



Uma Aplicação da Teoria de Valores Extremos para Avaliação do Risco de Contratos de Resseguro

Eduardo Fraga Lima de Melo¹

Doutorando e Mestre em Administração (Finanças) pelo Instituto COPPEAD de Administração/UFRJ, Graduado *Magna Cum Laude* em Ciências Atuariais pelo Instituto de Matemática da UFRJ, do Departamento Técnico Atuarial da SUSEP.

eduardoflm@susep.gov.br

Resumo

O objetivo deste artigo é apresentar modelos para a avaliação do risco de um contrato de resseguro, baseado em excesso de danos, utilizando a teoria de valores extremos. Dentro da abordagem da teoria do risco, foram propostos modelos para os excessos de sinistros de incêndio, raio e explosão em função de um determinado limiar (limite de retenção de contratos hipotéticos). Para estimação da severidade foram utilizadas as distribuições de valores extremos Pareto Generalizada (GPD) e Pareto Generalizada Modificada (MGPD), além de outras consagradas no mercado como a Gamma e a Log-Normal. A distribuição ajustada para o número de sinistros foi a Poisson. Obtidas as estimativas dos parâmetros que maximizavam a função de verossimilhança, as distribuições de valores extremos apresentaram melhores ajustes para limites de retenção altos. Por fim, a distribuição do montante de sinistros agregados foi gerada por meio de convoluções e comparada com a distribuição empírica. É importante notar que os modelos aqui desenvolvidos podem ser facilmente aplicados para avaliação do risco de subscrição de companhias resseguradoras.

Palavras-Chave

Resseguro; teoria de valores extremos; convolução; máxima verossimilhança.

Sumário

1. Introdução; 2. Teorias do Risco e de Valores Extremos; 3. Dados utilizados e modelos propostos; 3.1. Dados utilizados; 3.2. Modelos propostos; 4. Resultados; 5. Considerações finais; 6. Referências bibliográficas; 7. Anexo.

¹ O autor agradece o apoio da FUNENSEG.

Abstract

An application of the extreme value theory to the risk valuation of reinsurance treaties

Eduardo Fraga Lima de Melo¹

MBA in Finance; currently taking a PhD in Finance Program at the Business School of COPPEAD/UFRJ (Federal University of Rio de Janeiro); BSc *Magna Cum Laude* in Actuary Sciences from Instituto de Matemática da UFRJ ('School of Mathematics of the Federal University of Rio de Janeiro'), Technical Department of Actuary Sciences of SUSEP (Private Insurance Superintendence)

eduardofilm@susep.gov.br

Plan of Contents

1. Introduction; 2. Risk and Extraneous Value Theories; 3. Compiled data and proposed models; 3.1. Compiled data; 3.2. Proposed Models; 4. Findings; 5. Final Considerations; 6. References; 7. Appendix.

Summary

The objective of this article is to present models for the risk valuation of reinsurance treaties based on damage excesses, with extreme value theory (EVT). Using the risk theory approach, we proposed some models for the excesses of fire claims above a defined threshold (retention limit of a hypothetical reinsurance treaty). The distribution used to estimate the number of claims was Poisson and the adjusted distributions for the claims severity were the Generalized Pareto Distribution (GPD), Modified Generalized Pareto Distribution (MGPD) and the well known Gamma and Log-Normal distributions. Once obtained the estimates of the parameters that maximized the likelihood function, the distributions of extreme values presented better adjustments for high retention limits. Finally, the distribution of the amount of aggregated claims was generated through convolutions and compared with the empiric distribution. It is important to note that the models developed here can be easily applied for the underwriting risk valuation of reinsurance companies.

Keywords

Reinsurance; extreme value theory; convolutions; maximum likelihood.

Table of Contents

1. Introdução; 2. Teorias do Risco e de Valores Extremos; 3. Dados utilizados e modelos propostos; 3.1. Dados utilizados; 3.2. Modelos propostos; 4. Resultados; 5. Considerações finais; 6. Referências bibliográficas; 7. Anexo.

¹ The author would like to acknowledge and thanks the support provided by FUNENSEG (National School of Insurance Foundation, Brazil)

1. Introdução

Acontecimentos recentes ocorridos no mundo, como grandes catástrofes, chamaram a atenção pela severidade das conseqüentes perdas. Eventos deste porte, por serem de baixa freqüência e grande severidade, são de difícil previsão por parte de qualquer agente econômico que se proponha a prover proteção ou a reter este risco. Resseguradores são exemplos destes agentes que retêm grandes responsabilidades e que são muitas vezes a ponta de uma cadeia de transferência ou pulverização de riscos. Além destes tipos de contratos, outros instrumentos podem ser considerados para a pulverização de riscos, dentre eles citamos o cosseguro e os derivativos que têm sido negociados na CBOT (*Chicago Board of Trade* – www.cbot.com). Nesta bolsa de valores, especuladores fazem “apostas” sobre o resultado incerto deste “jogo”, mas estes instrumentos não são o foco deste artigo. Para maiores detalhes sobre o papel destes instrumentos veja D’arcy, France & Gorvett (1999).

Resseguradores são um importante personagem na manutenção da solvência e confiabilidade do sistema de seguros. O negócio principal destas instituições é segurar grandes responsabilidades. Uma possível insolvência da cessionária pode trazer sérios riscos ao equilíbrio do mercado, uma vez que o “default” pode gerar um efeito dominó nas seguradoras cedentes, comprometendo a confiança dos consumidores no instrumento “seguro”. Neste contexto, é de suma importância que os riscos a que estão expostos sejam devidamente avaliados e corretamente precificados ou dimensionados. Como uma ferramenta poderosa para inferir caudas de distribuições de probabilidades, a teoria de valores extremos é de grande valor para avaliação do risco de contratos de resseguro ou cálculo de prêmios.

Por isso, uma outra aplicação possível da metodologia exposta neste artigo é na avaliação do *risco de subscrição referente a prêmio* de resseguradores. Para maiores detalhes sobre risco de subscrição de seguradores, consulte Fraga (2005). Esta modelagem do risco de subscrição pode ajudar a empresa a ter ciência melhor das necessidades de capital, identificar e avaliar os riscos e determinar as implicações de cada categoria de risco no capital. Então, a estratégia de negócios pode e deve estar alinhada com os resultados apontados pelo modelo.

Uma vez que há a possibilidade de não recuperar algumas perdas, em caso de falha da cessionária, a cedente enfrenta um risco de crédito. Logo, o risco definido pelo ponto de vista da resseguradora representa também um risco para o segurador. Uma das várias decisões que a seguradora deve tomar é o grau de retenção ou de cessão de riscos. Este grau está diretamente relacionado ao valor do limite de retenção a ser estabelecido e ao apetite da seguradora de aumentar a produção. Portanto, dentro das alternativas de contratos de resseguro com diferentes limites de retenção e prêmios de resseguro, deve ser feita uma análise de qual contrato efetivamente agrega mais valor à companhia, dado um determinado nível de risco. Para esta análise, a teoria de valores extremos também pode ser uma boa ferramenta para simular os resultados operacionais futuros.

Como característica natural dos contratos de resseguro baseados em excesso de danos, a companhia resseguradora fica responsável pelo pagamento de sinistros que ultrapassem um determinado limiar, também denominado limite de retenção. Neste tipo de contrato não proporcional, a responsabilidade de uma seguradora fica limitada pelo Limite de Retenção (LR).

