

**A influência de valores moderados  
de perdas individuais nos valores  
extremos da perda agregada com  
aplicação em risco operacional.**

Plinio Lucas Dias Andrade

DISSERTAÇÃO APRESENTADA  
AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA  
UNIVERSIDADE DE SÃO PAULO  
PARA  
OBTENÇÃO DO TÍTULO  
DE  
MESTRE EM CIÊNCIAS

Área de Concentração: Estatística  
Orientador: Prof. Dr. Vladimir Belitsky

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES e CNPq

São Paulo, fevereiro de 2009

*”A percepção do desconhecido é a mais fascinante das experiências. O homem que não tem os olhos abertos para o misterioso passará pela vida sem ver nada.”*

(Albert Einstein)



# Dedicatória

*À Deus,  
aos meus pais Manoel e Inês,  
aos meus irmãos Rodrigo e Carolina.*



## Agradecimentos

Agradeço a todos, que de forma direta ou indireta, tornaram possível a conclusão deste trabalho. Em especial:

Primeiramente aos meus pais e irmãos pelo apoio incondicional.

Aos Professores Dr. Luiz Carlos Baida e Dr. Fernando Ferrari por me proporcionarem um primeiro contato com a Teoria de Probabilidades e Estatística.

À Profa. Dra. Laura Leticia Ramos Rifo por me fazer acreditar que podemos transcender as barreiras que surgem em nossos caminhos.

Aos grandes amigos Andressa, Evandro, Hugo, James e Paula por muitas vezes acreditarem mais em mim do que eu mesmo.

Ao meu orientador Prof. Dr. Vladimir Belitsky pelo voto de confiança, paciência, disponibilidade e por me conduzir brilhantemente na execução e conclusão deste trabalho.

Finalmente, à CAPES e ao CNPq pelo apoio financeiro.



## Resumo

O problema ao qual nos dedicamos no presente trabalho surgiu na aplicação do modelo de perda agregada para estimação em risco operacional. Na área de risco operacional, nosso interesse é a estimação de altos quantís da função de distribuição da perda agregada. Estes quantís dependem fortemente do comportamento da cauda direita da distribuição de perdas individuais. Os valores moderados de perdas individuais podem ou não influenciar os quantís de interesse.

Estudamos esta influência e propomos estimativas apropriadas, para posteriormente verificarmos nossos resultados teóricos em um conjunto de dados. Esta verificação demanda de uma análise detalhada de métodos matemáticos envolvidos no cálculo da perda acumulada no modelo de perda agregada.

A análise desenvolvida por nós fornece resultados que podem ser de interesse para o desenvolvimento deste modelo no tratamento de risco operacional.

**Palavras-chave:** Distribuição de perda agregada, método de Monte Carlo, método de Panjer, modelo de Cramér-Lundberg, peaks over threshold, teoria de valores extremos, valor em risco.





## Abstract

The problem addressed in the present work arises in application of the aggregate loss model for estimating operational risk. In the operational risk area, the estimation is focused on high quantiles of the distribution function of the aggregate loss. These quantiles depend strongly on the behavior of the right tail of the distribution of the individual loss. The moderate values of individual losses may or may not influence the quantiles in interest.

We study this influence and provide appropriate estimates. We then verify our theoretical results on a set of data. The verification required a detailed analysis of mathematical methods involved in the calculation of the aggregate loss in the aggregate loss model.

The analysis carried out by us here provide results that may be of independent interest for the employment of this model in treatment of operational risk.

**Keywords:** Aggregate loss distribution, Monte Carlo method, Panjer recursive formula, Cramér-Lundberg model, peaks over threshold, extreme value theory, value at risk.



# Sumário

<b>Lista de Abreviaturas</b>	<b>xiii</b>
<b>Lista de Figuras</b>	<b>xv</b>
<b>Lista de Tabelas</b>	<b>xvii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivos . . . . .	4
1.2 Contribuições . . . . .	4
1.3 Organização do Trabalho . . . . .	5
<b>2 Conceitos</b>	<b>7</b>
2.1 O Modelo de Cramér-Lundberg . . . . .	7
2.2 Resultados Assintóticos . . . . .	11
2.3 Método de Monte Carlo . . . . .	21
2.4 Método Recursivo de Panjer . . . . .	22
2.5 Valor em Risco . . . . .	25
2.6 Modelagem de Frequência . . . . .	26
2.7 Modelagem de Severidade . . . . .	31
2.7.1 O método Peaks Over Threshold - POT . . . . .	33
<b>3 Aplicação e Conclusões</b>	<b>43</b>
3.1 Apresentação dos dados . . . . .	43
3.2 Planejamento e Execução . . . . .	45
3.2.1 Resultados . . . . .	47
3.3 Considerações Finais . . . . .	65
3.4 Sugestões para Pesquisas Futuras . . . . .	66
<b>A Resultados Analíticos para a Perda Agregada</b>	<b>69</b>

**Referências Bibliográficas**

**73**

## Lista de Abreviaturas

f.d.a	Função de distribuição acumulada.
i.i.d	Independente e idênticamente distribuídas.
MMC	Método de Monte Carlo.
MRP	Método recursivo de Panjer.
POT	Peaks over threshold.
v.a	Variável aleatória.



## Lista de Figuras

2.1	Gráfico da distribuição exponencial com o gráficos da função escada sobreposto (oriundos da discretização da distribuição exponencial) para diferentes valores de $h$ . . . . .	24
2.2	Histograma de dados Poisson para diferentes valores de $\lambda$ . . . . .	28
2.3	Histograma de dados Binomial Negativo para diferentes valores de $r$ e $\beta$ . . . . .	29
2.4	Gráfico da razão $kn_k/n_{k-1}$ versus $k$ . . . . .	31
2.5	Distribuições de Fréchet ( $\alpha = 1$ ), Weibull ( $\alpha = -1$ ) e Gumbel . . . . .	35
3.1	Histograma e Boxplot do conjunto de dados de perdas operacionais. . . . .	44
3.2	Discretização da severidade para os diferentes valores do passo $h$ considerando a distribuição empírica como geradora das perdas individuais. . . . .	48
3.3	Distribuição de perdas agregadas obtidas pelo Método Recursivo de Panjer para os diferentes valores de $h$ considerando a distribuição empírica como geradora das perdas individuais. . . . .	49
3.4	Distribuição de perdas agregadas obtidas pelo Método de Monte Carlo considerando a distribuição empírica como geradora das perdas individuais. . . . .	50
3.5	Gráfico da média amostral dos excessos para o conjunto de dados. A linha vertical indica o ponto onde a abscissa assume valor 1000. . . . .	51
3.6	Gráficos de diagnósticos sobre a qualidade do ajuste GPD para os excessos acima do limiar 1000 obtidos via máxima verossimilhança. . . . .	52
3.7	Gráfico da função de distribuição empírica com curva GPD sobreposta obtida pelo Método POT. . . . .	53
3.8	Gráfico da função de distribuição empírica com gráfico da função $\hat{F}_X(\cdot)$ dada por (3.1) sobreposto. . . . .	54
3.9	Discretização da severidade para os diferentes valores do passo $h$ para a função $\hat{F}_X(\cdot)$ dada por (3.1). . . . .	55
3.10	Distribuição de perdas agregadas obtidas pelo Método Recursivo de Panjer para os diferentes valores de $h$ considerando que os dados foram gerados por $\hat{F}_X(\cdot)$ dada por (3.1). . . . .	56



3.11	Distribuição de perdas agregadas obtidas pelo Método de Monte Carlo considerando que os dados foram gerados por $\hat{F}_X(\cdot)$ dada por (3.1). . . . .	58
3.12	Distribuição de perdas agregadas obtidas pelo Método de Panjer para os valores abaixo do limiar $u = 1000$ com curva normal sobreposta. . . . .	59
3.13	Distribuição de perdas agregadas obtidas pelo Método de Monte Carlo para os valores abaixo do limiar $u = 1000$ com curva normal e curva da função de distribuição obtida via MRP sobrepostos. . . . .	60
3.14	Histograma da amostra gerada pelo Método de Monte Carlo com curva normal aproximada sobreposta. . . . .	61
3.15	Ajuste GPD feito para os excessos acima do limiar $u = 1000$ . A linha contínua representa a curva paramétrica da função $G_{\hat{\xi}, \hat{\sigma}, 0}(\cdot)$ e os pontos a função de distribuição empírica obtida com base na amostra de excessos. . . . .	62
3.16	Discretização feita para a GPD $G_{\hat{\xi}, \hat{\sigma}, 0}(\cdot)$ que modela os excessos. . . . .	63
3.17	Distribuição de perdas agregadas obtidas para valores acima do limiar $u = 1000$ . . .	63
A.1	Distribuição de perda agregada teórica do exemplo 1. . . . .	70

## Lista de Tabelas

2.1	Valores de $a$ e $b$ dos modelos Poisson e Binomial Negativo. . . . .	23
2.2	Algumas quantidade de interesse para recursões. . . . .	25
2.3	Perfis de acidentes. . . . .	30
3.1	Principais estatísticas descritivas do conjunto de dados de perdas operacionais. A unidade monetária é o real (R\$). . . . .	43
3.2	Valores em Risco obtidos através do Método Recursivo de Panjer para diferentes valores de $h$ usados na discretização do modelo para a severidade considerando a distribuição empírica como geradora das perdas individuais. . . . .	47
3.3	Valores em Risco obtidos através do Método de Monte Carlo considerando a distribuição empírica como geradora das perdas individuais. . . . .	50
3.4	Resultados obtidos pela aplicação do método POT ao conjunto de dados. As estimativas para os parâmetros foram obtidas via método da máxima verosimilhança e os valores entre parêntesis representam o erro padrão destes estimadores. . . . .	53
3.5	Valores em Risco obtidos através do Método Recursivo de Panjer para diferentes valores de $h$ usados na discretização do modelo para a severidade e considerando que os dados foram gerados por $\hat{F}_X(\cdot)$ dada por (3.1). . . . .	57
3.6	Valores em Risco obtidos através do Método de Monte Carlo considerando que os dados foram gerados por $\hat{F}_X(\cdot)$ dada por (3.1). . . . .	57
3.7	Valores em Risco obtidos do modelo para $S = S^A + S^B$ considerando a distribuição empírica como modelo para as severidades abaixo do limiar $u = 1000$ . Neste caso obtidos via MRP. . . . .	64
3.8	Valores em Risco obtidos via Método de Panjer do modelo para $S = S^A + S^B$ considerando o modelo normal para $S^A$ . . . . .	64
3.9	Valores em Risco obtidos do modelo para $S = S^A + S^B$ considerando a distribuição empírica como modelo para as severidades abaixo do limiar $u = 1000$ . Neste caso obtidos via Método de Monte Carlo. . . . .	64

3.10	Valores em Risco obtidos via Método de Monte Carlo do modelo para $S = S^A + S^B$ considerando o modelo normal para $S^A$ . . . . .	64
3.11	Valores em Risco obtidos pelas abordagens em estudo. . . . .	65
3.12	Valores em Risco obtidos via Método de Panjer apenas para $S^D$ . . . . .	65

## Capítulo 1

### Introdução

Alguns acontecimentos recentes, ditos catastróficos, chamaram a atenção pela severidade que suas perdas causaram. Presenciamos recentemente no mercado financeiro eventos que abalaram, e continuam abalando, a economia mundial. Devido à baixa frequência e grande severidade destes eventos, seu controle e previsão tornam-se difíceis para qualquer agente que se proponha a prover proteção ou reter perdas muito severas e inesperadas. Cruz (2002) [6] ilustra vários acontecimentos que culminaram no desenvolvimento da atual teoria do risco e consolidação do chamado Risco Operacional.

Em termos de uma definição formal, o Risco Operacional (RO) foi definido pelo comitê Basileia como *"o risco de perdas resultantes de processos internos, de pessoal e de sistemas inadequados ou falhos, bem como de eventos extremos"*. Neste contexto, notamos a importância da avaliação, precificação e dimensionamento dos riscos aos quais as instituições financeiras (bancos e seguradoras) estão expostas. Atualmente, a Teoria de Valores Extremos (TVE) vem tornando-se uma ferramenta poderosa para inferir caudas de distribuições de probabilidades, reduzindo a dificuldade na avaliação do risco.

De acordo com a teoria do risco coletivo a soma de perdas individuais, conhecida como perda agregada, é dada por (para maiores detalhes sugerimos Klugman et al (2004) [12]):

$$S = \begin{cases} \sum_{i=1}^N X_i & , N = 1, 2, \dots \\ 0 & , N = 0 \end{cases}$$

onde  $N$  é o número aleatório de perdas e  $X_i$  a severidade de cada perda  $i$ , também chamada de perda individual.

Em geral assumimos que  $N$  é independente de cada um dos  $X_i$ 's e tem distribuição Poisson ou

Binomial Negativa. Além disso,  $X_1, X_2, \dots$ , são variáveis aleatórias (v.a.s) contínuas independente e identicamente distribuídas (i.i.d).

O problema teórico e prático que iremos abordar neste trabalho surgiu da necessidade de se obter estimativas para altos quantis da função de distribuição da variável aleatória  $S$  acima. Antes de estarmos aptos a estimar estes quantis, devemos obter a função de distribuição de  $S$ , o que em geral não é uma tarefa simples devido ao grande número de convoluções que deve ser calculado. Iremos propor algumas abordagens que tornam a obtenção da função de distribuição de  $S$  factível computacionalmente e que preserva as propriedades intrínsecas apresentadas pelos dados. O objetivo destas técnicas é o de prover uma reserva de capital capaz de reter possíveis perdas extremas e assim o emprego da TVE, mais precisamente do método Peaks Over Threshold (POT), torna-se necessário. As dificuldades que surgem no emprego do método POT em conjunto com as dificuldades da obtenção da distribuição de perda agregada serão abordadas no presente trabalho.

Iremos obter uma estimativa para a função de distribuição de perda agregada através de um método conhecido como Método Recursivo de Panjer (MRP), que faz uso da propriedade que algumas variáveis aleatórias discretas possuem de gerar suas probabilidades recursivamente. Felizmente os modelos Poisson e Binomial Negativo satisfazem esta propriedade e uma vez que possuímos um modelo para frequência de eventos  $N$  e para as severidades  $X_i$ , podemos empregar este método para se obter a distribuição de  $S$ .

O emprego do método POT consiste basicamente em se obter um modelo paramétrico para a cauda direita da distribuição que gerou as perdas individuais. Logo, é de nosso interesse estudar o impacto de valores moderados, isto é, valores abaixo do comportamento de cauda da distribuição que gerou as perdas individuais na cauda direita da distribuição de perda agregada e, conseqüentemente, nos altos quantis da distribuição de  $S$ .

Como uma alternativa para o problema acima, poderíamos encontrar um modelo para os dados abaixo de um limiar determinado, onde acreditamos ser o início do comportamento de cauda direita da distribuição que gerou as perdas individuais, bem como um modelo para a frequência destes dados e usar o método POT para modelar os dados acima deste limiar e encontrar um modelo para a frequência destes dados restantes. Usaríamos o MRP em duas etapas e assim, estaríamos aptos a calcular a convolução das duas distribuições resultantes obtendo a distribuição de  $S$ .

Formalmente teríamos:

$$S = S^A + S^D$$

onde  $S^A = \sum_{i=1}^{N^A} X_i$  é a v.a que representa a perda agregada considerando que cada um dos  $X_i$ 's são *menores* que o limiar determinado e  $N^A$  é a variável que representa a frequência de ocorrências *abaixo* deste limiar. Analogamente,  $S^D = \sum_{i=1}^{N^D} X_i$  é a v.a que representa a perda agregada considerando que cada um dos  $X_i$ 's são *maiores* que o limiar determinado e  $N^D$  a frequência de ocorrência *acima* deste limiar.

Assim, poderíamos obter uma aproximação para  $F_S(\cdot)$ , a função de distribuição de  $S$ , através da seguinte convolução:

$$F_S(\cdot) = (F_{S^A} * F_{S^D})(\cdot)$$

Obviamente este tipo de abordagem torna o problema dispendioso em termos computacionais.

Poderíamos simplesmente obter um modelo paramétrico ou não-paramétrico para toda a distribuição que gerou as perdas, tornando desnecessário a análise do presente trabalho. No entanto, o método POT vem se consagrando pelo fato de inferir com precisão a cauda de distribuições de probabilidades, proporcionando mais segurança ao se estimar quantis extremos.

Como o método POT requer que seja escolhido um limiar suficientemente alto, para que as propriedades de seu resultado sejam satisfeitas, temos que um grande número de observações fica abaixo deste limiar, sugerindo que a aproximação normal pode ser uma alternativa para a distribuição de  $S^A$  tornando desnecessário o emprego do MRP para sua obtenção. Assim, teríamos que  $S^A \sim N(\mu, \sigma^2)$  e uma pergunta natural seria: Qual o erro cometido ao se assumir esta normalidade? Uma possível resposta será dada pelo resultado 2 a ser apresentado no capítulo 2.

Os resultados que serão apresentados no presente trabalho serão aplicados em um conjunto de dados referentes às perdas operacionais no período de um ano de um banco brasileiro. Estes mesmos dados foram utilizados em Belitsky e Moreira (2007) [1]. Também iremos comparar as respostas obtidas através do MRP com um método de simulação, conhecido como Método de Monte Carlo (MMC).

## 1.1 Objetivos

Em suma, os principais objetivos deste trabalho, como dito anteriormente, são:

1. Apresentar o modelo de Cramér-Lundberg e discorrer sobre o erro cometido ao se assumir a normalidade da distribuição de perda agregada;
2. Estudar o Método Recursivo de Panjer (MRP) e apontar as principais dificuldades de seu emprego;
3. Apresentar a Teoria de Valores Extremos, o método Peaks Over Threshold (POT) e o conceito de Valor em Risco (VaR);
4. Aplicar a teoria apresentada em um conjunto de dados de perdas operacionais de um banco brasileiro através de diferentes abordagens.

## 1.2 Contribuições

No final do século passado, várias técnicas estatísticas tendo como base a Teoria de Valores Extremos (TVE) passaram a ser aceitas na prática, especialmente nas áreas de finanças e atuária. A popularização destas técnicas ocorreu devido à qualidade desta abordagem, capaz de quantificar com maior precisão as probabilidades de ocorrência de eventos raros ou que ainda não ocorreram. Além disso, o emprego de técnicas alternativas como o MRP para obtenção de aproximações adequadas para a distribuição de perda agregada é pouco difundida no país, assim como sua aplicação em conjunto com a TVE, sendo esta uma das contribuições deste trabalho que visa apresentar, através de técnicas modernas, uma forma alternativa de modelagem em Risco Operacional.

Além disto, podemos encontrar alguns resultados acerca da ordem de convergência da distribuição de perda agregada para a distribuição normal, sendo uma referência essencial o trabalho de Landers e Rogge (1976) [14]. No entanto, um resultado do tipo Berry-Esséen para a soma de um número aleatório de variáveis aleatórias ainda é uma incógnita e a aproximação que obtemos aqui pode contribuir para futuras pesquisas e refinamento do resultado.

### 1.3 Organização do Trabalho

No capítulo 2 apresentamos o modelo de Cramér-Lundberg, alguns resultados assintóticos da distribuição de perda agregada, pressupostos teóricos para a obtenção de modelos de frequência e severidade para posterior obtenção da distribuição de perda agregada através dos métodos de Monte Carlo e Panjer, além do conceito de valor em risco. No capítulo 3 apresentaremos um conjunto de dados ao qual serão aplicados vários conceitos estudados no capítulo 2 e de onde iremos tirar algumas conclusões a respeito das abordagens propostas.





## Capítulo 2

### Conceitos

#### 2.1 O Modelo de Cramér-Lundberg

Nesta seção queremos apresentar um modelo para a distribuição de perda agregada, que será definido posteriormente e que, por equanto, pode ser entendido como a soma de todas as perdas que ocorreram em um determinado intervalo de tempo fixado. Estas perdas ocorridas são chamadas perdas individuais ou severidades, e ocorrem em determinados períodos de tempo (meses, semestres, anos) e suas ocorrências são observadas de forma irregular, isto é, em intervalos de tempo aleatórios. Logo, é apropriado incorporar um processo de "chegadas" ao modelo de perda agregada e modelar estas perdas através de um processo caracterizado por uma frequência aleatória de eventos e uma magnitude aleatória de seu efeito.

