	0,091
	I _{MEX} (2var)	μ	274,41	[236,48; 312,43]	1,509	1,000	0,154
		σ	92,19	[57,00; 135,26]	1,599	1,110	0,378
		ξ	0,47	[0,09; 0,84]	0,252	1,000	0,949
	II _{MEX} (2var)	μ	274,24	[237,24; 311,37]	0,676	1,010	0,613
		σ	92,03	[56,42; 134,76]	0,848	1,010	0,954
		ξ	0,47	[0,11; 0,85]	0,469	1,010	0,857
	I _{MEX} (4var)	μ	277,69	[237,85; 316,53]	0,763	1,070	0,574
		σ	95,03	[55,46; 135,80]	0,533	1,020	0,833
		ξ	0,49	[0,10; 0,94]	1,897	1,000	0,079
	II _{MEX} (4var)	μ	277,59	[238,44; 317,55]	0,005	1,030	0,433
		σ	94,97	[56,86; 139,12]	0,357	1,000	0,243
		ξ	0,49	[0,07; 0,92]	1,035	0,991	0,102
	I _{MEX} (6var)	μ	279,39	[242,00; 323,75]	0,874	0,979	0,540
		σ	97,14	[56,45; 142,99]	0,761	1,040	0,102
		ξ	0,51	[0,11; 0,96]	1,382	1,020	0,714
	II _{MEX} (6var)	μ	279,04	[238,97; 321,73]	0,797	1,100	0,514
		σ	96,51	[56,88; 141,24]	0,048	1,030	0,684
		ξ	0,50	[0,09; 0,98]	0,301	0,991	0,307

Fonte: Elaborada pela autora

Tabela 7 – Estimativas dos parâmetros μ , σ e ξ , da distribuição GEV via inferência Bayesiana com informação *a priori* da média entre o risco país da Argentina e do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	Parâmetro	Estimativa	HPD _{95%}	Critérios de Convergência		
					CG	R-L	H-W
GEV	I _{MED} (1var)	μ	209,89	[169,61; 245,27]	0,133	1,000	0,412
		σ	124,76	[50,17; 215,43]	0,811	0,979	0,469
		ξ	0,20	[-0,18; 0,56]	0,385	1,070	0,272
	II _{MED} (1var)	μ	210,87	[174,63; 250,92]	0,682	0,951	0,372
		σ	119,22	[45,71; 205,71]	1,486	1,120	0,402
		ξ	0,22	[-0,14; 0,60]	1,149	0,951	0,711
	I _{MED} (2var)	μ	242,32	[208,45; 274,67]	0,750	1,000	0,480
		σ	84,98	[44,40; 141,32]	0,628	0,958	0,764
		ξ	0,42	[0,04; 0,81]	0,299	1,070	0,976
	II _{MED} (2var)	μ	242,13	[208,03; 273,91]	0,181	1,050	0,561
		σ	84,28	[45,96; 141,13]	1,555	0,991	0,328
		ξ	0,42	[0,03; 0,80]	1,619	1,050	0,308
	I _{MED} (4var)	μ	257,95	[226,73; 290,75]	0,529	1,070	0,533
		σ	83,42	[48,98; 122,06]	0,513	1,000	0,926
		ξ	0,49	[0,12; 0,88]	0,225	1,020	0,463
	II _{MED} (4var)	μ	257,78	[223,39; 288,63]	0,7443	0,991	0,250
		σ	83,13	[49,31; 122,64]	0,221	0,991	0,683
		ξ	0,49	[0,12; 0,89]	1,566	0,991	0,321
	I _{MED} (6var)	μ	263,96	[228,94; 296,99]	0,746	1,000	0,737
		σ	86,11	[51,95; 125,57]	1,345	1,040	0,792
		ξ	0,51	[0,12; 0,92]	0,149	1,000	0,950
	II _{MED} (6var)	μ	263,68	[229,10; 296,87]	1,378	0,951	0,334
		σ	85,99	[52,36; 125,91]	1,541	1,010	0,236
		ξ	0,51	[0,12; 0,93]	1,277	1,080	0,176

Fonte: Elaborada pela autora

APÊNDICE B - Resultados dos critérios para os níveis de retorno dos ajustes da distribuição Gumbel e GEV para *priori* não informativa e *prioris* informativas com estruturas I e II com informação do risco país da Argentina, do México e da média entre os dois países.