A Teoria de Valores Extremos (TVE) é um ramo da probabilidade que estuda o comportamento estocástico de extremos associados a um conjunto de variáveis aleatórias (ou vetores aleatórios) com distribuição comum. Dentro da denominação geral de extremos inclui-se o máximo e o mínimo, estatísticas de ordem extremas e excessos acima (ou abaixo) de limiares altos (ou baixos). O importante é que as características e propriedades das distribuições desses extremos aleatórios são determinadas pelas caudas extremas (inferior e superior) da distribuição dos dados. Restringir a atenção às caudas de uma distribuição apresenta a vantagem de termos diversos modelos estatísticos adequados para as mesmas. Estes modelos baseados em excessos além de limiares nos permitirão fazer inferência mais precisa sobre as caudas e parâmetros da distribuição dos dados.

Nesta última década do século XX, vimos várias técnicas estatísticas, desenvolvidas sob a denominação geral de teoria de valores extremos, experimentarem uma aceitação prática crescente, especialmente nas áreas de finanças e atuária. Esta popularidade é devida à sua habilidade de melhor quantificar as probabilidades de ocorrência de eventos raros, ou mesmo nunca antes observados, de um processo. Exemplos de tais aplicações podem ser consultados em McNeil (1997), com uma revisão dos resultados principais da TVE e ilustrações usando dados atuariais e Smith (1999) com revisões de técnicas estatísticas usadas na TVE e exemplos de suas utilidades.

Apesar de ser bastante útil ao ambiente de seguros, não existem muitos trabalhos no Brasil usando esta técnica. Portanto, neste artigo, aplicaremos a teoria de valores extremos à avaliação do risco de um contrato de resseguro de excesso de danos para sinistros de incêndio, raio e explosão. Para isto, será simulada a distribuição agregada dos sinistros por meio de convoluções e comparada a *performance* dos modelos utilizando variadas distribuições para a severidade.

O artigo está organizado da seguinte forma: na seção 2 serão apresentados conceitos de teoria de risco, valores extremos e possíveis métodos para inferência do montante de sinistros; na seção 3 são apresentados os dados e os modelos propostos para avaliação dos contratos de resseguro; na seção 4 é feita a análise dos resultados. Por fim, na seção 5, são feitas as considerações finais.

2. Teorias do Risco e de Valores Extremos

De acordo com a teoria do risco coletivo, a soma dos valores dos sinistros (sinistro agregado) é denotada por (Bowers et al. 1997):

$$S_{col} = \sum_{i=1}^N X_i \quad (1)$$

Na equação (1), N representa o número (aleatório) de sinistros e X_i a severidade de cada sinistro i . É comum assumir-se que N seja proveniente de uma distribuição Poisson ou Binomial Negativa independente da variável X_i . Grande parte da literatura em risco atuarial estuda a distribuição de S_{col} . Contudo, raramente estes estudos abordam o problema da influência de grandes sinistros no comportamento de S_{col} .

Para a avaliação de contratos de resseguro, uma tarefa relevante é a análise da distribuição de grandes perdas usando-se dados históricos (Hogg & Klugman, 1984). Para os contratos baseados em excesso de danos, estamos interessados em mensurar o número de ocorrências e a severidade de sinistros acima de um determinado limiar, chamado limite de retenção. Eventos catastróficos podem resultar no pagamento de vários sinistros de alto valor. Assim, torna-se necessário o cálculo ou a estimação apropriada do prêmio de resseguro para cobrir este risco. Estes cálculos, geralmente associados a eventos de baixa probabilidade, podem tornar-se mais precisos usando-se modelos adequados da TVE (Teoria de Valores Extremos). Isso significa obter uma estimação acurada da parte extrema da cauda da distribuição de S_{col} .

Sejam os sinistros de um contrato de seguro denotados por X_1, X_2, \dots , podemos assumir que eles são variáveis aleatórias independentes e identicamente distribuídas. Podemos, também, representar a sua distribuição acumulada por $F(x) = Pr(X \leq x)$, onde x é positivo. Agora, se estivermos interessados nos sinistros que excedam um determinado limiar (u), isto significa dizer que a perda Y_i para o sinistro X_i é dada por:

$$Y_i = \begin{cases} 0 & \text{se } 0 < X_i < u \\ X_i - u & \text{se } u \leq X_i \end{cases} \quad (2)$$

Uma questão aparece aqui de imediato, como escolher o limiar “ u ”? Para responder a esta questão, precisamos fixar um período de observação dos dados e ter informação acerca da frequência dos sinistros ocorridos em um período de tempo. Considere N o número de sinistros acima do limiar “ u ”. Então, o montante agregado de sinistros é:

$$Z = \sum_{i=1}^N Y_i \quad (3)$$

O problema de precificação pode ser resolvido de inúmeras maneiras. Uma fórmula comum para se tarifar é $P = E[Z] + k \cdot \text{sigma}[Z]$, onde o preço é o valor esperado da perda mais um carregamento de segurança que é k vezes o desvio padrão da perda agregada. O valor esperado de Z ($E[Z]$) é conhecido como prêmio puro e pode ser demonstrado que $E[Z] = E[Y] \cdot E[N]$. De acordo com Dhaene & Vyncke (2002), outras possibilidades para se inferir o montante de sinistros agregados são: (i) por meio do uso de funções geratrizes de momentos das variáveis do problema; (ii) pela utilização de aproximações embasadas pelo Teorema do Limite Central, em que se encaixam a Gamma Transladada e a NP (Normal Power); (iii) por meio de fórmulas recursivas como a de Panjer; e, por fim, a que será usada neste artigo, (iv) pelo emprego de convoluções.



Conforme dito acima, estamos interessados na distribuição da quantidade total de sinistros acima de um determinado limiar (u), que chamamos de Z . Assumindo que as distribuições de X_i são independentes, a distribuição da soma pode ser calculada pelo uso de convoluções. A operação “convolução” calcula a função de distribuição $A + B$ pelo uso das distribuições de A e B , como segue:

$$F_{A+B}(s) = \Pr(A + B \leq s) = \int F_B(s-a) dF_A(a) =: F_A * F_B(s) \quad (4)$$

A distribuição acumulada $F_A * F_B(s)$ é chamada convolução da distribuição de A com a distribuição de B . Para a distribuição de $A + B + C$, não interessa em qual ordem efetuamos as convoluções, temos:

$$(F_A * F_B) * F_C \equiv F_A * (F_B * F_C) \equiv F_A * F_B * F_C \quad (5)$$

Para a soma de “ n ” variáveis aleatórias independentes e identicamente distribuídas com distribuição marginal F , a distribuição acumulada da “ n -ésima” convolução de F é escrita como:

$$F * F * F * \dots * F \equiv F^{*n} \quad (6)$$

Voltando ao problema da escolha do limiar “ u ”, uma possível solução seria definir “ u ” tal que $\Pr(Z > 0) < p$ para algum valor de p . Isto equivale a dizer que “ u ” é determinado de tal forma que, no período de cobertura, um sinistro que exceda o limite “ u ” ocorre com probabilidade menor que p . A questão está em inferir um alto quantil da distribuição de severidade, momento em que o ressegurador é acionado. Para isto, precisamos de uma boa estimativa da distribuição da severidade para altos valores, ou seja, na cauda. Nós também precisamos ter uma boa estimativa do número de sinistros, mas este não é objeto de discussão neste artigo.

A teoria de valores extremos descreve o comportamento de grandes observações que excedem um limiar alto. Esta característica nos leva a estender sua aplicação diretamente para contratos de resseguro. A questão principal é: dado que uma observação é alta, o quão extrema ela é? A distribuição que se ajusta bem a este tipo de dado é a Pareto Generalizada (GPD). A GPD é geralmente expressa da seguinte forma (Mendes, 2004):

$$G_{\varepsilon, \sigma}(x) = \begin{cases} 1 - \left(1 + \frac{\varepsilon x}{\sigma}\right)^{-1/\varepsilon}; & \text{se } \varepsilon \neq 0 \\ 1 - \exp(-x/\sigma); & \text{se } \varepsilon = 0 \end{cases} \quad (7)$$

onde $\sigma > 0$, e o suporte é $x \geq 0$ quando $\varepsilon \geq 0$ e $0 \leq x \leq -\sigma/\varepsilon$ quando $\varepsilon < 0$. A GPD é a generalização de algumas distribuições. Quando $\varepsilon > 0$, temos uma reparametrização da distribuição Pareto com escala $\alpha = 1/\varepsilon$. Se $\varepsilon < 0$, temos uma distribuição Pareto do tipo II. Para $\varepsilon = 0$, temos a distribuição exponencial. Podemos estender a família ao adicionar um parâmetro de locação μ .