Tal processo será obtido com o auxílio de um modelo desenvolvido por Filip Lundberg em 1903 e aprimorado, através de conceitos de processos estocásticos, por Harald Cramér. Este modelo é conhecido por *Modelo de Cramér-Lundberg* e possui a seguinte estrutura.

**Definição** (Modelo de Cramér-Lundberg): O modelo de Cramér-Lundberg é dado pelas seguintes condições:

(a) Perdas individuais:

As perdas  $\{X_i, i \geq 1\}$  são v.a.s positivas e i.i.d tendo distribuição comum  $F_X(\cdot)$ .

(b) Tempo de ocorrência:

As perdas ocorrem nos instantes aleatórios  $0 < T_1 < T_2 < \dots$

(c) Número de ocorrências:

O número de perdas que ocorrem no intervalo  $[0, t]$  é denotado por:

$$N(t) = \sup\{n \geq 1 : T_n \leq t\}, \quad t \geq 0,$$

onde, por convenção,  $\sup\emptyset = 0$ .

(d) Tempo entre ocorrências:

$$Y_1 = T_1, \quad Y_k = T_k - T_{k-1}, \quad k = 2, 3, \dots,$$

são v.a.s i.i.d exponenciais com média finita  $E(Y_1) = 1/\lambda$

(e) As sequências  $\{X_i, i \geq 1\}$  e  $\{Y_i, i \geq 1\}$  são independentes.

Maiores detalhes sobre este modelo podem ser obtidos em Embrechts et al (1997) [7].

Nas condições do modelo de Cramér Lundberg definido acima, a perda agregada é definida como um processo de perdas totais  $(S(t))_{t \geq 0}$  e dada pela expressão:

$$S(t) = \begin{cases} \sum_{i=1}^{N(t)} X_i & , N(t) = 1, 2, \dots \\ 0 & , N(t) = 0 \end{cases} \quad (2.1)$$

Na prática, o valor de  $T$  será fixo e isto nos permite suprimir  $t$  na expressão (2.1). Logo, quando nos referirmos à perda agregada, estaremos interessados na seguinte variável aleatória<sup>1</sup>:

$$S = \begin{cases} \sum_{i=1}^N X_i & , N = 1, 2, \dots \\ 0 & , N = 0 \end{cases} \quad (2.2)$$

Como já foi mencionado, as perdas individuais  $X_i$ , também conhecidas como severidades, são supostas independente e idênticamente distribuídas e ocorrem de acordo com a variável de contagem  $N$  também aleatória e conhecida como variável de frequência que, condicional a  $N = n$ , será suposta independente de cada um dos  $X_i$ 's.

Um dos objetivos de se obter a distribuição de perda agregada é o cálculo do *Valor em Risco* que será definido em (2.5), mas trata-se de um quantil da função distribuição de  $S$  do modelo (2.2). Por exemplo, poderíamos estar interessados em estimar a soma de todas as perdas que um banco pode sofrer no decorrer de, digamos, um ano com uma dada probabilidade.

Iremos agora apresentar alguns resultados analíticos da variável aleatória  $S$ .

Seja  $F_X(x) = P(X \leq x)$  a função de distribuição acumulada (f.d.a) comum para as  $X_i$ 's e  $F_S(\cdot)$  a f.d.a de  $S$ , logo

$$\begin{aligned} F_S(x) &= P(S \leq x) = \sum_{n=0}^{\infty} P(S \leq x, N = n) \\ &= \sum_{n=0}^{\infty} P(N = n)P(S \leq x | N = n) \\ &= \sum_{n=0}^{\infty} P(N = n)F_X^{*n}(x) \end{aligned}$$

---

<sup>1</sup>Posteriormente, iremos considerar que  $N$ , terá distribuição Poisson de parâmetro  $n$ . Este valor de  $n$  irá influenciar em resultados a serem apresentados na subseção (2.2)

Condicional a  $N = n$ ,  $F_X^{*n}(\cdot)$  é a  $n$ -ésima convolução da f.d.a da variável aleatória  $X$ , isto é,  $F^{*n}(x) = P(X_1 + \dots + X_n \leq x)$ . Além disso, temos que

$$F_X^{*0}(x) = \begin{cases} 0 & , x < 0 \\ 1 & , x \geq 0 \end{cases}$$

e

$$F_X^{*k}(x) = \int_{-\infty}^{\infty} F_X^{*(k-1)}(x-y) dF_X(y), \text{ para } k = 1, 2, \dots$$

Se  $X$  for uma v.a contínua e não-negativa então a densidade da  $k$ -ésima convolução será dada por:

$$f_X^{*k}(x) = \int_0^x f_X^{*(k-1)}(x-y) f_X(y) dy, \text{ para } k = 1, 2, \dots$$

E com isso, obtemos a função densidade de probabilidade da distribuição de  $S$

$$f_S(x) = \sum_{n=0}^{\infty} P(N = n) f_X^{*n}(x) \quad (2.3)$$

A esperança e a variância de  $S$  são dadas em termos da esperança e variância de  $X$  e  $N$ , a saber

$$E(S) = E(N)E(X) \quad (2.4)$$

$$Var(S) = E(N)Var(X) + Var(N)E^2(X) \quad (2.5)$$

Não existe, exceto para poucos casos simples (apresentados no apêndice (A) deste trabalho), uma expressão analítica para a função distribuição de  $S$ , o que nos leva a recorrer a métodos alternativos para a obtenção de uma aproximação para tal distribuição.

Neste momento, uma pergunta natural que poderia surgir ao se estudar o modelo (2.2) seria a viabilidade de se aproximar a distribuição de  $S$  pela distribuição normal, já que trata-se de uma soma de v.a.s apesar de que, neste caso, esta soma ser de um número aleatório de perdas individuais.

Klugman et al (2004) [12] afirmam que a distribuição normal é uma boa aproximação para a distribuição de perdas agregadas quando  $E(N)$  é grande. A próxima subseção irá fornecer alguns resultados ao se assumir a normalidade da variável aleatória  $S$ .

## 2.2 Resultados Assintóticos

Nesta etapa do trabalho, queremos encontrar uma estimativa para o erro cometido quando assumimos a normalidade da distribuição de perdas agregadas. O resultado obtido será aplicável para o caso particular onde  $N \sim Poisson(n)$ .

Antes disso, iremos recordar um resultado famoso da Teoria de Probabilidades conhecido como *Teorema de Berry-Esseen* que nos fornece um limite superior do desvio da distribuição normal no Teorema Central do Limite. Este resultado será utilizado posteriormente na formulação do *Resultado 2* que é o objetivo desta subseção.

**Teorema** (Berry-Esseen): Seja  $\{X_n, n \geq 1\}$  uma seqüência de variáveis aleatórias independentes tais que  $E(X_n) = 0$ ,  $E(X_n^2) = \sigma_n^2$ ,  $s_n^2 = \sum_{i=1}^n \sigma_i^2 > 0$  e  $\Gamma_n^{2+\delta} := \sum_{i=1}^n E|X_i|^{2+\delta} < \infty$  para algum  $\delta \in (0, 1]$ . E seja  $S_n = \sum_{i=1}^n X_i$ ,  $n \geq 1$ . Então existe uma constante universal  $c_\delta$  tal que:

$$\sup_{x \in \mathbb{R}} \left| P \left( \frac{S_n}{s_n} \leq x \right) - \Phi(x) \right| \leq c_\delta \left( \frac{\Gamma_n}{s_n} \right)^{2+\delta} \quad (2.6)$$

onde  $\Phi(\cdot)$  denota a f.d.a da distribuição normal padrão.

No caso de variáveis aleatórias independentes e idênticamente distribuídas, o teorema se reduz em:

**Teorema** (Berry-Esseen para v.a's i.i.d's): Seja  $\{X_n, n \geq 1\}$  uma seqüência de variáveis aleatórias independentes e idênticamente distribuídas tais que  $E(X_n) = 0$ ,  $E(X_n^2) = \sigma^2 > 0$ ,  $\gamma^{2+\delta} := E|X_n|^{2+\delta} < \infty$  para algum  $\delta \in (0, 1]$ . Então existe uma constante universal  $c_\delta$  tal que:

$$\sup_{x \in \mathbb{R}} \left| P \left( \frac{S_n}{s_n} \leq x \right) - \Phi(x) \right| \leq \frac{c_\delta}{n^{\frac{\delta}{2}}} \left( \frac{\gamma}{\sigma} \right)^{2+\delta} \quad (2.7)$$

Van Beek (1972) [23] mostrou que  $c_1 \leq 0.7975$  enquanto outros valores para  $c_\delta$  podem ser encontrados em Tysiak (1983) [22]. Em nosso trabalho, podemos considerar  $\delta = 1$ , pois queremos aproximar pela distribuição normal os dados abaixo de um dado limiar. Desta forma, o terceiro momento absoluto sempre será finito já que os dados são truncados abaixo deste limiar.

Iremos agora enunciar um resultado retirado da observação 5 de Landers e Rogge (1976) [14] ao qual iremos nos referir como *Resultado 1*. Este resultado, juntamente com o Teorema de Berry-Esseen, também irá nos auxiliar na posterior obtenção do *Resultado 2* a ser apresentado a seguir.

**Resultado 1:**(Ordem de aproximação no Teorema Central do Limite para Somas Aleatórias). Seja  $\{X_i, i \geq 1\}$  uma seqüência de variáveis aleatórias independentes e idênticamente distribuídas com variância  $\sigma^2 > 0$ . Seja, para cada  $n \in \mathbb{N}$ ,  $N_n$  uma variável independente da seqüência  $\{X_i, i \geq 1\}$ . Seja  $\{\varepsilon_n, n \geq 1\}$  uma seqüência de números reais tais que  $\frac{1}{\sqrt{n}} \leq \varepsilon_n \rightarrow 0$ , quando  $n \rightarrow \infty$  e uma constante  $\tau > 0$ . Se

$$P \left( \left| \frac{N_n}{n\tau} - 1 \right| > \varepsilon_n \right) = O(\varepsilon_n)$$

então

$$d_n = \sup_{t \in \mathbb{R}} |P(S_n \leq t) - \Phi(t)| = O(\varepsilon_n), \quad (2.8)$$

onde

$$S_n = \frac{\sum_{i=1}^{N_n} (X_i - E(X_i))}{\sigma \sqrt{n\tau}}. \quad (2.9)$$

Note que este resultado apenas nos fornece a ordem de convergência da distribuição de  $S_n$  para

a distribuição normal padrão. No que segue, iremos obter um resultado do tipo Berry-Esseen a partir do resultado 1. Para isto, iremos mostrar que a hipótese do resultado 1 será satisfeita quando  $N_n \sim Poisson(n)$  para algum  $\varepsilon_n$  escolhido de tal forma que  $\frac{1}{\sqrt{n}} \leq \varepsilon_n \rightarrow 0$  quando  $n \rightarrow \infty$ .

Nestas condições, devemos mostrar que:

$$P\left(\left|\frac{N_n}{n\tau} - 1\right| > \varepsilon_n\right) \leq d\varepsilon_n \text{ para algum } d > 0. \quad (2.10)$$

Antes disso, uma observação deve ser feita acerca da grandeza  $\tau$  presente no resultado 1.

**Observação:** Queremos assumir que a variável aleatória  $S$  da expressão (2.2) tem distribuição normal. Logo, se  $S$  é a variável aleatória que representa a perda agregada temos que

$$S = \sum_{i=1}^{N_n} X_i$$

onde  $\{X_i, i \geq 1\}$  são as v.a's i.i.d's que representam as perdas individuais (severidades), enquanto  $N_n$  é a v.a. contadora onde supomos que  $N_n \sim Poisson(n)$ .

Na prática observa-se que, quando  $n$  é grande, temos que

$$S \sim N(E(S), Var(S))$$

Logo, usaremos a distribuição normal com média igual a  $E(S)$  e variância igual a  $Var(S)$  para aproximar a verdadeira distribuição de  $S$  considerando apenas os valores de perdas individuais. Vale lembrar que na situação descrita acima e no modelo (2.2), temos que as expressões (2.4) e (2.5) nos fornecem que:

$$E(S) = E(N_n)E(X) = nE(X_i)$$

e

$$Var(S) = E(N_n)Var(X_i) + Var(N_n)E^2(X_i) = n(Var(X_i) + E^2(X_i))$$



Assim, se  $\{X_i, i \geq 1\}$  é uma seqüência de v.a.'s independentes com  $E(X_i) = \mu$  e  $Var(X_i) = \sigma^2 < \infty$  e, para cada  $n \in \mathbb{N}$ ,  $N_n \sim Poisson(n)$  independente das  $X_i$ 's  $i \geq 1$ , temos que:

$$S_n = \frac{\sum_{i=1}^{N_n} X_i - n\mu}{\sqrt{n}(\mu^2 + \sigma^2)^{1/2}} = \frac{\sum_{i=1}^{N_n} X_i - n\mu}{\sigma\sqrt{n} \left(1 + \left(\frac{\mu}{\sigma}\right)^2\right)^{1/2}} = \frac{\sum_{i=1}^{N_n} X_i - n\mu}{\sigma\sqrt{n}\tau}$$

Comparando a ultima expressão acima com a expressão (2.9) concluimos, neste caso, que:

$$\tau = \left(1 + \left(\frac{\mu}{\sigma}\right)^2\right)^{1/2} \quad (2.11)$$

Isto conclui a presente observação.

Iremos agora mostrar que a condição (2.10) é satisfeita para o caso onde  $N_n \sim Poisson(n)$ .

Note que

$$\begin{aligned} P\left(\left|\frac{N_n}{n\tau} - 1\right| > \varepsilon_n\right) &= P(N_n > n\tau + n\tau\varepsilon_n) + P(N_n < n\tau - n\tau\varepsilon_n) \\ &\approx P(N_n \geq n\tau + n\tau\varepsilon_n) + P(N_n \leq n\tau - n\tau\varepsilon_n) \end{aligned}$$

já que para  $n$  suficientemente grande,  $P(N_n = n\tau + n\tau\varepsilon_n) \approx 0$

Logo, estamos interessados no comportamento de cauda direita e esquerda da distribuição de  $N_n$ . Pela observação feita anteriormente podemos considerar, sem perda de generalidade, que  $\tau = 1$ , já que se tivermos  $\tau > 1$  o resultado também será satisfeito.

Para mostrar a validade da condição (2.10) iremos utilizar dois fatos, a saber:

**Fato 1:**(Estimativas de caudas direita e esquerda para a distribuição Poisson) Observe que se  $X \sim Poisson(\lambda)$ , temos que

$$P(X \geq k) = \sum_{n=k}^{\infty} P(X = n) = \sum_{n=k}^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda} \left[ \frac{\lambda^k}{k!} + \frac{\lambda^{k+1}}{(k+1)!} + \dots \right]$$

$$\begin{aligned}
&\leq e^{-\lambda} \left[ \frac{\lambda^k}{k!} + \frac{\lambda^{k+1}}{k k!} + \frac{\lambda^{k+2}}{k^2 k!} \dots \right] = \frac{e^{-\lambda} \lambda^k}{k!} \left[ 1 + \frac{\lambda}{k} + \left( \frac{\lambda}{k} \right)^2 + \dots \right] \\
&= \frac{e^{-\lambda} \lambda^k}{k!} \frac{1}{1 - \lambda/k} = \frac{e^{-\lambda} \lambda^k}{k!} \frac{k}{k - \lambda}
\end{aligned}$$

Se  $k$  é um ponto na cauda esquerda da distribuição de  $X$  tal que  $k < E(X)$ , vamos ter que

$$P(X \leq k) \leq kP(X = k), \quad k < E(X).$$

**Fato 2:** (Fórmula de Stirling) Para  $n \rightarrow \infty$ , temos que:

$$n! \downarrow \sqrt{2\pi n} n^{n+1/2} e^{-n}$$

Em particular  $n! \geq \sqrt{2\pi n} n^{n+1/2} e^{-n}$ .

Com o auxílio dos fatos 1 e 2 acima e sabendo que  $N_n \sim \text{Poisson}(n)$  vamos obter que

$$\begin{aligned}
P(N_n \geq n + n\varepsilon_n) &\leq \frac{e^{-n} n^{n+n\varepsilon_n}}{(n + n\varepsilon_n)!} \left( \frac{n + n\varepsilon_n}{n\varepsilon_n} \right) \\
&\leq \frac{e^{-n} n^{n+n\varepsilon_n}}{\sqrt{2\pi} (n + n\varepsilon_n)^{n+n\varepsilon_n+1/2} e^{-n-n\varepsilon_n}} (1 + \varepsilon_n^{-1}) \\
&= \frac{1}{\sqrt{2\pi}} (1 + \varepsilon_n^{-1}) \frac{e^{n\varepsilon_n}}{(n + n\varepsilon_n)^{1/2}} \left( \frac{n}{n + n\varepsilon_n} \right)^{n+n\varepsilon_n} \\
&= \frac{e^{n\varepsilon_n}}{\sqrt{2\pi}} \frac{(1 + \varepsilon_n^{-1})}{(n + n\varepsilon_n)^{1/2}} \left( 1 - \frac{1}{1 + \varepsilon_n^{-1}} \right)^{n+n\varepsilon_n} \\
&= \frac{e^{n\varepsilon_n}}{\sqrt{2\pi}} \frac{(1 + \varepsilon_n^{-1})}{(n + n\varepsilon_n)^{1/2}} \left( 1 - \frac{1}{1 + \varepsilon_n^{-1}} \right)^{\varepsilon_n^{-1}(n\varepsilon_n + n\varepsilon_n^2)}
\end{aligned}$$

e

$$P(N_n \leq n - n\varepsilon_n) \leq (n - n\varepsilon_n) \frac{e^{-n} n^{n-n\varepsilon_n}}{(n - n\varepsilon_n)!}$$

$$\begin{aligned}
&\leq (n - n\varepsilon_n) \frac{e^{-n} n^{n-n\varepsilon_n}}{\sqrt{2\pi} (n - n\varepsilon_n)^{n-n\varepsilon_n+1/2} e^{-n+n\varepsilon_n}} \\
&= \frac{e^{-n\varepsilon_n}}{\sqrt{2\pi}} \frac{(n - n\varepsilon_n)}{(n - n\varepsilon_n)^{1/2}} \left( \frac{n}{n - n\varepsilon_n} \right)^{n-n\varepsilon_n} \\
&= \frac{e^{-n\varepsilon_n}}{\sqrt{2\pi}} (n - n\varepsilon_n)^{1/2} \left( 1 + \frac{1}{\varepsilon_n^{-1} - 1} \right)^{\varepsilon_n^{-1}(n\varepsilon_n - n\varepsilon_n^2)}
\end{aligned}$$

Se tomarmos  $\varepsilon_n = n^{-\frac{1}{3}}$  teremos:

$$\begin{aligned}
P(N_n \geq n + n\varepsilon_n) &= P(N_n \geq n + n^{2/3}) \leq \frac{e^{n^{2/3}} (1 + n^{1/3})}{\sqrt{2\pi} (n + n^{2/3})^{1/2}} \left( 1 - \frac{1}{1 + n^{1/3}} \right)^{n^{1/3}(n^{2/3} + n^{1/3})} \\
&= \frac{1}{\sqrt{2\pi}} \frac{(1 + n^{1/3})}{(n^{2/3}(1 + n^{1/3}))^{1/2}} e^{n^{2/3}} \left( 1 - \frac{1}{1 + n^{1/3}} \right)^{n^{1/3}(n^{2/3} + n^{1/3})} \\
&= \frac{1}{\sqrt{2\pi}} n^{-1/3} (1 + n^{1/3})^{1/2} e^{n^{2/3}} \left( 1 - \frac{1}{1 + n^{1/3}} \right)^{n^{1/3}(n^{2/3} + n^{1/3})} := \frac{1}{\sqrt{2\pi}} p_n
\end{aligned}$$

Assim, para  $p_n$  definido acima, temos:

$$\begin{aligned}
\frac{p_n}{\varepsilon_n} &= \frac{n^{1/3}}{\sqrt{2\pi}} n^{-1/3} (1 + n^{1/3})^{1/2} e^{n^{2/3}} \left( 1 - \frac{1}{1 + n^{1/3}} \right)^{n^{1/3}(n^{2/3} + n^{1/3})} \\
&= \frac{(1 + n^{1/3})^{1/2}}{\sqrt{2\pi}} e^{n^{2/3}} \left( 1 - \frac{1}{1 + n^{1/3}} \right)^{n^{1/3}(n^{2/3} + n^{1/3})}
\end{aligned}$$

Note que, quando  $n \rightarrow \infty$ ,

$$\left( 1 - \frac{1}{1 + n^{1/3}} \right)^{n^{1/3}} \rightarrow e^{-1} \Rightarrow \left( 1 - \frac{1}{1 + n^{1/3}} \right)^{n^{1/3}(n^{2/3} + n^{1/3})} e^{n^{2/3}} \rightarrow 0$$

Portanto,  $p_n \leq \varepsilon_n$ .

