Asymptotic Analysis of Ruin Probabilities in Insurance Risk Management

Heavy-tailed Distribution

保险风险管理中的破产渐近分析

——重尾分布

杨 洋 王开永 著

科学出版社

保险风险管理中的破产渐近分析

——重尾分布

杨　洋　王开永　著

（国家自然科学基金资助项目）

科　学　出　版　社

北　京

内 容 简 介

本书以重大灾害风险相关理论为依据，以全球重大灾害风险发展形势及损失概算为着眼点，从减少重大灾害风险对保险公司造成的损失、维护保险公司正常运营的角度出发，研究重大灾害下保险公司的风险模型。本书针对经典的更新风险模型，以及一些与金融保险业息息相关的复杂化了的非经典模型，研究相应破产概率的渐近性问题，其中既包括当初始资本趋于无穷时各种风险模型下破产概率的渐近结果，也有经典的更新风险模型中当初始资本固定时有限时破产概率的渐近性结果。重尾分布是保险领域刻画索赔额的重要模型，可为保险公司对其业务进行风险评估和科学决策提供参考。本书详细介绍了各种重尾分布的定义、性质及分类，并以此为基础，研究了各种相依结构的重尾风险模型中破产概率的渐近性问题。

本书适合保险公司的研究和管理人员、从事概率统计和风险理论研究的科研人员，以及高等院校相关专业的师生参考阅读。

图书在版编目(CIP)数据

保险风险管理中的破产渐近分析：重尾分布/杨洋，王开永著.—北京：科学出版社，2013

ISBN 978-7-03-034301-7

Ⅰ.①保… Ⅱ.①杨… ②王… Ⅲ.①保险企业-风险管理-研究 Ⅳ.F840.32

中国版本图书馆 CIP 数据核字（2012）第 094866 号

责任编辑：王 钰 杨 阳 / 责任校对：王万红
责任印制：吕春珉 / 封面设计：耕者设计工作室

科学出版社 出版
北京东黄城根北街16号
邮政编码：100717
http://www.sciencep.com
双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2013 年 3 月第 一 版 开本：B5（720×1000）
2013 年 3 月第一次印刷 印张：20
字数：392 000

定价：70.00 元

（如有印装质量问题，我社负责调换〈双青〉）
销售部电话 010-62135157 编辑部电话 010-62135517-2038

前　　言

"… was it heavy? Did it achieve total heaviosity?"

——Alvie (Woody Allen) to Annie (Diane Keaton) in Annie Hall, 1977.

重尾分析是极值理论中的分支之一，用于研究由重大事件造成的极端现象。它既包括概率模型，也需要统计分析，其数学工具主要建立在分布理论正则变换、渐近理论和概率测度极限理论上。

近十年来，重尾分析广泛应用于保险风险管理中，其最主要的原因是，大量极端事件的发生。例如，2001 年的"9.11"恐怖袭击、2004 年的印度洋海啸、2005 年的 Katrina 飓风、2008 年的汶川大地震、2010 年的海地大地震、2011 年的日本大地震，特别是近年来的金融危机等。这些极端事件往往带来重尾索赔额等。数据显示，经典的轻尾分布刻画的经验公式存在着明显的偏差，重尾分析获得的保险公司破产概率的渐近结果更为合理。本书以重大灾害风险相关理论为依据，主要研究重尾风险模型，包括经典的更新风险模型和一些与保险业息息相关的复杂化的非经典模型，以及相应破产概率的渐近性问题。在一些非经典的风险模型中，作为主要对象的索赔额过程，它们之间不必是相互独立的，如可以是某种负相依关系或其他地相依关系；相应地，索赔间隔时间过程也可以不必相互独立；但我们仍然要求索赔额过程与索赔间隔时间过程彼此是相互独立的。尽管本书研究了各种经典与非经典的风险模型，但它们都有两点共同之处：其一，每个索赔额到达时，造成保险公司的索赔额或者净损失分布都是重尾的，特别是次指数或带有控制变换尾的。在保险业，特别是财产保险业中，许多重大的风险都是由一个（或一些）大额索赔造成，其分布往往是重尾而非轻尾的。因此，本书将重尾风险模型作为主要的研究对象。其二，各种破产概率渐近性的研究，与极限理论中的大偏差理论，随机游动理论和分布理论等紧密相关。因此，本书将以此作为重尾风险理论研究的主要工具。

本书的结构如下： 第 1 章给出本书的研究背景和经典的风险模型，包括本书常用的一些风险量，并给出 Cramér-Lundberg 风险模型中轻尾索赔额时破产概率的估计；第 2 章详细介绍重尾分布的特征和常用的重尾分布子族，如正则变换、

长尾分布、控制变换尾分布、次指数分布等；第 3 章和第 4 章分别介绍带利率和不带利率的多种重尾风险模型，包括各种相依的风险模型、复合更新风险模型、带红利干扰的风险模型等，研究这些经典或者非经典风险模型中有限时破产概率和无限时破产概率的渐近性以及一致渐近性；第 5 章以随机加权和尾概率的研究为基础，讨论带保险风险和金融风险的离散时风险模型；第 6 章简单介绍大偏差理论，包括精致大偏差和粗略大偏差，并且研究独立更新风险模型中，带有固定初始资本的有限时破产概率的渐近性。

