

## SOME ASPECTS OF INSURANCE MATHEMATICS\*

P. EMBRECHTS<sup>†</sup> AND C. KLÜPPELBERG<sup>†</sup>

**Abstract.** This paper was written on the invitation of the editors of this journal and is, in general, a review. Its aim is to show how the theory of probability and mathematical statistics are applied for solving problems of the insurance field. In § 2 a description of the basic probabilistic models in risk theory is given and the problems connected with the structure of the insurance payments are considered. Section 3 is devoted to income aspect of insurance activities and the sizes of the insurance payments are under consideration. Statistical aspects of insurance mathematics are considered in § 4.

**Key words.** insurance mathematics, risk theory, Poisson and compound Poisson models, Cox processes, the Cramér–Lundberg theorem, martingales, optional stopping theorem, piecewise deterministic Markov processes, general insurance risk model

**1. Introduction.** Approximately 300 years ago Edward Lloyd, the owner of a coffee house in London, realized the need for insurance covering transport risks (shipping industry). Today Lloyd's of London has a premium income of more than \$ 20 million each workday. The existence of risks to humans, property, environment, etc., has resulted in the establishment of an enormous world industry that offers financial cover to losses due to risk exposure. A key item in all of this is the inherent presence of randomness. The goal of our paper is to show how probability theory and mathematical statistics are used to solve problems in the realm of insurance. The amalgamation of relevant theory from diverse fields has now resulted in the emergence of a full-bodied branch of science called "insurance mathematics." There is no way in which we will be able to review the entire field. A glance at the variety of topics included in this theory reveals such names as risk theory, life insurance mathematics, premium rating, credibility theory, pension funding, solvency studies, population theory, IBNR modeling, reserving, insurance and mathematical theory of finance, reinsurance, survival modeling, and loss distributions. Books have been written on each of these topics. This clearly shows the vast amount of material now available under the overall umbrella of insurance mathematics. To a great extent, however, classical topics like risk theory (from a probabilistic point of view) and credibility theory (from a statistical one) still form a core of methodology on which many other techniques are built.

In § 2 of the paper, we set up the basic probabilistic models in risk theory and concentrate in particular on the claim structure. Various approximations for the aggregate claim distribution will be given, concentrating both on the claim arrival process as well as on the claim size process. Special attention will be given to the modeling of large claims (catastrophic events). The latter is becoming more and more a central theme to be considered in various classical models. In § 3 we turn to the income side of the company by looking at the definition of premiums. By splitting the gross premium in a net part and a so-called loading component, a natural way to include martingale methodology will be given. Once these two aspects (assets (income) versus liabilities (outgoing)) have been discussed we turn to the basic risk process and indicate how modern martingale theory yields various generalizations of the classical Cramér–Lundberg model, allowing for economic factors (such as inflation,

---

\*Received by the editors December 3, 1992.

<sup>†</sup>Department of Mathematics, ETH-Zentrum, CH-8092 Zürich, Switzerland.

borrowing, dividend payment, investment, etc.) to be modeled. The incorporation of large claim structures in these models will be discussed, as well as the use of diffusion approximations. In § 4 we turn to some of the statistical aspects related to the models treated earlier. Such themes as claim size fitting, the estimation of the adjustment coefficient, ruin estimation and applications of bootstrap techniques to insurance data will be included. Some recent work on the modeling of large claims via a large claim index also figures. Although we concentrate to a large extent on parts of the theory that are closely related to our own research interests, throughout the paper we try to give hints for further reading so that the reader still obtains an overall picture on how insurance mathematics is currently evolving.

There are a few books that attempt to cover insurance mathematics as a whole. One interesting attempt is [12]. Over the recent years, many books have been written on risk theory. The classic by Bühlmann [15] is still worth reading. The more modern approach based on martingales is to be found in [41] and [45]. The analysis of risk processes on the basis of queueing methodology is nicely presented in [6]. The link between the more practical background to risk theory and its mathematical counterpart is well-presented in [9]. From a teaching point of view, [48], [72], and [52] are to be recommended. The latter text investigates in more detail relevant statistical questions in nonlife insurance mathematics. An interesting step further in this direction is [43] where the main topics are ordering of risks, credibility theory, and IBNR-techniques. An excellent recent review article on credibility theory is [57]. The more data-analytic (statistical) aspects of nonlife insurance are discussed in [54], whereas the more numerical analysis aspects are summarized in [37]. The above list clearly indicates that the nonlife side of insurance mathematics is well covered in textbook format. With respect to life insurance, the main transitions go from (classical) deterministic theory, over probabilistically founded presentations to those using the theory of stochastic processes for a more dynamic modeling. A useful textbook in the spirit of the middle approach, keeping the link with classical terminology, is [42]. The more dynamic modeling, via the theory of stochastic processes, has not yet entered the textbook stage. One of the many interesting papers on this topic is [53]. See also [66] and [77]. Finally, the transition from classical actuaries of the “first and second kind” toward actuaries of the “third kind” (Bühlmann terminology) manifests itself in the increasing effort to stress the modeling of the asset side of the risk process using the modern theory of mathematical finance. We shall not deal with these matters in our paper; an annotated list of references in finance, useful for insurance mathematicians, is to be found in [3]. For a discussion on the interplay between insurance and finance, see [71] and [64].

