1

Risk Theory

For most of the problems treated in insurance mathematics, risk theory still provides the quintessential mathematical basis. The present chapter will serve a similar purpose for the rest of this book. The basic risk theory models will be introduced, stressing the instances where a division between small and large claims is relevant. Nowadays, there is a multitude of textbooks available treating risk theory at various mathematical levels. Consequently, our treatment will not be encyclopaedic, but will focus more on those aspects of the theory where we feel that, for modelling extremal events, the existing literature needs complementing. Readers with a background in finance rather than insurance may use this chapter as a first introduction to the stochastic modelling of claim processes.

After the introduction of the basic risk model in Section 1.1, we derive in Section 1.2 the classical Cramér–Lundberg estimate for ruin probabilities in the infinite horizon case based on a small claim condition. Using the Cramér–Lundberg approach, a first estimation of asymptotic ruin probabilities in the case of regularly varying claim size distributions is obtained in Section 1.3.1. The natural generalisation to subexponentially distributed claim sizes is given in Sections 1.3.2, 1.3.3 and further discussed in Section 1.4. The latter section, together with Appendix A3, contains the basic results on regular variation and subexponentiality needed further in the text.

P. Embrechts et al., *Modelling Extremal Events*, Stochastic Modelling and Applied Probability 33, DOI 10.1007/978-3-642-33483-2_2,

1.1 The Ruin Problem

The basic insurance risk model goes back to the early work by Filip Lundberg [431] who in his famous Uppsala thesis of 1903 laid the foundation of actuarial risk theory. Lundberg realised that Poisson processes lie at the heart of non–life insurance models. Via a suitable time transformation (so–called operational time) he was able to restrict his analysis to the homogeneous Poisson process. This "discovery" is similar to the recognition by Bachelier in 1900 that Brownian motion is the key building block for financial models. It was then left to Harald Cramér and his Stockholm School to incorporate Lundberg's ideas into the emerging theory of stochastic processes. In doing so, Cramér contributed considerably to laying the foundation of both non–life insurance mathematics as well as probability theory. The basic model coming out of these first contributions, referred to in the sequel as *the Cramér–Lundberg model*, has the following structure:

Definition 1.1.1 (The Cramér–Lundberg model, the renewal model)
The Cramér–Lundberg model *is given by conditions* (a)–(e)*:*

(a) *The* claim size process:
the claim sizes $(X_k)_{k\in\mathbb{N}}$ *are positive iid rvs having common non–lattice df* F*, finite mean* $\mu = EX_1$*, and variance* $\sigma^2 = \mathrm{var}(X_1) \le \infty$*.*

(b) *The* claim times:
the claims occur at the random instants of time

$$0 < T_1 < T_2 < \cdots \quad \text{a.s.}$$

(c) *The* claim arrival process:
the number of claims in the interval $[0,t]$ *is denoted by*

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

where, by convention, $\sup\emptyset = 0$*.*

(d) *The* inter–arrival times

$$Y_1 = T_1\,, \quad Y_k = T_k - T_{k-1}\,, \quad k = 2, 3, \ldots, \tag{1.1}$$

are iid exponentially distributed with finite mean $EY_1 = 1/\lambda$*.*

(e) *The sequences* (X_k) *and* (Y_k) *are independent of each other.*

The renewal model *is given by* (a)–(c), (e) *and*

(d') *the inter–arrival times* Y_k *given in* (1.1) *are iid with finite mean* $EY_1 = 1/\lambda$*.* □

Remarks. 1) A consequence of the above definition is that $(N(t))$ is a homogeneous Poisson process with intensity $\lambda > 0$ (for a definition we refer to Example 2.5.2). Hence

$$P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \ldots .$$

2) The renewal model is a slight generalisation of the Cramér–Lundberg model which allows for renewal counting processes (see Section 2.5.2). The latter are more general than the Poisson process for the claim arrivals. □

The *total claim amount process* $(S(t))_{t \geq 0}$ of the underlying portfolio is defined as

$$S(t) = \begin{cases} \sum_{i=1}^{N(t)} X_i \,, & N(t) > 0 \,, \\ 0 \,, & N(t) = 0 \,. \end{cases} \tag{1.2}$$

The general theory of random sums will be discussed in Section 2.5. It is clear that in the important case of the Cramér–Lundberg model more detailed information about $(S(t))$ can be obtained. We shall henceforth treat this case as a basic example on which newly introduced methodology can be tested. An important quantity in this context is the *total claim amount distribution* (or *aggregate claim* (*size*) *distribution*)

$$G_t(x) = P(S(t) \leq x) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} F^{n*}(x) \,, \qquad x \geq 0 \,, \quad t \geq 0 \,, \tag{1.3}$$

where $F^{n*}(x) = P(\sum_{i=1}^{n} X_i \leq x)$ is the n–fold convolution of F. Throughout the text, for a general df H on $(-\infty, \infty)$,

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

The resulting *risk process* $(U(t))_{t \geq 0}$ is now defined as

$$U(t) = u + ct - S(t) \,, \qquad t \geq 0 \,. \tag{1.4}$$

In (1.4), $u \geq 0$ denotes the *initial capital* and $c > 0$ stands for the *premium income rate*. The choice of c is discussed below; see (1.7). For an explanation on why in this case a deterministic (linear) income rate makes sense from an actuarial point of view; see for instance Bühlmann [98]. In Figure 1.1.2 some realisations of $(U(t))$ are given in the case of exponentially distributed claim sizes.

In the classical Cramér–Lundberg set–up, the following quantities are relevant for various insurance–related problems.

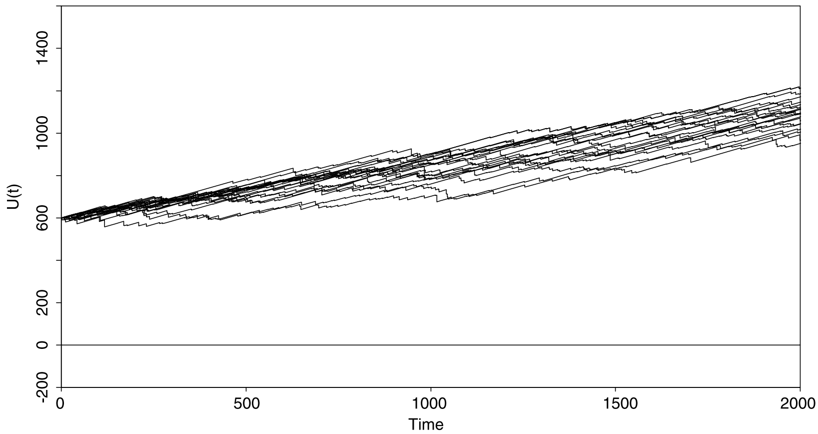


Figure 1.1.2 *Some realisations of* $(U(t))$ *for exponential claim sizes.*

Definition 1.1.3 (Ruin)
The ruin probability in finite time (*or* with finite horizon) :

$$\psi(u,T) = P(U(t) < 0 \quad \textit{for some } t \le T)\,, \quad 0 < T < \infty\,, \quad u \ge 0\,.$$

The ruin probability in infinite time (*or* with infinite horizon) :

$$\psi(u) = \psi(u,\infty)\,, \quad u \ge 0\,.$$

The ruin times*:*

$$\tau(T) = \inf\{t : 0 \le t \le T\,,\; U(t) < 0\}\,, \qquad 0 < T \le \infty\,,$$

where, by convention, $\inf \emptyset = \infty$*. We usually write* $\tau = \tau(\infty)$ *for the* ruin time with infinite horizon. □

The following result is elementary.

Lemma 1.1.4 *For the renewal model,*

$$EU(t) = u + ct - \mu EN(t)\,. \tag{1.5}$$

For the Cramér–Lundberg model,

$$EU(t) = u + ct - \lambda\mu t\,. \tag{1.6}$$

Proof. Since $EU(t) = u + ct - ES(t)$, and

$$\begin{aligned} ES(t) &= \sum_{k=0}^{\infty} E(S(t) \mid N(t) = k) \; P(N(t) = k) \\ &= \sum_{k=1}^{\infty} E\left(\sum_{i=1}^{N(t)} X_i \middle| N(t) = k \right) \; P(N(t) = k) \\ &= \sum_{k=1}^{\infty} E\left(\sum_{i=1}^{k} X_i \right) \; P(N(t) = k) \\ &= \mu \sum_{k=1}^{\infty} k \; P(N(t) = k) \\ &= \mu EN(t)\,, \end{aligned}$$

relation (1.5) follows. Because $EN(t) = \lambda t$ for the homogeneous Poisson process, (1.6) follows immediately. □

This elementary lemma yields a first guess of the premium rate c in (1.1). The latter is a major problem in insurance to which, at least for more general models, a vast amount of literature has been devoted; see for instance Goovaerts, De Vylder and Haezendonck [279]. We shall restrict our discussion to the above models. The determination of a suitable insurance premium rate obviously depends on the criteria used in order to define "suitable". It all depends on the measure of solvency we want to optimise over a given time period. The obvious (but by no means the only) measures available to us are the ruin probabilities $\psi(u,T)$ for $T \le \infty$. The premium rate c should be chosen so that a small $\psi(u,T)$ results for given u and T. A first step in this direction would be to require that $\psi(u) < 1$, for all $u \ge 0$. However, since $\psi(u) = P(\tau < \infty)$, this is equivalent to $P(\tau = \infty) > 0$: the company is given a strictly positive probability of infinitely long survival. Clearly, adjustments to this strategy have to be made before real premiums can be cashed. Anyhow, to set the stage, the above criterion is a useful one.

It follows immediately from (1.5) and Proposition 2.5.12 that in the renewal model, for $t \to \infty$,

$$\begin{aligned} EU(t) &= u + (c - \lambda\mu)\; t\,(1 + o(1)) \\ &= u + \left(\frac{c}{\lambda\mu} - 1 \right) \lambda\mu t\,(1 + o(1))\,. \end{aligned}$$

Therefore, $EU(t)/t \to c - \lambda\mu$, and an obvious condition towards solvency is $c - \lambda\mu > 0$, implying that $(U(t))$ has a positive drift for large t. This leads to the basic *net profit condition* in the renewal model:

$$\rho = \frac{c}{\lambda\mu} - 1 > 0\,. \tag{1.7}$$

The constant ρ is called the *safety loading*, which can be interpreted as a *risk premium rate*; indeed, the premium income over the period $[0,t]$ equals $ct = (1+\rho)\lambda\mu t$.

By definition of the risk process, ruin can occur only at the claim times T_i, hence for $u \geq 0$,

$$\begin{aligned}
\psi(u) &= P\left(u + ct - S(t) < 0 \quad \text{for some } t \geq 0\right) \\
&= P\left(u + cT_n - S\left(T_n\right) < 0 \quad \text{for some } n \geq 1\right) \\
&= P\left(u + \sum_{k=1}^{n} (cY_k - X_k) < 0 \quad \text{for some } n \geq 1\right) \\
&= P\left(\sup_{n\geq 1} \sum_{k=1}^{n} (X_k - cY_k) > u\right)\,.
\end{aligned}$$

Therefore, $\psi(u) < 1$ is equivalent to the condition

$$1 - \psi(u) = P\left(\sup_{n\geq 1} \sum_{k=1}^{n} (X_k - cY_k) \leq u\right) > 0\,, \qquad u \geq 0\,. \tag{1.8}$$

From (1.8) it follows that, in the renewal model, the determination of the non–ruin probability $1-\psi(u)$ is reduced to the study of the df of the ultimate maximum of a random walk. Indeed, consider the iid sequence

$$Z_k = X_k - cY_k\,, \quad k \geq 1\,,$$

and the corresponding random walk

$$R_0 = 0\,, \quad R_n = \sum_{k=1}^{n} Z_k\,, \quad n \geq 1\,. \tag{1.9}$$

Notice that $EZ_1 = \mu - c/\lambda < 0$ is just the net profit condition (1.7). Then the non–ruin probability is given by

$$1 - \psi(u) = P\left(\sup_{n\geq 1} R_n \leq u\right)\,.$$

This probability can for instance be determined via Spitzer's identity (cf. Feller [235], p. 613), which, for a general random walk, gives the distribution of its ultimate supremum. An application of the latter result allows us to express the non–ruin probability as a compound geometric df, i.e.

$$1 - \psi(u) = (1-\alpha) \sum_{n=0}^{\infty} \alpha^n \, H^{n*}(u) \tag{1.10}$$

for some constant $\alpha \in (0,1)$ and a df H. As before, H^{n*} denotes the nth convolution of H. Both α and H can in general be determined via the classical Wiener–Hopf theory; see again Feller [235], Sections XII.3 and XVIII.3, and Resnick [531].