本书的第 1~4 章和第 6 章由杨洋编写，第 5 章由王开永编写。第 3 章和第 6 章的内容包含我的博士毕业论文，在此对我的博士生导师苏州大学王岳宝教授致以深深的感谢，感谢他悉心的指导和热情的鼓励；本书的第 3 和第 4 章是我在立陶宛 Vilnius University 访学期间完成，这两章内容得到了 Remigijus Leipus 教授和 Jonas Šiaulys 教授的讨论和帮助，在此对他们表示衷心的感谢；同时我还要深深的感谢南京审计学院方习年教授和东南大学林金官教授，他们在很多方面给予了我关心帮助和大力支持；最后感谢国家自然科学基金（11001052）、江苏省自然科学基金（BK2010480）、江苏省青蓝工程和南京审计学院学术专著出版资助项目对本书的资助。

由于编者水平有限，书中纰漏在所难免，恳求广大读者批评指正。

杨 洋

2012 年 2 月

目　　录

前言
第 1 章　引论 ······ 1
1.1　风险过程与破产概率 ······ 3
1.2　索赔额分布 ······ 7
1.2.1　轻尾分布 ······ 8
1.2.2　重尾分布 ······ 9
1.2.3　常见的重尾分布子族及其性质 ······ 10
1.3　索赔到达过程 ······ 14
1.4　Cramér-Lundberg 估计 ······ 18
第 2 章　重尾分布 ······ 27
2.1　重尾分布族及其性质 ······ 27
2.2　正则变换 ······ 30
2.3　长尾函数与长尾分布 ······ 34
2.4　次指数分布 ······ 42
2.5　控制变换尾分布与 $\mathscr{O}$ 正则函数 ······ 57
2.6　重尾分布间的控制关系 ······ 61
第 3 章　不带利率的重尾风险模型 ······ 73
3.1　Veraverbeke 定理 ······ 73
3.2　独立增量随机游动极大值尾概率的估计 ······ 80
3.3　两类相依风险模型的无限时破产概率 ······ 88
3.3.1　带有被调节索赔额过程破产概率的渐近性 ······ 91
3.3.2　带有负上象限相依索赔额过程破产概率的渐近性 ······ 96
3.4　带有次指数索赔额独立风险模型的有限时破产概率 ······ 98
3.5　带有负下象限相依索赔时间间隔风险模型的有限时破产概率 ······ 104
3.6　红利干扰模型中的无限时破产概率 ······ 117
3.6.1　随机游动极大值尾概率的渐近性 ······ 117
3.6.2　负相协更新门限超出概率的渐近性 ······ 123
3.6.3　红利干扰风险模型中破产概率的渐近性 ······ 127

第 4 章　带利率的重尾风险模型 ······ 131
4.1　独立风险模型中的有限时破产概率 ······ 132
4.2　负相依风险模型中的有限时破产概率 ······ 140
4.3　负相依复合更新风险模型中的有限时破产概率 ······ 157
4.3.1　控制变换情形下随机和尾概率的估计 ······ 158
4.3.2　Gumbel 最大值吸引场情形下随机和尾概率的估计 ······ 163
4.3.3　相依复合更新风险模型中破产概率的渐近性 ······ 168
4.4　宽象限相依更新风险模型中有限时破产概率的一致渐近性 ······ 170
第 5 章　随机加权和 ······ 191
5.1　独立随机加权和 ······ 191
5.1.1　有界权重情形 ······ 191
5.1.2　一般权重情形 ······ 209
5.2　相依随机加权和 ······ 222
5.2.1　上尾独立情形 ······ 222
5.2.2　准渐近独立情形 ······ 233
第 6 章　大偏差理论 ······ 247
6.1　大偏差理论简介及回顾 ······ 248
6.2　独立次指数随机变量差的精致大偏差及应用 ······ 251
6.2.1　精致大偏差结果 ······ 252
6.2.2　随机游动首次上穿时的尾渐近性 ······ 264
6.2.3　固定的初始资本下有限时破产概率的渐近性 ······ 271
6.3　带控制变换尾实值随机变量和的精致大偏差 ······ 274
6.3.1　负相协随机变量的基本更新定理 ······ 274
6.3.2　控制变换尾分布族随机变量和的精致大偏差 I ······ 279
6.3.3　控制变换尾分布族随机变量和的精致大偏差 II ······ 283
6.4　粗略大偏差 ······ 291
6.4.1　非负随机变量和的粗略大偏差 ······ 291
6.4.2　实值随机变量和的粗略大偏差 ······ 294
数学符号和缩写 ······ 301
主要参考文献 ······ 303