Tabela 8 – Resultados dos critérios de convergência para os níveis de retorno para os tempos de retorno de 2 e 3 semestres do ajuste das distribuições Gumbel e GEV com estimação por inferência Bayesiana: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

Modelo	Tempos de retorno	Níveis de retorno	Critérios de Convergência		
			CG	R-L	H-W
Gumbel	2	354,36	0,847	1,020	0,467
	3	426,40	0,895	1,000	0,487
GEV	2	322,60	1,589	1,070	0,676
	3	399,10	1,070	1,070	0,926

Fonte: Elaborada pela autora

Tabela 9 – Estimativas dos níveis de retorno para os tempos de retorno (TR) de 2 e 3 semestres, para os modelos de distribuição Gumbel e GEV via inferência Bayesiana com informação *a priori* do risco país da Argentina para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	TR	Níveis de Retorno	Critérios de Convergência		
				CG	R-L	H-W
Gumbel	I _{ARG} (1var)	2	151,49	0,494	1,030	0,990
		3	185,89	0,611	1,030	0,969
	II _{ARG} (1var)	2	152,03	0,719	1,010	0,241
		3	180,60	0,716	0,997	0,268
	I _{ARG} (2var)	2	167,45	1,019	0,992	0,834
		3	207,08	0,960	1,010	0,926
	II _{ARG} (2var)	2	169,62	1,075	0,997	0,233
		3	203,72	0,706	0,991	0,189
	I _{ARG} (4var)	2	190,51	1,305	1,020	0,476
		3	235,50	1,335	1,020	0,351
	II _{ARG} (4var)	2	194,44	0,508	1,020	0,879
		3	234,09	0,556	1,010	0,858
I _{ARG} (6var)	2	208,01	0,744	1,000	0,884	
	3	255,83	0,728	1,020	0,858	
II _{ARG} (6var)	2	212,64	1,255	1,010	0,100	
	3	255,25	0,798	1,030	0,252	
GEV	I _{ARG} (1var)	2	218,50	1,624	1,070	0,540
		3	350,80	1,259	1,090	0,510
	II _{ARG} (1var)	2	215,80	1,241	1,080	0,248
		3	344,90	1,154	1,080	0,303
	I _{ARG} (2var)	2	216,90	0,665	0,979	0,603
		3	343,20	0,326	1,040	0,716
	II _{ARG} (2var)	2	215,20	0,568	0,991	0,743
		3	339,10	0,721	1,010	0,880
	I _{ARG} (4var)	2	219,80	1,488	0,970	0,548
		3	336,90	1,931	0,970	0,334
	II _{ARG} (4var)	2	185,70	0,395	1,020	0,982
		3	269,70	0,144	1,040	0,297
I _{ARG} (6var)	2	226,40	0,203	1,010	0,647	
	3	334,00	0,337	0,970	0,819	
II _{ARG} (6var)	2	303,60	0,113	1,020	0,254	
	3	366,50	0,329	0,979	0,157	

Fonte: Elaborada pela autora

Tabela 10 – Estimativas dos níveis de retorno para os tempos de retorno (TR) de 2 e 3 semestres, para os modelos de distribuição Gumbel e GEV via inferência Bayesiana com informação *a priori* do risco país do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberger e Welch (H-W).