Estimation of $\psi(u)$ can be worked out for a large category of models by applying a variety of (mostly analytic) techniques to functional relationships like (1.10). It is beyond the scope of this text to review these methods in detail. Besides the Wiener–Hopf methodology for the calculation of $\psi(u)$, renewal theory also yields relevant estimates, as we shall show in the next section. In doing so we shall concentrate on the Cramér–Lundberg model, first showing what typical estimates in a "small claim regime" look like. We then discuss what theory may be used to yield estimates for "large claims".

Notes and Comments

In recent years a multitude of textbooks on risk theory has been published. The interested reader may consult for instance Bowers et al. [85], Bühlmann [97], Gerber [256], Grandell [282], Straub [609], or Beard, Pentikäinen and Pesonen [54]. The latter book has recently appeared in a much updated form as Daykin, Pentikäinen and Pesonen [167]. In the review paper Embrechts and Klüppelberg [211] further references are to be found. A summary of Cramér's work on risk theory is presented in Cramér [140]; see also the recently published collected works of Cramér [141, 142] edited by Anders Martin–Löf. For more references on the historical background to this earlier work, together with a discussion on "where risk theory is evolving to" see Embrechts [202]. A proof of Spitzer's identity, which can be used in order to calculate the probability in (1.8), can be found in any basic textbook on stochastic processes; see for instance Chung [120], Karlin and Taylor [371], Prabhu [505], Resnick [531]. A classical source on Wiener–Hopf techniques is Feller [235], see also Asmussen [27]. An elementary proof of the Wiener–Hopf factorisation, relating the so–called ladder–height distributions of a simple random walk to the step distribution, is to be found in Kennedy [376]. A detailed discussion, including the estimation of ruin probabilities as an application, is given in Prabhu [507]; see also Prabhu [506, 508]. A comment on the relationship between the net profit condition and the asymptotic behaviour of the random walk (1.9) is to be found in Rogozin [548]. For a summary of the Wiener–Hopf theory relevant for risk theory see for instance Asmussen [28], Bühlmann [97] or Embrechts and Veraverbeke [218].

In Section 8.3 we come back to ruin probabilities. There we describe $\psi(u)$ via the distribution of the ultimate supremum of a random walk. Moreover, we characterise a sample path of the risk process leading to ruin.

1.2 The Cramér–Lundberg Estimate

In the previous section we mentioned a general method for obtaining estimates of the ruin probability $\psi(u)$ in the renewal model. If we further restrict ourselves to the Cramér–Lundberg model we can obtain a formula for $\psi(u)$ involving the claim size df F explicitly. Indeed, for the Cramér–Lundberg model under the net profit condition $\rho = c/(\lambda\mu) - 1 > 0$ one can show that

$$1 - \psi(u) = \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} (1+\rho)^{-n} \, F_I^{n*}(u) \,, \tag{1.11}$$

where

$$F_I(x) = \frac{1}{\mu} \int_0^x \overline{F}(y)\, dy \,, \quad x \ge 0 \,, \tag{1.12}$$

denotes *the integrated tail distribution* and

$$\overline{F}(x) = 1 - F(x) \,, \quad x \ge 0 \,,$$

denotes the tail of the df F. Later we shall show that formula (1.11) is the key tool for estimating ruin probabilities under the assumption of large claims. Also a proof of (1.11) will be given in Theorem 1.2.2 below.

In the sequel, the notion of Laplace–Stieltjes transform plays a crucial role.

Definition 1.2.1 (Laplace–Stieltjes transform)
Let H be a df concentrated on $(0,\infty)$, then

$$\widehat{h}(s) = \int_0^{\infty} e^{-sx} \, dH(x) \,, \quad s \in \mathbb{R} \,,$$

denotes the Laplace–Stieltjes transform of H. □

Remark. 1) Depending on the behaviour of $\overline{H}(x)$ for x large, $\widehat{h}(s)$ may be finite for a larger set of s–values than $s \ge 0$. In general, $\widehat{h}(s) < \infty$ for $s > -\gamma$ say, where $0 \le \gamma < \infty$ is the abscissa of convergence for $\widehat{h}(s)$. □

The following Cramér–Lundberg estimates of the ruin probability $\psi(u)$ are fundamental in risk theory.

Theorem 1.2.2 (Cramér–Lundberg theorem)
Consider the Cramér–Lundberg model including the net profit condition $\rho > 0$. *Assume that there exists a* $\nu > 0$ *such that*

$$\widehat{f}_I(-\nu) = \int_0^\infty e^{\nu x}\, dF_I(x) = \frac{c}{\lambda\mu} = 1 + \rho\,. \tag{1.13}$$

Then the following relations hold.

(*a*) *For all* $u \geq 0$,

$$\psi(u) \leq e^{-\nu u}\,. \tag{1.14}$$

(*b*) *If, moreover,*

$$\int_0^\infty x\, e^{\nu x}\, \overline{F}(x)\, dx < \infty\,, \tag{1.15}$$

then

$$\lim_{u\to\infty} e^{\nu u}\, \psi(u) = C < \infty\,, \tag{1.16}$$

where

$$C = \left[\frac{\nu}{\rho\mu} \int_0^\infty x\, e^{\nu x}\, \overline{F}(x)\, dx\right]^{-1}. \tag{1.17}$$

(*c*) *In the case of an exponential df* $F(x) = 1 - e^{-x/\mu}$, (1.11) *reduces to*

$$\psi(u) = \frac{1}{1+\rho}\, \exp\left\{-\frac{\rho}{\mu(1+\rho)}\, u\right\}, \quad u \geq 0\,. \tag{1.18}$$

Remarks. 2) The fundamental, so–called *Cramér–Lundberg condition* (1.13), can also be written as

$$\int_0^\infty e^{\nu x}\overline{F}(x)\, dx = \frac{c}{\lambda}\,.$$

3) It follows immediately from the definition of Laplace–Stieltjes transform that, whenever ν in (1.13) exists, it is uniquely determined; see also Grandell [282], p. 58.

4) Although the above results can be found in any basic textbook on risk theory, it is useful to discuss the proof of (b) in order to indicate how renewal–theoretic arguments enter (we have summarised the necessary renewal theory in Appendix A4). More importantly, we want to explain why the condition (1.13) has to be imposed. Very readable accounts of the relevant arguments are Feller [235], Sections VI.5, XI.7a, and Grandell [282]. □

Proof of (b). Denote $\delta(u) = 1 - \psi(u)$. Recall from (1.8) that $\delta(u)$ can be expressed via the random walk generated by $(X_i - cY_i)$. Then

$$
\begin{aligned}
&\delta(u)\\
&= P(S(t) - ct \le u \text{ for all } t > 0)\\
&= P\left(\sum_{k=1}^{n}(X_k - cY_k) \le u \text{ for all } n \ge 1\right)\\
&= P\left(\sum_{k=2}^{n}(X_k - cY_k) \le u + cY_1 - X_1 \text{ for all } n \ge 2\,,\ X_1 - cY_1 \le u\right)\\
&= P\left(S'(t) - ct \le u + cY_1 - X_1 \text{ for all } t > 0\,,\ X_1 - cY_1 \le u\right)\,,
\end{aligned}
$$

where S' is an independent copy of S. Hence

$$
\begin{aligned}
&\delta(u)\\
&= E\left(P\left(S'(t) - ct \le u + cY_1 - X_1 \text{ for all } t > 0\,,\ X_1 - cY_1 \le u \middle|\, Y_1\,,\ X_1\right)\right)\\
&= \int_0^\infty \int_0^{u+cs} P\left(S'(t) - ct \le u + cs - x \text{ for all } t > 0\right) dF(x) \lambda e^{-\lambda s} ds\\
&= \int_0^\infty \lambda e^{-\lambda s} \int_0^{u+cs} \delta(u + cs - x) dF(x) ds\\
&= \frac{\lambda}{c}\, e^{u\lambda/c} \int_u^\infty e^{-\lambda z/c} \left[\int_0^z \delta(z - x)\, dF(x)\right] dz\,, \qquad (1.19)
\end{aligned}
$$

where we used the substitution $u + cs = z$. The reader is urged to show explicitly where the various conditions in the Cramér–Lundberg model were used in the above calculations! This shows that δ is absolutely continuous with density

$$
\delta'(u) = \frac{\lambda}{c}\, \delta(u) - \frac{\lambda}{c} \int_0^u \delta(u - x)\, dF(x)\,. \qquad (1.20)
$$

From this equation for $1 - \psi(u)$ the whole theory concerning ruin in the classical Cramér–Lundberg model can be developed. A key point is that the integral in (1.20) is of *convolution type*; this opens the door to renewal theory. Integrate (1.20) from 0 to t with respect to Lebesgue measure to find

$$
\begin{aligned}
\delta(t) &= \delta(0) + \frac{\lambda}{c} \int_0^t \delta(u)\, du - \frac{\lambda}{c} \int_0^t\!\!\int_0^u \delta(u - x)\, dF(x)\, du\\
&= \delta(0) + \frac{\lambda}{c} \int_0^t \delta(t - u)\, du - \frac{\lambda}{c} \int_0^t \delta(t - x)\, F(x)\, dx\,.
\end{aligned}
$$

We finally arrive at the solution,

$$\delta(t) = \delta(0) + \frac{\lambda}{c} \int_0^t \delta(t-x)\overline{F}(x)\,dx\,. \tag{1.21}$$

Note that $\delta(0)$ is still unknown. However, letting $t \uparrow \infty$ in (1.21) and using the net profit condition (yielding $\delta(\infty) = 1 - \psi(\infty) = 1$) one finds $1 = \delta(0) + \mu\lambda/c$, hence $\delta(0) = 1 - \mu\lambda/c = \rho/(1+\rho)$. Consequently,

$$\delta(t) = \frac{\rho}{1+\rho} + \frac{1}{1+\rho} \int_0^t \delta(t-x)\,dF_I(x)\,, \tag{1.22}$$

where the integrated tail distribution F_I is defined in (1.12). Note that from (1.22), using Laplace–Stieltjes transforms, formula (1.11) immediately follows. The reader is advised to perform this easy calculation as an exercise and also to derive at this point formula (1.18). Equation (1.22) *looks* like a renewal equation; there is however one crucial difference and *this is exactly the point in the proof where a small claim condition of the type* (1.13) *enters.* First, rewrite (1.22) as follows in terms of $\psi(u) = 1 - \delta(u)$, setting $\alpha = 1/(1+\rho) < 1$,

$$\psi(u) = \alpha\overline{F}_I(u) + \int_0^u \psi(u-x)\; d\left(\alpha F_I(x)\right)\,. \tag{1.23}$$