第1章　引　　论

本书主要研究保险风险管理中破产概率的渐近性态, 同时介绍各种不同的非经典相依重尾风险模型。为什么要关注重尾风险模型? 这是因为保险公司的破产往往不是因为大量日积月累的小额索赔, 而是某些大额索赔, 甚至是某个创纪录的保险损失而造成的。为什么这么说? 我们先看一组数据。如 Sigma (1996) 所述, 在 1995 年全球约 1500 亿美元的损失中, 有一半以上是由 Kobe 地震造成的, 重大的自然灾害直接造成了 124 亿美元的保险损失, 其中一半以上是由四次独立的 10 亿美元以上损失的灾难带来的。我们称这类事件为极端事件 (extremal event)。无论其确切定义是什么, 我们至少可以认为表 1.1 [参见 Sigma (1996)] 包含这些所谓的极端事件, 而且近年来这类事件也频繁发生。例如, 2001 年的"9.11"恐怖袭击、2004 年的印度洋海啸、2005 年的 Katrina 飓风、2008 年的汶川大地震、2010 年的海地大地震、2011 年的日本大地震以及近年来的金融危机等。自 20 世纪 70 年代以来, 据有关统计, 全球范围内自然重大灾害和人为重大灾害的发生频率和造成损失程度有明显上升趋势。据瑞士再保险公司对自然重大灾害的数据统计结果显示, 20 世纪 70 年代全球平均每年发生 38 次, 80 年代平均每年发生 65 次, 90 年代每年的发生频率上升到 128 次, 2000 年以来平均每年重大灾害的发生次数高达 135 次。联合国的统计资料表明, 20 世纪以来全世界 54 个最严重的自然灾害事件中有 8 个发生在我国。据《中华人民共和国跨世纪减灾规划 (1998 — 2010 年)》中指出, 我国是世界上灾害频发且损失严重的少数国家之一。由于我国领土广阔, 各地气候错综复杂, 地理条件多样, 导致我国重大灾害风险分布地域广、种类多、发生频率高。例如, 洪水、地震、旱灾、台风等重大灾害几乎年年发生。据统计, 在过去的 40 年中, 我国平均每年出现的较大气象灾害为 24.5 次。全球热带海洋上每年生成 80 多个台风, 在我国登陆的有 7 个, 灾害发生频率约为美国的 4 倍、日本的 2 倍、菲律宾的 1.5 倍、俄罗斯的 30 倍。因此, 我们国家更有必要分析处理这些极端事件带来的保险损失。

表1.1　1970～1995 年理赔最大的 30 次保险损失(以 1992 年物价水平, 百万美元计)

损失额	日期 (月/日/年)	事　件	地　点
16 000	08/24/1992	Andrew 飓风	美国
11 838	01/17/1994	加利福尼亚州北岭地震	美国
5 724	09/27/1991	Mireille 龙卷风	日本
4 931	01/25/1990	Daria 暴风雪	欧洲
4 749	09/15/1989	Hugo 飓风	波多黎各

续表

损失额	日期 (月/日/年)	事 件	地 点
4 528	10/17/1989	Loma Prieta 地震	美国
3 427	02/26/1990	Vivian 暴风雪	欧洲
2 373	07/06/1988	Piper Alpha 海上石油铅井平台大爆炸	英国
2 282	01/17/1995	神户阪神大地震	日本
1 938	10/04/1995	Opal 飓风	美国
1 700	03/10/1993	东海岸暴风雪	美国
1 600	09/11/1992	Iniki 飓风	美国
1 500	10/23/1989	飞利浦石油大爆炸	美国
1 453	09/03/1979	Frederic 龙卷风	美国
1 422	09/18/1974	Fifi 龙卷风	洪都拉斯
1 320	09/12/1988	Gilbert 飓风	牙买加
1 238	12/17/1983	暴雪、严寒天气	美国
1 236	10/20/1991	蔓延到城市的森林火灾	美国
1 224	04/02/1974	14 个州的龙卷风	美国
1 172	08/04/1970	Celia 龙卷风	美国
1 168	04/25/1973	中西部地区密西西比河洪水	美国
1 048	05/05/1995	大风、冰雹、洪水	美国
1 005	01/02/1976	西北欧暴风雪	欧洲
950	08/17/1983	Alicia 飓风	美国
923	01/21/1995	西北欧暴风雨、洪水	欧洲
923	10/26/1993	蔓延到城市的森林火灾	美国
894	02/03/1990	Herta 飓风	欧洲
870	09/03/1993	Yancy 台风	日本
865	08/18/1991	Bob 飓风	美国
851	02/16/1980	加利福尼亚和亚利森那州大洪水	美国

回到表 1.1 所提供的数据，这些极端事件无一例外，都造成单个或者一组远远超出某个保险公司实际理赔能力的一些大额理赔，其直接决定了该保险公司的破产与否。这些极端事件造成大额理赔的一个共同特点就是具有重尾性，粗略地说就是大值出现的概率相对较大 (尽管其自身发生的概率很小)。在财产保险业中, 重尾分布, 特别是次指数分布, 已经被越来越多的学者认为是个体索赔额的标准模型 [参见 Embrechts 等 (1997)]。因此，研究重尾风险模型，对保险公司特别是中国的保险公司具有重要的现实意义。

在保险数学研究的大部分问题中，风险理论提供了经典的数学基础。本章将为本书的其他章节服务，介绍风险理论中的一些经典的风险模型，包括小额及大额理赔的例子。当然，我们研究的并不是如百科全书般的，我们更多的是关注与极端事件相关的重尾模型中破产概率的渐近结果，特别是一些最新的结果。

1.1　风险过程与破产概率

基本的保险风险模型要追溯到 Filip Lundberg 于 1903 年完成的著名 *Uppsala* 论文，其奠定了精算风险理论的基础。Lundberg 认为，非寿险模型的核心是 Poisson 过程，通过适当的时间变换，可以将研究限制在齐次 Poisson 过程。Harald Cramér 将 Lundberg 的思想融入了当时新兴的随机过程理论，奠定了非寿险数学和概率论的基础。基本的模型，即 Cramér-Lundberg 风险模型有如下结构：