Por outro lado,

$$P(N_n \leq n - n\varepsilon_n) = P(n - n^{2/3}) \leq \frac{e^{-n^{2/3}}}{\sqrt{2\pi}} (n - n^{2/3})^{1/2} \left(1 + \frac{1}{n^{1/3} - 1}\right)^{n^{1/3}(n^{2/3} - n^{1/3})} := \frac{1}{\sqrt{2\pi}} q_n$$

Assim, para  $q_n$  definido acima, temos:

$$\begin{aligned} \frac{q_n}{\sqrt{\varepsilon_n}} &= \frac{n^{1/3}}{\sqrt{2\pi}} (n - n^{2/3})^{1/2} \left(1 + \frac{1}{n^{1/3} - 1}\right)^{n^{1/3}(n^{2/3} - n^{1/3})} e^{-n^{2/3}} \\ &= \frac{(n^{5/3} - n^{4/3})^{1/2}}{\sqrt{2\pi}} \left(1 + \frac{1}{n^{1/3} - 1}\right)^{n^{1/3}(n^{2/3} - n^{1/3})} e^{-n^{2/3}} \end{aligned}$$

Note que, quando  $n \rightarrow \infty$ ,

$$\left(1 + \frac{1}{n^{1/3} - 1}\right)^{n^{1/3}} \rightarrow e \Rightarrow \left(1 + \frac{1}{n^{1/3} - 1}\right)^{n^{1/3}(n^{2/3} - n^{1/3})} e^{n^{-2/3}} \rightarrow 0$$

Portanto,  $q_n \leq \varepsilon_n$ .

Assim, podemos concluir que para  $\varepsilon_n = n^{-1/3}$ , a condição (2.10) é satisfeita e

$$P\left(\left|\frac{N_n}{n} - 1\right| > n^{-1/3}\right) \leq \frac{2}{\sqrt{2\pi}} n^{-1/3}$$

Logo, podemos aplicar o resultado 1 quando  $N_n \sim \text{Poisson}(n)$  e  $\varepsilon_n = n^{-1/3}$  e utilizar a expressão (2.8) para construir um resultado do tipo Berry-Esseen para este caso particular.

Com o intuito de diminuir o erro cometido, isto é, o valor de  $\varepsilon_n$ , fizemos os mesmo cálculos para  $\varepsilon_n = n^{-1}$  e  $\varepsilon_n = n^{-1/2}$ . No entanto, para estes valores, a condição (2.10) não será satisfeita.

Iremos agora encontrar um limite superior para o desvio da distribuição normal com base nas informações obtidas até o presente momento.

Vamos supor, sem perda de generalidade, que  $E(X_i) = 0$  e  $\text{Var}(X_i) = 1 \forall i \geq 1$ . Lembrando que  $\gamma^3 := E|X_i|^3 < \infty$  e, para cada  $n \in \mathbb{N}$ ,  $N_n \sim \text{Poisson}(n)$  independente de cada  $X_i$ ,  $i \geq 1$ .

Suponha que para um certo  $\varepsilon_n$  tal que  $\frac{1}{\sqrt{n}} \geq \varepsilon_n \rightarrow 0$  quando  $n \rightarrow \infty$ , a condição (2.10) seja satisfeita, isto é

$$P\left(\left|\frac{N_n}{n\tau} - 1\right| > \varepsilon_n\right) \leq \frac{2}{\sqrt{2\pi}}\varepsilon_n \quad (2.12)$$

Logo, teremos  $d_n$  escrito como segue

$$d_n = \sup_{t \in \mathbb{R}} \left| P\left(\frac{1}{\sqrt{n\tau}} \sum_{i=1}^{N_n} X_i \leq t\right) - \Phi(t) \right|$$

Seja  $I_n = \{k \in \mathbb{N} : n\tau - n\tau\varepsilon_n \leq k \leq n\tau + n\tau\varepsilon_n\}$

$$\begin{aligned} d_n &= \sup_{t \in \mathbb{R}} \left| P\left(\frac{1}{\sqrt{n\tau}} \sum_{i=1}^{N_n} X_i \leq t\right) - \Phi(t) \right| \\ &= \sup_{t \in \mathbb{R}} \left| \sum_{k=0}^{\infty} P(N_n = k) P\left(\frac{1}{\sqrt{n\tau}} \sum_{i=1}^k X_i \leq t\right) - \Phi(t) \right| \\ &\leq \sup_{t \in \mathbb{R}} \left| \sum_{k \in I_n} P(N_n = k) P\left(\frac{1}{\sqrt{n\tau}} \sum_{i=1}^k X_i \leq t\right) - \Phi(t) \right| + \frac{2}{\sqrt{2\pi}}\varepsilon_n \text{ (justificado por (2.12))} \\ &\leq \sum_{k \in I_n} P(N_n = k) \sup_{t \in \mathbb{R}} \left| P\left(\frac{1}{\sqrt{n\tau}} \sum_{i=1}^k X_i \leq t\right) - \Phi(t) \right| + \frac{2}{\sqrt{2\pi}}\varepsilon_n \\ &= \sum_{k \in I_n} P(N_n = k) \sup_{t \in \mathbb{R}} \left| P\left(\frac{1}{\sqrt{n\tau}} \sum_{i=1}^k X_i \leq t\right) - \Phi\left(t\sqrt{\frac{n\tau}{k}}\right) + \Phi\left(t\sqrt{\frac{n\tau}{k}}\right) - \Phi(t) \right| + \frac{2}{\sqrt{2\pi}}\varepsilon_n \\ &\leq \sum_{k \in I_n} P(N_n = k) \sup_{t \in \mathbb{R}} \left| P\left(\frac{1}{\sqrt{k}} \sum_{i=1}^k X_i \leq t\sqrt{\frac{n\tau}{k}}\right) - \Phi\left(t\sqrt{\frac{n\tau}{k}}\right) \right| \\ &+ \sum_{k \in I_n} P(N_n = k) \sup_{t \in \mathbb{R}} \left| \Phi\left(t\sqrt{\frac{n\tau}{k}}\right) - \Phi(t) \right| + \frac{2}{\sqrt{2\pi}}\varepsilon_n \\ &= a_n + b_n + \frac{2}{\sqrt{2\pi}}\varepsilon_n \end{aligned}$$

Aplicando o Teorema de Berry-Esseen, temos que

$$\sup_{t \in \mathbb{R}, k \in I_n} \left| P \left( \frac{1}{\sqrt{k}} \sum_{i=1}^k X_i \leq t \sqrt{\frac{n\tau}{k}} \right) - \Phi \left( t \sqrt{\frac{n\tau}{k}} \right) \right| \leq \frac{0.7975\gamma^3}{\sqrt{k}}$$

Logo,  $a_n \leq 0.7975\gamma^3/\sqrt{k}$  e como  $n\tau - n\tau\varepsilon_n \leq k \leq n\tau + n\tau\varepsilon_n$ , temos que

$$a_n \leq \frac{0.7975\gamma^3}{\sqrt{n\tau - n\tau\varepsilon_n}}$$

Para encontrarmos um limite superior para  $b_n$  iremos recorrer ao lema 5.2 pg. 153 de Petrov (1995) [16] que nos diz que se  $\Phi(\cdot)$  representa a função distribuição acumulada da variável aleatória normal-padrão, então:

$$\sup_x |\Phi(px) - \Phi(x)| \leq \frac{(p-1)}{\sqrt{2\pi e}}, \quad \text{se } p \geq 1;$$

$$\sup_x |\Phi(px) - \Phi(x)| \leq \frac{(p^{-1}-1)}{\sqrt{2\pi e}}, \quad \text{se } p < 1.$$

Logo, usando este resultado, temos que

$$\begin{aligned} \sup_{t \in \mathbb{R}, k \in I_n} \left| \Phi \left( t \sqrt{\frac{n\tau}{k}} \right) - \Phi(t) \right| &\leq \sup_{t \in \mathbb{R}, k \leq n\tau} \left| \Phi \left( t \sqrt{\frac{n\tau}{k}} \right) - \Phi(t) \right| \\ &+ \sup_{t \in \mathbb{R}, k > n\tau} \left| \Phi \left( t \sqrt{\frac{n\tau}{k}} \right) - \Phi(t) \right| \\ &\leq \frac{1}{\sqrt{2\pi e}} \left( \sqrt{\frac{n\tau}{k}} - 1 \right) + \frac{1}{\sqrt{2\pi e}} \left( \sqrt{\frac{k}{n\tau}} - 1 \right) \end{aligned}$$

Como  $n\tau - n\tau\varepsilon_n \leq k \leq n\tau + n\tau\varepsilon_n$  então, teremos através de algumas manipulações algébricas simples que:

$$\sqrt{\frac{n\tau}{k}} \leq \frac{1}{\sqrt{1 - \varepsilon_n}} = (1 - \varepsilon_n)^{-1/2}$$

e

$$\sqrt{\frac{k}{n\tau}} \leq (1 + \varepsilon_n)^{1/2}$$

Portanto,

$$b_n \leq \frac{1}{\sqrt{2\pi e}} [(1 - \varepsilon_n)^{-1/2} + (1 + \varepsilon_n)^{1/2} - 2]$$

Finalmente obtemos o resultado de nosso interesse,

$$d_n \leq a_n + b_n + \frac{1}{\sqrt{2\pi}} \varepsilon_n \leq \frac{0.7975\gamma^3}{\sqrt{n\tau - n\tau\varepsilon_n}} + \frac{1}{\sqrt{2\pi e}} [(1 - \varepsilon_n)^{-1/2} + (1 + \varepsilon_n)^{1/2} - 2] + \frac{2}{\sqrt{2\pi}} \varepsilon_n$$

Formulamos então o seguinte resultado, ao qual iremos nos referir como *Resultado 2* e que será aplicado em nosso conjunto de dados no terceiro capítulo deste trabalho.

**Resultado 2:** Seja  $\{X_i, i \geq 1\}$  uma seqüência de variáveis aleatórias i.i.d tal que  $E(X_i) = 0$  e  $Var(X_i) = 1, \forall i$ . Seja, para cada  $n \in \mathbb{N}$ ,  $N_n$  uma variável independente da seqüência  $\{X_i, i \geq 1\}$ . Seja  $\{\varepsilon_n, n \geq 1\}$  uma seqüência de números reais tais que  $\frac{1}{\sqrt{n}} \leq \varepsilon_n \rightarrow 0$ , quando  $n \rightarrow \infty$  e uma constante  $\tau > 0$ . Se

$$P\left(\left|\frac{N_n}{n\tau} - 1\right| > \varepsilon_n\right) \leq d\varepsilon_n$$

para algum  $d > 0$ , então

$$\sup_{t \in \mathbb{R}} |P(S_n \leq t) - \Phi(t)| \leq \frac{0.7975\gamma^3}{\sqrt{n\tau - n\tau\varepsilon_n}} + \frac{1}{\sqrt{2\pi e}} [(1 - \varepsilon_n)^{-1/2} + (1 + \varepsilon_n)^{1/2} - 2] + d\varepsilon_n \quad (2.13)$$

onde

$$S_n = \sum_{i=1}^{N_n} X_i$$

Quando aplicarmos este resultado ao nosso conjunto de dados, iremos tomar  $\varepsilon_n = n^{-1/3}$  e  $d = 2/\sqrt{2\pi}$ .

### 2.3 Método de Monte Carlo

O emprego do método de Monte Carlo (MMC) na estimação de quantis da distribuição de  $S$  é bastante comum dada a sua simplicidade e fácil implementação. Esta abordagem simula um número arbitrariamente grande de cenários de perdas agregadas com base nos modelos de frequência e severidade obtidos a priori. O algoritmo do MMC funciona da seguinte forma:

- 1:** Gere um número arbitrariamente grande de realizações da distribuição de frequências usada para modelar a distribuição de  $N$ , obtendo a seqüência  $n_1, n_2, \dots, n_k$  que representa os cenários para o número total de perdas no período determinado;
- 2:** Para cada cenário  $k$ , simule  $n_k$  valores da distribuição usada para modelar a severidade.

Este passo é executado como segue:

- 2.1:** Se  $n_k = i$ , gere  $i$  valores da distribuição uniforme em  $[0, 1]$ . Teremos então a seqüência  $p_1, \dots, p_i$  de valores em  $[0, 1]$ ;



**2.2:** Para cada um dos  $i$  valores gerados, calcule o  $p_i$ -ésimo quantil da distribuição usada para modelar a severidade.

**3:** Some cada vetor de severidades obtidas no passo 2 para obter uma amostra de perdas agregadas que representam vários cenários de perdas agregadas no intervalo de tempo determinado.

Com este algoritmo, podemos estimar os quantis da distribuição de  $S$  através dos quantis da distribuição empírica dos dados gerados. A distribuição empírica será apresentada posteriormente.

Um observação relevante deve ser feita:

**Observação:** Um crítica comum ao emprego do método MMC é que este pode não gerar valores suficientes na cauda direita da distribuição de perdas individuais o que pode causar uma subestimação da cauda direita da distribuição de perda agregada resultante. Por este motivo, o MMC pode não ser uma boa alternativa nestes casos, o que poderia aumentar muito o custo computacional na geração destes valores na cauda direita da distribuição de perdas individuais, além de que não há como determinar quantas simulações devem ser feitas para gerar valores extremos suficientes, o que geralmente implica na subestimação de quantis de nível muito alto ao se usar este método.

## 2.4 Método Recursivo de Panjer

O método recursivo que será introduzido a seguir foi proposto por *Harry H. Panjer* [15] em 1981 em um artigo da *Astin Bulletin* e devido a isso, iremos nos referir a este método como *Método Recursivo de Panjer* ou abreviadamente MRP.

Para aplicarmos o MRP, a distribuição de  $N$  deve satisfazer a seguinte propriedade:

$$P(N = n) = \left(a + \frac{b}{n}\right) P(N = n - 1) \text{ para } n = 1, 2, \dots \quad (2.14)$$

para algumas constantes  $a$  e  $b$ .

Se a distribuição de uma v.a discreta satisfaz esta propriedade, então podemos obter as probabilidades desta distribuição de forma recursiva, sendo este o pilar do MRP. Felizmente os modelos *Poisson* e *Binomial Negativo* satisfazem a propriedade acima e, além destes dois modelos, as distribuições *Binomial* e *Geométrica* também satisfazem esta relação.

A tabela a seguir dá os valores de  $a$  e  $b$  para os modelos Poisson e Binomial Negativo:

<i>Modelo</i>	Poisson( $\lambda$ )	Binomial Negativo ( $r, \beta$ )
$a$	0	$\beta/(1 + \beta)$
$b$	$\lambda$	$(r - 1)\beta/(1 + \beta)$

Tabela 2.1: Valores de  $a$  e  $b$  dos modelos Poisson e Binomial Negativo.

Pelo fato dos modelos Poisson e Binomial Negativo serem amplamente utilizados para modelar a frequência de perdas, o MRP tem sido uma alternativa mais atraente para se obter uma estimativa para a distribuição de  $S$ .

O método recursivo de Panjer, a ser introduzido a seguir, requer que tanto frequência quanto severidade sejam variáveis aleatórias discretas. No caso da frequência, está sempre será discreta pois trata-se de uma variável contadora (Poisson, Binomial Negativa). No entanto, as perdas individuais em geral são v.a.s contínuas ou não-aritméticas (como a empírica) tornando sua discretização necessária. Introduziremos agora o *método do arredondamento* capaz de discretizar distribuições contínuas ou "rediscretizar" as não-aritméticas.

**Método do arredondamento:** Seja  $F_X(\cdot)$  a função de distribuição de uma variável aleatória  $X$ . Seja  $f_X^d(j)$  a probabilidade no ponto  $jh$ ,  $j = 0, 1, 2, \dots$ . Defina<sup>2</sup>

$$f_X^d(0) = P(X < h/2) = F_X(h/2 - 0)$$

e

$$\begin{aligned} f_X^d(j) &= P(jh - h/2 \leq X < jh + h/2) \\ &= F_X(jh + h/2 - 0) - F_X(jh - h/2 - 0) \text{ para } j = 1, 2, \dots \end{aligned}$$

<sup>2</sup>A notação  $F_X(x-0)$  indica que a probabilidade discreta em  $x$  não deve ser incluída. Para distribuições contínuas, isto não fará diferença.

Os valores  $f_X^d(0), f_X^d(h), f_X^d(2h), \dots$ , são probabilidades da distribuição da v.a  $X$  agora discretizada.

O valor  $h$  que aparece no método do arredondamento, é definido como o passo (ou span) da discretização e deve ser escolhido de acordo com a necessidade do especialista. É claro que quanto menor o valor de  $h$  mais próximos estaremos da distribuição da v.a original  $X$  conservando suas propriedades. A figura (2.1) mostra a discretização feita para a distribuição exponencial de parâmetro unitário. É possível notar a perda de detalhes da distribuição original a medida que se aumenta o valor do passo  $h$ .

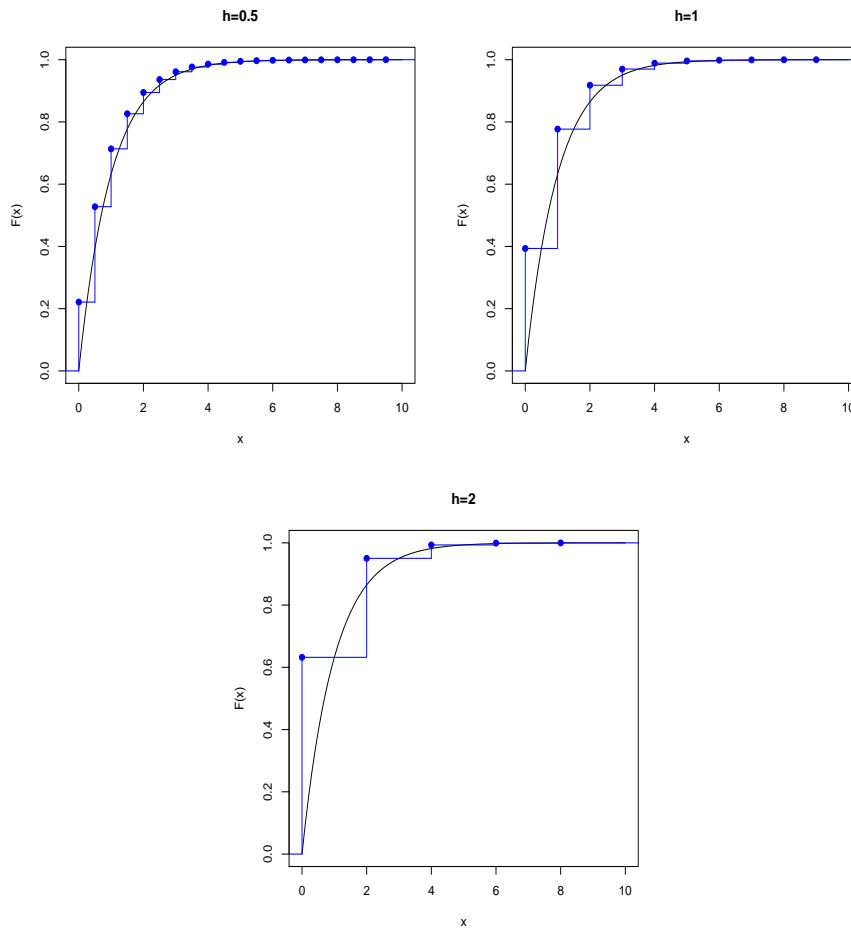


Figura 2.1: Gráfico da distribuição exponencial com o gráficos da função escada sobreposto (oriundos da discretização da distribuição exponencial) para diferentes valores de  $h$ .