Considere x_0 o ponto à direita final da distribuição F . Isto equivale a dizer que $x_0 = \sup \{x \in R : F(x) < 1\} < \infty$. A distribuição dos excessos acima de um limiar é definida por:

$$F_u(x) = \Pr(X - u \leq x | X > u) = \frac{F(x+u) - F(u)}{1 - F(u)} \text{ para } 0 \leq x < x_0 - u \quad (8)$$

O teorema de Pickands-Balkema-de Haan (Balkema & De Haan 1974, Pickands 1975) mostra que, sob condições específicas (domínio de atração do máximo), a distribuição Pareto generalizada é a distribuição limite para a distribuição dos excessos, quando o limiar tende para o ponto à direita final. Este teorema sugere que, para um limiar suficientemente alto, a distribuição dos excessos pode ser aproximada por $G_{\varepsilon, \sigma}(x)$ para alguns valores de ε e σ . A relevância estatística do resultado é que podemos ajustar distribuições GPD aos dados que excedem limiares altos.

Se pudermos ajustar a GPD à distribuição condicional dos excessos acima de um limiar alto, nós podemos ajustá-la à cauda da distribuição original acima do limiar alto (Reiss & Thomas 1997). Para $x \geq u$, ou seja, pontos na cauda da distribuição, tem-se:

$$F(x) = \Pr(X \leq x) = (1 - \Pr(X \leq x))F_u(x - u) + \Pr(X \leq u) \quad (9)$$

Sabe-se que podemos estimar $F_u(x - u)$ por $G_{\varepsilon, \sigma}(x - u)$ para u grande. Pode-se estimar $\Pr(X \geq u)$ dos dados pela distribuição empírica acumulada no ponto u ($F_n(u)$). Isto significa que para $x \geq u$, podemos usar a seguinte estimativa para a cauda:

$$\overline{F(x)} = (1 - F_n(u))G_{\varepsilon, u, \sigma}(x) + F_n(u) \quad (10)$$

Pode-se mostrar que $\overline{F(x)}$ também é uma distribuição Pareto generalizada, com o mesmo parâmetro ε , mas com parâmetro de escala $\overline{\sigma} = \sigma(1 - F_n(u))^\varepsilon$ e parâmetro de locação $\overline{\mu} = u - \overline{\sigma}((1 - F_n(u))^{-\varepsilon} - 1)/\varepsilon$. Existe, ainda, o caso particular em que se modela a distribuição da exponencial dos sinistros com uma Pareto generalizada. Esta distribuição, Anderson & Dancy (1992) chamam de MGPD (Modified Generalized Pareto Distribution). O complementar de sua distribuição acumulada é:

$$\Pr(S > s) = \begin{cases} (1 + \varepsilon \lambda s^\gamma)^{-1/\varepsilon}; & \text{se } \varepsilon \neq 0 \\ \exp(-\lambda s^\gamma) & ; \text{se } \varepsilon = 0 \end{cases} \quad (11)$$

Para detalhes sobre as distribuições Poisson, Gamma e Log-Normal utilizadas, consultar DeGroot (1987). Neste trabalho, à exceção da distribuição Gamma, que teve seus parâmetros estimados com estimadores empíricos, para inferir os parâmetros das distribuições, utilizamos o método de máxima verossimilhança. Para maiores detalhes sobre a aplicação deste método para teoria de valores extremos, veja Hosking & Wallis (1987).

3. Dados utilizados e modelos propostos

3.1 Dados utilizados

Os dados utilizados nesta aplicação foram os sinistros diretos ocorridos de incêndio, raio e explosão dos ramos compreensivos condominiais (comercial, hotéis e shopping) e empresariais (comércio, indústria e hotéis) para o período de análise de Janeiro/2003 a Dezembro/2003 (1 ano) de um conjunto selecionado de seguradoras.

Tabela 1 – Estatísticas descritivas dos sinistros no ano de 2003

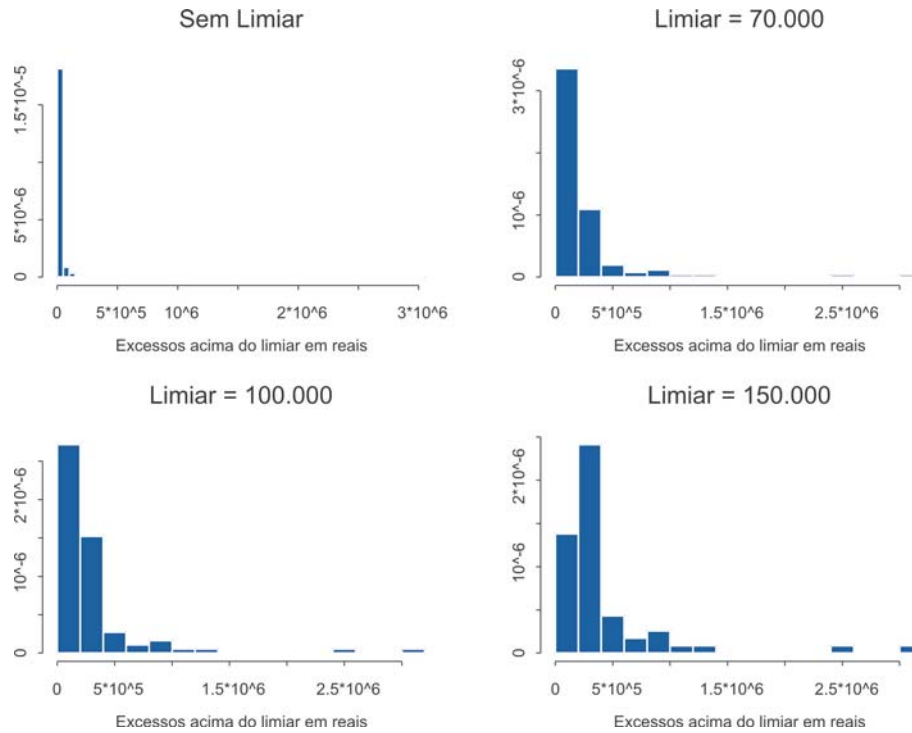
| <i>Estatística</i> | <i>Valor</i> |
|--------------------|--------------|
| Mínimo | 1,00 |
| 1º Quartil | 500,00 |
| Mediana | 2.150,00 |
| 3º Quartil | 10.829,00 |
| Máximo | 3.041.790,00 |
| Média | 24.543,00 |
| Desvio Padrão | 119.095 |
| Assimetria | 16.16 |
| Curtose Absoluta | 347.70 |

* Número de Sinistros: 1808.

O Gráfico 1 mostra os histogramas dos excessos dos sinistros sob consideração neste artigo para os diferentes níveis de retenção. No primeiro histograma, em que não há limite de retenção, as altas severidades de alguns sinistros fazem o histograma ficar mais difícil de analisar ou de se verificar alguma distribuição de probabilidade.



Gráfico 1 – Histograma dos sinistros individuais de incêndio, raio e explosão para os diferentes níveis de retenção (sem LR, 70.000, 100.000 e 150.000)



Na Tabela 2, pode ser observada a evolução do montante de sinistros pelos meses de 2003 para cada um dos diferentes níveis de retenção. É com base nestes valores que será feita a comparação dos resultados obtidos pelos modelos ajustados com a distribuição empírica.