定义 1.1　Cramér-Lundberg风险模型满足如下条件：

(a) 索赔额 $\{X_n,\ n\geqslant 1\}$ 是一列独立同分布(i.i.d.) 的非负随机变量，具有共同的非格点分布函数 F，有限均值 $\mu_F=\mathbb{E}X_1$ 和方差 $\sigma_F^2=\mathbb{D}X_1\leqslant\infty$。

(b) 索赔发生在随机时刻

$$0<\sigma_1<\sigma_2<\cdots\ \text{a.s.}$$

(c) 记时间间隔 $[0,t]$ 上的索赔数为

$$N(t)=\sup\{n\geqslant 1:\sigma_n\leqslant t\},\quad t\geqslant 0$$

约定，$\sup\varnothing=0$。

(d) 索赔时间间隔

$$\theta_1=\sigma_1,\quad \theta_k=\sigma_k-\sigma_{k-1},\quad k\geqslant 2 \tag{1.1}$$

是一列 i.i.d.的指数分布随机变量，具有有限均值 $\mathbb{E}\theta_1=1/\lambda$。

(e) 索赔额 $\{X_n,\ n\geqslant 1\}$ 与索赔时间间隔 $\{\theta_n,\ n\geqslant 1\}$ 相互独立。

更新风险模型满足条件(a)~(c), (e)和 (e′)。

(e′) 式(1.1) 中定义的索赔时间间隔 $\{\theta_n,\ n\geqslant 1\}$ 是一列 i.i.d.的非负随机变量，具有有限均值 $\mathbb{E}\theta_1=1/\lambda$。

上述更新风险模型是对 Cramér-Lundberg 风险模型的一般化，其是由 Sparre-Andersen 于 1957 年提出的 [参见 Andersen (1957)]，所以又称之为 Sparre-Andersen 风险模型。注意到，在 Cramér-Lundberg 风险模型中，计数过程 $\{N(t),\ t\geqslant 0\}$ 是一个强度为 λ 的齐次 Poisson 过程，即

$$\mathbb{P}(N(t)=k)=\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!},\quad k\geqslant 0$$

在更新风险模型中的 $N(t)$ 通常称为更新计数过程。

定义索赔累积过程为 $S(t)=\sum\limits_{k=1}^{N(t)}X_k,\ t\geqslant 0$。在 Cramér-Lundberg 风险模型中 $\{S(t),\ t\geqslant 0\}$ 又称为复合 Poisson 过程。图 1.1 和图 1.2 分别给出齐次 Poisson 过

程 $\{N(t),\ t \geqslant 0\}$ 和复合 Poisson 过程 $\{S(t),\ t \geqslant 0\}$ 的样本轨道，这里记部分和 $S_n = \sum_{k=1}^{n} X_k,\ n \geqslant 1,\ S_0 \equiv 0$。