### O Método Recursivo de Panjer (MRP)

De posse de um modelo para  $N$  (frequência) e um modelo devidamente discretizado para  $X$  (severidade), estamos aptos a empregar o MRP e obter a distribuição de  $S$ . Segundo o método, a função densidade de probabilidade das perdas agregadas  $f_S(\cdot)$  é dada pela seguinte fórmula recursiva:

$$f_S(s) = \frac{1}{1 - af_X(0)} \sum_{y=1}^s \left( a + \frac{by}{s} \right) f_X(y) f_S(s-y) \quad (2.15)$$

Onde  $a$  e  $b$  são constantes obtidas pela relação

$$P(N = n) = \left( a + \frac{b}{n} \right) P(N = n - 1), \quad n = 1, 2, \dots$$

e  $f_S(0) = M_N(f_X(0))$ , onde  $M_N(\cdot)$  é a função geradora de momentos de  $N$ .

É claro que  $f_X(\cdot)$  deverá ser substituída por  $f_X^d(\cdot)$  na fórmula acima, pois a recursão só funciona com distribuições discretas.

A tabela (2.2) abaixo sumariza algumas quantidades de interesse:

Modelo	<i>Poisson</i> ( $\lambda$ )	<i>BN</i> ( $r, \beta$ )
$a$	0	$\beta/(1 + \beta)$
$b$	$\lambda$	$(r - 1)\beta/(1 + \beta)$
$P(N = 0)$	$e^{-\lambda}$	$(1 + \beta)^{-r}$
$f_S(0)$	$e^{\lambda(f_X(0)-1)}$	$[1 + \beta(1 - f_X(0))]^{-r}$

Tabela 2.2: Algumas quantidade de interesse para recursões.

## 2.5 Valor em Risco

Tendo como objetivo o controle de risco e quantificação de capital de reserva para cobertura de movimentos adversos do mercado, as instituições financeiras (bancos e seguradoras principalmente) utilizam o *Valor em Risco* (*Value at Risk*, *VaR*) como medida. Tal medida é interpretada como

a possível perda que poderia acontecer a uma carteira ou portfólio em um horizonte de tempo  $T$  com probabilidade  $\alpha$ .

**Definição:** (Valor em Risco) Dada uma variável aleatória  $X$  com função de distribuição  $F_X(\cdot)$  e uma vez que o horizonte de tempo  $T$  é fixado, o valor em risco de nível  $\alpha$ , denotado por  $VaR_\alpha(X)$  é definido como o  $\alpha$ -quantil da distribuição de  $X$ . Em termos formais temos que:

$$VaR_\alpha(X) = F_X^{-1}(\alpha) \quad (2.16)$$

Em geral  $\alpha$  é um valor muito próximo de 1, ficando entre 0.950 e 0.999 sendo que o *comitê Basileia II* regulamenta que o valor de  $\alpha$  a ser escolhido pelos bancos para a estimação do capital de reserva para perdas operacionais é 0.999. Com isso, nosso interesse fica centrado na estimação de altos quantis de uma distribuição de probabilidades, e esta estimativa será tão boa ou ruim quanto a nossa capacidade de estimar a cauda direita de uma distribuição de probabilidade.

Podemos então interpretar o  $VaR$  da seguinte forma:

*O risco de um investimento perder mais do que 100.VaR $_\alpha$ % de seu valor durante o período  $T$  é de  $1 - \alpha$ .*

No contexto de reserva de capital para cobrir possíveis perdas extremas, o  $VaR$  calculado indica *o montante a ser reservado para cobrir, com probabilidade  $\alpha$ , todas as possíveis perdas que podem ocorrer no período  $T$ .*

Portanto, será de nosso interesse calcular uma estimativa do  $VaR_\alpha$  da função de distribuição de  $S$  da expressão (2.2), para estimar a quantidade de capital a ser reservado no período de um ano (em nosso caso particular).

## 2.6 Modelagem de Frequência

Nesta seção apresentaremos duas distribuições usadas para a modelagem de frequência. Trata-se dos modelos *Poisson* e *Binomial Negativo*, que são amplamente utilizados na modelagem de risco operacional. Por serem modelos simples e conhecidos, iremos nos focar na questão de qual dos dois modelos se aplica melhor em cada situação. Para detalhes sobre outros modelos, sugerimos Klugmann et al (1997) [12] e Cruz (2002) [6].

### Modelo Poisson

Devido a sua simplicidade e o fato de se ajustar muito bem na maioria dos conjuntos de dados de frequência, o modelo Poisson tem sido amplamente utilizado.

Assim, se  $N$  é uma variável aleatória de *Poisson* de parâmetro  $\lambda$  (denotamos por  $N \sim \text{Poisson}(\lambda)$ ), temos que:

$$P(N = k) = \frac{e^{-\lambda} \lambda^k}{k!}, k = 0, 1, 2, \dots, \lambda > 0$$

A esperança e variância do modelo Poisson são dadas, respectivamente por

$$E(N) = \lambda$$

$$\text{Var}(N) = \lambda$$

Abaixo temos alguns gráficos que mostram o comportamento da variável de Poisson para alguns valores de  $\lambda$ .

### Modelo Binomial Negativo

Este é o segundo modelo mais utilizado depois do modelo Poisson. Trata-se de um modelo alternativo devido ao fato de possuir dois parâmetros, dando mais flexibilidade à forma da distribuição o que não ocorre no modelo Poisson.

Dizemos que uma variável aleatória  $N$  tem distribuição *Binomial Negativa* com parâmetros  $r$  e  $\beta$  (notação:  $N \sim \text{BN}(r, \beta)$ ) se

$$P(N = k) = \frac{(k+r-1)!}{(k-1)!r!} \left( \frac{1}{1+\beta} \right)^r \left( \frac{\beta}{1+\beta} \right)^k$$

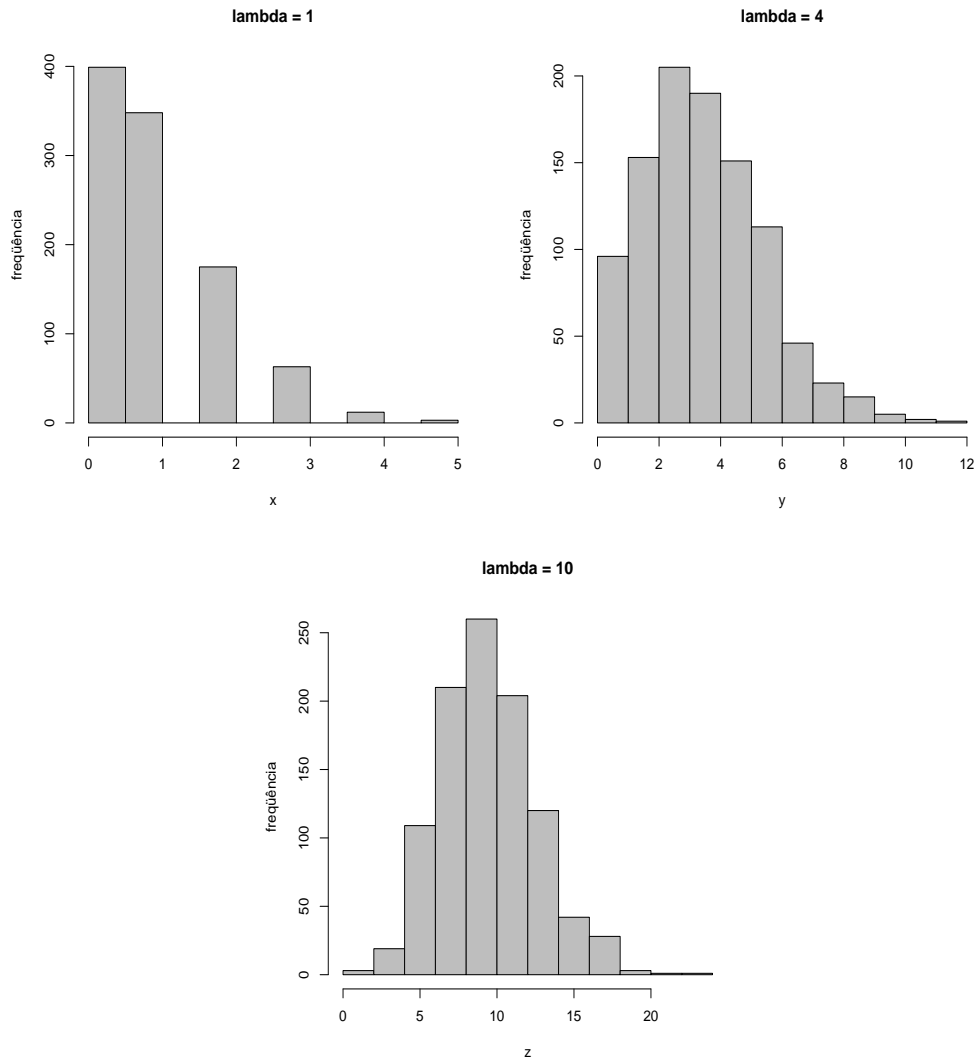


Figura 2.2: Histograma de dados Poisson para diferentes valores de  $\lambda$ .

onde  $k = 0, 1, \dots, n$ ,  $r > 0$  e  $\beta > 0$ .

A esperança e variância do modelo Binomial Negativo são dadas, respectivamente por

$$E(N) = r\beta$$

$$Var(N) = r\beta(1 + \beta)$$

Abaixo temos alguns gráficos com o comportamento desta distribuição para diferentes valores para  $r$  e  $\beta$ .

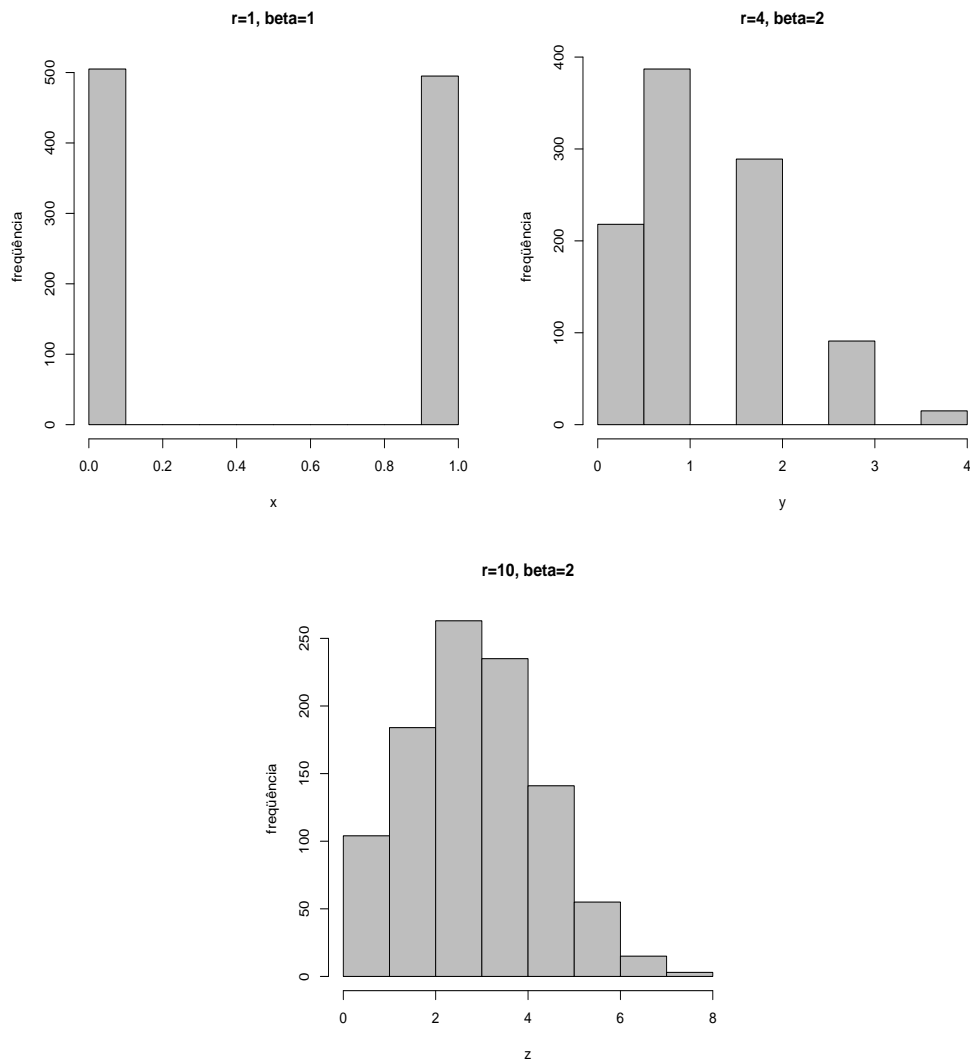


Figura 2.3: Histograma de dados Binomial Negativo para diferentes valores de  $r$  e  $\beta$ .

### Poisson X Binomial Negativo: Qual modelo escolher?

Alguns recursos exploratórios podem nos auxiliar na resposta desta pergunta.



Note que no modelo Poisson temos uma propriedade interessante: esperança e variância são iguais, isto é,  $E(N) = Var(N) = \lambda$ . Assim, se a amostra de frequências possuir média e variância amostrais aproximadamente iguais, teremos evidências de que o ajuste a ser feito é Poisson. Note que no caso do modelo Binomial Negativo,  $E(N) < Var(N)$ .

Uma outra alternativa, mais sofisticada, usa a propriedade que estas duas distribuições possuem de gerar suas probabilidades recursivamente. Assim, dada uma amostra de frequências, seja  $n_k$  a frequência observada no período  $k$ , calculamos  $kn_k/n_{k-1}$  e plotamos estes valores contra  $k$ .

Os pontos do gráfico deverão ter um comportamento aproximadamente linear. O seguinte procedimento deve ser adotado:

- Se a inclinação da reta (imaginária) que se ajusta aos pontos for positiva, então há evidências de que o modelo a ser ajustado é Binomial Negativo;
- Se a inclinação for 0 (ou muito próxima de 0), então o modelo mais adequado é Poisson.

Infelizmente, pode ocorrer casos onde a inclinação da reta é praticamente nula, mas não dispomos de técnicas para constatar o quão significativa é esta inclinação. No entanto, utilizar esta técnica com o fato de que para o caso Poisson  $E(N) = Var(N)$ , auxiliará na escolha do modelo.

**Exemplo:** Considere dos dados abaixo retirados de *Klugmann et al* (1997).

$n^\circ$ de acidentes	$n^\circ$ de apólices	$kn_k/n_{k-1}$
0	7840	—
1	1317	0.17
2	239	0.36
3	42	0.53
4	14	1.33
5	4	1.43
6	4	6.00
7	1	1.75
8+	0	—
<i>Total</i>	9461	

Tabela 2.3: Perfis de acidentes.

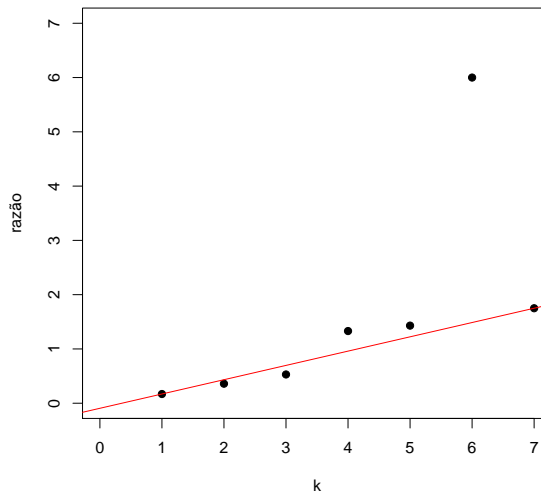


Figura 2.4: Gráfico da razão  $kn_k/n_{k-1}$  versus  $k$ .

Pela figura (2.4) podemos notar que a linha (imaginária) que se ajusta aos pontos tem uma leve inclinação positiva mas, como discutido anteriormente, não sabemos o quão significativa é esta inclinação. No entanto, para este conjunto de dados, a média e variância amostral calculadas foram 0.2144 e 0.2889 respectivamente, o que nos leva a considerar o modelo Binomial Negativo pelo fato da variância amostral ser maior que a média amostral.

Ao longo deste trabalho, iremos apenas considerar o modelo Poisson, já que os resultados assintóticos que queremos utilizar posteriormente requerem que sejam verificadas certas propriedades do modelo de frequência adotado. Para o caso particular que iremos estudar, o modelo Poisson é adequado e satisfaz as propriedades em questão.

## 2.7 Modelagem de Severidade

Modelar as perdas individuais ou severidades pode ser uma tarefa simples ou extremamente complicada, já que os dados podem ser mal coletados, possuir censura (truncamento) ou serem limitados de alguma forma.

Os dados podem ser modelados através de duas abordagens principais:

**A** Não paramétrica;

**B** Paramétrica.

Na abordagem não paramétrica, dispomos de ferramentas simples para se obter um modelo para os dados. No entanto, esta abordagem só se torna eficiente em geral quando dispomos de um conjunto de dados muito grande.

Um exemplo de modelo não paramétrico é a distribuição empírica, onde dada uma amostra  $x_1, \dots, x_n$ , a função de distribuição que gerou tal amostra é estimada por:

$$E_X(x) = \frac{\text{número de valores} \leq x}{n} \quad (2.17)$$

Na abordagem paramétrica podemos encontrar uma grande variedade de modelos disponíveis. Em geral os modelos para a severidade são baseados em distribuições contínuas e alguns exemplos são dados abaixo:

- 1 Distribuições comuns: Entre elas a exponencial, lognormal, Weibull, gama, beta, Pareto, normal, entre outras;
- 2 Mistura de distribuições;
- 3 Distribuições alfa-estáveis;
- 4 Distribuições de valores extremos;
- 5 Distribuições truncadas.

Neste trabalho iremos nos concentrar apenas no item 1 para modelar a severidade como já discutido anteriormente. A distribuição empírica será utilizada na construção do método POT, a ser discutido a seguir, e em outras situações que serão citadas quando oportuno.

### 2.7.1 O método Peaks Over Threshold - POT

Esta seção pretende apresentar o método *Peaks Over Threshold* (ou abreviadamente POT) que será usado para modelar as perdas individuais (severidades) acima de um limiar, que também será determinado pelo método.

Tal método surgiu na década de 80 na abordagem de problemas na área de hidrologia e que vem sendo amplamente utilizado nas mais diversas áreas como, por exemplo, na medicina onde os excessos acima de um limiar alto podem representar o tempo de vida restante, enquanto que no contexto de seguros, os excessos podem representar perdas em excesso, isto é, riscos com os quais as empresas preferem transferir (resseguro) ou criar uma proteção (reserva de capital).

O método POT soluciona o chamado *problema de estimação da cauda de uma distribuição probabilística através de sua amostra* ou, simplesmente, *problema de estimação da cauda*, ou seja, o método resolve o seguinte problema:

**(P)** Dada uma amostra aleatória  $x_1, \dots, x_n$  gerada de uma distribuição cuja função de distribuição seja  $F(\cdot)$  (desconhecida). Queremos reconstruir, através de uma boa aproximação, a cauda de  $F(\cdot)$  à direita de um dado limiar  $u$ . Tal limiar deve ser escolhido de tal forma que uma boa precisão seja obtida na aproximação da cauda e que seja coerente no contexto da aplicação desta aproximação.

O desenvolvimento de tal método tornou-se possível através de resultados provenientes da teoria de valores extremos, cujos primeiros resultados importantes datam de 1928 com os estudos de Fisher e Tippett acerca do comportamento assintótico do máximo de variáveis aleatórias independentes e idênticamente distribuídas. A teoria apresentada a seguir será necessária para justificar o método POT.

#### Teoria de Valores Extremos

Sejam  $X, X_1, X_2, \dots$  variáveis aleatórias (v.a.s) independentes e idênticamente distribuídas (i.i.d) com função distribuição comum  $F_X(\cdot)$ . Seja  $M_n = \max(X_1, \dots, X_n)$ .

**Definição** (Max-estabilidade): Dizemos que  $F_X(\cdot)$  é max-estável se satisfaz a igualdade em distribuição:

$$\max(X_1, \dots, X_n) \stackrel{d}{=} c_n X + d_n$$

para constantes  $c_n > 0$  e  $d_n \in \mathbb{R}$  apropriadas.