Tabela 2 – Evolução do montante de sinistros pelos meses do ano de 2003 para cada um dos diferentes níveis de retenção

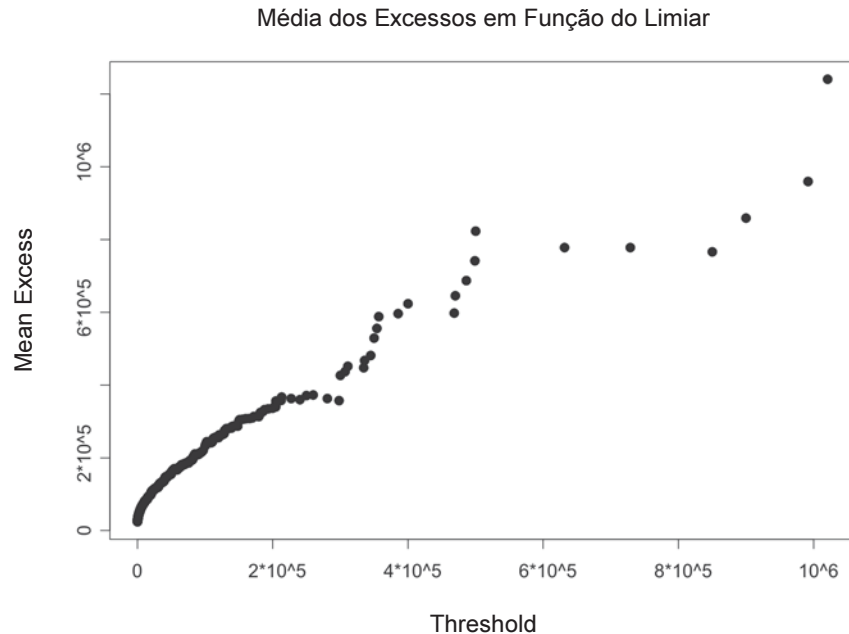
| Meses | Limite de Retenção | | | |
|--------|--------------------|--------------|--------------|--------------|
| | LR = 0 | LR = 70.000 | LR = 100.000 | LR = 150.000 |
| | Montante* | Montante* | Montante* | Montante* |
| Jan/03 | 4.418.656,59 | 2.217.537,42 | 1.853.889,21 | 1.461.479,21 |
| Fev/03 | 4.384.984,08 | 2.442.033,21 | 2.119.584,21 | 1.691.979,21 |
| Mar/03 | 6.343.084,36 | 3.974.978,85 | 3.690.856,89 | 3.380.565,45 |
| Abr/03 | 3.069.782,69 | 1.094.986,01 | 784.994,01 | 488.637,01 |
| Mai/03 | 2.415.757,57 | 674.707,00 | 474.447,00 | 312.519,00 |
| Jun/03 | 3.397.857,12 | 1.915.421,90 | 1.685.926,65 | 1.393.850,65 |
| Jul/03 | 4.121.676,59 | 2.207.318,76 | 1.948.987,56 | 1.598.987,56 |
| Ago/03 | 5.539.043,96 | 3.733.894,75 | 3.443.127,86 | 3.145.383,86 |
| Set/03 | 1.661.200,27 | 437.481,04 | 267.565,00 | 148.222,00 |
| Out/03 | 2.547.769,78 | 904.094,62 | 630.710,62 | 449.382,62 |
| Nov/03 | 2.146.147,33 | 685.603,03 | 475.603,03 | 300.000,00 |
| Dez/03 | 4.328.535,22 | 2.931.914,00 | 2.631.914,00 | 2.167.950,00 |

* Montante de excessos de sinistros acima do limite de retenção LR.

3.2. Modelos propostos Foram propostos 4 modelos distintos com diferentes distribuições assumidas para a severidade dos sinistros, quais sejam: (i) Gamma, (ii) Log-Normal, (iii) Pareto generalizada (GPD) e (iv) Pareto Generalizada Modificada (MGPD). Cada um desses quatro modelos foi aplicado aos dados, considerando diferentes níveis de limite de retenção (R\$ 0,00 / R\$ 70.000 / R\$ 100.000,00 / R\$ 150.000,00). Estes níveis de limite de retenção são considerados franquias dedutíveis ou, no linguajar da teoria de valores extremos, limiares. Para todos estes modelos, foi considerada como distribuição do número de sinistro a Poisson.

À exceção da distribuição Gamma, que foi ajustada com estimadores empíricos, as outras distribuições foram ajustadas usando estimadores de máxima verossimilhança, inclusive a Poisson. Depois de estimados os parâmetros, foram feitas as convoluções por meio de simulações para obtenção da distribuição dos sinistros agregados por mês. Os ajustes e as simulações foram executados no pacote estatístico SPlus®.

Gráfico 2 – Média dos excessos em função do limiar



A interpretação do gráfico de média dos excessos (Gráfico 2) pode ser encontrada em Hogg & Klugman (1984). Se os pontos mostrarem uma tendência de subida, então é um sinal de um comportamento de cauda pesada. Dados exponencialmente distribuídos resultariam em um gráfico com uma linha horizontal e dados de uma distribuição de cauda curta resultariam em um gráfico com linha decrescente. Em particular, os dados neste artigo apresentam um gráfico de média dos excessos com coeficiente angular positivo para um determinado limiar u . Isto pode ser uma indicação de que os dados seguem uma distribuição Pareto generalizada.

4. Resultados

Para ajustar as distribuições MGPD, GPD e Log-Normal aos dados, foram utilizados os estimadores de máxima verossimilhança. Entretanto, não foi possível executar a otimização para se encontrar as estimativas que maximizavam a função de verossimilhança da distribuição Gamma, que teve seus parâmetros estimados com estimadores empíricos. Depois de feitas as estimativas dos parâmetros das distribuições, foi utilizado o algoritmo de bootstrap (Cribari Neto, 2000) para se acessar o erro padrão destas estimativas. Na Tabela 3, pode-se observar o valor de cada uma das estimativas das distribuições para cada limite de retenção considerado e o respectivo erro padrão.

Tabela 3 – Estimativa e erro padrão dos parâmetros para as distribuições ajustadas para severidade para os diferentes níveis de limite de retenção

| | | Distribuições para Severidade* | | | | | | | | |
|--------------|-------------|--------------------------------|----------|-----------|-------|---------|-------|------------|-------|------------|
| | | Gamma** | | MGPD | | GPD | | Log-Normal | | |
| | | α | β | λ | ξ | β | ξ | β | μ | σ^2 |
| LR*** = 0 | Estimativa | 0,042 | 1,73e-06 | 1,010 | 1,641 | 1.754 | 1,615 | 1.642 | 7,84 | 2,113 |
| | Erro Padrão | 0,017 | 8,94e-07 | 0,007 | 0,064 | 100 | 0,064 | 93 | 0,05 | 0,034 |
| LR = 70.000 | Estimativa | 0,238 | 1,26e-06 | 0,920 | 0,393 | 42.167 | 0,455 | 112.518 | 11,04 | 1,588 |
| | Erro Padrão | 0,107 | 8,12e-07 | 0,013 | 0,123 | 6.841 | 0,134 | 21.057 | 0,14 | 0,143 |
| LR = 100.000 | Estimativa | 0,278 | 1,19e-06 | 0,930 | 0,369 | 62.363 | 0,422 | 148.663 | 11,31 | 1,615 |
| | Erro Padrão | 0,141 | 1,01e-06 | 0,015 | 0,146 | 11.984 | 0,156 | 31.468 | 0,17 | 0,142 |
| LR = 150.000 | Estimativa | 0,339 | 1,13e-06 | 1,110 | 0,446 | 851.737 | 0,378 | 201.292 | 11,73 | 1,437 |
| | Erro Padrão | 0,184 | 9,82e-07 | 0,014 | 0,193 | 234.524 | 0,177 | 48.021 | 0,19 | 0,174 |

* Erro Padrão calculado por meio de algoritmo de *bootstrap*.