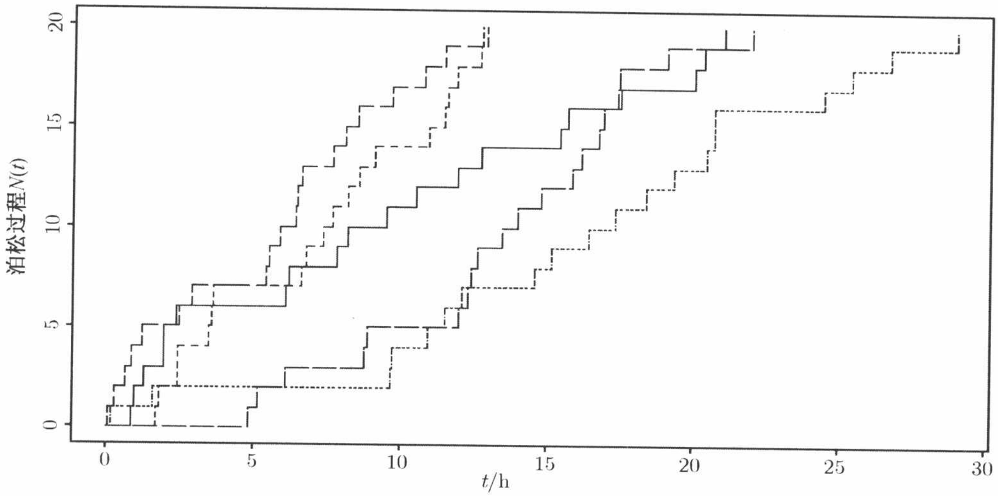

图 1.1 强度为 1 的 Poisson 过程 $N(t)$ 的 5 个样本轨道

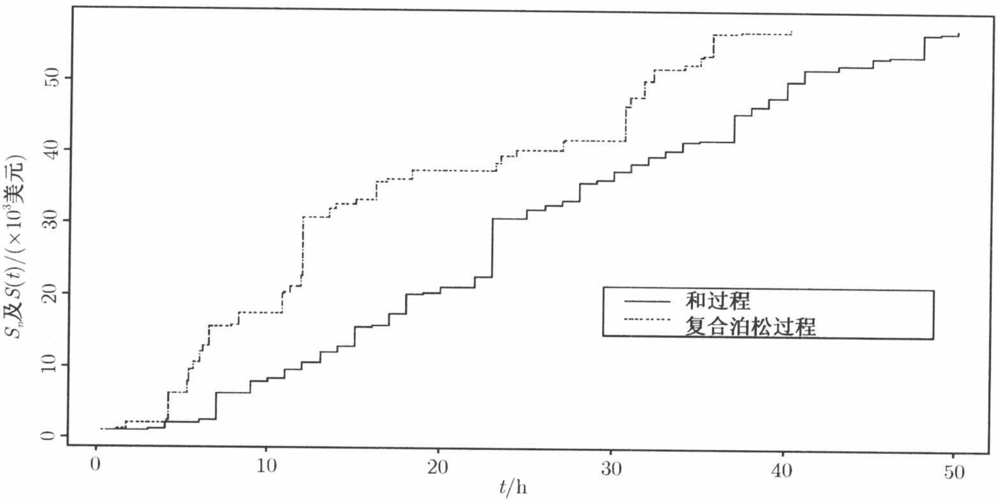

图 1.2 S_n 和复合 Poisson 过程 $S(t)$ 的 1 个样本轨道

注：$\lambda = 1$；X_1 为标准指数分布。

在 Cramér-Lundberg 风险模型中，索赔累积过程 $S(t)$ 的分布函数可以表示为

复合和的形式

$$\mathbb{P}(S(t) \leqslant x) = \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!} F^{n*}(x), \quad x \geqslant 0,\ t \geqslant 0 \tag{1.2}$$

其中，$F^{n*}(x) = \mathbb{P}(S_n \leqslant x)$ 为 F 的 n 重卷积。定义保险公司到时刻 $t \geqslant 0$ 的风险过程为

$$R(x,t) = x + ct - S(t) \tag{1.3}$$

其中，$x \geqslant 0$ 为保险公司的初始资本；$c > 0$ 为保费收入率，其需要满足一定的安全负荷条件。在某些非标准的风险模型中，保费收入过程 $C(t)$ 也可以是一般的非负非降随机过程，而不仅仅是本节讨论的线性过程 ct，相关的讨论见第 4 章。图 1.3 分别给出 Cramér-Lundberg 风险模型中指数索赔额的风险过程 $R(x,t)$ 一个和多个实现值。

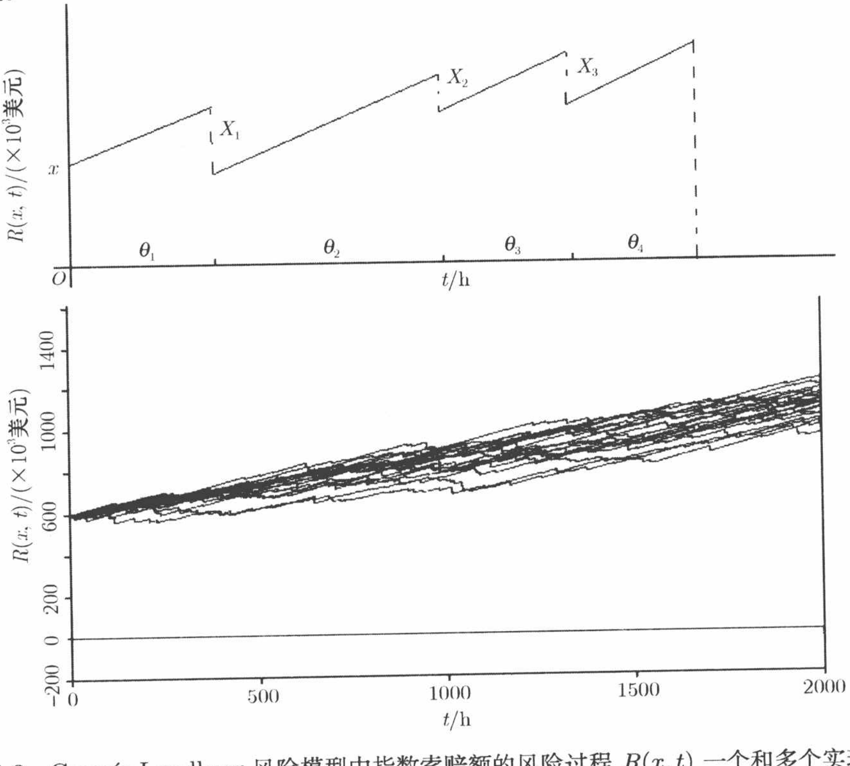

图 1.3　Cramér-Lundberg 风险模型中指数索赔额的风险过程 $R(x,t)$ 一个和多个实现值

在保险风险管理领域，下面两个量是理论和实际工作者共同关心的，也是本书重点研究分析的目标之一。

定义 1.2　有限时破产概率：对某个固定的 $T \geqslant 0$，在 $[0,T]$ 内破产的概率

$$\psi(x,T) = \mathbb{P}\left(\inf_{0 \leqslant t \leqslant T} R(x,t) < 0 \,\middle|\, R(x,0) = x \right)$$

无限时破产概率（最终破产概率）

$$\psi(x)=\psi(x,\infty)=\mathbb{P}\left(\inf_{t\geqslant 0}R(x,t)<0\,\middle|\,R(x,0)=x\right)$$