O Teorema Central do Limite estabelece que dentro das funções distribuições não-degeneradas apenas as distribuições estáveis podem ser distribuições limite para a soma de  $n$  v.a.s. De forma análoga, o teorema a seguir irá estabelecer que apenas as distribuições max-estáveis podem ser distribuições limite para o máximo de  $n$  variáveis aleatórias.

**Teorema** (Fisher-Tippett) Se existirem seqüências  $\{c_n, n \geq 1\}$  e  $\{d_n, n \geq 1\}$  de números reais com  $c_n > 0$  e uma função distribuição  $H(\cdot)$  não-degenerada tais que

$$P\left(\frac{M_n - d_n}{c_n} \leq x\right) \xrightarrow{n \rightarrow \infty} H(x)$$

então  $H(\cdot)$  pertence a uma das três famílias de distribuições:

- Tipo I ou Gumbel:

$$H_{I,\mu,\sigma}(x) = H_I\left(\frac{x - \mu}{\sigma}\right), \quad x, \mu \in \mathbb{R}, \quad \sigma > 0$$

- Tipo II ou Fréchet:

$$H_{II,\mu,\sigma,\alpha}(x) = \begin{cases} 0, & (x - \mu) \leq 0 \\ \exp\{-((x - \mu)/\sigma)^{-\alpha}\}, & (x - \mu) > 0 \end{cases} \quad \alpha > 0$$

- Tipo III ou Weibull:

$$H_{III,\mu,\sigma,\alpha}(x) = \begin{cases} \exp\{-(-(x - \mu)/\sigma)^{-\alpha}\}, & (x - \mu) \leq 0 \\ 1, & (x - \mu) > 0 \end{cases} \quad \alpha < 0$$

Para uma prova detalhada deste resultado, sugerimos Embrechts, Kluppelberg e Mikosh (1997).

A figura (2.5) abaixo ilustra as três famílias de distribuições de valores extremos.

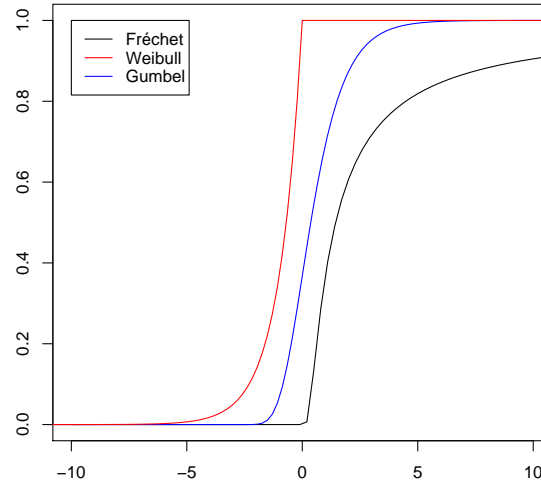


Figura 2.5: Distribuições de Fréchet ( $\alpha = 1$ ), Weibull ( $\alpha = -1$ ) e Gumbel

As distribuições do tipo I, II e III admitem uma forma padrão, isto é, sem os parâmetros de locação ( $\mu$ ) e escala ( $\sigma$ ) e são dadas por:

$$H_I(x) = \exp\{-e^{-x}\}, \quad x \in \mathbb{R}$$

$$H_{II,\alpha}(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \quad \alpha > 0$$

$$H_{III,\alpha}(x) = \begin{cases} \exp\{-(-x)^{-\alpha}\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha < 0$$

Sendo que as respectivas funções densidade de probabilidade são dadas por:

- Gumbel:  $h_I(x) = H_I(x)e^{-x}$ ,  $x \in \mathbb{R}$
- Fréchet:  $h_{II,\alpha}(x) = \alpha H_{II,\alpha}(x)x^{-(1+\alpha)}$ ,  $x \geq 0$ ,  $\alpha > 0$
- Weibull:  $h_{III,\alpha}(x) = |x|H_{III,\alpha}(x)(-x)^{-(1+\alpha)}$ ,  $x \leq 0$ ,  $\alpha < 0$

As distribuições dos tipos I, II e III podem ser representadas por uma única família de distribuições conhecida como Distribuição de *Valores Extremos Generalizada* (*Generalized Extreme Value*, GEV). A distribuição padrão, isto é, sem parâmetros de locação e escala, denotada por  $H_\xi$  é dada por:

$$H_\xi(x) = \begin{cases} \exp\{-(1 + \xi x)^{-1/\xi}\}, & \text{se } \xi \neq 0 \\ \exp\{-\exp\{-y\}\}, & \text{se } \xi = 0 \end{cases} \quad 1 + \xi x > 0$$

Para se aplicar o método POT, um outro conceito deve ser introduzido. Trata-se da definição de *domínio de atração* que será utilizado na formulação do teorema no qual o método POT se baseia.

**Definição** (Domínio de atração): Dizemos que  $F_X(\cdot)$  pertence ao domínio máximo de atração da distribuição de valores extremos  $H$  (Notação:  $F_X \in MDA(H)$ ) se, para as constantes apropriadas  $c_n > 0$  e  $d_n \in \mathbb{R}$ ,

$$\frac{\max(X_1, \dots, X_n) - d_n}{c_n} \xrightarrow{n \rightarrow \infty} H\left(\frac{x - d_n}{c_n}\right)$$

### Observações acima de um limiar (excedentes)

Sejam  $X_1, X_2, \dots, X_n$  v.a.s i.i.d com função distribuição comum  $F_X$ . Iremos denotar por  $x_F$  o limite superior da distribuição  $F_X$ , onde podemos ter  $x_F = \infty$  no caso de distribuições com cauda

infinita<sup>3</sup>, sendo este o caso mais comum em aplicações. Seja  $u$  um valor pré-determinado tal que  $u < x_F$ . Chamamos de *excedentes* acima de um limiar  $u$  os valores  $X_i$  tais que  $X_i > u$  e  $N_u$  o número de excedentes acima do limiar  $u$ . Os *excessos* acima do limiar  $u$  são dados por

$$Y_i = X_i - u \geq 0, \quad i = 1, \dots, N_u$$

A função de distribuição dos excessos da distribuição  $F_X(\cdot)$  acima do limiar  $u < x_F$  é denotada por  $F_u(\cdot)$  sendo obtida pela fórmula de probabilidade condicional a saber:

$$F_u(x) = P(X - u | X > u), \quad x \in \mathbb{R}, u < x_F \quad (2.18)$$

### Família de distribuições de Pareto Generalizadas

Uma distribuição de Pareto generalizada (abreviadamente GPD do inglês *Generalized Pareto Distribution*) é uma função que possui uma das seguintes formas:

$$G_{0,\sigma,\mu}(x) = \begin{cases} 1 - e^{-\frac{x-\mu}{\sigma}}, & x \geq \mu \\ 0, & x < \mu \end{cases} \quad \xi = 0 \quad (2.19)$$

$$G_{\xi,\sigma,\mu}(x) = \begin{cases} 1 - \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}, & x \geq \mu \\ 0, & x < \mu \end{cases} \quad \xi > 0 \quad (2.20)$$

$$G_{\xi,\sigma,\mu}(x) = \begin{cases} 0, & x < \mu \\ 1 - \left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi}, & x \in \left[\mu, \mu - \frac{\sigma}{\xi}\right] \\ 1, & x > \mu - \frac{\sigma}{\xi} \end{cases} \quad \xi < 0 \quad (2.21)$$

onde  $\xi$ ,  $\sigma$  e  $\mu$  são os parâmetros de forma, escala e localização respectivamente.

---

<sup>3</sup>Dizemos que uma função de distribuição  $F_X(\cdot)$  tem cauda finita se  $x_F < \infty$  e  $F_X(x) = 1$  para  $x \geq x_F$ . Dizemos que tem cauda infinita quando  $x_F = \infty$  e  $\lim_{x \rightarrow x_F} F_X(x) = 1$ .



Dentro da classe de funções  $G_{\xi,\sigma,\mu}(\cdot)$ , as distribuições de maior interesse em aplicações são aquelas onde  $\xi \geq 0$ .

### O resultado de Pickands - A base do método POT

O teorema, a ser apresentado a seguir, estabelece que dado uma função de distribuição que satisfaça certas condições, a sua função de distribuição dos excessos correspondente pode ser aproximada por uma GPD e que o erro cometido por esta aproximação tende a zero conforme o limiar cresce. Formalmente temos:

**Teorema** (Pickands): Seja  $F_X(\cdot)$  uma função de distribuição. Se  $F_X \in MDA(H_\xi)$  então

$$\lim_{u \uparrow x_F} \sup_{0 < x < x_F - u} |F_u(x) - G_{\xi,\sigma(u),0}(x)| = 0$$

A recíproca do resultado acima é válida sendo útil para determinar em qual domínio de atração encontra-se  $F_X(\cdot)$  graças à relação que existe entre a GPD e a GEV a saber:

$$G_\xi(x) = 1 + \ln(H_\xi(x)), \quad \ln(H_\xi(x)) > -1$$

A relação existente entre a cauda direita de uma função de distribuição e a distribuição de seus excessos acima de uma limiar é o que irá consolidar o método POT a partir do resultado de Pickands. Em outras palavras, queremos modelar a cauda direita de uma distribuição  $F_X(\cdot)$  desconhecida e para isto usaremos a relação (a ser mostrada a seguir) entre  $F_X(\cdot)$  e  $F_u(\cdot)$ . Como  $F_u(\cdot)$  também é uma distribuição desconhecida, podemos utilizar o resultado de Pickands para estimar  $F_u(\cdot)$  por uma GPD adequada.

Partindo da relação dada em (2.18), temos que

$$F_u(x) = \begin{cases} 0, & x < 0 \\ \frac{F_X(x+u) - F_X(u)}{1 - F_X(u)}, & x \geq 0 \end{cases} \quad u < x_F$$

Assim, para  $u < x_F$ , temos que

$$F_X(x) = F_X(u) + (1 - F_X(u))F_u(x - u), \quad x \geq u$$

Logo, se o resultado de Pickands for aplicável e para  $u$  suficientemente grande, podemos considerar que

$$F_X(x) \approx F_X(u) + (1 - F_X(u))G_{\xi, \sigma(u), 0}(x - u), \quad x \geq u$$

Posto isto, temos as seguintes perguntas a serem respondidas:

- 1 - Qual limiar  $u$  devemos escolher?
- 2 - Qual a estimativa para  $F_X(u)$ ?
- 3 - Para que o resultado de Pickands seja aplicável,  $F_X(\cdot)$  deve pertencer ao domínio de atração de alguma distribuição GEV. Sempre será satisfeita a condição  $F_X(\cdot) \in MDA(H_\xi)$ ?

A primeira pergunta é motivada pelo fato de que o resultado de Pickands nos diz que podemos aproximar  $F_u(\cdot)$  por  $G_{\xi, \sigma(u), 0}(\cdot)$  quando  $u \uparrow x_F$ . Como  $\xi \geq 0$  na maioria dos problemas práticos, temos que  $x_F = \infty$ . Assim, nos deparamos com um problema de *trade-off* entre vício e variância, onde um valor  $u$  muito alto implicará em um número pequeno de observações na cauda, podendo resultar em uma variabilidade maior dos estimadores, enquanto que um limiar não suficientemente alto não irá satisfazer as condições do teorema de Pickands, podendo resultar em estimativas distorcidas para a GPD que aproxima os excessos.

A esperança condicional dos excessos, também conhecida como função média dos excessos, no fornecerá uma ferramenta intuitiva para a escolha de um limiar "ótimo", ou seja, um limiar não muito alto e que satisfaça as condições do teorema de Pickands.

Temos que a esperança condicional dos excessos é definida por:

$$e(u) := E[X - u | X > u], \quad u < x_F$$

$$\begin{aligned}
&= \int_0^\infty \frac{x}{1 - F_X(u)} dF_X(x + u) \\
&= \int_0^\infty \frac{x f_X(x + u)}{1 - F_X(u)} dx \quad (\text{Caso } X \text{ seja contínua})
\end{aligned}$$

Se  $X \sim G_{\xi, \sigma(u), 0}$  com  $\xi < 1$ , então para  $u < x_F$ ,

$$e(u) = \frac{\sigma + u\xi}{1 - \xi}, \quad \sigma + u\xi > 0$$

Note que a  $e(u)$  de uma GPD é uma função linear em  $u$ . Se  $\xi > 0$ ,  $e(u)$  tem inclinação positiva enquanto que  $\xi < 0$  indica uma inclinação negativa para  $e(u)$ . Este fato irá motivar a escolha do limiar "ótimo" do método POT.

Seja  $x_1, x_2, \dots, x_n$  uma amostra aleatória. Um estimador para a função média dos excessos, denotada por  $\hat{e}(u)$ , é dado por

$$\hat{e}(u) := \frac{1}{N_u} \sum_{i=1}^{N_u} y_i$$

Lembrando que  $N_u$  é o número de excedentes acima de  $u$  e  $y_i$  são os excessos (amostrais) acima de  $u$ , isto é,  $y_i = x_i - u$  para cada  $x_i > u$ .

Usando o fato de que  $e(u)$  é linear quando  $X \sim G_{\xi, \sigma(u), 0}$ , podemos plotar o gráfico de  $\hat{e}(u)$  versus  $u$  e procurar, indo da direita para a esquerda no gráfico de  $\hat{e}(u)$ , uma região que apresente comportamento linear e que não seja muito irregular. O valor da abscissa do extremo esquerdo deste trecho será declarado como o limiar "ótimo" e será denotado por  $u^{opt}$ .

A resposta para a segunda pergunta parece natural. Iremos estimar  $F_X(u)$  pela distribuição empírica construída com base na amostra retirada de  $F_X(\cdot)$  avaliada no ponto  $u$ . Tal estimativa de  $F_X(u)$  será denotada por  $\hat{F}_X(u^{opt})$  e dada por

$$\begin{aligned}
\hat{F}_X(u^{opt}) &= \frac{\text{número de pontos da amostra que não excedem } u^{opt}}{\text{número total de pontos da amostra}} \\
&= \frac{n - N_{u^{opt}}}{n}
\end{aligned}$$

Por último, a resposta para a terceira pergunta é dada por Pickands (1975) [18] onde ele afirma que praticamente todas as distribuições conhecidas pertencem ao domínio de atração de alguma distribuição de valores extremos. Tal pressuposto pode ser testado via métodos estatísticos que não serão abordados no presente trabalho. Iremos apenas aceitar que a função distribuição  $F_X(\cdot)$ , da qual conhecemos apenas uma amostra, está no domínio de atração de alguma GEV.

Finalmente estamos aptos a formular o método POT através das 3 etapas seguintes:

- **(1)** Encontrar o menor limiar possível tal que a amostra dos excessos acima deste limiar possa ser bem aproximada por uma GPD com parâmetro de locação nulo. Este limiar será denominado *limiar ótimo* e denotado por  $u^{opt}$ ;
- **(2)** Encontrar uma GPD que melhor se adere à amostra dos excessos acima de  $u^{opt}$  através da estimação dos parâmetros de forma ( $\xi$ ) e escala ( $\sigma$ ), denotados por  $\hat{\xi}$  e  $\hat{\sigma}$  respectivamente;
- **(3)** Construir a seguinte função que irá aproximar a cauda de  $F_X$  acima de  $u^{opt}$ :

$$F_X^*(x) = \hat{F}_X(u^{opt}) + (1 - \hat{F}_X(u^{opt}))G_{\hat{\xi}, \hat{\sigma}(u^{opt}), 0}(x - u^{opt}), \quad x \geq u^{opt} \quad (2.22)$$

onde  $\hat{F}_X(u^{opt}) = \frac{n - N_{u^{opt}}}{n}$ .



## Capítulo 3

# Aplicação e Conclusões

### 3.1 Apresentação dos dados

Este conjunto de dados, retirado de Belitsky e Moreira (2007) [1], foi cedido por um banco brasileiro sendo composto de 354 observações de perdas referentes à fraudes ocorridas em um determinado setor deste banco no período de um ano obtidos até julho de 2006. Não possuímos autorização do banco para divulgação destes dados, mas a tabela (3.1) nos fornece as principais estatísticas descritivas.

<b>Principais estatísticas</b>	
Número de observações	354
Soma das perdas individuais	388.396,80
Mínimo	0,01
Máximo	33.390,00
Primeiro Quartil	0,30
Mediana	16,83
Média	1.097,00
Terceiro Quartil	298,20
Variância	13.696.069,00
Desvio padrão	3.700,82
Segundo momento	14.861.152,00

Tabela 3.1: Principais estatísticas descritivas do conjunto de dados de perdas operacionais. A unidade monetária é o real (R\$).

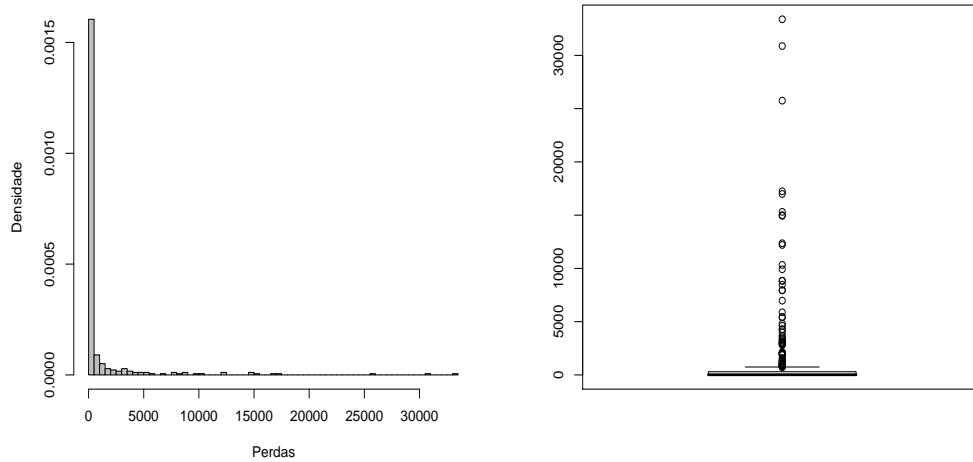


Figura 3.1: Histograma e Boxplot do conjunto de dados de perdas operacionais.

Pela tabela (3.1) e pelos gráficos da figura (3.1) podemos observar o comportamento característico deste tipo de conjunto de dados. Uma grande parte dos dados (em torno de 70%) é constituída de valores considerados pequenos sendo que grandes perdas, apesar de raras, são de fato significativas. Note que a maior perda foi de aproximadamente 33 mil reais e se compararmos média amostral com mediana amostral iremos notar a presença destes extremos na amostra.

Antes de prosseguirmos, iremos assumir alguns pressupostos:

1. Vamos assumir que todos os dados referem-se a perdas que ocorreram em um mesmo setor do banco, já que na prática os bancos separam as perdas operacionais em vários setores tratando-os separadamente. Logo, existe uma função de distribuição  $F_X(\cdot)$  tal que toda perda ocorrida é uma realização de uma variável aleatória  $X$  com função de distribuição  $F_X(\cdot)$ ;
2. Assumiremos que as realizações são independentes;
3. A função de distribuição  $F_X(\cdot)$  que gerou a amostra pertence ao domínio de atração de alguma distribuição de valores extremos. Testes apropriados para se verificar este pressuposto podem ser obtidos em Embrechts et al (1997) [7] ou em Reiss e Thomas (2001) [20]. Iremos aceitar este pressuposto sem testá-lo por este não ser o foco do presente trabalho e assim, estarmos

aptos a aplicar o método POT obtendo uma estimativa para a cauda direita de  $F_X(\cdot)$  que será denotada por  $\hat{F}_X^*(\cdot)$ .

### 3.2 Planejamento e Execução

Nosso objetivo a partir de agora é obter um modelo para as severidades (perdas individuais) e para a frequência aleatória de eventos que produziu cada perda, aplicar os métodos de Panjer e Monte Carlo para obter uma estimativa para a função de distribuição de perda agregada e finalmente estimar o valor em risco (VaR) para diferentes níveis.

No caso da frequência, a tarefa será mais simples. Queremos calcular o VaR anual de nível  $\alpha$  do modelo de perda agregada e como possuímos 354 observações, vamos considerar que elas ocorreram no período de um ano. Desta forma, vamos considerar que  $N$  do modelo (2.2) é Poisson com parâmetro  $\hat{\lambda} = 354$ .