Because $0 < \alpha < 1$, this is a so–called *defective renewal equation* (for instance Feller [235], Section XI.7). In order to cast it into the standard renewal set–up of Appendix A4, we define the following *exponentially tilted* or *Esscher transformed* df $F_{I,\nu}$:

$$dF_{I,\nu}(x) = e^{\nu x} d\left(\alpha F_I(x)\right)\,,$$

where ν is the exponent appearing in the condition (1.13). Using this notation, (1.23) becomes

$$e^{\nu u}\psi(u) = \alpha\, e^{\nu u}\,\overline{F}_I(u) + \int_0^u e^{\nu(u-x)}\psi(u-x)\,dF_{I,\nu}(x)$$

which, by condition (1.13), is a standard renewal equation. A straightforward application of the key renewal theorem (Theorem A4.3(b)) yields

$$\lim_{u\to\infty} e^{\nu u}\psi(u) = \left[\frac{\nu}{\rho\mu}\int_0^\infty x\, e^{\nu x}\overline{F}(x)\,dx\right]^{-1}$$

which is exactly (1.16)–(1.17). The conditions needed for applying Theorem A4.3 are easily checked. By partial integration and using (1.15),

$$\alpha e^{\nu u}\overline{F}_I(u) = \int_u^\infty e^{\nu x} d\left(\alpha F_I(x)\right) - \nu\int_u^\infty \alpha\overline{F}_I(x)e^{\nu x}dx\,.$$

Hence $\alpha e^{\nu u}\overline{F}_I(u)$ is the difference of two non–increasing Riemann integrable functions, and therefore it is directly Riemann integrable. Moreover,

$$\int_0^\infty \alpha\, e^{\nu u}\overline{F}_I(u)\,du = \alpha\,\frac{1-\widehat{f}_I(-\nu)}{-\nu} = \frac{\rho}{\nu(1+\rho)} < \infty\,,$$

and

$$\int_0^\infty x\,dF_{I,\nu}(x) = \frac{1}{\mu(1+\rho)}\int_0^\infty x\,e^{\nu x}\overline{F}(x)\,dx < \infty\,,$$

by (1.15). □

Because of the considerable importance for insurance the solution ν of (1.13) gained a special name:

Definition 1.2.3 (Lundberg exponent)
Given a claim size df F, the constant $\nu > 0$ satisfying

$$\int_0^\infty e^{\nu x}\overline{F}(x)\,dx = \frac{c}{\lambda}\,,$$

is called the Lundberg exponent *or* adjustment coefficient *of the underlying risk process.* □

Returning to (1.13), clearly the existence of ν implies that $\widehat{f}_I(s)$ has to exist in a non–empty neighbourhood of 0, implying that the tail $\overline{F}_I$ of the integrated claim size df, and hence also the tail $\overline{F}$, is exponentially bounded. Indeed, it follows from Markov's inequality that

$$\overline{F}(x) \le e^{-\nu x}\,E\,e^{\nu X_1}\,,\quad x > 0\,.$$

This inequality means that large claims are very unlikely (exponentially small probabilities!) to occur. For this reason (1.13) is often called a *small claim condition.*

The Cramér–Lundberg condition can easily be discussed graphically. The existence of ν in (1.13) crucially depends on the left abscissa of convergence $-\gamma$ of $\widehat{f}_I$. Various situations can occur as indicated in Figure 1.2.4. The most common case, and indeed the one fully covered by Theorem 1.2.2, corresponds to Figure 1.2.4(1). Typical claim size dfs and densities (denoted by f) covered by this regime are given in Table 1.2.5. We shall not discuss in detail the intermediate cases, unimportant for applications, of Figure 1.2.4(2) and (3).

If one scans the literature with the following question in mind:

Which distributions do actually fit claim size data?

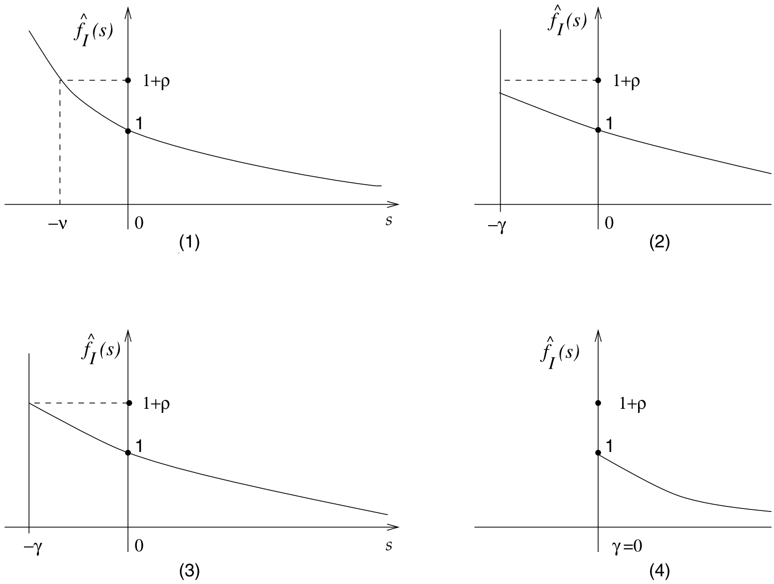


Figure 1.2.4 *Special cases in the Cramér–Lundberg condition.*

then most often one will find one of the dfs listed in Table 1.2.6. All the dfs in Table 1.2.5 allow for the construction of the Lundberg exponent. For the ones in Table 1.2.6 however, this exponent does not exist. Indeed, the case (4) in Figure 1.2.4 applies. For that reason we have labelled the two tables with "small claims", respectively "large claims". A more precise discussion of these distributions follows in Section 1.4. A detailed study of the properties of the distributions listed in Table 1.2.6 with special emphasis on insurance is to be found in Hogg and Klugman [330]. A wealth of material on these and related classes of dfs is presented in Johnson and Kotz [358, 359, 360].

For the sake of argument, assume that we have a portfolio following the Cramér–Lundberg model for which individual claim sizes can be modelled by a Pareto df

$$\overline{F}(x) = (1+x)^{-\alpha}\,, \qquad x \geq 0\,, \quad \alpha > 1\,.$$

It then follows that $EX_1 = \int_0^\infty (1+x)^{-\alpha}\,dx = (\alpha-1)^{-1}$ and the net profit condition amounts to $\rho = c(\alpha-1)/\lambda - 1 > 0$. Question:

Can we work out the exponential Cramér–Lundberg estimate in this case, for a given premium rate c satisfying the above condition?

The answer to this question is *no*. Indeed, in this case, for every $\nu > 0$

$$\int_0^\infty e^{\nu x}(1+x)^{-\alpha}\,dx = \infty\,,$$

Name	Tail $\overline{F}$ or density f	Parameters
Exponential	$\overline{F}(x) = e^{-\lambda x}$	$\lambda > 0$
Gamma	$f(x) = \dfrac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta x}$	$\alpha,\, \beta > 0$
Weibull	$\overline{F}(x) = e^{-cx^\tau}$	$c > 0,\, \tau \geq 1$
Truncated normal	$f(x) = \sqrt{\frac{2}{\pi}}\, e^{-x^2/2}$	—
Any distribution with bounded support		

Table 1.2.5 *Claim size dfs: "small claims". All dfs have support* $(0,\infty)$.

i.e. there is *no* exponential Cramér–Lundberg estimate in this case. We are in the regime of Figure 1.2.4(4): zero is an essential singularity of $\widehat{f}_I$, this means that $\widehat{f}_I(-\varepsilon) = \infty$ for every $\varepsilon > 0$.

However, it turns out that most individual claim size data are modelled by such dfs; see for instance Hogg and Klugman [330] and Ramlau–Hansen [522, 523] for very convincing empirical evidence on this. In Chapter 6 we shall analyse insurance data and come to the conclusion that also in these cases (1.13) is violated. So clearly, classical risk theory has to be adjusted to take this observation into account. In the next section we discuss in detail the class of *subexponential distributions* which will be *the* candidates for loss distributions in the heavy–tailed case. A detailed discussion of the theory of subexponential distributions is rather technical, so we content ourselves with an overview of that part of the theory which is most easily applicable within risk theory in particular and insurance and finance in general. In Section 1.3 we present the large–claims equivalent of the Cramér–Lundberg estimate; see for instance Theorem 1.3.6.

Notes and Comments

The reader interested in the various mathematical approaches to calculating ruin probabilities should consult any of the standard textbooks on risk theory; see the Notes and Comments of Section 1.1. A short proof of (1.14) based on martingale techniques is for instance discussed in Grandell [283]; see also Gerber [255, 256]. An excellent review on the subject is Grandell [283]. Theo-

Name	Tail $\overline{F}$ or density f	Parameters
Lognormal	$f(x) = \dfrac{1}{\sqrt{2\pi}\,\sigma x} e^{-(\ln x-\mu)^2/(2\sigma^2)}$	$\mu \in \mathbb{R}$, $\sigma > 0$
Pareto	$\overline{F}(x) = \left(\dfrac{\kappa}{\kappa + x}\right)^{\alpha}$	$\alpha, \kappa > 0$
Burr	$\overline{F}(x) = \left(\dfrac{\kappa}{\kappa + x^{\tau}}\right)^{\alpha}$	$\alpha, \kappa, \tau > 0$
Benktander–type–I	$\overline{F}(x) = (1 + 2(\beta/\alpha)\ln x)$ $e^{-\beta(\ln x)^2-(\alpha+1)\ln x}$	$\alpha, \beta > 0$
Benktander–type–II	$\overline{F}(x) = e^{\alpha/\beta} x^{-(1-\beta)} e^{-\alpha\, x^{\beta}/\beta}$	$\alpha > 0$ $0 < \beta < 1$
Weibull	$\overline{F}(x) = e^{-cx^{\tau}}$	$c > 0$ $0 < \tau < 1$
Loggamma	$f(x) = \dfrac{\alpha^{\beta}}{\Gamma(\beta)} (\ln x)^{\beta-1}\, x^{-\alpha-1}$	$\alpha, \beta > 0$
Truncated α–stable	$\overline{F}(x) = P(\lvert X\rvert > x)$ where X is an α–stable rv (see Definition 2.2.1)	$1 < \alpha < 2$

Table 1.2.6 *Claim size dfs: “large claims”. All dfs have support* $(0,\infty)$ *except for the Benktander cases and the loggamma with* $(1,\infty)$*.*

rem 1.2.2 can also be formulated for the renewal model; a detailed analysis of the Wiener–Hopf technique together with relevant renewal–type arguments can be worked out; see for instance Embrechts and Veraverbeke [218] for details and further references. Useful textbooks containing a discussion on the link between the asymptotic behaviour of the tail of a measure to properties of its Laplace–Stieltjes transform in a neighbourhood of zero are Bingham, Goldie and Teugels [72] (see for instance Section 1.7 in the latter), Feller [235], Section XIII.5, and Widder [642].

Using Wiener–Hopf theory, a theorem similar to Theorem 1.2.2 can be proved in the renewal model with F supported by $(-\infty,+\infty)$. For details see

Embrechts and Veraverbeke [218] and Thorin [622]. An interesting survey paper is Thorin [624].

Exponential–type ruin estimates hold for much wider classes of risk processes; see for instance Embrechts, Grandell and Schmidli [208], Grandell [282] and the references therein. The latter references also concentrate in detail on ruin estimation in finite time. For an approach based on diffusion approximations see Example 2.5.18.

A detailed discussion of ruin estimation under the various regimes given in Figure 1.2.4 is to be found in Embrechts and Veraverbeke [218]; see also Embrechts [201] for an example based on the generalised inverse Gaussian distribution. A useful review of the various claim size models used in non–life insurance is Hogg and Klugman [330]. The reader should be aware that for most models there is no standard notation or indeed parametrisation. We shall on some occasions say that "under the assumption of a Pareto distribution", meaning that the exact parametrisation is not important for that particular discussion. If however the specific parameter values are of interest, we will always make this clear, in many cases by explicitly stating which functional form of the density f or the df F is used.