下面的引理是非常基本的。

引理 1.1 *在更新风险模型中*，$\mathbb{E}S(t)=\mu_F\mathbb{E}N(t)$。

利用这个基本的引理，我们可以对式 (1.3) 中定义的保费收入率 c 提出一定的限制。保费收入率 c 应该使对给定的 x 和 $T\leqslant\infty$，破产概率 $\psi(x,T)$ 较小。首要的是要求对所有的 $x\geqslant 0$，$\psi(x)<1$，这才能使得保险公司有严格正的概率永久存在不破产。这一准则是非常有效的。

由引理 1.1 及下述定理 1.2 (1) 知，在更新风险模型中，当 $t\to\infty$ 时

$$\mathbb{E}R(x,t)=x+(c-\lambda\mu_F)t(1+o(1))$$

因此，$\mathbb{E}R(x,t)/t\to c-\lambda\mu_F$，$t\to\infty$。显然，保险公司确保偿付能力的条件是 $c-\lambda\mu_F>0$，即当 t 充分大时，风险过程 $\{R(x,t),\ t\geqslant 0\}$ 具有正漂移。这就得到了更新风险模型中一个重要的安全负荷条件

$$\rho=\frac{c}{\lambda\mu_F}-1>0 \tag{1.4}$$

常数 ρ 称为安全负荷，也可以解释为风险保费率。事实上，$[0,t]$ 上的保费收入为 $ct=(1+\rho)\lambda\mu_F t$。

由于破产只可能发生在索赔到达时刻 σ_n，对任意的 $x\geqslant 0$, 有

$$\psi(x)=\mathbb{P}\left(\sup_{n\geqslant 1}\sum_{k=1}^{n}(X_k-c\theta_k)>x\right) \tag{1.5}$$

由式 (1.5) 知，在更新风险模型中，对无限时破产概率 $\psi(x)$ 的研究，本质上是对随机游动极大值尾概率的研究。事实上，考虑一个 i.i.d. 随机变量列

$$Z_k=X_k-c\theta_k,\quad k\geqslant 1 \tag{1.6}$$

以及由 $\{Z_k,\ k\geqslant 1\}$ 产生的随机游动

$$S_n^Z=\sum_{k=1}^{n}Z_k, n\geqslant 1, S_0^Z\equiv 0$$

注意到，由安全负荷条件 (1.4) 知，$\mathbb{E}Z_1=\mu_F-c/\lambda<0$. 因此，保险公司最终不破产的概率

$$1-\psi(x)=\mathbb{P}\left(\sup_{n\geqslant 1}S_n^Z\leqslant x\right)$$

可以通过 Spitzer 等式具体表示 [参见 Feller (1971)], 即

$$1-\psi(x)=(1-p)\sum_{n=0}^{\infty}p^nH^{n*}(x) \tag{1.7}$$

其中, 常数 $p\in(0,1)$, H 为某个分布函数, p 和 H 一般可以通过经典的 Wiener-Hopf 定理确定, H^{n*} 表示 H 的 n 重卷积。

关于破产概率 $\psi(x)$ 的估计, 除了 Wiener-Hopf 方法外, 更新理论也是常用的估计方法之一, 我们将在 1.4 节讨论 Cramér-Lundberg 风险模型中带有小额索赔的无限时破产概率的渐近性问题, 对大额索赔的相应结果, 我们将在第 3 章讨论。

注 1.1　关于风险理论的专著, 有兴趣的读者可以参见Bowers等(1987)、Bühlmann (1996)、Gerber (1979)、Grandell (1991 a) 等。关于Cramér在风险理论中的工作, 可以参见Cramér (1954)。Spitzer等式的证明, 其可以用于计算(1.7)中的破产概率, 可以参见一些随机过程的专著, 如Chung (1974) 等。经典的Wiener-Hopf方法源自于Feller (1971), 也可参见Asmussen (2003)。Wiener-Hopf分解相应于所谓的一个简单随机游动的梯高分布, 其证明参见Kennedy (1994), 进一步的讨论, 包括对破产概率的估计及其应用, 参见Prabhu (1980)。Wiener-Hopf理论在风险理论中的应用参见Asmussen (2000)、Bühlmann (1996)、Embrechts和Veraverbeke (1982)。

1.2　索赔额分布

本节我们将简要介绍一些目前常常用于刻画索赔额分布 F 的分布族。粗略地说, 我们分为两大类, 即轻尾分布 (light-tailed distribution, 有时也称为 Cramér 型条件的分布) 和重尾分布 (heavy-tailed distribution)。在不同的领域, 重尾分布这个术语可以理解为不同的含义, 但是通常是指不存在任何指数阶矩的分布。所谓轻尾是指尾分布 $\mathbb{P}(\xi>x)=\overline{V}(x)=1-V(x)>0$ 满足 $\overline{V}(x)=O(\mathrm{e}^{-sx})$, $x\to\infty$, 对某个 $s>0$; 或者等价的, $\mathbb{E}\mathrm{e}^{s\xi}=\int_0^\infty \mathrm{e}^{sx}V(\mathrm{d}x)<\infty$ 对某个 $s>0$。与之相对应的, 如果对所有的 $s>0$, $\mathbb{E}\mathrm{e}^{s\xi}=\infty$, 则称 $[0,\infty)$ 上的随机变量 ξ (或者对应的分布 V) 是重尾的。分别记重尾分布族和轻尾分布族为 $\mathscr{K}$ 和 $\mathscr{K}^c$。对于一般的支撑在 $\mathbb{R}$ 上的随机变量 ξ, 则当 ξ^+ 及其分布是重尾时, 将它和它的分布称为重尾的 [该术语, 据我们所知, 最早可参阅 Willekens (1986)]。一般来说, 在实际操作中, 对于某个给定的保险公司, 如果占总索赔次数 20% 的索赔额之和达到了公司总索赔额的 80% 以上, 则索赔额分布 F 可以视为重尾的, 即