**Observação:** Se tivéssemos dados mensais, a estimativa para o parâmetro  $\lambda$  escolhido seria a média do número de perdas ocorridas em cada mês. E o VaR a ser calculado seria o VaR mensal de nível  $\alpha$ . Em nosso caso particular não dispomos de informação sobre qual perda ocorreu em cada mês nem sobre a quantidade de perdas ocorridas em cada mês.

A questão sobre o modelo a ser usado para as severidades é um dos pontos-chaves deste trabalho e será tratado com muito mais atenção. Iremos executar oito abordagens sendo que cada uma delas considera um modelo diferente para a severidade ou um método computacional diferente para a obtenção da distribuição da perda agregada.

As abordagens em questão são descritas abaixo:

1. Considerar a distribuição empírica como estimativa para a função distribuição que gerou os dados de severidades, aplicando o método recursivo de Panjer (MRP) e o método de Monte Carlo (MMC) para a obtenção da função de distribuição da perda agregada;
2. Considerar a distribuição empírica abaixo de um limiar pré-determinado e a distribuição de Pareto generalizada (GPD), dada pelo método POT acima deste limiar;



Formalmente teríamos:

$$\hat{F}_X(x) = \begin{cases} E_X(x) & , x < u \\ F_X^*(x) & , x \geq u \end{cases} \quad (3.1)$$

onde  $E_X(\cdot)$  é a função distribuição empírica dada por (2.17) obtida com base na amostra e  $\hat{F}_X^*(\cdot)$  é a GPD fornecida pelo método POT que aproxima a cauda da distribuição adjacente  $F_X(\cdot)$ . Novamente usaremos o MRP e o MMC para obtermos a função de distribuição da perda agregada;

3. A ultima etapa é dividida em duas como segue:

- (a) Aplicar o método de Panjer para os dados abaixo do limiar pré-determinado considerando a distribuição empírica como modelo para a severidade. A estimativa do parâmetro  $\lambda$  para o modelo de freqüência será o número de ocorrências abaixo deste limiar e será denotado como  $\hat{\lambda}_A$  (de antes). A v.a  $S^A$  representa a perda agregada abaixo do limiar escolhido e terá a estimativa para sua função distribuição denotada por  $\hat{F}_{S^A}(\cdot)$  e obtida via MRP e MMC.

Em seguida, aplicar o método de Panjer para os dados acima do limiar pré-determinado considerando a GPD dada pelo método POT, que modela os excessos acima deste limiar, como modelo para a severidade. A estimativa do parâmetro  $\lambda$  para o modelo de freqüência será o número de ocorrências acima deste limiar e será denotado como  $\hat{\lambda}_D$  (de depois). A v.a  $S^D$  representa a perda agregada abaixo do limiar escolhido e terá como estimativa para sua função distribuição denotada por  $\hat{F}_{S^D}(\cdot)$ .

A distribuição de  $S$  será dada por  $S^A + S^D$  onde assumiremos a independência entre  $S^A$  e  $S^D$ . A estimativa para sua função distribuição será denotada por  $\hat{F}_S(\cdot)$  e será obtida através da convolução das duas distribuições obtidas, isto é

$$\hat{F}_S(\cdot) = (\hat{F}_{S^A} * \hat{F}_{S^D})(\cdot) \quad (3.2)$$

- (b) Iremos considerar que  $S^A$  tem distribuição normal com média e variância dadas respecti-

vamente por (2.4) e (2.5) e que o modelo para  $S^D$  é o mesmo obtido no sub-item anterior. Novamente a função distribuição  $F_S(\cdot)$  será obtida através da convolução dada em (3.2).

Para cada abordagem, iremos calcular os respectivos valores em risco anual de nível  $\alpha \in [0.950, 0.999]$ .

### 3.2.1 Resultados

#### Modelo: Empírico

Aplicamos o método recursivo de Panjer (MRP) considerando que  $N \sim Poisson(\hat{\lambda} = 354)$  e usamos a função de distribuição empírica como estimativa para a função de distribuição que gerou os dados. Usamos o método do arredondamento apresentado na sub-seção (2.4) usando três passos diferentes (denotado por  $h$ ). Apesar da distribuição empírica ser discreta ela não é aritmética, isto é, os intervalos entre os saltos da função de distribuição empírica não são iguais, logo devemos "rediscretizá-la". O intervalo de valores discretizados foi  $[0, 100000]$ . Observamos que adotar um passo menor que 5 acarreta em uma demora demasiada para a obtenção dos resultados.

As figuras (3.2) e (3.3) ilustram, respectivamente, as discretizações obtidas e os gráficos das distribuições de perdas agregadas para os diferentes valores de  $h$ .

Os valores em risco anuais obtidos para os diferentes níveis e diferentes passos são dados na tabela (3.2)

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$h = 5$	513.630,00	572.590,00	594.970,00	642.445,00
$h = 10$	513.550,00	572.510,00	594.890,00	642.360,00
$h = 100$	511.700,00	570.700,00	593.100,00	640.600,00

Tabela 3.2: Valores em Risco obtidos através do Método Recursivo de Panjer para diferentes valores de  $h$  usados na discretização do modelo para a severidade considerando a distribuição empírica como geradora das perdas individuais.

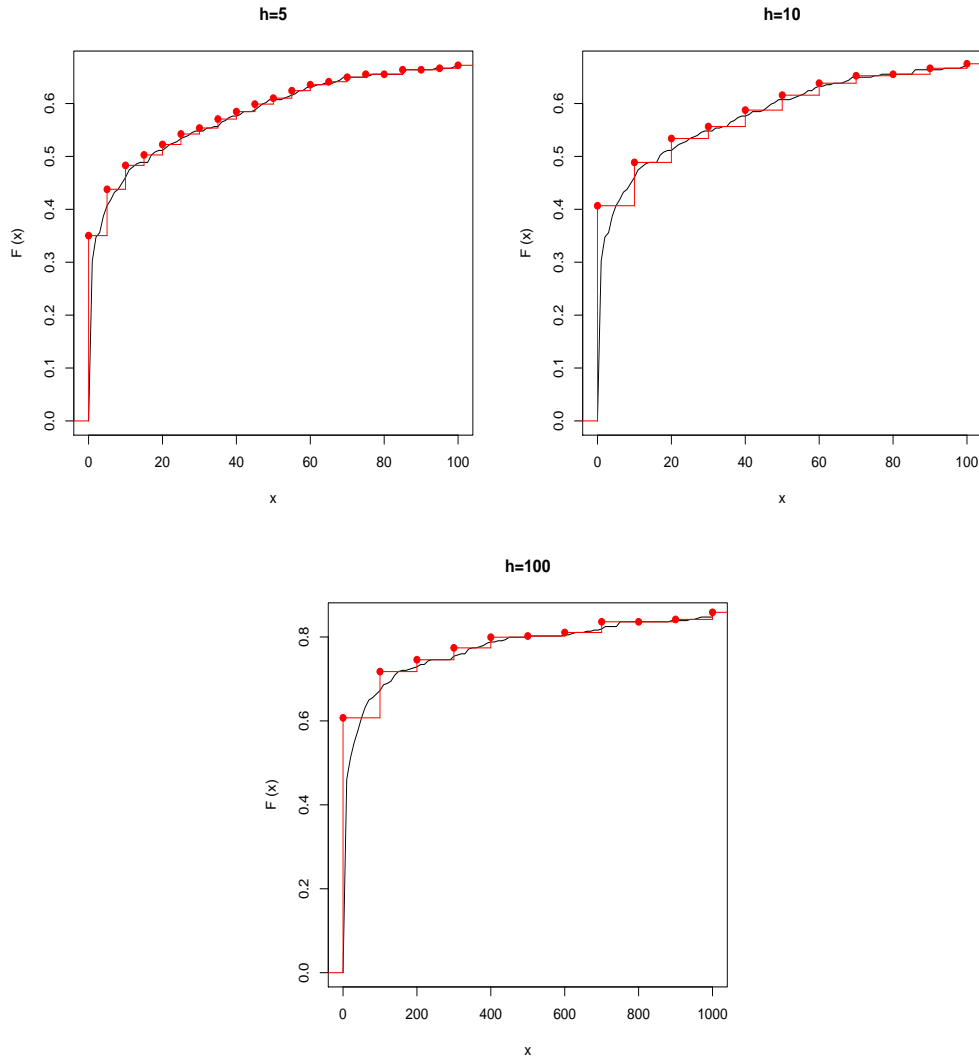


Figura 3.2: Discretização da severidade para os diferentes valores do passo  $h$  considerando a distribuição empírica como geradora das perdas individuais.

Observe que a diferença dos resultados levando em conta os diferentes passos da discretização não influencia significativamente no resultado obtido pelo cálculo do VaR neste caso. No entanto, a diferença de tempo útil para se aplicar o método para  $h = 5$  e  $h = 100$  é realmente digna de nota. Enquanto o primeiro caso ( $h = 5$ ) levou pouco mais de 1 hora de cálculo, o segundo caso ( $h = 100$ ) retorna o resultado em alguns segundos.

Em seguida aplicamos o método de Monte Carlo (MMC) sob as mesmas condições adotadas

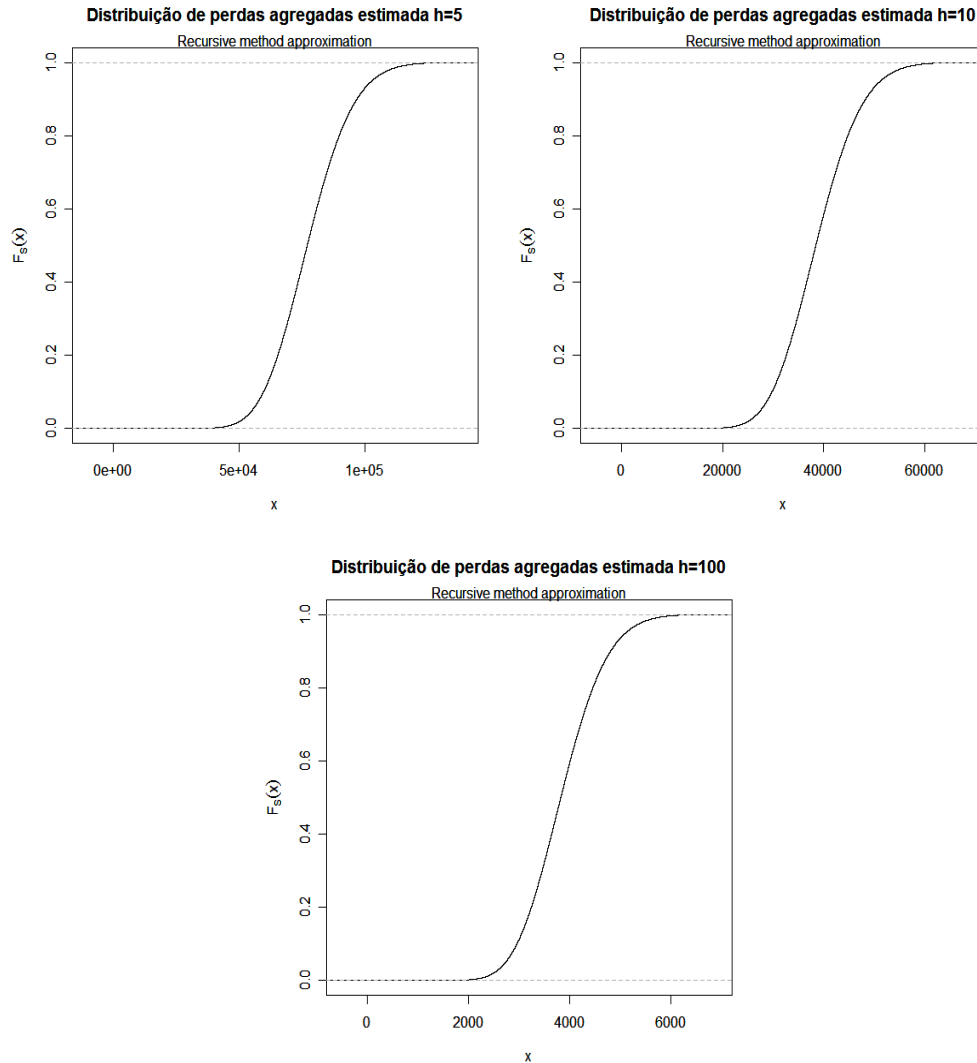


Figura 3.3: Distribuição de perdas agregadas obtidas pelo Método Recursivo de Panjer para os diferentes valores de  $h$  considerando a distribuição empírica como geradora das perdas individuais.

anteriormente, isto é,  $N \sim Poisson(\hat{\lambda} = 354)$  e severidades modeladas pela distribuição empírica. Pelo fato do valor de  $\hat{\lambda}$  ser muito grande, apenas foi possível realizar 20000 (vinte mil) simulações já que o MMC cria uma matriz de cenários muito grande que a memória física dos computadores (mesmo os mais modernos) não consegue alocar. Obviamente a cada nova simulação os resultados mudam, mas não ficam muito distantes dos valores obtidos na tabela (3.3).

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$VaR_\alpha$	490.703,30	548.450,00	569.873,60	616.392,70

Tabela 3.3: Valores em Risco obtidos através do Método de Monte Carlo considerando a distribuição empírica como geradora das perdas individuais.

Observe que neste caso o MMC subestima o VaR se comparado com o MRP por não conseguir gerar muitos cenários de perdas individuais extremas. Este fato foi observado na sub-seção (2.3) quando introduzimos o MMC.

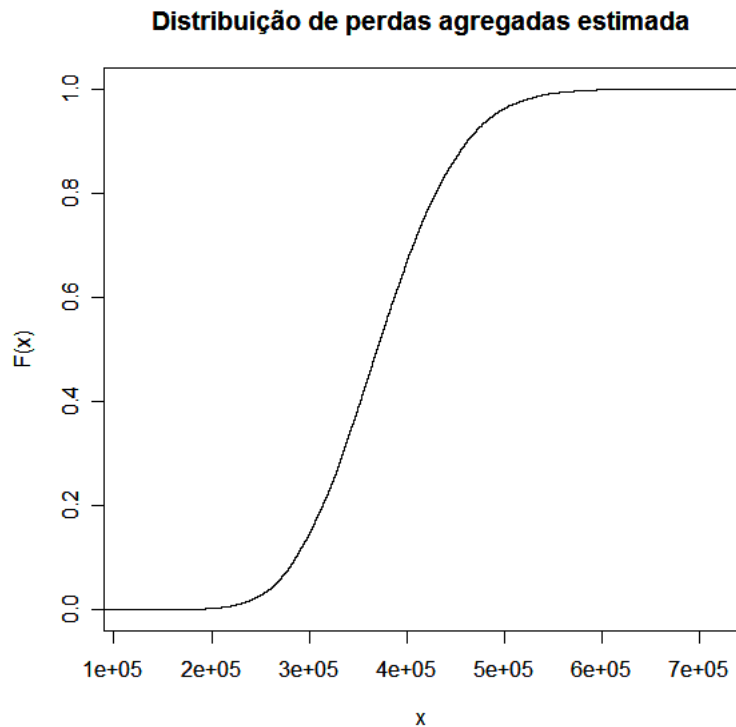


Figura 3.4: Distribuição de perdas agregadas obtidas pelo Método de Monte Carlo considerando a distribuição empírica como geradora das perdas individuais.

**Aplicação do Método POT ao conjunto de dados.**

Antes de continuarmos com as próximas abordagens devemos voltar a nossa atenção para a execução do método POT na modelagem da cauda direita da distribuição que gerou os dados de perdas individuais, bem como os seus excessos acima de um limiar. Lembramos que o método foi apresentado na sub-seção (2.7.1) e seguiremos a estratégia em três etapas lá apresentada.

Para obtermos o limiar ótimo (denotado por  $u^{opt}$ ), iremos usar o gráfico da média amostral dos excessos procurando uma região com comportamento linear e pouco irregular. Pelo gráfico da figura (3.5) optamos por escolher  $u^{opt} = 1000$ , devido ao fato de que o trecho do gráfico pouco acima da abscissa 1000 apresenta um comportamento linear crescente e pouco irregular. Abaixo dele notamos um comportamento convexo.

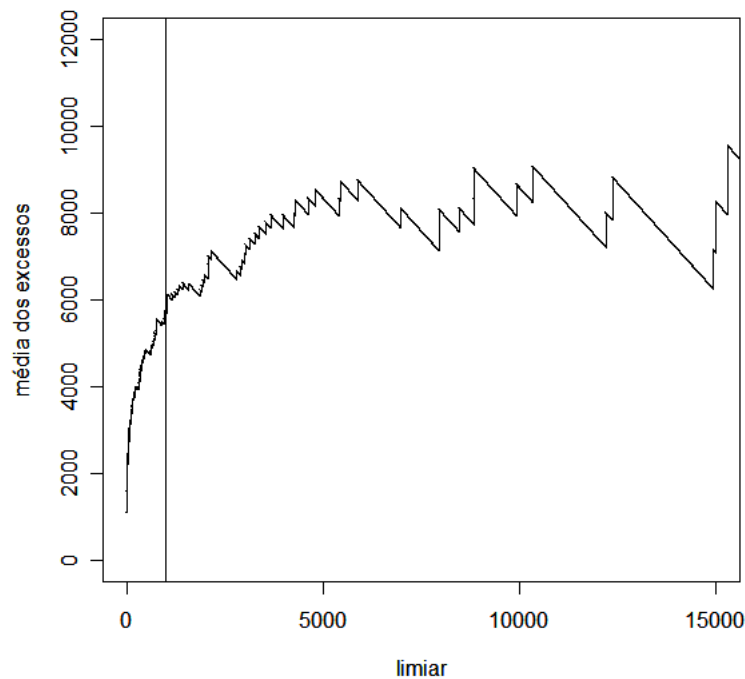


Figura 3.5: Gráfico da média amostral dos excessos para o conjunto de dados. A linha vertical indica o ponto onde a abscissa assume valor 1000.

Belistky e Moreira (2007) [1] discorrem sobre a possível escolha do limiar de valor 6000 por este

também iniciar um trecho aproximadamente linear, mas descartam esta escolha usando uma ferramenta alternativa que não será apresentada aqui. Iremos aceitar a escolha do limiar de valor 1000 observando os gráficos da figura (3.6) que evidenciam a qualidade do ajuste feito através do método da máxima verossimilhança. Além disso, tal escolha retorna uma estimativa para o parâmetro de forma  $\xi$  positivo o que é muito mais comum em caso de dados financeiros, enquanto que a escolha do limiar 6000 retorna uma estimativa negativa para  $\xi$ .

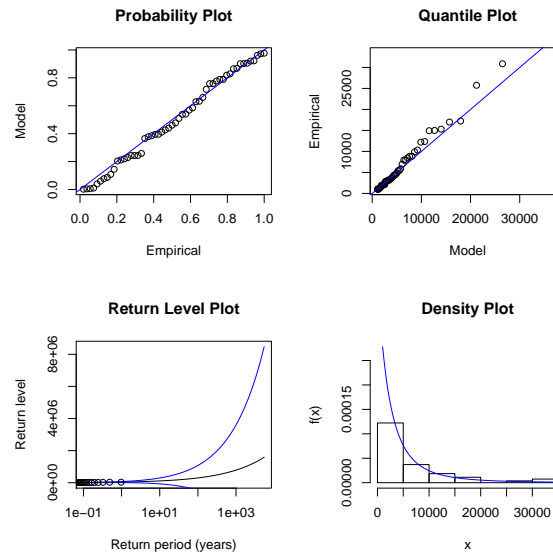


Figura 3.6: Gráficos de diagnósticos sobre a qualidade do ajuste GPD para os excessos acima do limiar 1000 obtidos via máxima verossimilhança.

A figura (3.7) mostra a função de distribuição empírica construída com base na amostra e a curva GPD dada por  $F_X^*(.)$  obtida pelo método POT e dada pela expressão (2.22). A tabela (3.4) sumariza as quantidades de interesse na aplicação do método em nosso caso particular.

$\hat{\xi}$	0.4128529 (0.2325208)
$\hat{\sigma}$	3593.708 (948.860778)
$u^{opt}$	1000
Número de excedentes	54

Tabela 3.4: Resultados obtidos pela aplicação do método POT ao conjunto de dados. As estimativas para os parâmetros foram obtidas via método da máxima verosimilhança e os valores entre parêntesis representam o erro padrão destes estimadores.