	Modelo	TR	Níveis de Retorno	Critérios de Convergência		
				CG	R-L	H-W
Gumbel	I _{MEX} (1var)	2	312,77	0,859	1,000	0,821
		3	382,26	0,723	0,988	0,750
	II _{MEX} (1var)	2	303,08	0,430	0,989	0,357
		3	351,12	0,460	0,970	0,558
	I _{MEX} (2var)	2	327,24	0,898	1,030	0,983
		3	396,62	0,887	1,020	0,969
	II _{MEX} (2var)	2	316,91	0,309	1,020	0,811
		3	371,05	0,091	1,000	0,771
	I _{MEX} (4var)	2	338,21	0,526	1,010	0,212
		3	408,11	0,730	1,010	0,125
	II _{MEX} (4var)	2	328,84	1,479	1,010	0,938
		3	388,48	1,305	0,998	0,987
I _{MEX} (6var)	2	342,89	0,806	1,020	0,540	
	3	413,24	0,719	1,020	0,721	
II _{MEX} (6var)	2	334,60	1,198	0,967	0,475	
	3	396,99	0,866	0,996	0,743	
GEV	I _{MEX} (1var)	2	303,60	0,113	1,020	0,254
		3	366,50	0,329	0,979	0,157
	II _{MEX} (1var)	2	303,10	0,542	1,010	0,754
		3	365,80	0,06	1,030	0,866
	I _{MEX} (2var)	2	311,30	1,714	1,040	0,148
		3	378,30	1,750	1,070	0,217
	II _{MEX} (2var)	2	311,10	0,841	0,991	0,721
		3	378,10	1,131	1,010	0,787
	I _{MEX} (4var)	2	315,90	0,637	1,040	0,707
		3	386,40	0,128	1,020	0,753
	II _{MEX} (4var)	2	315,80	0,058	1,050	0,226
		3	386,30	0,017	1,050	0,310
I _{MEX} (6var)	2	318,60	0,251	1,090	0,387	
	3	391,60	0,539	1,070	0,218	
II _{MEX} (6var)	2	317,90	0,578	1,100	0,518	
	3	390,30	0,312	1,050	0,628	

Fonte: Elaborada pela autora

Tabela 11 – Estimativas dos níveis de retorno para os tempos de retorno (TR) de 2 e 3 semestres, para os modelos de distribuição Gumbel e GEV via inferência Bayesiana com informação *a priori* da média do risco país da Argentina e do México para as estruturas I e II, seus respectivos intervalos HPD com 95% de credibilidade e resultados dos critérios de convergência: módulo da estatística de Geweke (CG), fator de Raftery-Lewis (R-L) e valor-p de Heidelberg e Welch (H-W).

	Modelo	TR	Níveis de Retorno	Critérios de Convergência		
				CG	R-L	H-W
Gumbel	I _{MED} (1var)	2	244,40	0,110	1,020	0,300
		3	295,58	0,505	1,010	0,336
	II _{MED} (1var)	2	247,97	0,373	1,020	0,861
		3	285,79	0,834	0,980	0,783
	I _{MED} (2var)	2	266,92	0,313	1,020	0,631
		3	321,46	0,600	1,020	0,651
	II _{MED} (2var)	2	269,29	1,158	1,010	0,944
		3	312,54	1,602	1,030	0,858
	I _{MED} (4var)	2	290,66	0,214	1,000	0,729
		3	348,33	0,289	1,020	0,771
	II _{MED} (4var)	2	290,74	0,797	1,020	0,756
		3	339,64	0,919	1,010	0,707
I _{MED} (6var)	2	259,10	0,587	1,010	0,906	
	3	363,70	1,001	0,999	0,909	
II _{MED} (6var)	2	302,06	0,986	1,020	0,569	
	3	354,33	0,566	1,000	0,871	
GEV	I _{MED} (1var)	2	256,90	1,082	1,020	0,664
		3	330,70	1,181	1,020	0,811
	II _{MED} (1var)	2	255,90	0,559	0,951	0,212
		3	327,50	1,257	0,970	0,244
	I _{MED} (2var)	2	275,80	0,305	0,979	0,519
		3	334,40	0,118	0,958	0,520
	II _{MED} (2var)	2	275,40	1,298	1,030	0,183
		3	333,90	1,210	1,010	0,071
	I _{MED} (4var)	2	291,40	0,269	1,020	0,610
		3	352,70	0,043	1,020	0,756
	II _{MED} (4var)	2	291,10	0,623	1,030	0,376
		3	352,30	0,748	0,991	0,527
I _{MED} (6var)	2	298,60	1,075	1,020	0,683	
	3	362,80	1,285	1,000	0,708	
II _{MED} (6var)	2	298,30	1,575	0,991	0,110	
	3	362,50	1,347	0,970	0,116	

Fonte: Elaborada pela autora