1.3 Ruin Theory for Heavy–Tailed Distributions

Throughout this section, all rvs are positive with infinite support, i.e. $F(x) < 1$ for all $x > 0$. We have already seen that Pareto distributions violate the Cramér–Lundberg condition (1.13) so that Theorem 1.2.2 is not applicable for such claim size distributions. What alternative methodology can be used? The answer lies in the representation (1.11), together with Lemma 1.3.1 below. As from Section 1.3.1 onwards, we shall extensively use the theory of regular variation. The reader unfamiliar with the latter theory is urged first to read Appendix A3.1 before proceeding. We denote by $\mathcal{R}_\alpha$ the class of regularly varying functions with index $\alpha \in \mathbb{R}$. The case $\alpha = 0$ corresponds to the so–called slowly varying functions.

From Section 1.3.2 onwards, the class of subexponential distributions will play a fundamental role. For the latter, no complete textbook treatment exists. Because of their importance for the modelling of large claims, we have included a more detailed analysis of their properties. The results immediately needed for proving ruin estimates in the heavy–tailed case are presented in this chapter. For some of the more technical theorems, the reader is referred to Appendix A3.2. The main ideas underlying subexponentiality are presented in Section 1.3.2; Section 1.4 may be skipped upon first reading.

1.3.1 Some Preliminary Results

We start our discussion with a convolution closure property for regularly varying dfs. From Appendix A3.1, recall that L belongs to $\mathcal{R}_0$, i.e. L is slowly varying, whenever for all $t > 0$

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1 .$$

The following result is to be found in Feller [235], p. 278.

Lemma 1.3.1 (Convolution closure of dfs with regularly varying tails)
If F_1, F_2 are two dfs such that $\overline{F}_i(x) = x^{-\alpha}L_i(x)$ for $\alpha \geq 0$ and $L_i \in \mathcal{R}_0$, $i = 1, 2$, then the convolution $G = F_1 * F_2$ has a regularly varying tail such that

$$\overline{G}(x) \sim x^{-\alpha}\left(L_1(x) + L_2(x)\right) , \qquad x \to \infty .$$

Proof. Let X_1, X_2 be independent rvs with dfs F_1, respectively F_2. Using $\{X_1 + X_2 > x\} \supset \{X_1 > x\} \cup \{X_2 > x\}$ one easily checks that

$$\overline{G}(x) \geq \left(\overline{F}_1(x) + \overline{F}_2(x)\right)(1 - o(1)) .$$

If $0 < \delta < 1/2$, then from

$$\{X_1 + X_2 > x\} \subset \{X_1 > (1-\delta)x\} \cup \{X_2 > (1-\delta)x\} \cup \{X_1 > \delta x, X_2 > \delta x\} ,$$

it follows that

$$\begin{aligned} \overline{G}(x) &\leq \overline{F}_1((1-\delta)x) + \overline{F}_2((1-\delta)x) + \overline{F}_1(\delta x)\,\overline{F}_2(\delta x) \\ &= \Big(\overline{F}_1((1-\delta)x) + \overline{F}_2((1-\delta)x)\Big)(1 + o(1)) . \end{aligned}$$

Hence

$$1 \leq \liminf_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}_1(x) + \overline{F}_2(x)} \leq \limsup_{x\to\infty} \frac{\overline{G}(x)}{\overline{F}_1(x) + \overline{F}_2(x)} \leq (1-\delta)^{-\alpha} ,$$

which proves the result upon letting $\delta \downarrow 0$. One easily shows that $L_1 + L_2$ is slowly varying. □

An alternative proof of this result can be based upon Karamata's Tauberian theorem (Theorem A3.9). An important corollary obtained via induction on n is the following:

Corollary 1.3.2 *If $\overline{F}(x) = x^{-\alpha}L(x)$ for $\alpha \geq 0$ and $L \in R_0$, then for all $n \geq 1$,*

$$\overline{F^{n*}}(x) \sim n\overline{F}(x) , \quad x \to \infty .$$ □

Suppose now that $X_1, \ldots, X_n$ are iid with df F as in the above corollary. Denote the partial sum of $X_1, \ldots, X_n$ by $S_n = X_1 + \cdots + X_n$ and their maximum by $M_n = \max(X_1, \ldots, X_n)$. Then for all $n \geq 2$,

$$\begin{aligned} P\,(S_n > x) &= \overline{F^{n*}}(x)\,, \\ P\,(M_n > x) &= \overline{F^{n}}(x) \\ &= \overline{F}(x) \sum_{k=0}^{n-1} F^k(x) \\ &\sim n\overline{F}(x)\,, \quad x \to \infty\,. \end{aligned} \tag{1.24}$$

Therefore, with the above notation, Corollary 1.3.2 can be reformulated as

$$\boxed{\begin{gathered} \overline{F} \in \mathcal{R}_{-\alpha}\,, \quad \alpha \geq 0, \\ \text{implies} \\ P(S_n > x) \sim P(M_n > x)\,, \quad x \to \infty. \end{gathered}}$$

This implies that for dfs with regularly varying tails, the tail of the df of the sum S_n is mainly determined by the tail of the df of the maximum M_n. This is exactly one of the intuitive notions of heavy–tailed distribution or large claims. Hence, stated in a somewhat vague way:

> Under the assumption of regular variation, the tail of the maximum determines the tail of the sum.

Recall that in the Cramér–Lundberg model the following relation holds; see (1.11):

$$\psi(u) = \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} (1+\rho)^{-n}\, \overline{F_I^{n*}}(u)\,, \quad u \geq 0\,,$$

where $F_I(x) = \mu^{-1} \int_0^x \overline{F}(y)\,dy$ is the integrated tail distribution. Under the condition $\overline{F}_I \in \mathcal{R}_{-\alpha}$ for some $\alpha \geq 0$, *we might hope that* the following asymptotic estimate holds:

$$\frac{\psi(u)}{\overline{F}_I(u)} = \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} (1+\rho)^{-n}\, \frac{\overline{F_I^{n*}}(u)}{\overline{F}_I(u)} \tag{1.25}$$

$$\to \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} (1+\rho)^{-n}\, n = \rho^{-1}\,, \quad u \to \infty\,. \tag{1.26}$$

The key problem left open in the above calculation is the step from (1.25) to (1.26).

Can one safely interchange limits and sums?

The answer is *yes*; see Theorem 1.3.6. Consequently, (1.26) is the natural ruin estimate whenever $\overline{F}_I$ is regularly varying. Below we shall show that a similar estimate holds true for a much wider class of dfs. In its turn, (1.26) can be reformulated as follows.

> For claim size distributions with regularly varying tails, ultimate ruin $\psi(u)$ for large initial capital u is essentially determined by the tail $\overline{F}(y)$ of the claim size distribution for large values of y, i.e.
> $$\psi(u) \sim \frac{1}{\rho\mu} \int_u^\infty \overline{F}(y)\, dy\,, \quad u \to \infty\,.$$

From Table 1.2.6 we obtain the following typical claim size distributions covered by the above result:

– Pareto
– Burr
– loggamma
– truncated stable distributions.

1.3.2 Cramér–Lundberg Theory for Subexponential Distributions

As stated above, the crucial step in obtaining (1.26) was the property $\overline{F_I^{n*}}(x) \sim n\overline{F}_I(x)$ for $x \to \infty$ and $n \geq 2$. This naturally leads us to a class of dfs which allows for a very general theory of ruin estimation for large claims. The main result in this set–up is Theorem 1.3.6 below.

Definition 1.3.3 (Subexponential distribution function)
A df F with support $(0,\infty)$ is subexponential*, if for all $n \geq 2$,*

$$\lim_{x\to\infty} \frac{\overline{F^{n*}}(x)}{\overline{F}(x)} = n\,. \tag{1.27}$$

The class of subexponential dfs will be denoted by $\mathcal{S}$. □

Remark. 1) Relation (1.27) yields the following intuitive characterisation of subexponentiality; see (1.24).

$$\boxed{\text{For all } n \geq 2, \quad P(S_n > x) \sim P(M_n > x)\,, \quad x \to \infty\,. \quad (1.28)}$$

$\square$

In order to check for subexponentiality, one does not need to show (1.27) for all $n \geq 2$.

Lemma 1.3.4 (A sufficient condition for subexponentiality)
If $\limsup_{x\to\infty} \overline{F^{2*}}(x)/\overline{F}(x) \leq 2$*, then* $F \in \mathcal{S}$*.*

Proof. As F stands for the df of a positive rv, it follows immediately that $F^{2*}(x) \leq F^2(x)$, i.e. $\overline{F^{2*}}(x) \geq \overline{F^2}(x)$ for all $x \geq 0$. Therefore, $\liminf_{x\to\infty} \overline{F^{2*}}(x)/\overline{F}(x) \geq 2$, so that because of the condition of the lemma, the limit relation (1.27) holds for $n = 2$. The proof is then by induction on n. For $x \geq y > 0$,

$$\begin{aligned}
\frac{\overline{F^{(n+1)*}}(x)}{\overline{F}(x)} &= 1 + \frac{F(x) - F^{(n+1)*}(x)}{\overline{F}(x)} \qquad (1.29)\\
&= 1 + \int_0^x \frac{\overline{F^{n*}}(x-t)}{\overline{F}(x)}\, dF(t)\\
&= 1 + \left(\int_0^{x-y} + \int_{x-y}^x\right)\left(\frac{\overline{F^{n*}}(x-t)}{\overline{F}(x-t)}\,\frac{\overline{F}(x-t)}{\overline{F}(x)}\right) dF(t)\\
&= 1 + I_1(x) + I_2(x)\,.
\end{aligned}$$

By inserting $-n+n$ in I_1 and noting that $(\overline{F^{n*}}(x-t)/\overline{F}(x-t) - n)$ can be made arbitrarily small for $0 \leq t \leq x-y$ and y sufficiently large, it follows that

$$I_1(x) = (n + o(1)) \int_0^{x-y} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t)\,.$$

Now

$$\begin{aligned}
\int_0^{x-y} \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t) &= \frac{F(x) - F^{2*}(x)}{\overline{F}(x)} - \int_{x-y}^x \frac{\overline{F}(x-t)}{\overline{F}(x)}\, dF(t)\\
&= \frac{F(x) - F^{2*}(x)}{\overline{F}(x)} - J(x,y)\\
&= (1 + o(1)) - J(x,y)\,,
\end{aligned}$$

where $J(x,y) \leq (F(x) - F(x-y))/\overline{F}(x) \to 0$ as $x \to \infty$ by Lemma 1.3.5 (a) below. Therefore $\lim_{x\to\infty} I_1(x) = n$.

Finally, since $\overline{F^{n*}}(x-t)/\overline{F}(x-t)$ is bounded for $x-y \le t \le x$ and $\lim_{x\to\infty} J(x,y)=0$, $\lim_{x\to\infty} I_2(x)=0$, completing the proof. $\square$

Remarks. 2) The condition in Lemma 1.3.4 is trivially necessary for $F \in \mathcal{S}$.

3) In the beginning of the above proof we used that for the df F of a positive rv, always $\liminf_{x\to\infty} \overline{F^{2*}}(x)/\overline{F}(x) \ge 2$. One easily shows in this case that, for all $n \ge 2$,

$$\liminf_{x\to\infty} \overline{F^{n*}}(x)/\overline{F}(x) \ge n\,.$$

Indeed $S_n \ge M_n$, hence $\overline{F^{n*}}(x) = P(S_n > x) \ge P(M_n > x) = \overline{F^n}(x)$. Therefore

$$\liminf_{x\to\infty} \frac{\overline{F^{n*}}(x)}{\overline{F}(x)} \ge \lim_{x\to\infty} \frac{\overline{F^n}(x)}{\overline{F}(x)} = n\,. \qquad \square$$

The following lemma is crucial if we want to derive (1.26) from (1.25) for subexponential F_I.