** Ajustada com estimadores empíricos para os parâmetros.

*** LR (limite de retenção).

Na Tabela 4, são mostradas as estatísticas de soma do logaritmo da verossimilhança com penalidade pelo número de parâmetros estimados. Sabe-se que quanto maior o número de parâmetros a ser estimado no modelo maior a verossimilhança. Então, como critério para escolha do melhor ajuste por máxima verossimilhança, deve-se colocar uma penalidade pelo número de parâmetros. Assim, podemos denotar a fórmula da estatística da seguinte forma:

$$\sum_{i=1}^n \log(f(x_i)) - 2 * k \quad (12)$$

onde $f(x_i)$ é a função de densidade de probabilidade da severidade sob consideração aplicada à observação x_i e k é o número de parâmetros estimados para a função de densidade de probabilidade. Sob este critério, quanto maior o valor desta estatística, melhor será o ajuste.

Tabela 4 – Estatísticas de soma do logaritmo da verossimilhança para cada uma das distribuições ajustadas para severidade e para os diferentes níveis de limite de retenção e percentual de número de sinistros que excederam o limiar em relação ao total de sinistros

| Soma do Log-Verossimilhança* | Distribuições | | | | % de sinistros que excederam o LR |
|------------------------------|---------------|------------------|------------------|-------------------|-----------------------------------|
| | Gamma** | MGPD | GPD | Log-Normal | |
| LR*** = 0 | -20.338,42 | -18.120,63 | -18.118,53 | -18.092,90 | 100% |
| LR = 70.000 | -1.683,74 | -1,643,86 | -1.643,98 | -1.643,99 | 7,02% |
| LR = 100.000 | -1.157,36 | -1,139,66 | -1.138,54 | -1.139,46 | 4,76% |
| LR = 150.000 | -757,70 | -748,74 | -745,84 | -746,64 | 3,04% |

* Soma do logaritmo da verossimilhança com penalidade pelo número de parâmetros.

** Ajustada com estimadores empíricos para os parâmetros.

*** LR (Limite de retenção).

Pela Tabela 4, podemos observar que, para limites de retenção altos, as distribuições de valores extremos (MGPD e GPD) têm uma melhor *performance* se comparadas às distribuições Gamma e Log-Normal. Para cada uma destas distribuições, foram simuladas as convoluções para que fosse obtida a distribuição do montante de sinistros agregados. Os Gráficos 6 a 9 no anexo apresentam as funções de densidade de probabilidade do montante de sinistros agregados estimadas para cada uma das distribuições de severidade para o limite de retenção de R\$ 150.000. Para estes gráficos foi feito um alisamento nos dados, uma vez que a simulação das convoluções gera a distribuição acumulada e não a densidade de probabilidade, conforme metodologia explicitada no Splus® (2003).

Depois de ajustadas as distribuições para severidade, foram simuladas as distribuições do montante de sinistros agregados por meio de convoluções. Nas Tabelas 5, 6, 7 e 8 são apresentados os resultados das comparações com as distribuições empíricas. Para se verificar o modelo mais adequado aos dados, foi efetuado o cálculo da soma dos erros ao quadrado.



Tabela 5 – Comparação do montante de sinistros agregados estimado com o observado no ano de 2003, sem considerar limite de retenção

| Probabilidade | Observado | Gamma* | MGPD | GPD | Log-Normal |
|---------------|--------------|------------------|----------------|----------------|--------------|
| 0,07692 | 1.661.200,27 | 1.810.000,00 | 3.200.000,00 | 3.170.000,00 | 1.840.000,00 |
| 0,15385 | 2.146.147,33 | 2.220.000,00 | 4.290.000,00 | 4.290.000,00 | 2.090.000,00 |
| 0,23077 | 2.415.757,57 | 2.540.000,00 | 5.600.000,00 | 5.500.000,00 | 2.300.000,00 |
| 0,30769 | 2.547.769,78 | 2.820.000,00 | 6.900.000,00 | 6.900.000,00 | 2.500.000,00 |
| 0,38462 | 3.069.782,69 | 3.090.000,00 | 8.800.000,00 | 8.600.000,00 | 2.690.000,00 |
| 0,46154 | 3.397.857,12 | 3.360.000,00 | 11.100.000,00 | 11.000.000,00 | 2.910.000,00 |
| 0,53846 | 4.121.676,59 | 3.640.000,00 | 14.400.000,00 | 14.300.000,00 | 3.150.000,00 |
| 0,61538 | 4.328.535,22 | 3.940.000,00 | 19.500.000,00 | 19.400.000,00 | 3.410.000,00 |
| 0,69231 | 4.384.984,08 | 4.280.000,00 | 27.800.000,00 | 27.900.000,00 | 3.730.000,00 |
| 0,76923 | 4.418.656,59 | 4.680.000,00 | 43.500.000,00 | 44.200.000,00 | 4.140.000,00 |
| 0,84615 | 5.539.043,96 | 5.200.000,00 | 83.300.000,00 | 83.200.000,00 | 4.790.000,00 |
| 0,92308 | 6.343.084,36 | 6.000.000,00 | 120.100.000,00 | 120.000.000,00 | 6.000.000,00 |
| SOMA** | | 8,139E+11 | 2,158E+16 | 2,154E+16 | 3,406E+12 |

* Ajustada com estimadores empíricos para os parâmetros.

** Soma dos quadrados das diferenças entre os montantes estimados e os observados.

Na Tabela 5, podemos notar que, por ter caudas mais pesadas que as distribuições Gamma e Log-Normal, as distribuições de valores extremos (MGPD e GPD) “erram” por um valor bastante alto. Ou seja, considerando um limiar igual a zero, ou sem considerar limite de retenção, os modelos com distribuições Gamma e Log-Normal para a severidade têm melhor *performance*. Na Tabela 6, expomos os resultados para um limite de retenção igual a R\$ 70.000,00.



Tabela 6 – Comparação do montante de sinistros agregados estimado com o observado no ano de 2003, considerando limite de retenção igual a R\$70.000,00

| Probabilidade | Observado* | Gamma** | MGPD | GPD | Log-Normal |
|---------------|--------------|------------------|--------------|--------------|--------------|
| 0,07692 | 437.481,04 | 340.000,00 | 620.000,00 | 680.000,00 | 520.000,00 |
| 0,15385 | 674.707,00 | 620.000,00 | 850.000,00 | 920.000,00 | 750.000,00 |
| 0,23077 | 685.603,03 | 860.000,00 | 1.040.000,00 | 1.140.000,00 | 940.000,00 |
| 0,30769 | 904.094,62 | 1.070.000,00 | 1.220.000,00 | 1.330.000,00 | 1.150.000,00 |
| 0,38462 | 1.094.986,01 | 1.290.000,00 | 1.410.000,00 | 1.520.000,00 | 1.360.000,00 |
| 0,46154 | 1.915.421,90 | 1.520.000,00 | 1.600.000,00 | 1.720.000,00 | 1.580.000,00 |
| 0,53846 | 2.207.318,76 | 1.760.000,00 | 1.810.000,00 | 1.940.000,00 | 1.820.000,00 |
| 0,61538 | 2.217.537,42 | 2.040.000,00 | 2.050.000,00 | 2.180.000,00 | 2.120.000,00 |
| 0,69231 | 2.442.033,21 | 2.360.000,00 | 2.330.000,00 | 2.460.000,00 | 2.490.000,00 |
| 0,76923 | 2.931.914,00 | 2.750.000,00 | 2.680.000,00 | 2.820.000,00 | 2.960.000,00 |
| 0,84615 | 3.733.894,75 | 3.270.000,00 | 3.200.000,00 | 3.320.000,00 | 3.690.000,00 |
| 0,92308 | 3.974.978,85 | 4.110.000,00 | 4.120.000,00 | 4.320.000,00 | 5.000.000,00 |
| SOMA*** | | 8,625E+11 | 1,056E+12 | 1,102E+12 | 1,536E+12 |

* Montantes de sinistros observados considerando limite de retenção de R\$ 70.000,00.