$$\frac{1}{\mu_F}\int_{u_{0.2}}^{\infty}xF\mathrm{d}x\geqslant 0.8$$

其中, $\overline{F}(u_{0.2})=0.2$, μ_F 是分布 F 的均值, 参见 Asmussen (2003)。所谓重尾分布, 实际上是规范分布 (右) 尾部变化性状的准则。当然, 任何重尾分布的支撑都是右

无界的。在第 2 章我们将会给出重尾分布或者某些特殊重尾分布的一些等价条件，从中可以看出，分布 V 是重尾的，当且仅当，V 的尾分布 $\overline{V}$ 不能被任何趋于零的指数函数所控制。

1.2.1 轻尾分布

例 1.1(指数分布) 指数分布具有密度函数

$$v(x) = \delta \mathrm{e}^{-\delta x}, \quad x \geqslant 0 \tag{1.8}$$

其中，参数 $\delta > 0$ 可以被视为常数故障率 $v(x)/\overline{V}(x)$。

在风险理论中，指数分布是最简单最容易处理的一种分布。在 Cramér-Lundberg 风险模型中，如果索赔额是指数分布，无限时破产概率 $\psi(x)$ 甚至有精确的表达式。

例 1.2(Gamma分布) 参数为 $p > 0$, $\delta > 0$ 的Gamma分布具有密度函数

$$v(x) = \frac{\delta^p}{\Gamma(p)} x^{p-1} \mathrm{e}^{-\delta x}, \quad x \geqslant 0 \tag{1.9}$$

注意到，当 $s < \delta$ 时，$\mathbb{E}\mathrm{e}^{s\xi} = \delta^p(\delta - s)^{-p} < \infty$。尾分布 $\overline{V}(x)$ 的精确表达式为

$$\overline{V}(x) = \frac{\Gamma(\delta x; p)}{\Gamma(p)}$$

其中，$\Gamma(x; p) = \displaystyle\int_x^\infty t^{p-1} \mathrm{e}^{-t} \mathrm{d}t$ 是 Gamma 函数。尾分布渐近等价于

$$\overline{V}(x) \sim \frac{\delta^{p-1}}{\Gamma(p)} x^{p-1} \mathrm{e}^{-\delta x}, \quad x \to \infty$$

在无穷可分分布理论中，Gamma 分布的密度 (1.9) 可以视为指数分布的密度 (1.8) 的 p 阶幂（或者 $1/p$ 阶根，如果 $p < 1$)。特别的，如果 p 为正整数，ξ 具有密度 (1.9)，则 $\xi \stackrel{\mathrm{d}}{=} \eta_1 + \cdots + \eta_p$，其中 $\{\eta_n, \ n \geqslant 1\}$ 是 i.i.d. 的随机变量，具有指数密度 (1.8)。此时，又称随机变量 ξ 服从 Erlang(p) 分布。将此特点联系到 Poisson 过程，可以将 $\overline{V}(x) = \mathbb{P}(\eta_1 + \cdots + \eta_p > x)$ 看成是在 $[0, x]$ 上至多有 $p-1$ 个 Poisson 事件发生的概率，故

$$\overline{V}(x) = \mathrm{e}^{-\delta x} \sum_{k=0}^{p-1} \frac{(\delta x)^k}{k!}$$

例 1.3(超指数分布) 超指数分布是指数分布的有限混合，具有密度函数

$$v(x) = \sum_{k=1}^{p} \alpha_k \delta_k \mathrm{e}^{-\delta_k x}, \quad x \geqslant 0$$

其中，$\displaystyle\sum_{k=1}^{p} \alpha_k = 1$, $0 \leqslant \alpha_k \leqslant 1$, $\delta_k > 0$, $k = 1, 2, \cdots, p$。

1.2.2 重尾分布

例 1.4(Weibull分布) Weibull 分布来源于可靠性理论，其中故障率 $\delta(x)=v(x)/\overline{V}(x)$（风险理论中也称为风险率）有着重要作用，指数分布是最简单的情形，其故障率 $\delta(x)$ 为常数。然而，实际中观察到的 $\delta(x)$ 一般来说是单调不降的或者单调不增的函数。取 $\delta(x)=c_Vr_Vx^{r_V-1}$, $c_V>0, r_V>0$，则我们得到Weibull分布的尾分布函数和密度函数分别为

$$\overline{V}(x)=\mathrm{e}^{-c_Vx^{r_V}}, \quad v(x)=c_Vr_Vx^{r_V-1}\mathrm{e}^{-c_Vx^{r_V}}$$