Lemma 1.3.5 (Some basic properties of subexponential distributions)

(a) *If $F \in \mathcal{S}$, then uniformly on compact y–sets of* $(0,\infty)$,

$$\lim_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} = 1\,. \tag{1.30}$$

(b) *If* (1.30) *holds then, for all $\varepsilon > 0$,*

$$e^{\varepsilon x}\overline{F}(x) \to \infty\,, \quad x \to \infty\,.$$

(c) *If $F \in \mathcal{S}$ then, given $\varepsilon > 0$, there exists a finite constant K so that for all $n \ge 2$,*

$$\frac{\overline{F^{n*}}(x)}{\overline{F}(x)} \le K(1+\varepsilon)^n\,, \quad x \ge 0\,. \tag{1.31}$$

Proof. (a) For $x \ge y > 0$, by (1.29),

$$\begin{aligned}\frac{\overline{F^{2*}}(x)}{\overline{F}(x)} &= 1 + \int_0^y \frac{\overline{F}(x-t)}{\overline{F}(x)}\,dF(t) + \int_y^x \frac{\overline{F}(x-t)}{\overline{F}(x)}\,dF(t)\\ &\ge 1 + F(y) + \frac{\overline{F}(x-y)}{\overline{F}(x)}\,\big(F(x)-F(y)\big)\,.\end{aligned}$$

Thus, for x large enough so that $F(x)-F(y) \ne 0$,

$$1 \le \frac{\overline{F}(x-y)}{\overline{F}(x)} \le \left(\frac{\overline{F^{2*}}(x)}{\overline{F}(x)} - 1 - F(y)\right)(F(x)-F(y))^{-1}\,.$$

In the latter estimate, the rhs tends to 1 as $x \to \infty$. Uniformity immediately follows from monotonicity in y. Alternatively, use that the property (1.30) can be reformulated as $\overline{F} \circ \ln \in \mathcal{R}_0$, then apply Theorem A3.2.

(b) By (a), $\overline{F} \circ \ln \in \mathcal{R}_0$. But then the conclusion that $x^\varepsilon \overline{F}(\ln x) \to \infty$ as $x \to \infty$ follows immediately from the representation theorem for $\mathcal{R}_0$; see Theorem A3.3.

(c) Let $\alpha_n = \sup_{x \geq 0} \overline{F^{n*}}(x)/\overline{F}(x)$. Using (1.29) we obtain, for every $T < \infty$,

$$\begin{aligned}
\alpha_{n+1} &\leq 1 + \sup_{0 \leq x \leq T} \int_0^x \frac{\overline{F^{n*}}(x-y)}{\overline{F}(x)}\, dF(y) \\
&\qquad + \sup_{x \geq T} \int_0^x \frac{\overline{F^{n*}}(x-y)}{\overline{F}(x-y)} \frac{\overline{F}(x-y)}{\overline{F}(x)}\, dF(y) \\
&\leq 1 + A_T + \alpha_n \sup_{x \geq T} \frac{F(x) - F^{2*}(x)}{\overline{F}(x)} ,
\end{aligned}$$

where $A_T = (\overline{F}(T))^{-1} < \infty$. Now since $F \in \mathcal{S}$ we can, given any $\varepsilon > 0$, choose T such that

$$\alpha_{n+1} \leq 1 + A_T + \alpha_n (1 + \varepsilon) .$$

Hence

$$\alpha_n \leq (1 + A_T)\, \varepsilon^{-1} (1 + \varepsilon)^n ,$$

implying (1.31). □

Remark. 4) Lemma 1.3.5(b) justifies the name subexponential for $F \in \mathcal{S}$; indeed $\overline{F}(x)$ decays to 0 slower than any exponential $e^{-\varepsilon x}$ for $\varepsilon > 0$. Furthermore, since for any $\varepsilon > 0$:

$$\int_y^\infty e^{\varepsilon x}\, dF(x) \geq e^{\varepsilon y}\, \overline{F}(y) , \quad y \geq 0 ,$$

it follows from Lemma 1.3.5(b) that for $F \in \mathcal{S}$, $\widehat{f}(-\varepsilon) = \infty$ for all $\varepsilon > 0$. Therefore the Laplace–Stieltjes transform of a subexponential df has an essential singularity at 0. This result was first proved by Chistyakov [115], Theorem 2. As follows from the proof of Lemma 1.3.5(b) the latter property holds true for the larger class of dfs satisfying (1.30). For a further discussion on these classes see Section 1.4. □

Recall that for a df F with finite mean μ, $F_I(x) = \mu^{-1} \int_0^x \overline{F}(y)\, dy$. An immediate, important consequence from the above result is the following.

Theorem 1.3.6 (The Cramér–Lundberg theorem for large claims, I)
Consider the Cramér–Lundberg model with net profit condition $\rho > 0$ and $F_I \in \mathcal{S}$, then

$$\psi(u) \sim \rho^{-1}\, \overline{F}_I(u)\,, \qquad u \to \infty\,. \tag{1.32}$$

Proof. Since $(1+\rho)^{-1} < 1$, there exists an $\varepsilon > 0$ such that $(1+\rho)^{-1}(1+\varepsilon) < 1$. Hence because of (1.31),

$$(1+\rho)^{-n}\,\frac{\overline{F_I^{n*}}(u)}{\overline{F}_I(u)} \le (1+\rho)^{-n}\,K(1+\varepsilon)^n\,, \quad u \ge 0\,,$$

which allows by dominated convergence the interchange of limit and sum in (1.25), yielding the desired result. □

In Figure 1.3.7 realisations of the risk process $(U(t))$ are given in the case of lognormal and Pareto claims. These realisations should be compared with the ones in Figure 1.1.2 (exponential claims).

This essentially finishes our task of finding a Cramér–Lundberg type estimate in the heavy–tailed case.

For claim size distributions with subexponential integrated tail distribution, ultimate ruin $\psi(u)$ is given by (1.32).

In addition to dfs with regularly varying tails, the following examples from Table 1.2.6 yield the estimate (1.32). This will be shown in Section 1.4.

– lognormal
– Benktander–type–I
– Benktander–type–II
– Weibull ($0 < \tau < 1$).

From a mathematical point of view, the result in Theorem 1.3.6 can be substantially improved. Indeed, Corollary A3.21 yields the following result.

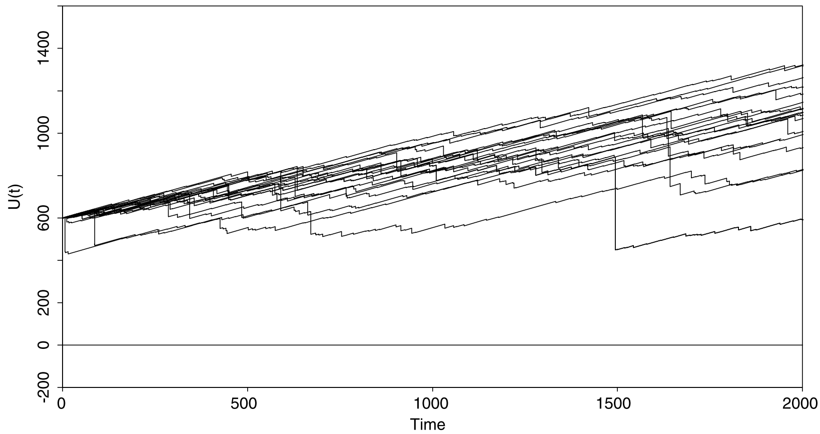
Theorem 1.3.8 (The Cramér–Lundberg theorem for large claims, II)
Consider the Cramér–Lundberg model with net profit condition $\rho > 0$. Then the following assertions are equivalent:

(*a*) $F_I \in \mathcal{S}$,

(*b*) $1 - \psi(u) \in \mathcal{S}$,

(*c*) $\lim_{u\to\infty}\ \psi(u)/\overline{F}_I(u) = \rho^{-1}$. □

Consequently, the estimate (1.32) is *only* possible under the condition $F_I \in \mathcal{S}$. In the case of the Cramér–Lundberg theory, $\mathcal{S}$ is therefore the natural class when it comes to ruin estimates whenever the Cramér–Lundberg condition



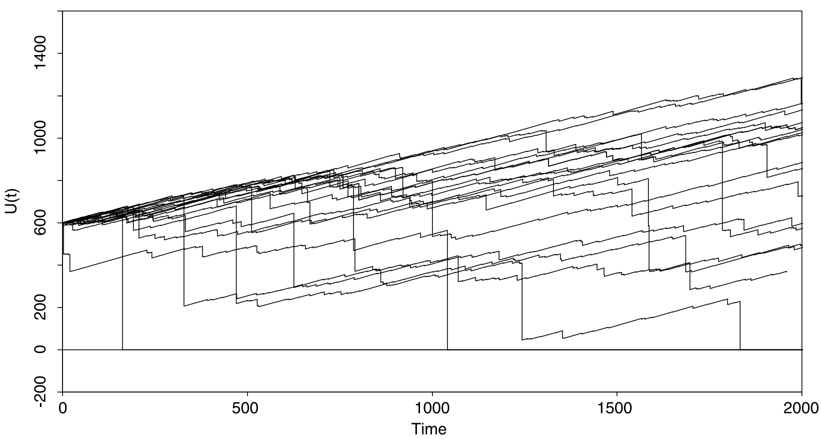


Figure 1.3.7 *Some realisations of the risk process* $(U(t))$ *for lognormal* (*top*) *and Pareto* (*bottom*) *claim sizes.*

(1.13) is violated. In Section 1.4 we shall come back to the condition $F_I \in \mathcal{S}$, relating it to conditions on F itself.

1.3.3 The Total Claim Amount in the Subexponential Case

In Section 1.3.2 we have stressed the importance of $\mathcal{S}$ for the estimation of ruin probabilities for large claims. From a mathematical point of view it is important that in the Cramér–Lundberg model, $1 - \psi(u)$ can be expressed as a compound geometric sum; see (1.11). The same methods used for proving

Theorem 1.3.6 yield an estimate of the total claim amount distribution for large claims. Indeed, in Section 1.1 we observed that, within the Cramér–Lundberg model, for all $t \geq 0$,

$$G_t(x) = P(S(t) \leq x) = \sum_{n=0}^{\infty} e^{-\lambda t} \, \frac{(\lambda t)^n}{n!} \, F^{n*}(x) \,, \quad x \geq 0 \,, \tag{1.33}$$

where $S(t) = \sum_{k=1}^{N(t)} X_k$ is the total (or aggregate) claim amount up to time t. The claim arrival process $(N(t))_{t \geq 0}$ in (1.33) is a homogeneous Poisson process with intensity $\lambda > 0$, hence

$$P(N(t) = n) = e^{-\lambda t} \, \frac{(\lambda t)^n}{n!} \,, \quad n \geq 0 \,. \tag{1.34}$$

The same calculation leading up to (1.33) would, for a general claim arrival process (still assumed to be independent of the claim size process (X_k)), yield the formula

$$G_t(x) = \sum_{n=0}^{\infty} p_t(n) \, F^{n*}(x) \,, \quad x \geq 0 \,, \tag{1.35}$$

where

$$p_t(n) = P(N(t) = n) \,, \quad n \in \mathbb{N}_0 \,,$$

defines a probability measure on $\mathbb{N}_0$. In the case of a subexponential df F the same argument as given for the proof of Theorem 1.3.6 yields the following result.