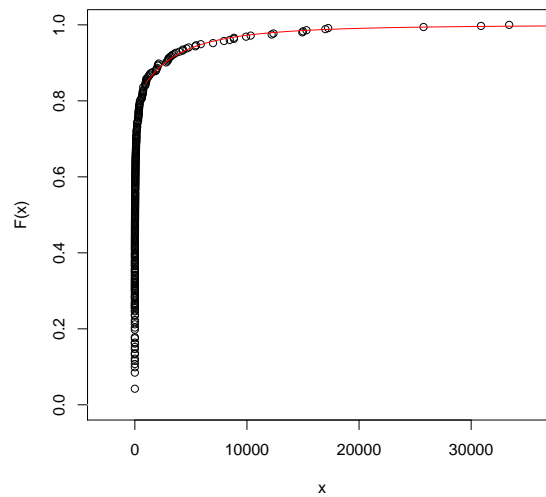


Figura 3.7: Gráfico da função de distribuição empírica com curva GPD sobreposta obtida pelo Método POT.

### Modelo: Empírico + Cauda GPD

Agora aplicamos o método recursivo de Panjer (MRP) considerando que  $N \sim Poisson(\hat{\lambda} = 354)$  e usamos a função de distribuição  $\hat{F}_X(\cdot)$  dada pela expressão (3.1) como estimativa para a função de distribuição que gerou os dados. A Figura (3.8) mostra a distribuição empírica construída com base na amostra com a curva teórica sobreposta dada por  $\hat{F}_X(\cdot)$ .



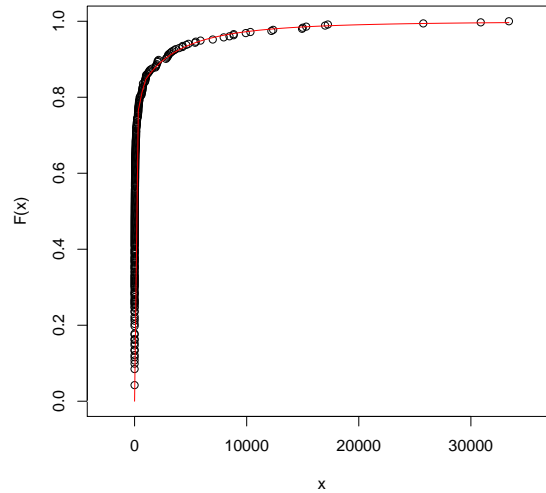


Figura 3.8: Gráfico da função de distribuição empírica com gráfico da função  $\hat{F}_X(\cdot)$  dada por (3.1) sobreposto.

Novamente usamos o método do arredondamento apresentado na sub-seção (2.4) usando três passos diferentes (denotado por  $h$ ). Diferente da primeira abordagem, o intervalo de valores discretizados foi  $[0, 800000]$ . O motivo deste intervalo ser oito vezes maior, se comparado com o intervalo adotado na primeira abordagem, se justifica pelo fato de usarmos uma GPD para modelar a cauda direita da distribuição que gerou os dados. Como a GPD possui uma cauda significativamente pesada, a discretização deve ser feita em um intervalo maior para que as probabilidades geradas somem aproximadamente 1.

As figuras (3.9) e (3.10) ilustram, respectivamente, as discretizações obtidas e os gráficos das distribuições de perdas agregadas para os diferentes valores de  $h$ .

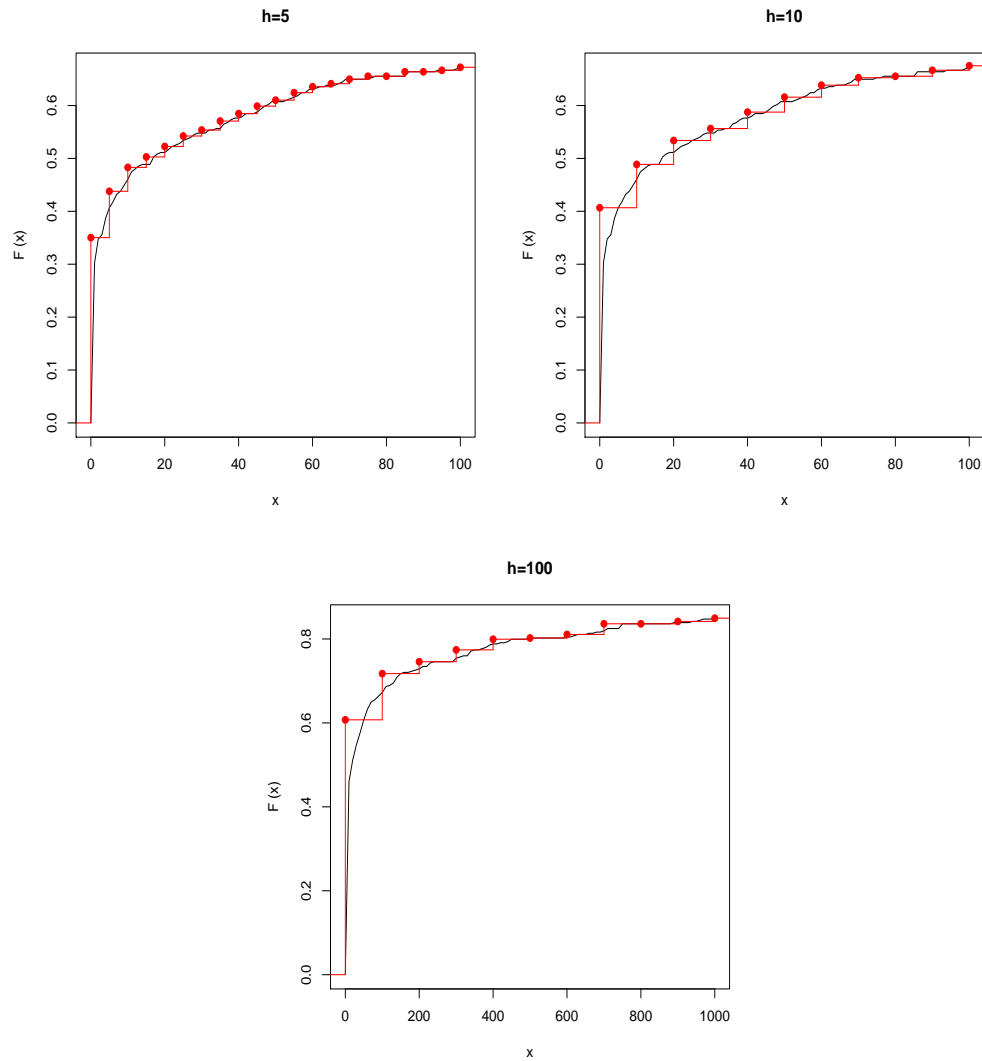


Figura 3.9: Discretização da severidade para os diferentes valores do passo  $h$  para a função  $\hat{F}_X(\cdot)$  dada por (3.1).

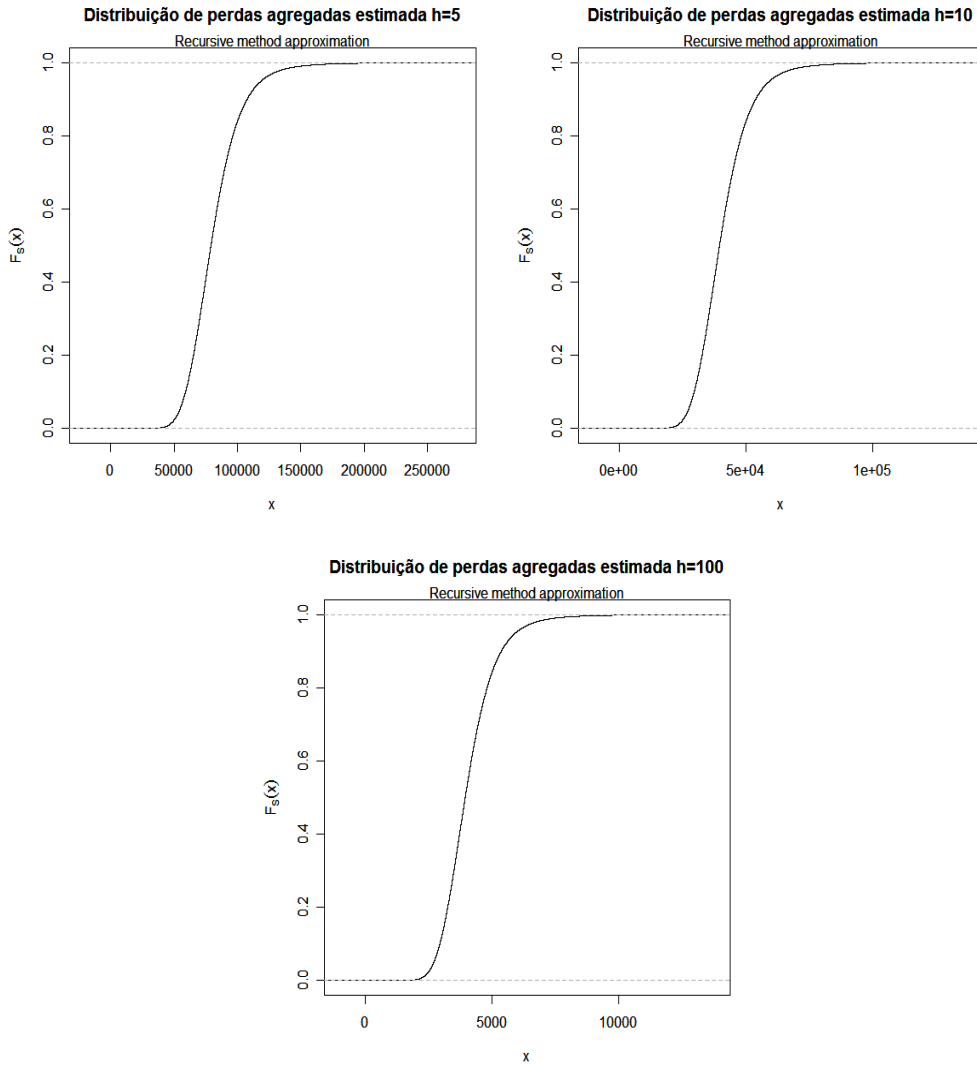


Figura 3.10: Distribuição de perdas agregadas obtidas pelo Método Recursivo de Panjer para os diferentes valores de  $h$  considerando que os dados foram gerados por  $\hat{F}_X(\cdot)$  dada por (3.1).

Os valores em risco anuais obtidos para os diferentes níveis e diferentes passos são dados na tabela (3.5)

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$h = 5$	595.420,00	753.790,00	849.220,00	1.279.565,00
$h = 10$	595.390,00	753.760,00	849.190,00	1.279.540,00
$h = 100$	593.600,00	752.000,00	857.862,50	1.277.900,00

Tabela 3.5: Valores em Risco obtidos através do Método Recursivo de Panjer para diferentes valores de  $h$  usados na discretização do modelo para a severidade e considerando que os dados foram gerados por  $\hat{F}_X(\cdot)$  dada por (3.1).

Fica claro que para níveis acima de 95% os valores obtidos para o VaR são substancialmente maiores que os obtidos na primeira abordagem.

Assim como na primeira abordagem, os resultados obtidos pelo MMC subestimam o VaR se comparado com o MRP como podemos observar na tabela (3.6) e comparar com a tabela (3.5).

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$VaR_\alpha$	589.161,00	742.466,00	847.400,00	1.257.322,50

Tabela 3.6: Valores em Risco obtidos através do Método de Monte Carlo considerando que os dados foram gerados por  $\hat{F}_X(\cdot)$  dada por (3.1).

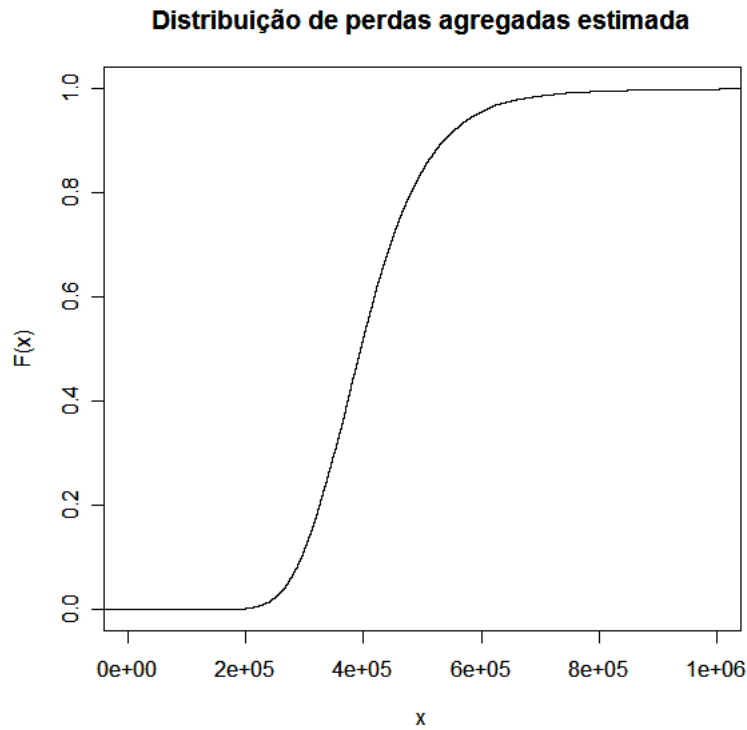


Figura 3.11: Distribuição de perdas agregadas obtidas pelo Método de Monte Carlo considerando que os dados foram gerados por  $\hat{F}_X(\cdot)$  dada por (3.1).

### Convolução

Como dito anteriormente, o que nos motivou a usar a aproximação normal para a distribuição de perda agregada abaixo de um limiar é que uma grande parte da amostra fica abaixo deste limiar. Em nosso caso particular temos que 300 (das 354) observações ficam abaixo do limiar escolhido ao se aplicar o método POT.

Usando uma versão amostral das expressões (2.4) e (2.5), obtemos a média e a variância da distribuição de perda agregada que serão usadas na aproximação normal. Logo

$$\hat{E}(S^A) = 26.698, 43$$

$$\widehat{Var}(S^A) = 12.895.102,00$$

Tomando  $\hat{\lambda}_A = 300$  e usando a distribuição empírica como modelo para as severidades, aplicamos o MRP para os dados abaixo do limiar  $u = 1000$  e comparamos a função de distribuição obtida pelo método com a função de distribuição da normal com média 26.698,43 e variância 12.895.102,00 calculados acima. A figura (3.12) sugere que a distribuição normal é uma boa aproximação.

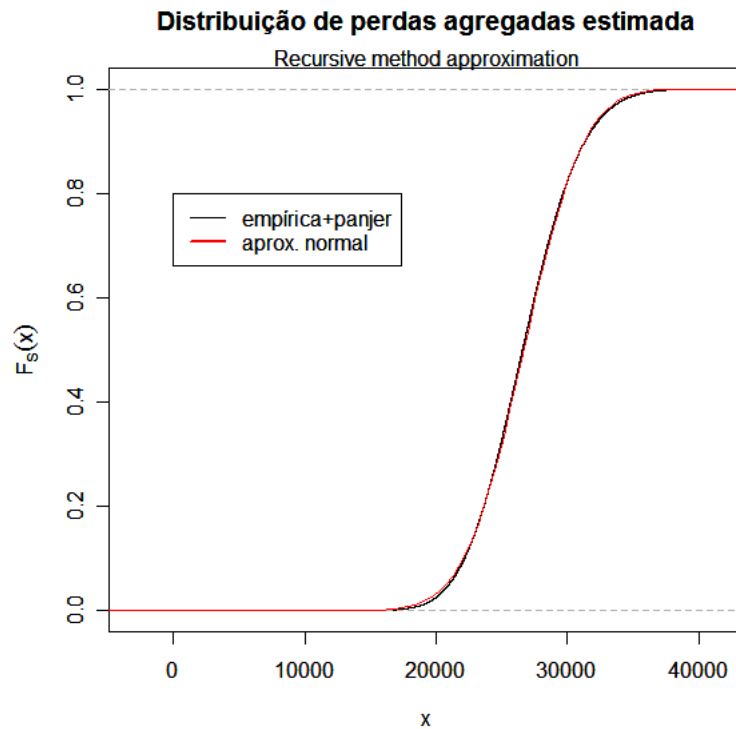


Figura 3.12: Distribuição de perdas agregadas obtidas pelo Método de Panjer para os valores abaixo do limiar  $u = 1000$  com curva normal sobreposta.

A média calculada através do MRP é 26.654,12 ficando bem próxima do valor 26.698,43 calculado anteriormente.

Analogamente, executamos o MMC sob as mesmas condições e comparamos os resultados obti-

dos com a distribuição normal. As figuras (3.13) e (3.14) mostram, respectivamente, as comparações entre as funções de distribuição e densidade.

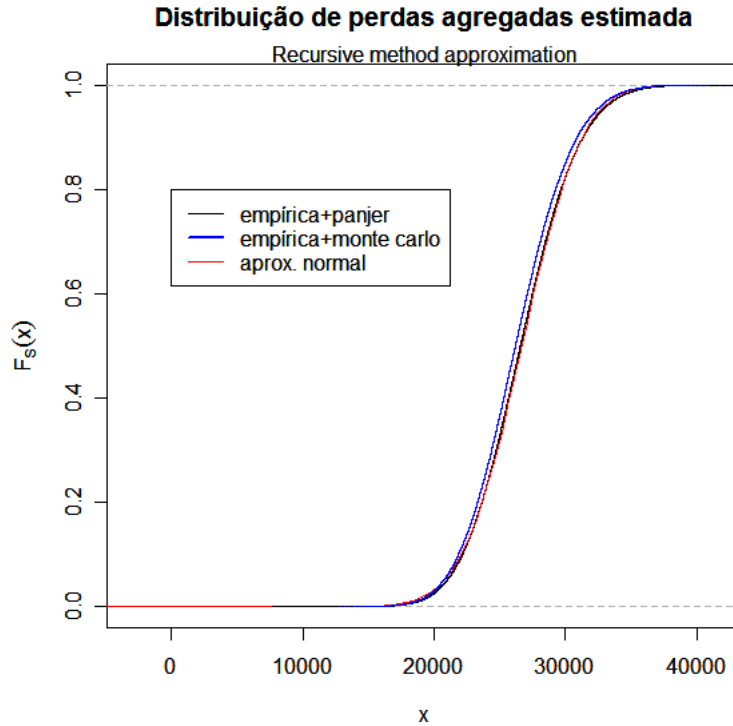


Figura 3.13: Distribuição de perdas agregadas obtidas pelo Método de Monte Carlo para os valores abaixo do limiar  $u = 1000$  com curva normal e curva da função de distribuição obtida via MRP sobrepostos.

A aproximação normal não parece ser tão boa se comparada com o resultado obtido pelo MMC. Note que a média dos valores gerados pelo MMC é dada por 26.320,25 ficando mais distante do valor 26.698,43. Isto se deve ao fato já mencionado de que o MMC subestima os quantis da distribuição de perda agregada por não gerar cenários extremos de perdas individuais suficientes.

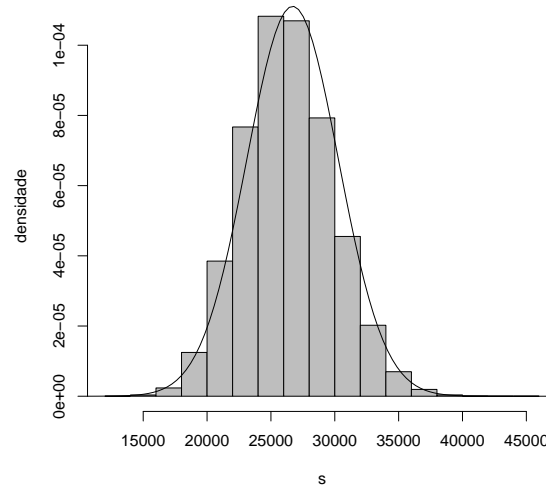


Figura 3.14: Histograma da amostra gerada pelo Método de Monte Carlo com curva normal aproximada sobreposta.

A próxima etapa consiste em se aplicar o MRP aos dados acima do limiar. Consideramos assim que  $\hat{\lambda}_D = 54$  e o modelo para as severidades é dado pela GPD que modela os excessos acima do limiar  $u = 1000$ , isto é, a GPD  $G_{\hat{\xi}, \hat{\sigma}, 0}(\cdot)$  da expressão (2.22).