当 $0<r_V<1$ 时，V 是重尾的，并且所有幂阶矩存在有限。

例 1.5(对数正态分布) 参数为 μ、σ^2 的对数正态分布定义为 e^{ξ} 的分布函数，其中 ξ 服从 $\mathrm{N}(\mu,\sigma^2)$。对数正态分布具有密度函数

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

尾分布渐近等价于

$$\overline{V}(x)\sim\frac{\sigma}{\sqrt{2\pi}\log x}\exp\left\{-\frac{1}{2}\left(\frac{\log x-\mu}{\sigma}\right)^2\right\}, \quad x\to\infty$$

对数正态分布具有所有的幂阶矩，特别的，一阶矩为 $\mathrm{e}^{\mu+\sigma^2/2}$，二阶矩为 $\mathrm{e}^{2\mu+2\sigma^2}$。

例 1.6(Pareto分布) 一般的Pareto分布具有密度函数

$$v(x)=\frac{\alpha}{(1+x)^{\alpha+1}}, \quad x\geqslant 0$$

有时也允许位置参数 $a>0$ 和尺度参数 $\beta>0$，则

$$v(x)=\frac{\alpha}{\beta[1+(x-a)/\beta]^{\alpha+1}}, \quad x\geqslant a$$

Pareto 分布具有有限的 $p<\alpha-1$ 阶矩，而所有的 $p>\alpha-1$ 阶矩都不存在。

例 1.7(对数Gamma分布) 参数为 p,δ 的对数Gamma分布定义为 e^{ξ} 的分布函数，其中 ξ 具有Gamma密度(1.9)。对数Gamma分布具有密度函数

$$v(x)=\frac{\delta^p(\log x)^{p-1}}{x^{\delta+1}\Gamma(p)}, \quad x>1$$

对数 Gamma 分布的 $p<\delta$ 阶矩存在，而 $p>\delta$ 阶矩不存在。当 $p=1$ 时，对数 Gamma 分布即 Pareto 分布。

其他常见的重尾分布还包括Burr分布，参数 $\alpha,\kappa,\tau>0$，其尾分布

$$\overline{V}(x)=\left(\frac{\kappa}{\kappa+x^{\tau}}\right)^{\alpha}, \quad x>0$$

Benktander-type-I 分布，参数 $\alpha,\beta>0$，其尾分布

$$\overline{V}(x)=\left(1+\frac{2\beta\log x}{\alpha}\right)\mathrm{e}^{-\beta\log^2 x-(\alpha+1)\log x},\quad x>1$$

Benktander-type-II 分布，参数 $\alpha>0$，$0<\beta<1$，其尾分布

$$\overline{V}(x)=x^{-(1-\beta)}\mathrm{e}^{\frac{\alpha}{\beta}-\frac{\alpha x^{\beta}}{\beta}},\quad x>1$$

1.2.3 常见的重尾分布子族及其性质

若无特殊说明，本节均假设随机变量是正的，并且具有右无界支撑，即对所有 $x>0$，$V(x)<1$。由于重尾分布族本身过于宽泛，所以为了对重尾分布族进行更为深入细致的研究，也出于保险等领域实际应用的需要，人们引入一些比较重要的重尾分布子族，它们在本书的工作中起着重要作用。一个重要的重尾分布子族是次指数分布族。次指数分布已经被越来越多的学者认为是个体索赔额的标准模型 [参见 Embrechts 等 (1997)]。次指数分布是由 Chistyakov (1964) 引入的。

定义 1.3(次指数分布) 支撑在 $[0,\infty)$ 上的分布 V 是次指数的(subexponential)，记作 $V\in\mathscr{S}$，如果对所有的 $n\geqslant 2$ 有

$$\lim_{x\to\infty}\frac{\overline{V^{n*}}(x)}{\overline{V}(x)}=n \tag{1.10}$$

注 1.2 次指数分布定义式(1.10)立即可以得到次指数性的特征：最大值与和等价，即假设 $\{\xi_k,\ k=1,2,\cdots,n\}$ 是一列 i.i.d.随机变量，具有共同的分布 $V\in\mathscr{S}$，则对所有的 $n\geqslant 2$，当 $x\to\infty$ 时有

$$\mathbb{P}(\xi_1+\cdots+\xi_n>x)\sim\mathbb{P}(\max(\xi_1,\cdots,\xi_n)>x) \tag{1.11}$$

事实上，对每个固定的 $n\geqslant 2$ 有

$$\begin{aligned}\mathbb{P}(\max(\xi_1,\cdots,\xi_n)>x)&=1-\mathbb{P}(\max(\xi_1,\cdots,\xi_n)\leqslant x)\\&=1-(1-\overline{V}(x))^n\\&=1-\mathrm{e}^{n\log(1-\overline{V}(x))}\\&\sim -n\log(1-\overline{V}(x))\\&\sim n\overline{V}(x),\quad x\to\infty\end{aligned}$$

式 (1.11) 意味着，i.i.d. 次指数随机变量和 $\xi_1+\cdots+\xi_n$ 超出某个大值 x 的概率渐近等价于 $\xi_1,\cdots,\xi_n$ 中最大的一个随机变量超出 x 的概率。这体现了所谓的一个大跳原则，反映了独立次指数随机变量和的概率性态。