Theorem 1.3.9 (The total claim amount in the subexponential case)
Consider (1.35) *with* $F \in \mathcal{S}$, *fix* $t > 0$, *and suppose that* $(p_t(n))$ *satisfies*

$$\sum_{n=0}^{\infty} (1 + \varepsilon)^n \, p_t(n) < \infty \tag{1.36}$$

for some $\varepsilon > 0$. *Then* $G_t \in \mathcal{S}$ *and*

$$\overline{G}_t(x) \sim EN(t) \, \overline{F}(x) \,, \quad x \to \infty \,. \tag{1.37}$$

$\square$

Remarks: 1) Condition (1.36) is equivalent to the fact that the probability generating function $\sum_{n=0}^{\infty} p_t(n) s^n$ is analytic in a neighbourhood of $s = 1$.

2) The most general formulation of Theorem 1.3.9 is to be found in Cline [124], Theorem 2.13. $\square$

Example 1.3.10 (The total claim amount in the Cramér–Lundberg model)
Suppose $(N(t))$ is a homogeneous Poisson process with individual probabilities (1.34) so that trivially $p_t(n)$ satisfies (1.36). Then, for $F \in \mathcal{S}$,

$$\overline{G}_t(x) \sim \lambda t \overline{F}(x) \,, \quad x \to \infty \,. \qquad \square$$

Example 1.3.11 (The total claim amount in the negative binomial case)
The negative binomial process is a claim arrival process satisfying

$$p_t(n) = \binom{\gamma+n-1}{n} \left(\frac{\beta}{\beta+t}\right)^{\gamma} \left(\frac{t}{\beta+t}\right)^{n}, \quad n \in \mathbb{N}_0, \quad \beta, \gamma > 0. \quad (1.38)$$

Seal [572, 573] stresses that, apart from the homogeneous Poisson process, this process is the main realistic model for the claim number distribution in insurance applications. One easily verifies that

$$EN(t) = \gamma t/\beta, \quad \operatorname{var}(N(t)) = \gamma t(1+t/\beta)/\beta.$$

Denoting $q = \beta/(\beta+t)$ and $p = t/(\beta+t)$, one obtains from (1.38), by using Stirling's formula $\Gamma(x+1) \sim \sqrt{2\pi x}\,(x/e)^x$ as $x \to \infty$, that

$$p_t(n) \sim p^n\, n^{\gamma-1}\, q^{\gamma}/\Gamma(\gamma), \quad n \to \infty.$$

Therefore the condition (1.36) is fulfilled, so that for $F \in \mathcal{S}$,

$$\overline{G}_t(x) \sim \frac{\gamma t}{\beta}\,\overline{F}(x), \quad x \to \infty.$$

Recall that in the homogeneous Poisson case, $EN(t) = \lambda t = \operatorname{var}(N(t))$. For the negative binomial process,

$$\operatorname{var}(N(t)) = \left(1 + \frac{t}{\beta}\right) EN(t) > EN(t), \quad t > 0. \quad (1.39)$$

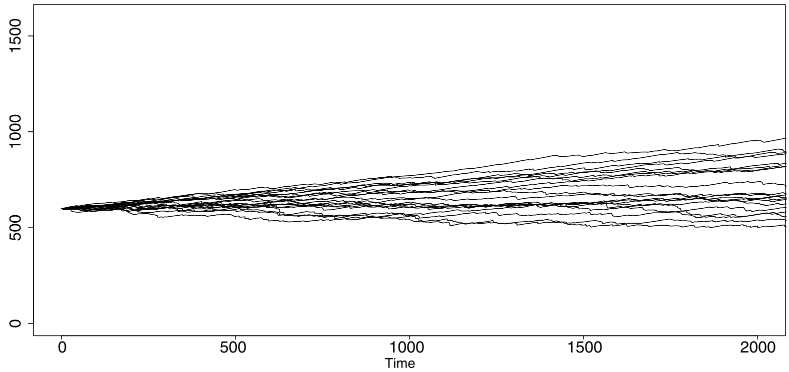
The condition (1.39) is referred to as *over–dispersion* of the process $(N(t))$; see for instance Cox and Isham [134], p. 12. As discussed in McCullagh and Nelder [449], p. 131, over–dispersion may arise in a number of different ways, for instance

(a) by observing a homogeneous Poisson process over an interval whose length is random rather than fixed,
(b) when the data are produced by a clustered Poisson process, or
(c) in behavioural studies and in accident–proneness when there is inter–subject variability.

It is mainly (c) which is often encountered in the analysis of insurance data. The features mentioned under (c) can be modelled by *mixed Poisson processes*. Their precise definition given below is motivated by the following example. Suppose Λ is a rv which is $\Gamma(\gamma, \beta)$ distributed with density

$$f_\Lambda(x) = \frac{\beta^\gamma}{\Gamma(\gamma)}\, x^{\gamma-1} e^{-\beta x}, \quad x > 0.$$

Then



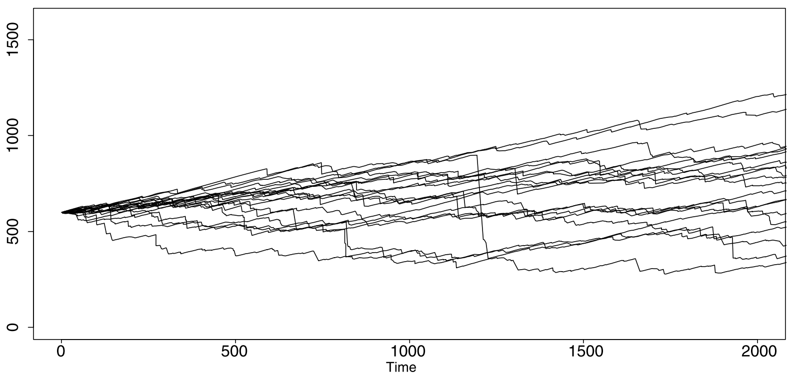


Figure 1.3.12 *Realisations of the risk process $(U(t))$ with linear premium income and total claim amount process $S(t) = \sum_{i=1}^{N(t)} X_i$, where $(N(t))$ is a negative binomial process. The claim size distribution is either exponential (top) or lognormal (bottom). Compared with Figure 1.1.2, the top figure clearly shows the over–dispersion effect. If we compare the bottom graph with the corresponding Figure 1.3.7 we notice the accumulation of many small claims.*

$$\int_0^\infty e^{-xt} \frac{(xt)^n}{n!} f_\Lambda(x)\, dx = \frac{\Gamma(n+\gamma)}{n!\Gamma(\gamma)} \left(\frac{\beta}{\beta+t}\right)^\gamma \left(\frac{t}{\beta+t}\right)^n , \quad n = 0, 1, 2, \ldots .$$

The latter formula equals $p_t(n)$ in (1.38). Hence we have obtained the negative binomial probabilities by *randomising* the Poisson parameter λ over a gamma distribution. This is exactly an example of what is meant by a mixed Poisson process. □

Definition 1.3.13 (Mixed Poisson process)
Suppose Λ is a rv with $P(\Lambda > 0) = 1$, and suppose $N = (N(t))_{t\geq 0}$ is a ho-

mogeneous Poisson process with intensity 1*, independent of* Λ*. The process* $(N(\Lambda t))_{t\geq 0}$ *is called* mixed Poisson. □

The rv Λ in the definition above can be interpreted as a random time change. Processes more general than mixed Poisson, for instance Cox processes, have belonged to the toolkit of actuaries for a long time. Mixed Poisson processes are treated in every standard text on risk theory. Recent textbook treatments are Grandell [284] and Panjer and Willmot [489]. The homogeneous Poisson process with intensity $\lambda > 0$ is obtained for Λ degenerate at λ, i.e. $P(\Lambda = \lambda) = 1$.

Notes and Comments

The class of subexponential distributions was independently introduced by Chistyakov [115] and Chover, Ney and Wainger [116], mainly in the context of branching processes. An early textbook treatment is given in Athreya and Ney [34], from which the proof of Lemma 1.3.5 is taken. Lemma 1.3.5(c) is attributed to Kesten. An independent introduction of $\mathcal{S}$ through questions in queueing theory is to be found in Borovkov [83, 84]; see also Pakes [488]. The importance of $\mathcal{S}$ as a useful class of heavy–tailed dfs in the context of applied probability in general, and insurance mathematics in particular, was realised early on by Teugels [617]. A recent survey paper is Goldie and Klüppelberg [274].

In the next section we shall prove that the condition $F_I \in \mathcal{S}$ is also satisfied for F lognormal. Whenever F is Pareto, it immediately follows that $F_I \in \mathcal{S}$. In these forms (i.e. Pareto, lognormal F), Theorem 1.3.6 has an interesting history, kindly communicated to us by Olof Thorin. In Thorin [623] the estimate

$$\psi(u) \sim k \int_u^\infty \overline{F}(x)\, dx\,, \qquad u \to \infty\,,$$

for some constant k was obtained for a wide class of distributions F assuming certain regularity conditions:

$$F(y) = \int_0^\infty \left(1 - e^{-yt}\right)\, V'(t)\, dt\,, \qquad y \geq 0\,,$$

with V' continuous, positive for $t > 0$, and having

$$V'(0) = 0 \quad \text{and} \quad \int_0^\infty V'(t)\, dt = 1\,.$$

An interesting special case is obtained by choosing $V'(t)$ as a gamma density with shape parameter greater than 1, giving the Pareto case. It was pointed out in Thorin and Wikstad [625] that the same method also works for the lognormal distribution. The Pareto case was obtained independently by von Bahr [633] and early versions of these results were previously discussed by Thorin at the 19th International Congress of Actuaries in Oslo (1972) and at the Wisconsin Actuarial Conference in 1971. Thorin also deals with the renewal case. Embrechts and Veraverbeke [218] obtained the full answer as presented in Theorem 1.3.8. In the latter paper these results were also formulated in the most general form for the renewal model allowing for real–valued, not necessarily positive claims. It turns out that also in that case, under the assumption $F_I \in \mathcal{S}$, the estimate $\psi(u) \sim \rho^{-1}\overline{F}_I(u)$ holds. In the renewal model however, we do not have the full equivalence relationships as in the Cramér–Lundberg case.

A recent overview concerning approximation methods for $\overline{G}_t(x)$ is given in Buchwalder, Chevallier and Klüppelberg [95]. The use of the fast Fourier transform method with special emphasis on heavy tails is highlighted in Embrechts, Grübel and Pitts [209]. A particularly important methodology for application is the so–called Panjer recursion method. For a discussion and further references on this topic; see Dickson [180] or Panjer and Willmot [489]. A light–tailed version of Example 1.3.11 is to be found in Embrechts, Maejima and Teugels [214]. An especially useful method in the light–tailed case is the so–called saddlepoint approximation; see Jensen [356] for an introduction including applications to risk theory. A very readable textbook treatment on approximation methods is Hipp and Michel [328]; see also Feilmeier and Bertram [232].

There are much fewer papers on statistical estimation of $\overline{G}_t$ than there are on asymptotic expansions. Clearly, one could work out a parametric estimation procedure or use non–parametric methods. The latter approach looks especially promising, as can be seen from Pitts [502, 503] and references therein.

1.4 Cramér–Lundberg Theory for Large Claims: a Discussion

1.4.1 Some Related Classes of Heavy–Tailed Distributions

In order to get direct conditions on F so that the heavy–tailed Cramér–Lundberg estimate in Theorem 1.3.6 holds, we first study some related classes of heavy–tailed dfs.