** Ajustada com estimadores empíricos para os parâmetros.

*** Soma dos quadrados das diferenças entre os montantes estimados e os observados.

Na Tabela 6, já podemos perceber a vantagem de se usar distribuições de valores extremos para limiares altos sobre a distribuição Log-Normal para os dados sob consideração. Neste caso, o modelo mais adequado, ou o que apresentou a menor soma dos erros quadrados, foi o que utilizou a Gamma para a distribuição da severidade.

Tabela 7 – Comparação do montante de sinistros agregados estimado com o observado no ano de 2003, considerando limite de retenção igual a R\$100.000,00

| Probabilidade | Observado* | Gamma** | MGPD | GPD | Log-Normal |
|---------------|--------------|--------------|--------------|------------------|--------------|
| 0,07692 | 267.565,00 | 210.000,00 | 390.000,00 | 430.000,00 | 310.000,00 |
| 0,15385 | 474.447,00 | 420.000,00 | 600.000,00 | 650.000,00 | 500.000,00 |
| 0,23077 | 475.603,03 | 620.000,00 | 790.000,00 | 840.000,00 | 680.000,00 |
| 0,30769 | 630.710,62 | 810.000,00 | 960.000,00 | 1.030.000,00 | 870.000,00 |
| 0,38462 | 784.994,01 | 1.020.000,00 | 1.140.000,00 | 1.200.000,00 | 1.070.000,00 |
| 0,46154 | 1.685.926,65 | 1.230.000,00 | 1.320.000,00 | 1.400.000,00 | 1.290.000,00 |
| 0,53846 | 1.853.889,21 | 1.470.000,00 | 1.520.000,00 | 1.600.000,00 | 1.550.000,00 |
| 0,61538 | 1.948.987,56 | 1.730.000,00 | 1.750.000,00 | 1.830.000,00 | 1.850.000,00 |
| 0,69231 | 2.119.584,21 | 2.040.000,00 | 2.030.000,00 | 2.120.000,00 | 2.240.000,00 |
| 0,76923 | 2.631.914,00 | 2.230.000,00 | 2.380.000,00 | 2.460.000,00 | 2.740.000,00 |
| 0,84615 | 3.443.127,86 | 2.900.000,00 | 2.880.000,00 | 2.960.000,00 | 3.520.000,00 |
| 0,92308 | 3.690.856,89 | 3.850.000,00 | 3.820.000,00 | 3.880.000,00 | 5.100.000,00 |
| SOMA*** | | 1,006E+12 | 1,054E+12 | 9,808E+11 | 2,459E+12 |

* Montantes de sinistros observados considerando limite de retenção de R\$ 100.000,00.

** Ajustada com estimadores empíricos para os parâmetros.

*** Soma dos quadrados das diferenças entre os montantes estimados e os observados.

Para o limite retenção de R\$ 100.000, a utilização da distribuição GPD para a severidade proveu o melhor ajuste dentre os tentados. Parece que com limiares altos as distribuições de valores extremos apresentam melhores *performances*. Apresentamos, abaixo, os Gráficos 3 a 5 que demonstram os resultados do ajuste da distribuição GPD para modelagem dos sinistros individuais. Podemos perceber o bom ajuste da distribuição GPD para os sinistros acima do limite retenção de R\$ 100.000.

Gráfico 3 – Distribuição empírica dos excessos de sinistros individuais, considerando um limiar de R\$ 100.000, comparada com a distribuição GPD ajustada com estimadores de máxima verossimilhança para este mesmo conjunto de dados.

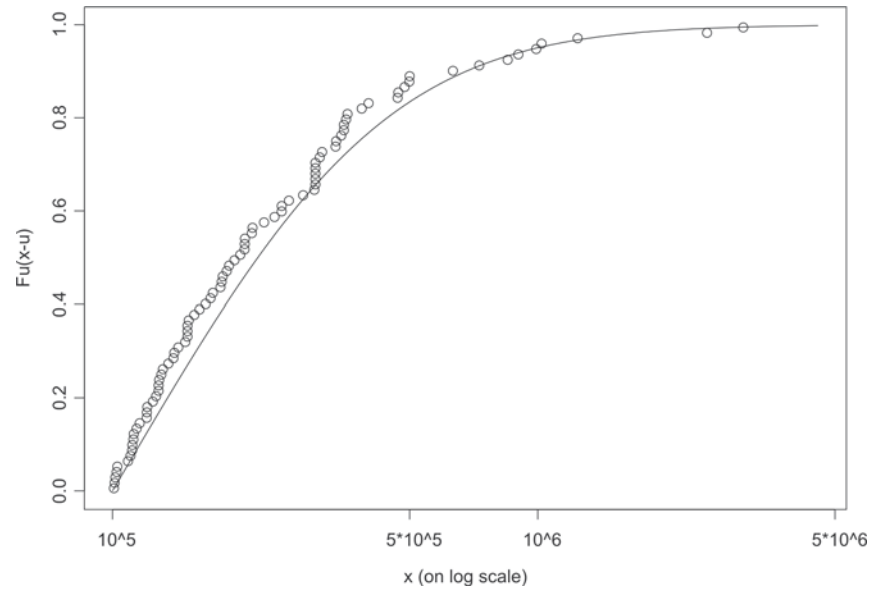


Gráfico 4 – Cauda da distribuição empírica dos excessos de sinistros individuais, considerando um limiar de R\$ 100.000, comparada com a cauda da distribuição GPD ajustada com estimadores de máxima verossimilhança para este mesmo conjunto de dados

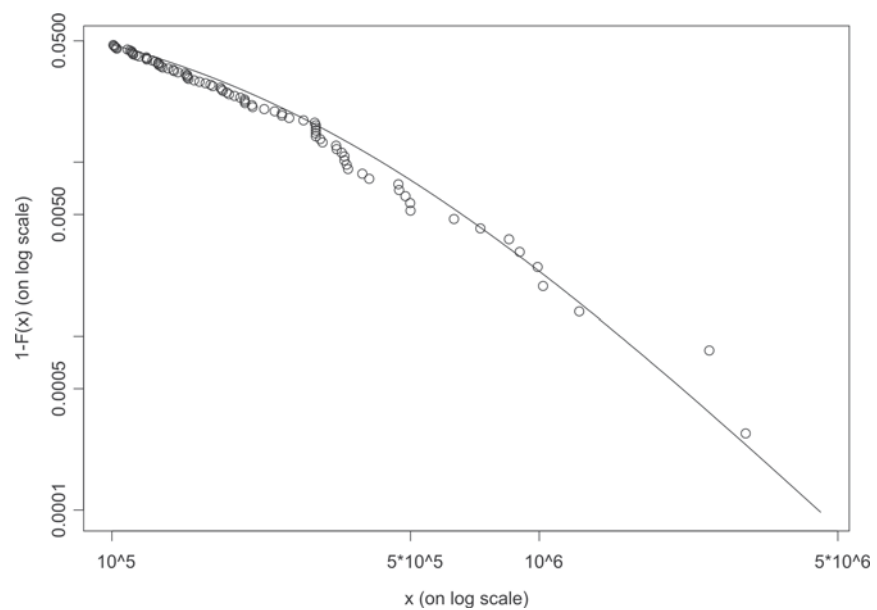


Gráfico 5 – QQ Plot dos resíduos do ajuste pela GPD nos sinistros individuais, para um limiar de R\$ 100.000