A figura (3.15) ilustra o ajuste feito aos excessos acima do limiar  $u = 1000$ . Os estimadores para  $\xi$  e  $\sigma$  são os mesmo da tabela (3.4).



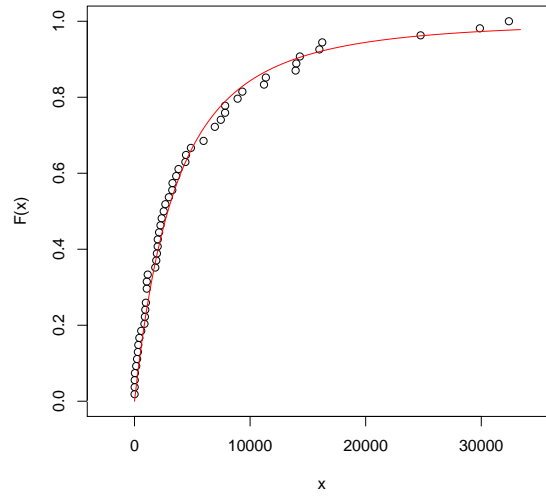


Figura 3.15: Ajuste GPD feito para os excessos acima do limiar  $u = 1000$ . A linha contínua representa a curva paramétrica da função  $G_{\hat{\xi}, \hat{\sigma}, 0}(\cdot)$  e os pontos a função de distribuição empírica obtida com base na amostra de excessos.

Esta GPD obtida que modela os excessos possui a cauda mais pesada que a GPD realocada e reescalada do método POT. Logo para aplicarmos o MRP, será necessário um intervalo maior para a discretização. Neste caso a discretização foi feita no intervalo  $[0, 10000000]$  e o salto utilizado foi  $h = 100$ . A figura (3.16) mostra um recorte da GPD que modela os excessos e podemos notar que mesmo com um salto relativamente grande a discretização ainda detalha bem a distribuição original.

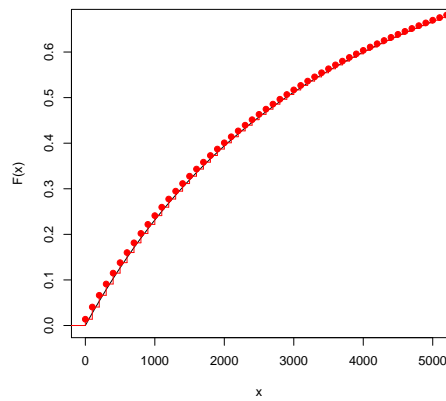


Figura 3.16: Discretização feita para a GPD  $G_{\hat{\xi},\hat{\sigma},0}(\cdot)$  que modela os excessos.

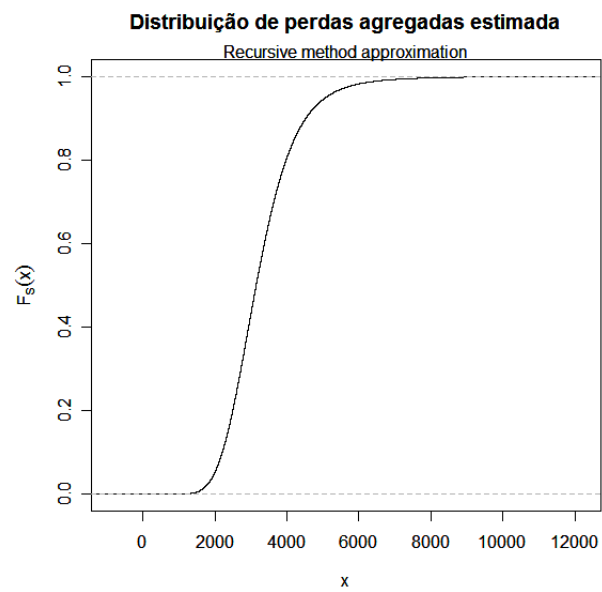


Figura 3.17: Distribuição de perdas agregadas obtidas para valores acima do limiar  $u = 1000$ .

Agora estamos aptos a calcular a convolução das duas distribuições obtidas, isto é, obter uma estimativa para  $F_S(\cdot)$  onde  $S = S^A + S^D$  e assim, estimar o valor em risco para diferentes níveis.

Os resultados obtidos pelo MRP são dados nas tabelas (3.7) e (3.8).

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$VaR_\alpha$	537.426,00	707.447,00	827.445,00	1.164.303,00

Tabela 3.7: Valores em Risco obtidos do modelo para  $S = S^A + S^B$  considerando a distribuição empírica como modelo para as severidades abaixo do limiar  $u = 1000$ . Neste caso obtidos via MRP.

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$VaR_\alpha$	535.841,00	694.508,00	790.690,00	1.161.009,00

Tabela 3.8: Valores em Risco obtidos via Método de Panjer do modelo para  $S = S^A + S^B$  considerando o modelo normal para  $S^A$ .

As tabelas (3.9) e (3.10) nos fornecem os respectivos resultados obtidos via MMC.

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$VaR_\alpha$	542.008,00	689.526,00	779.012,00	1.120.673,00

Tabela 3.9: Valores em Risco obtidos do modelo para  $S = S^A + S^B$  considerando a distribuição empírica como modelo para as severidades abaixo do limiar  $u = 1000$ . Neste caso obtidos via Método de Monte Carlo.

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$VaR_\alpha$	534.950,00	687.297,00	772.938,00	1.146.161,00

Tabela 3.10: Valores em Risco obtidos via Método de Monte Carlo do modelo para  $S = S^A + S^B$  considerando o modelo normal para  $S^A$ .

## Sumário de Resultados

A tabela (3.11) sumariza os resultados obtidos nas abordagens apresentadas. Para o caso onde diferentes saltos de discretização foram usados, iremos apenas explicitar os resultados obtidos ao

se utilizar o menor salto para simplificar a tabela. Lembrando que os resultados obtidos para os saltos maiores não alteraram significativamente os resultados dentro de um contexto de aplicação em Risco Operacional.

Abordagem	$VaR_\alpha$			
	95,0%	99,0%	99,5%	99,9%
MRP - empírica	513.630,00	572.590,00	594.970,00	642.445,00
MMC - empírica	490.703,30	548.450,00	569.873,60	616.392,70
MRP - empírica c/ cauda GPD	595.420,00	753.790,00	849.220,00	1.279.565,00
MMC - empírica c/ cauda GPD	589.161,00	742.466,00	847.400,00	1.257.322,50
MRP - convolução empírica+GPD	537.426,00	707.447,00	827.445,00	1.164.303,00
MRP - convolução normal+GPD	535.841,00	694.508,00	790.690,00	1.161.009,00
MMC - convolução empírica+GPD	542.008,00	689.526,00	779.012,00	1.120.673,00
MMC - convolução normal+GPD	534.950,00	687.297,00	772.938,00	1.146.161,00

Tabela 3.11: Valores em Risco obtidos pelas abordagens em estudo.

### 3.3 Considerações Finais

A tabela (3.12) abaixo fornece os valores em risco da distribuição de perda agregada obtidos via Método de Panjer e considerando apenas os valores de perdas individuais em excesso acima do limiar  $u = 1000$ , isto é, os valores em risco da distribuição da v.a  $S^D$ .

Nível ( $\alpha$ )	95,0%	99,0%	99,5%	99,9%
$VaR_\alpha$	509.100,00	667.700,00	763.900,00	1.129.000,00

Tabela 3.12: Valores em Risco obtidos via Método de Panjer apenas para  $S^D$ .

Obviamente, os resultados obtidos pelas duas primeiras abordagens diferem significativamente dos demais. Considerar o modelo empírico para as perdas individuais implica em simplesmente ignorar que possíveis perdas, bem mais drásticas que as estimadas, podem ocorrer já que pouca atenção é dada aos eventos presentes na cauda da distribuição que gerou as perdas individuais. No entanto, a soma de todas as perdas individuais do ano estudado, dada na tabela (3.1), é R\$ 388.396,80 sendo que a reserva estipulada considerando o modelo empírico foi R\$ 642.445,00 ( $VaR_{99,9\%}$ ) obtido pelo MRP, o que poderia indicar que este valor seja suficiente para cobrir as perdas do próximo ano caso não ocorra nenhum evento incomum.

A melhoria da cauda da distribuição empírica, através do método POT, nos fornece uma estimativa muito maior para a reserva de capital para níveis maiores, e certamente nos fornece uma proteção maior frente a eventos adversos que podem ocorrer.

Da mesma forma, as abordagens que consideraram a convolução de distribuições acima e abaixo de um limiar determinado (seguindo os passos descritos na seção (3.2)), também retorna valores maiores para a reserva de capital e também fornece uma maior segurança contra eventos extremos.

Nas abordagens onde o método POT foi considerado, isto é, onde existe um tratamento melhor para a cauda direita da distribuição que gerou os dados de perdas individuais, é possível notar a influência dos valores abaixo do limiar determinado, já que nestes casos, a estimativa da reserva de capital foi maior que aquela onde simplesmente obtemos a distribuição de perdas agregadas considerando as perdas em excesso acima de um limiar. Podemos visualizar esta influência na tabela (3.12) se comparada com a tabela (3.11). Isto indica que os valores moderados, isto é, os valores que se encontram abaixo do comportamento de cauda da distribuição de severidade, influencia de forma significativa os valores da cauda da distribuição de perda agregada neste caso particular.

Recorde que na seção (2.2) obtemos um limite superior para o erro ao se assumir a normalidade da distribuição de perdas agregadas (através do resultado 2). Em nosso caso particular, a aproximação normal foi feita abaixo do limiar  $u = 1000$  dado pelo método POT. Nestas condições, o erro cometido calculado foi de aproximadamente 0,1767, um valor alto apesar da figura (3.12) sugerir que o ajuste pela normal é razoável. Concluímos com isso que o tamanho da amostra de valores que estão abaixo do limiar  $u = 1000$  não é suficiente para usarmos a estimativa obtida pela expressão (2.13), além de que a grandeza  $\varepsilon_n = n^{1/3}$  pode não ser a mais adequada, sendo necessário a obtenção de valor melhor.

### 3.4 Sugestões para Pesquisas Futuras

O presente trabalho viabiliza a pesquisa dos seguintes assuntos:

- O estudo comparativo de resultados obtidos pelo Método Recursivo de Panjer e o Método da *Transformada Rápida de Fourier* (detalhes sobre este método podem ser encontrados em Klugman et al (2004) [12]). Além do estudo comparativo, um estudo sobre a eficiência das duas técnicas em termos computacionais pode ser considerada;
- Como no item anterior, um mesmo estudo pode ser feito comparando resultados obtidos pelo

Método Recursivo de Panjer ou Transformada Rápida de Fourier e um método não-paramétrico que vem sendo atualmente empregado em atuária conhecido como *Triângulo de Run-off* ou Método Triangular (veja Schmidt e Zocher (2008) [21]);

- Finalmente, muito ainda pode ser feito acerca de resultados assintóticos da distribuição de perdas agregadas sobre diferentes suposições, facilitando a estimação de quantis extremos. O resultado apresentado na seção (2.2) deste trabalho pode ser um ponto de partida para futuros resultados mais apurados.



## Apêndice A

### Resultados Analíticos para a Perda Agregada

Apresentaremos aqui, alguns exemplos de distribuições de perda agregada que admitem forma analítica para alguma escolha adequada de modelos para frequência e severidade.

**Exemplo 1:** (Composição Geométrica-Exponencial) Suponha que  $X_1, X_2, \dots$  são v.a.s i.i.d com distribuição comum exponencial com média  $\theta$  e função geradora de momentos  $M_X(t) = (1 - \theta t)^{-1}$ . Suponha que  $N$  tem distribuição geométrica com parâmetro  $\beta$  e função geradora de probabilidade  $P_N(t) = [1 - \beta(t - 1)]^{-1}$ . Vamos determinar a distribuição de  $S = X_1 + X_2 + \dots + X_N$ .

Segundo Klugman et al (2004) [12] a função geradora de momentos de  $S$ , denotada por  $M_S(\cdot)$  pode ser dada por:  $M_S(t) = P_N(M_X(t))$ , assim

$$\begin{aligned} M_S(t) &= P_N(M_X(t)) = P_N[(1 - \theta t)^{-1}] = \{1 - \beta[(1 - \theta t)^{-1} - 1]\} \\ &= \left[1 - \beta \left(\frac{1}{1 - \theta t} - 1\right)\right]^{-1} = \left(1 - \beta \frac{1 - 1 + \theta t}{1 - \theta t}\right)^{-1} \\ &= \frac{1 + \beta}{1 + \beta} \frac{1 - \theta t}{1 - \theta t(1 + \beta)} = \frac{1 - \theta(1 + \beta)t + \beta}{(1 + \beta)(1 - \theta(1 + \beta)t)} \\ &= \frac{1}{1 + \beta} + \frac{\beta}{1 + \beta} [1 - \theta(1 + \beta)t]^{-1} = M_I(t) + M_{II}(t) \end{aligned}$$

onde  $M_I(\cdot)$  é a função geradora de momentos de uma distribuição degenerada com massa concentrada em zero e  $M_{II}(\cdot)$  é a função geradora de momentos de uma distribuição exponencial com média  $\theta(1 + \beta)$ .



Para  $N \sim Geom(p = 0.5)$  e  $X \sim Exp(1/\theta = 0.1)$  temos o seguinte gráfico:

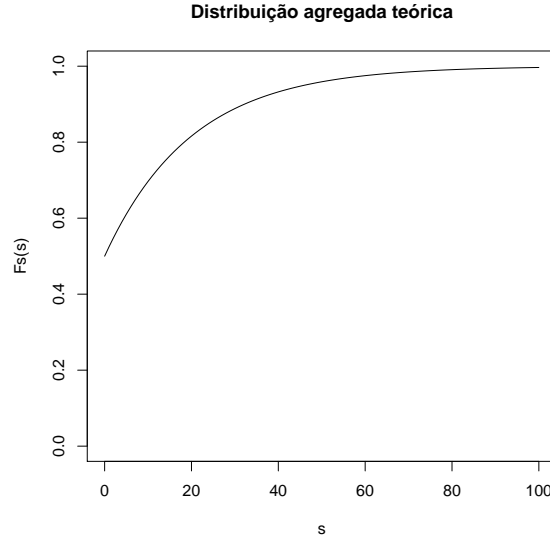


Figura A.1: Distribuição de perda agregada teórica do exemplo 1.

**Exemplo 2:** (Severidade Exponencial) Para qualquer v.a. contadora  $N$  com  $X_1, X_2, \dots$  i.i.d exponencial com média  $\theta$ , temos que a função de distribuição de perda agregada  $F_S(\cdot)$  de  $S$  será dada por:

$$F_S(x) = 1 - e^{-x/\theta} \sum_{j=0}^{\infty} \bar{P}_j \frac{(x/\theta)^j}{j!}, \quad x \geq 0$$

onde  $\bar{P}_j = \sum_{n=j+1}^{\infty} P(N = n)$  para  $j = 0, 1, \dots$

**Exemplo 3:** (Composição Binomial Negativa-Exponencial) Suponha que  $X_1, X_2, \dots$  são v.a.s i.i.d com distribuição comum exponencial com média  $\theta$  e que  $N$  tem distribuição Binomial Negativa com parâmetro  $r$  inteiro. A função de distribuição de perda agregada  $F_S(\cdot)$  de  $S$  será dada por:

$$F_S(x) = 1 - \left[ \sum_{n=1}^r C_{r,n} \left( \frac{\beta}{1+\beta} \right)^n \left( \frac{1}{1+\beta} \right)^{r-n} \right] \left[ \sum_{j=0}^{n-1} \frac{[x\theta^{-1}(1+\beta)^{-1}] e^{-x\theta^{-1}(1+\beta)^{-1}}}{j!} \right]$$

onde

$$C_{r,n} = \frac{r!}{(r-n)!n!}$$



## Referências Bibliográficas

- [1] BELITSKY, V., MOREIRA, F. M. *Emprego do método "Peaks-over-Threshold" na estimação de risco; uma exposição abrangente, detalhada mas simples (A thorough yet simple exposition of the Peaks-over-threshold method and its employment for risk estimation)* - Short Course. Maresias, March 25-30, 2007.
- [2] BEIRLANT, J., GOEGEBEUR, Y., SEGERS, J., TEUGELS, J. *Statistics of Extremes - Theory and Applications*. Chichester - John Wiley & Sons, 2004.
- [3] BLUM, J. R., HANSON, D. L., ROSEMBLATT, J. I. *On the Central Limit Theorem for the Sum of a Random Number of Independent Random Variables*. - Z. Wahrscheinlichkeitstheorie verm. Gebiete, 1, 1963, pp. 389-393.
- [4] CALLAERT, H., JANSSEN, P. *A Note on the Convergence Rate of Random Sums*. - Rev. Roum. Math. Pures et Appl., 28, 1983, pp. 147-151.
- [5] CHERNOBAI, A. S., FABOZZI, F. J., RACHEV, S. T. *Operational Risk - A Guide to Basel II Capital Requirements, Models, and Analysis*. - Hoboken, John Wiley & Sons, 2007.
- [6] CRUZ, M. G. *Modeling Measuring and Hedging Operational Risk*. - Chichester, John Wiley & Sons, 2002.
- [7] EMBRECHTS, P., KLUPPELBERG, C., MICKOSCH, T. *Modeling Extremal Events for Insurance and Finance*. - Berlin, Springer, 1997.
- [8] FELLER, W. *An Introduction to Probability Theory and its Applications*. - vol 1. New York, John Wiley & Sons, 1950.
- [9] HAIGHT, F. A. *Publications in Operational Research No 11 - Handbook of the Poisson Distribution* - New York, John Wiley & Sons, 1967.
- [10] JAMES, B. R. *Probabilidade: Um Curso em Nível Intermediário - 3ª edição*, Rio de Janeiro, Projeto Euclides, 2004.

- [11] KLÄVER, H. *Ein Berry-Esseen-Satz für Zufallssummen.* - Diploma Thesis, Univ. Münster, 2002.
- [12] KLUGMAN, S. A., PANJER, H. H., WILLMOT, G. E. *Loss Models - From Data to Decisions.* - New Jersey, John Wiley & Sons, 2004.
- [13] KRAJKA, A., RYCHLIK, Z. *The Order of Approximation in the Central Limit Theorem for Random Summation.* - Acta math. Sci. Hung., 51 (1-2), 1988, pp. 109-115.
- [14] LANDERS, D., ROGGE, L. *The Exact Approximation Order in the Central-Limit-Theorem for Random Summation.* - Z. Wahrscheinlichkeitstheorie verm. Gebiete, 36, 1976, pp. 269-283.
- [15] PANJER, H. H. *Recursive Evaluation of a Family of Compound Distributions.* - Astin Bulletin, 12, 1981, pp. 22-26.
- [16] PETROV, V. V. *Limit Theorems of Probability Theory - Sequences of Independent Random Variables* - New York, Oxford, 1995.
- [17] PETROV, V. V. *Sums of Independent Random Variables.* - New Jersey, Springer-Verlag, 1975.
- [18] PICKANDS, J. *Statistical Inference Using Extreme Order Statistics.* - Ann. Statist., 3, 1975, pp. 119-131.
- [19] RÉNYI, A. *On the Central Limit Theorem for the Sum of a Random Number of Independent Random Variables.* - Acta math. Sci. Hung., 11, 1960, pp. 97-102.
- [20] REISS, R. D., THOMAS, M. *Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields* - 2nd edition, Birkhäuser Verlag, 2001.
- [21] SCHMIDT, K. D., ZOCHER M. *The Bornhuetter-Ferguson Principle* - Casualty Actuarial Society, 2, 2008, pp. 85-110.
- [22] TYSLIAK, W. *Gleichmäßige Berry-Esseen-Abschätzungen* - Ph. Thesis, Univ. Wuppertal, 1983.
- [23] VAN BEEK, P. *An Application of Fourier Methods to the Problem of Sharpening the Berry-Esseen Inequality.* - Z. Wahrscheinlichkeitstheorie verm. Gebiete, 23, 1972, pp. 187-196. [3](#), [43](#), [51](#) [1](#), [26](#) [8](#), [44](#) [1](#), [11](#), [26](#), [66](#), [69](#) [4](#), [12](#) [22](#) [19](#) [41](#) [44](#) [67](#) [12](#) [12](#)