The class of *dominatedly varying distributions* is defined as

$$\mathcal{D} = \left\{ F \text{ df on } (0,\infty) : \limsup_{x\to\infty} \overline{F}(x/2)/\overline{F}(x) < \infty \right\}.$$

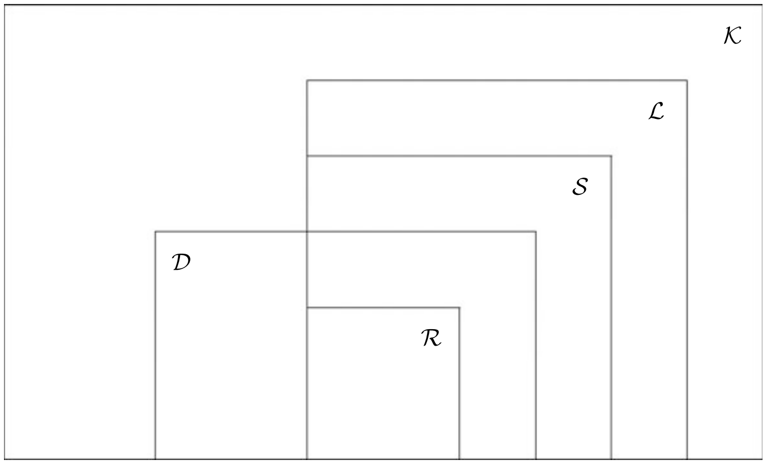


Figure 1.4.1 *Classes of heavy–tailed distributions.*

We have already encountered members of the following three families:

$$\mathcal{L} = \left\{F \text{ df on } (0,\infty) : \lim_{x\to\infty} \overline{F}(x-y)/\overline{F}(x) = 1 \quad \text{for all } y > 0\right\},$$

$$\mathcal{R} = \left\{F \text{ df on } (0,\infty) : \overline{F} \in \mathcal{R}_{-\alpha} \quad \text{for some } \alpha \geq 0\right\},$$

$$\mathcal{K} = \left\{F \text{ df on } (0,\infty) : \widehat{f}(-\varepsilon) = \int_0^\infty e^{\varepsilon x} dF(x) = \infty \quad \text{for all } \varepsilon > 0\right\}.$$

From the definition of slowly varying functions, it follows that $F \in \mathcal{L}$ if and only if $\overline{F} \circ \ln \in \mathcal{R}_0$. The following relations hold:

$$\boxed{\begin{array}{ll} \text{(a)} & \mathcal{R} \subset \mathcal{S} \subset \mathcal{L} \subset \mathcal{K} \quad \text{and} \quad \mathcal{R} \subset \mathcal{D}, \\ \text{(b)} & \mathcal{L} \cap \mathcal{D} \subset \mathcal{S}, \\ \text{(c)} & \mathcal{D} \not\subset \mathcal{S} \quad \text{and} \quad \mathcal{S} \not\subset \mathcal{D}. \end{array}}$$

The situation is summarised in Figure 1.4.1. A detailed discussion of these interrelationships is to be found in Embrechts and Omey [216]; see also Klüppelberg [387]. Most of the implications are easy, and indeed some of them we have already proved ($\mathcal{R} \subset \mathcal{S}$ in Corollary 1.3.2, $\mathcal{L} \subset \mathcal{K}$ in Remark 4 after the proof of Lemma 1.3.5). A mistake often encountered in the literature is the claim that $\mathcal{D} \subset \mathcal{S}$; the following df provides a counterexample.

Example 1.4.2 (The Peter and Paul distribution)
Consider a game where the first player (Peter) tosses a fair coin until it falls

head for the first time, receiving from the second player (Paul) 2^k Roubles, if this happens at trial k. The df of Peter's gain is

$$F(x) = \sum_{k:2^k \le x} 2^{-k}\,, \quad x \ge 0\,.$$

The problem underlying this game is the famous St. Petersburg paradox; see for instance Feller [234], Section X.4. It immediately follows that for all $k \in \mathbb{N}$, $\overline{F}(2^k-1)/\overline{F}(2^k) = 2$ so that $F \notin \mathcal{L}$ and a fortiori $F \notin \mathcal{S}$. On the other hand, one easily shows that $F \in \mathcal{D}$. For a full analysis see Goldie [271]. $\square$

The result $\mathcal{S} \ne \mathcal{L}$ is non–trivial; relevant examples are to be found in Embrechts and Goldie [204] and Pitman [501]. Concerning the relationship between $\mathcal{L}$ and $\mathcal{S}$, consider for $x \ge 0$,

$$\frac{\overline{F^{2*}}(x)}{\overline{F}(x)} = 1 + \int_0^x \frac{\overline{F}(x-y)}{\overline{F}(x)}\, dF(y)\,.$$

By definition, $F \in \mathcal{L}$ implies that for every y fixed the integrand above converges to 1. By the uniform convergence theorem for slowly varying functions (Theorem A3.2), this convergence holds also uniformly on compact y–intervals. In order however to interchange limits and integrals one needs some sort of uniform integrability condition (dominated convergence, monotonicity in $x, \ldots$). In general (i.e. for $F \in \mathcal{L}$) these conditions fail.

Let us first look at $\mathcal{S}$–membership in general. We have already established $\mathcal{R} \subset \mathcal{S}$ and $\mathcal{S} \subset \mathcal{L}$ (Lemma 1.3.5(a)), the latter implying that for all $\varepsilon > 0$, $\exp\{\varepsilon x\}\overline{F}(x) \to \infty$ as $x \to \infty$. From this it immediately follows that the exponential df $F(x) = 1 - \exp\{-\lambda x\}$ does *not* belong to $\mathcal{S}$. One could of course easily verify this directly, or use the fact that $\mathcal{S} \subset \mathcal{L}$ and immediately note that $F \notin \mathcal{L}$.

So by now we know that the dfs with power law tail behaviour (i.e. $F \in \mathcal{R}$) belong to $\mathcal{S}$. On the other hand, exponential distributions (and indeed dfs F with faster than exponential tail behaviour) do not belong to $\mathcal{S}$. What can be said about classes "in between", such as for example the important class of Weibull type variables where $\overline{F}(x) \sim \exp\{-x^\tau\}$ with $0 < \tau < 1$? Proposition A3.16, formulated in terms of the density f of F, the hazard rate $q = f/\overline{F}$ and the hazard function $Q(x) = \int_0^x q(y)dy$, immediately yields the following examples in $\mathcal{S}$. Note that, using the above notation, $\overline{F}(x) = \exp\{-Q(x)\}$.

Example 1.4.3 (Examples of subexponential distributions)

(a) Take F a Weibull distribution with parameters $0 < \tau < 1$ and $c > 0$, i.e.

$$\overline{F}(x) = e^{-c\,x^\tau}\,, \quad x \ge 0\,.$$

Then $f(x) = c\tau x^{\tau-1} e^{-cx^\tau}$, $Q(x) = cx^\tau$ and $q(x) = c\tau x^{\tau-1}$ which decreases to 0 if $\tau < 1$. We can immediately apply Proposition A3.16(b) since

$$x \mapsto e^{x\,q(x)} f(x) = e^{c\,(\tau-1)x^\tau}\, c\,\tau\, x^{\tau-1}$$

is integrable on $(0,\infty)$ for $0 < \tau < 1$. Therefore $F \in \mathcal{S}$.

(b) Using Proposition A3.16, one can also prove for

$$\overline{F}(x) \sim e^{-x(\ln x)^{-\beta}}, \quad x \to \infty, \quad \beta > 0,$$

that $F \in \mathcal{S}$. This example shows that one can come fairly close to exponential tail behaviour while staying in $\mathcal{S}$.

(c) At this point one could hope that for

$$\overline{F}(x) \sim e^{-x^\tau L(x)}, \quad x \to \infty, \quad 0 \le \tau < 1, \quad L \in \mathcal{R}_0,$$

F would belong to $\mathcal{S}$. Again, in this generality the answer to this question is *no*. One can construct examples of $L \in \mathcal{R}_0$ so that the corresponding F does not even belong to $\mathcal{L}$! An example for $\tau = 0$ was communicated to us by Charles Goldie; see also Cline [123] where counterexamples for $0 < \tau < 1$ are given. □

A particularly useful result is the following.

Proposition 1.4.4 (Dominated variation and subexponentiality)

(a) *If $F \in \mathcal{L} \cap \mathcal{D}$, then $F \in \mathcal{S}$.*

(b) *If F has finite mean μ and $F \in \mathcal{D}$, then $F_I \in \mathcal{L} \cap \mathcal{D}$. Consequently, because of* (a)*, $F_I \in \mathcal{S}$.*

Proof. (a) Because of (A.17), for $x \ge 0$,

$$\begin{aligned}\frac{\overline{F^{2*}}(x)}{\overline{F}(x)} &= 2\int_0^{x/2} \frac{\overline{F}(x-y)}{\overline{F}(x)}\, dF(y) + \frac{\left(\overline{F}(x/2)\right)^2}{\overline{F}(x)} \\ &= 2\int_0^{x/2} \frac{\overline{F}(x-y)}{\overline{F}(x)}\, dF(y) + o(1),\end{aligned}$$

where the $o(1)$ is a consequence of $F \in \mathcal{D}$. Now for all $0 \le y \le x/2$,

$$\frac{\overline{F}(x-y)}{\overline{F}(x)} \le \frac{\overline{F}(x/2)}{\overline{F}(x)},$$

so that because $F \in \mathcal{D}$, we can apply dominated convergence, yielding for $F \in \mathcal{L}$ the convergence of the integral to 1. Hence $F \in \mathcal{S}$.

(b) For ease of notation, and without loss of generality, we put $\mu = 1$. Since, for all $x \ge 0$,

$$\overline{F}_I(x) = \int_x^\infty \overline{F}(y)\,dy \geq \int_x^{2x} \overline{F}(y)\,dy \geq x\overline{F}(2x)\,, \tag{1.40}$$

it follows from $F \in \mathcal{D}$ that

$$\limsup_{x\to\infty} \frac{x\overline{F}(x)}{\overline{F}_I(x)} \leq \limsup_{x\to\infty} \frac{\overline{F}(x)}{\overline{F}(2x)} < \infty\,.$$

Moreover,

$$\begin{aligned}
\frac{\overline{F}_I(x/2)}{\overline{F}_I(x)} &= \frac{\displaystyle\int_{x/2}^\infty \overline{F}(y)\,dy}{\displaystyle\int_x^\infty \overline{F}(y)\,dy} = 1 + \frac{\displaystyle\int_{x/2}^x \overline{F}(y)\,dy}{\displaystyle\int_x^\infty \overline{F}(y)\,dy} \leq 1 + \frac{\overline{F}(x/2)x/2}{\overline{F}_I(x)} \\
&= 1 + 2^{-1}\frac{\overline{F}(x/2)}{\overline{F}(x)}\,\frac{x\overline{F}(x)}{\overline{F}_I(x)}\,,
\end{aligned}$$

whence

$$\limsup_{x\to\infty} \frac{\overline{F}_I(x/2)}{\overline{F}_I(x)} < \infty\,,$$

i.e. $F_I \in \mathcal{D}$. Take $y > 0$, then for $x \geq 0$

$$1 = \frac{\displaystyle\int_x^{x+y} \overline{F}(u)\,du + \int_{x+y}^\infty \overline{F}(u)\,du}{\overline{F}_I(x)}\,,$$

hence, by (1.40),

$$1 \leq \frac{y\overline{F}(x)}{\overline{F}_I(x)} + \frac{\overline{F}_I(x+y)}{\overline{F}_I(x)} \leq \frac{y\overline{F}(x)}{x\overline{F}(2x)} + \frac{\overline{F}_I(x+y)}{\overline{F}_I(x)}\,.$$

The first term in the latter sum is $o(1)$ as $x \to \infty$, since $F \in \mathcal{D}$. Therefore,

$$1 \leq \liminf_{x\to\infty} \frac{\overline{F}_I(x+y)}{\overline{F}_I(x)} \leq \limsup_{x\to\infty} \frac{\overline{F}_I(x+y)}{\overline{F}_I(x)} \leq 1\,,$$

i.e. $F_I \in \mathcal{L}$. □

1.4.2 The Heavy–Tailed Cramér–Lundberg Case Revisited

So far we have seen that, from an analytic point of view, the classes $\mathcal{R}$ and $\mathcal{S}$ yield natural models of claim size distributions for which the Cramér–Lundberg condition (1.13) is violated.