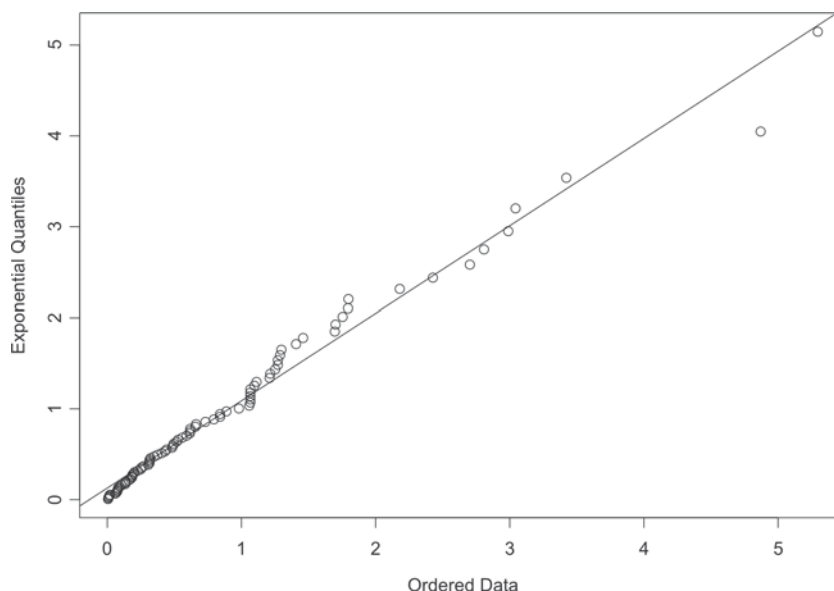


Tabela 8 – Comparação do montante de sinistros agregados estimado com o observado no ano de 2003, considerando limite de retenção igual a R\$ 150.000,00

| Probabilidade | Observado* | Gamma** | MGPD | GPD | Log-Normal |
|---------------|--------------|--------------|------------------|--------------|--------------|
| 0,07692 | 148.222,00 | 80.000,00 | 240.000,00 | 200.000,00 | 160.000,00 |
| 0,15385 | 300.000,00 | 200.000,00 | 430.000,00 | 380.000,00 | 300.000,00 |
| 0,23077 | 312.519,00 | 380.000,00 | 610.000,00 | 540.000,00 | 450.000,00 |
| 0,30769 | 449.382,62 | 540.000,00 | 770.000,00 | 700.000,00 | 600.000,00 |
| 0,38462 | 488.637,01 | 730.000,00 | 940.000,00 | 850.000,00 | 760.000,00 |
| 0,46154 | 1.393.850,65 | 920.000,00 | 1.120.000,00 | 1.020.000,00 | 940.000,00 |
| 0,53846 | 1.461.479,21 | 1.140.000,00 | 1.320.000,00 | 1.210.000,00 | 1.130.000,00 |
| 0,61538 | 1.598.987,56 | 1.390.000,00 | 1.540.000,00 | 1.420.000,00 | 1.370.000,00 |
| 0,69231 | 1.691.979,21 | 1.690.000,00 | 1.820.000,00 | 1.690.000,00 | 1.680.000,00 |
| 0,76923 | 2.167.950,00 | 2.060.000,00 | 2.160.000,00 | 2.030.000,00 | 2.080.000,00 |
| 0,84615 | 3.145.383,86 | 2.570.000,00 | 2.650.000,00 | 2.530.000,00 | 2.680.000,00 |
| 0,92308 | 3.380.565,45 | 3.430.000,00 | 3.500.000,00 | 3.400.000,00 | 3.860.000,00 |
| SOMA*** | | 7,995E+11 | 7,950E+11 | 8,879E+11 | 9,380E+11 |

* Montantes de sinistros observados considerando limite de retenção de R\$ 150.000,00.

** Ajustada com estimadores empíricos para os parâmetros.

*** Soma dos quadrados das diferenças entre os montantes estimados e os observados.

Para o limite de retenção de R\$ 150.000, o modelo considerando a distribuição para severidade MGPD foi o que apresentou o melhor resultado. Com base nos resultados apresentados, pode-se supor que, para limites de retenção elevados, o uso de distribuições de valores extremos para a modelagem da severidade traz benefícios para a estimação do montante de sinistros agregados.

5. Considerações finais Neste artigo foram apresentados modelos para a avaliação do risco de contratos de resseguro utilizando a teoria de valores extremos. Para isto, foi feita uma aplicação com dados reais de sinistros diretos de incêndio / raio / explosão de seguradoras que operam os ramos compreensivos condominiais e empresariais. Assim, foram ajustadas diversas distribuições (Gamma, Log-Normal, GPD e MGPD) à severidade dos sinistros, considerando diferentes níveis de retenção de um contrato de resseguro de excesso de danos hipotético. O ajuste da distribuição do número de sinistros foi feito usando uma Poisson. À exceção da distribuição Gamma, que teve seus parâmetros estimados empiricamente, todos os parâmetros das outras distribuições foram estimados por máxima verossimilhança.

Para a obtenção da distribuição dos sinistros agregados mensais foram utilizadas convoluções com a ferramenta de simulação no pacote estatístico SPlus®. Cada nível de limite de retenção (R\$ 0,00 / R\$ 70.000,00 / R\$ 100.000,00 / R\$ 150.000,00) foi considerado uma franquia dedutível ou, no linguajar da teoria de valores extremos, um limiar. Como método de comparação dos modelos, foi efetuada a soma das diferenças (“erros”) entre as distribuições ajustadas e a empírica ao quadrado.

Como resultados deste artigo, podemos destacar o melhor ajuste de distribuições de valores extremos, como GPD e MGPD, a dados de sinistros de grandes riscos. Esta melhoria fica ainda mais evidenciada quanto maior o limiar ou limite de retenção. Este resultado era de se esperar por causa da POT (*Peaks Over Threshold*), estabilidade que as distribuições Pareto generalizadas apresentam (MENDES, 2004). Portanto, a teoria de valores extremos é uma ferramenta poderosa para avaliação do risco de contratos de resseguro.

Os resultados da aplicação apresentados neste artigo são de grande relevância para a avaliação do risco de contratos de resseguro. Outra observação que pode ser feita é a extensão direta destes modelos à avaliação do risco de subscrição de prêmio de uma resseguradora. Uma vez que o método de obtenção da distribuição dos sinistros agregados é feito por convolução, ou seja, é obtida a distribuição completa desta variável, pode-se inferir, com base em uma probabilidade crítica, o nível de capital que uma resseguradora deve manter para reter um determinado risco.