In Seal [573], for instance, the numerical calculation of $\psi(u)$ is discussed for various classes of claim size dfs. After stressing the fact that the mixed

Poisson process in general, and the homogeeous Poisson process and the negative binomial process in particular, are the *only* claim arrival processes which fit real insurance data, Seal continues by saying

Types of distributions of independent claim sizes are just as limited, for apart from the Pareto *and* lognormal *distributions, we are* not *aware that any has been fitted successfully to actual claim sizes in actuarial history.*

Although perhaps formulated in a rather extreme form, more than ten years later the main point of this sentence still stands; see for instance Schnieper [568], Benabbou and Partrat [59] and Ramlau–Hansen [522, 523] for a more recent account on this theme. Some examples of insurance data will be discussed in Chapter 6.

In this section we discuss $\mathcal{S}$–membership with respect to standard classes of dfs as given above. We stick to the Cramér–Lundberg model for purposes of illustration on how the new methodology works. Recall in the Cramér–Lundberg set–up the main result of Section 1.3, i.e. Theorem 1.3.6:

$$\boxed{\text{If } F_I \in \mathcal{S} \text{ then } \psi(u) \sim \rho^{-1}\, \overline{F}_I(u),\ u \to \infty\,.}$$

The exponential Cramér–Lundberg estimates (1.14), (1.16) under the small claim condition (1.13) yield surprisingly good estimates for $\psi(u)$, even for moderate to small u. The large claim estimate $\psi(u) \sim \rho^{-1}\overline{F}_I(u)$ is however mainly of theoretical value and can indeed be further improved; see the Notes and Comments. A first problem with respect to applicability concerns the condition $F_I \in \mathcal{S}$. A natural question at this point is:

(1) *Does $F \in \mathcal{S}$ imply that $F_I \in \mathcal{S}$?*

And, though less important for our purposes:

(2) *Does $F_I \in \mathcal{S}$ imply that $F \in \mathcal{S}$?*

It will turn out that, in general, the answer to both questions (unfortunately) is *no*. This leads us immediately to the following task:

Give sufficient conditions for F such that $F_I \in \mathcal{S}$.

Concerning the latter problem, there are numerous answers to be found in the literature. We shall discuss some of them. The various classes of dfs introduced in the previous section play an important role here.

An immediate consequence of Proposition 1.4.4 is the following result.

Corollary 1.4.5 (The Cramér–Lundberg theorem for large claims, III) *Consider the Cramér–Lundberg model with net profit condition $\rho > 0$ and $F \in \mathcal{D}$, then*

$$\psi(u) \sim \rho^{-1}\overline{F}_I(u)\,, \quad u \to \infty\,. \qquad \square$$

The condition $F \in \mathcal{D}$ is readily verified for all relevant examples; this is in contrast to the non–trivial task of checking $F_I \in \mathcal{S}$. It is shown in Seneta [575], Appendix A3, that any $F \in \mathcal{D}$ has the property that there exists a $k \in \mathbb{N}$ so that $\int_0^\infty x^k \, dF(x) = \infty$, i.e. there always exist divergent (higher) moments. It immediately follows from Karamata's theorem (Theorem A3.6) that $F \in \mathcal{R}$ implies $F_I \in \mathcal{R}$ and hence $F_I \in \mathcal{S}$. For detailed estimates in the heavy–tailed Cramér–Lundberg model see Klüppelberg [388]. In the latter paper, also, various sufficient conditions for $F_I \in \mathcal{S}$ are given in terms of the *hazard rate* $q(x) = f(x)/\overline{F}(x)$ for F with density f or the *hazard function* $Q = -\ln \overline{F}$; see also Cline [123].

Lemma 1.4.6 (Sufficient conditions for $F_I \in \mathcal{S}$)
If one of the following conditions holds, then $F_I \in \mathcal{S}$.

(a) $\limsup_{x\to\infty} \; x\, q(x) < \infty$,
(b) $\lim_{x\to\infty} q(x) = 0$, $\lim_{x\to\infty} x\, q(x) = \infty$, *and one of the following conditions holds*:
  (*1*) $\limsup_{x\to\infty} \; x\, q(x)/Q(x) < 1$,
  (*2*) $q \in \mathcal{R}_\delta$, $-1 \le \delta < 0$,
  (*3*) $Q \in \mathcal{R}_\delta$, $0 < \delta < 1$, *and q is eventually decreasing,*
  (*4*) *q is eventually decreasing to* 0, $q \in \mathcal{R}_0$ *and* $Q(x) - x\, q(x) \in \mathcal{R}_1$. □

In Klüppelberg [387], Theorem 3.6, a Pitman–type result (see Proposition A3.16) is presented, characterising $F_I \in \mathcal{S}$ for certain absolutely continuous F with hazard rate q decreasing to zero.

Example 1.4.7 (Examples of $F_I \in \mathcal{S}$)
Using Lemma 1.4.6(b)(2), it is not difficult to see that $F_I \in \mathcal{S}$ in the following cases:

– Weibull with parameter $\tau \in (0, 1)$
– Benktander–type–I and –II
– lognormal. □

Corollary 1.4.8 (The Cramér–Lundberg theorem for large claims, IV)
Consider the Cramér–Lundberg model with net profit condition $\rho > 0$ and F satisfying one of the conditions (a), (b) *of Lemma 1.4.6, then*

$$\psi(u) \sim \rho^{-1}\, \overline{F}_I(u)\,, \quad u \to \infty\,. \qquad \square$$

We still have left the questions (1) and (2) above unanswered. Concerning question (2) (does $F_I \in \mathcal{S}$ imply that $F \in \mathcal{S}$?), on using Proposition 1.4.4(b) we find that a straightforward modification of the Peter and Paul distribution yields an example of a df F with finite mean such that $F_I \in \mathcal{S}$ but $F \notin \mathcal{S}$.

For details see Klüppelberg [387]. The latter paper also contains a discussion of question (1) (does $F \in \mathcal{S}$ with finite mean imply that $F_I \in \mathcal{S}$?).

At this point, the reader may have become rather bewildered concerning the properties of $\mathcal{S}$. On the one hand, we have shown that it is the right class of dfs to consider in risk theory under large claim conditions; see Theorem 1.3.8, (c) implies (a). On the other hand, one has to be extremely careful in making general statements concerning $\mathcal{S}$ and its relationship to other classes of dfs.

For our immediate purposes it suffices to notice that for distributions F with finite mean belonging to the families: *Pareto, Weibull* ($\tau < 1$)*, lognormal, Benktander–type–I and –II, Burr, loggamma*,

$$\boxed{F \in \mathcal{S} \quad \text{and} \quad F_I \in \mathcal{S}\,.}$$

For further discussions on the applied nature of classes of heavy–tailed distributions, we also refer to Benabbou and Partrat [59], Conti [132], Hogg and Klugman [330] and Schnieper [568]. In the latter paper the reader will find some critical remarks on the existing gap between theoretical and applied usefulness of families of claim size distributions. It also contains some examples of relevant software for the actuary.

Notes and Comments

The results presented so far only give a first, though representative, account of ruin estimation in the heavy–tailed case. The reader should view them also as examples of how the class $\mathcal{S}$, and its various related classes, offer an appropriate tool towards a "heavy–tailed calculus".

Nearly all of the results can be extended. For instance Veraverbeke [630] considers the following model, first introduced by Gerber:

$$U_B(t) = u + ct - S(t) + B_t\,, \quad t \geq 0\,,$$

where u, c and $S(t)$ are defined within the Cramér–Lundberg model, and B is a Wiener process (see Section 2.4), independent of the process S. The process B can be viewed as describing small perturbations (i.e. B_t is distributed as a normal rv with mean 0 and variance $\sigma_B^2 t$) around the risk process U in (1.4). In [630], Theorem 1, it is shown that an estimate similar to the one obtained in the Cramér–Lundberg model for subexponential claim size distributions holds. These results can also be generalised to the renewal model set–up, as noted by Furrer and Schmidli [248]. Subexponentiality is also useful beyond these models as for instance shown by Asmussen, Fløe Henriksen and Klüppelberg [31]. In the latter paper, a risk process, evolving in an environment given by a Markov process with a finite state space, is studied. An appealing example of this type of process with exponential claim sizes is given in Reinhard [524].

Asymptotic estimates for the ruin probability change when the company receives interest on its reserves. For $F \in \mathcal{R}$ and a positive force of interest δ the corresponding ruin probability satisfies

$$\psi_\delta(u) \sim k_\delta \overline{F}(u)\,, \quad u \to \infty\,,$$

i.e. it is tail–equivalent to the claim size df itself. This has been proved in Klüppelberg and Stadtmüller [399]. The case of subexponential claims has been treated in Asmussen [29].

Concerning the definition of $\mathcal{S}$, there is no *a priori* reason for assuming that the limit of $\overline{F^{2*}}(x)/\overline{F}(x)$ equals 2; an interesting class of distributions results from allowing this limit to be any value greater than 2.

Definition 1.4.9 *A df F on $(0,\infty)$ belongs to the class $\mathcal{S}(\gamma)$, $\gamma \geq 0$, if*

(a) $\lim_{x\to\infty} \overline{F^{2*}}(x)/\overline{F}(x) = 2d < \infty$,

(b) $\lim_{x\to\infty} \overline{F}(x-y)/\overline{F}(x) = e^{\gamma y}, \quad y \in \mathbb{R}$. □

One can show that $d = \int_0^\infty e^{\gamma y}\, dF(y) = \widehat{f}(-\gamma)$, so that $\mathcal{S} = \mathcal{S}(0)$. These classes of dfs turn out to cover exactly the situations illustrated in Figures 1.2.4(2) and (3), in between the light–tailed Cramér–Lundberg case and the heavy–tailed (subexponential) case. A nice illustration of this, using the class of generalised inverse Gaussian distributions, is to be found in Embrechts [201]; see also Klüppelberg [389] and references therein.

For a critical assessment of the approximation $\psi(u) \sim k\overline{F}_I(u)$ for some constant k and $u \to \infty$ see De Vylder and Goovaerts [179]. Further improvements can be obtained only if conditions beyond $F_I \in \mathcal{S}$ are imposed. One such set of conditions is higher–order subexponentiality, or indeed higher–order regular variation. In general $G \in \mathcal{S}$ means that $\overline{G^{2*}}(x) \sim 2\overline{G}(x)$ for $x \to \infty$; higher–order versions of $\mathcal{S}$ involve conditions on the asymptotic behaviour of $\overline{G^{2*}}(x) - 2\overline{G}(x)$ for $x \to \infty$. For details on these techniques the reader is referred to Omey and Willekens [486, 487], and also Bingham et al. [72], p. 185. With respect to the heavy–tailed ruin estimate $\psi(u) \sim \rho^{-1}\overline{F}_I(u)$, second–order assumptions on F lead to asymptotic estimates for $\psi(u) - \rho^{-1}\overline{F}_I(u)$ for $u \to \infty$. A numerical comparison of such results, together with a detailed simulation study of rare events in insurance, is to be found in Asmussen and Binswanger [30] and Binswanger [74].