6. Referências bibliográficas

- ANDERSON, C. W.; DANCY, G. P. The severity of extreme events. **Research Report 92/593** Department of Probability and Statistic, Univ. of Sheffield., 1992.
- BALKEMA, A. A. ; DE HAAN, L. Residual Life Time at Great Age. **Ann. Probability**, n. 2, p. 792-804, 1974.
- BOWERS, N.L. /et al/. **Actuarial Mathematics**. 2.ed.. SOA – Society of Actuaries, 1997.
- CRIBARI NETO, F.. **Bootstrap and applications**. Departamento de Estatística, Universidade Federal de Pernambuco, 2000. Disponível em: www.de.ufpe.br.
- D'ARCY, S.P.; FRANCE, V. G.; GORVETT, R. W. **Pricing Catastrophe Risk**: Could CAT futures have coped with Andrew ? 1999 Casualty Actuarial Society "Securitization of Risk" Discussion Paper Program, 1999.
- DEGROOT, M. H. **Probability and Statistics**. 2. ed. Addison-Wesley Publishing Company, 1987.
- DHAENE, J.; VYNCKE, D. The individual risk theory. **Working Paper**, 2002. Disponível em: www.econ.kuleuven.ac.be.
- FRAGA, E. Avaliação do risco de subscrição de prêmio utilizando inferência Bayesiana. **Revista Brasileira de Risco e Seguro**, v.1, n.1, p. 64-83, 2005. Disponível em: www.rbrs.com.br
- HOGG, R. V.; KLUGMAN, S. A.. **Loss Distributions**. Wiley, New York, 1984
- HOSKING, J. R. M.; WALLIS, J. R. Parameter and Quantile estimation for the Generalized Pareto Distribution. **Technometrics**, n. 29, p.339-349, 1987.
- MCNEIL, A. J. Estimating the Tails of Loss Severity Distribution using Extreme Value Theory. **ASTIN Bulletin**, n. 27, p. 117-137, 1997.
- MENDES, B. V. M **Uma introdução à análise de eventos extremos**. E-papers Serviços Editoriais, 2004.
- PICKANDS, J. III Statistical Inference using extreme order Statistics. **Annals of Statistics**, n. 3, p. 119-131, 1975.
- REGRESSION and Smoothing for Continuous Response Data. In: **S-PLUS Guide to Statistical and Mathematical Analysis**, 2003.
- REISS, R. D.; THOMAS, M. **Statistical Analysis of Extreme Values**, Birkhäuser, Springer-Verlag, 1997.
- SMITH, R. L. Measuring Risk with Extreme Value Theory. **Working paper**, Depart. Of Statistics, University of North Carolina, Chapel Hill, 1999.

7. Anexo

Gráfico 6 – Gráfico da densidade do montante de sinistros agregados estimados para um mês utilizando a distribuição Gamma para a severidade, considerando limite de retenção igual a R\$ 150.000,00

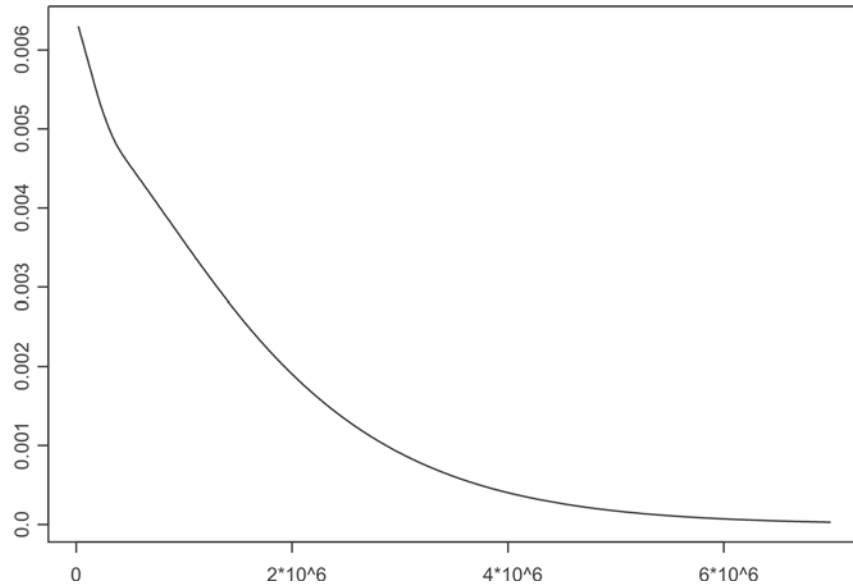


Gráfico 7 – Gráfico da densidade do montante de sinistros agregados estimados para um mês utilizando a distribuição Log-Normal para a severidade, considerando limite de retenção igual a R\$ 150.000,00

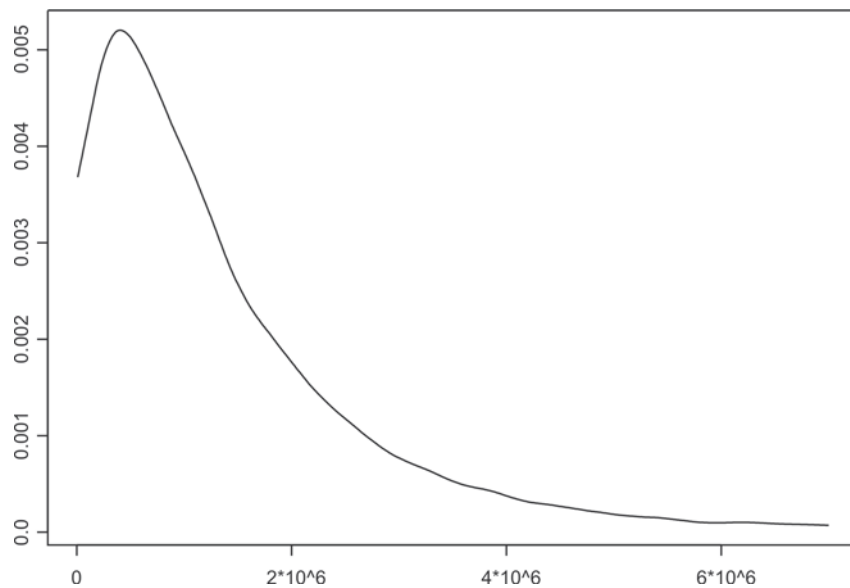


Gráfico 8 – Gráfico da densidade do montante de sinistros agregados estimados para um mês utilizando a distribuição GPD para a severidade, considerando limite de retenção igual a R\$ 150.000,00

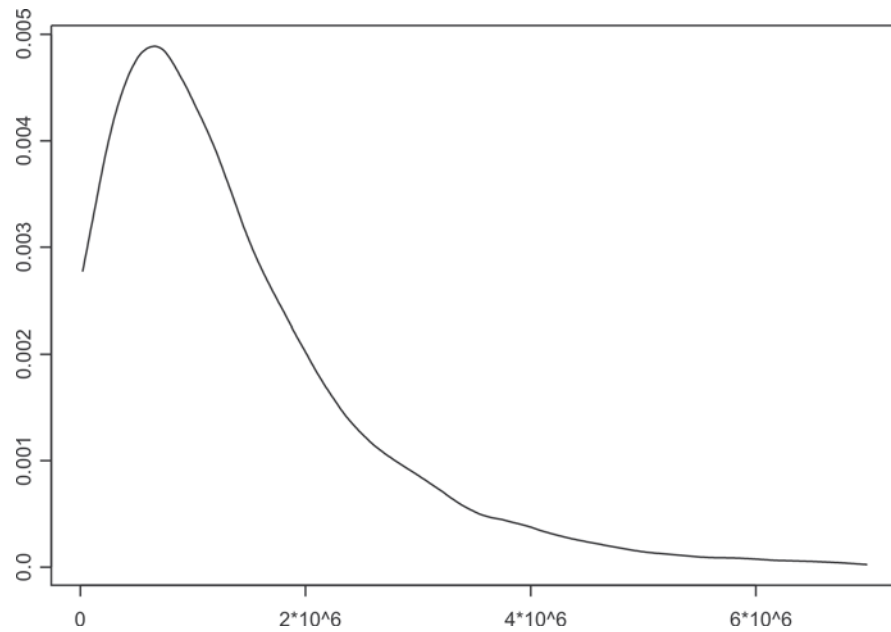


Gráfico 9 – Gráfico da densidade do montante de sinistros agregados estimados para um mês utilizando a distribuição MGPD para a severidade, considerando limite de retenção igual a R\$ 150.000,00