Finally, we would like to stress that there are various research journal in the field of insurance mathematics. The following are the more interesting ones for probabilists and statisticians: *The Journal of Risk and Insurance*, *Insurance: Mathematics and Economics*, *The ASTIN Bulletin*, *The Scandinavian Actuarial Journal* (previously known as: *Skandinavisk Aktuarietidskrift*), *Transactions of the International Congresses of Actuaries*, *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*.

**2. The accumulated claims process.** We consider the collective model where the whole portfolio of a certain insurance business represents the source of risk. We formulate the classical risk model in § 2.1 and study generalizations of the underlying point process in § 2.2. This approach is mainly based on the excellent monograph [45].

**2.1. The compound Poisson model.** In the classical risk model the claim

sizes  $(Y_k)_{k \in \mathbf{N}}$  are a sequence of independent and identically distributed random variables having common distribution function  $F$  with  $F(0) = 0$ , mean value  $\mu$ , and variance  $\sigma^2$ . The point process  $N = (N(t))_{t \geq 0}$  with  $N(0) = 0$ , which represents the number of claims in the time interval  $[0, t]$ , is assumed to be a (homogeneous) Poisson process with intensity  $\kappa$ . Furthermore,  $N$  and  $(Y_k)_{k \in \mathbf{N}}$  are assumed to be independent. Then  $S(t) = \sum_{i=1}^{N(t)} Y_i$  represents the *accumulated claims up to time t* and has distribution function

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

where  $F^{n*}(x) = \mathbf{P}\{\sum_{i=1}^n Y_i \leq x\}$  is the  $n$ -fold convolution of  $F$ .

$N$  is a point process on  $\mathbf{R}^+$  and at each point of  $N$  the company must pay out a stochastic amount of money. The assumption that  $N$  is a Poisson process is equivalent to independent identically distributed exponential interarrival times  $(T_k)_{k \in \mathbf{N}}$  of the claims.

The *risk process*  $U$  is defined by

$$U(t) = u + ct - S(t),$$

where  $u$  is the *initial capital* and  $c > 0$  is the constant *premium rate*.

For  $N$  with intensity  $\kappa$  we have

$$\mathbf{E} N(t) \equiv \kappa t.$$

Then the “profit” over the interval  $(0, t]$  is

$$Q(t) = ct - S(t)$$

with

$$\mathbf{E} Q(t) = ct - \mathbf{E} N(t) \mathbf{E} Y_1 = t(c - \kappa\mu).$$

The *relative safety loading*  $\rho$  is defined by

$$\rho = \frac{c - \kappa\mu}{\kappa\mu} = \frac{c}{\kappa\mu} - 1.$$

If the risk process  $U$  has positive safety loading  $\rho > 0$ , then  $U(t)$  almost surely drifts to  $+\infty$ . The condition  $\rho > 0$  is also restated as the so-called *net profit condition*  $c > \kappa\mu$ .

As a measure for the long-term stability of the risk process the ruin probability  $\psi(u)$  as a function of the initial capital  $u$  has been considered. It is defined as

$$\psi(u) = \mathbf{P}(U(t) < 0 \text{ for some } t > 0).$$

A similar definition can be given for a ruin in a finite time interval  $[0, T]$ ,

$$\psi(u, T) = \mathbf{P}(U(t) < 0 \text{ for some } t \leq T).$$

Stability is achieved by the requirement that  $\psi(u, T)$  does not fall below a certain level  $\varepsilon$  (e.g., 0.1%), for a given value of  $T$ . For most of this paper, we restrict our attention, however, to  $\psi(u)$ , i.e.,  $T = \infty$ . This probability of ruin is calculable for positive safety loading  $\rho$ . An asymptotic expression for  $u \rightarrow \infty$  can be obtained by solving an integral equation either by means of the Wiener–Hopf technique, which goes back to Cramér [17], or using renewal arguments as in [39]. An upper bound can be obtained

by martingale methods [40], [41]. These results are exponential type estimates and require the existence of exponential moments. If no exponential moments exist — a case that corresponds to the occurrence of large claims — a special theory can be developed as in [34] and we review these results in § 2.5.

We first derive the famous result of Cramér using the renewal arguments from [39]. Our approach follows closely Grandell's [45].

Denote by

$$\varphi(u) = 1 - \psi(u)$$

the nonruin probability and note that  $\varphi(u) = 0$  for  $u < 0$ . Furthermore,

$$\varphi(u) = \mathbf{P}(U(t) \geq 0, \forall t \geq 0) = \mathbf{P}(S(t) - ct \leq u, \forall t \geq 0)$$

and hence  $\varphi$  is increasing in  $u$ . Since  $\mathbf{E}Q(t) = t(c - \kappa\mu) > 0$ ,  $Q(t)/t \xrightarrow{\text{a.s.}} c - \kappa\mu > 0$ ,  $t \rightarrow \infty$ , by the strong law of large numbers (SLLN). This implies that  $Q(t) > 0 \forall t > T$ , where  $T$  is some random variable that is a function of  $N$  and  $(Y_k)_{k \in \mathbf{N}}$ . Since only finitely many claims occur before time  $T$ , it follows that  $\inf_{t > 0} Q(t)$  is finite with probability one and thus  $\varphi(\infty) = 1$ .

Moreover, by a renewal argument,  $\varphi$  satisfies the renewal equation

$$(2.1) \quad \varphi(u) = \varphi(0) + \frac{\kappa}{c} \int_0^u \varphi(u-z) \bar{F}(z) dz,$$

where  $\bar{F} = 1 - F$ , and if we denote

$$F_I(x) = \mu^{-1} \int_0^x \bar{F}(z) dz,$$

the integrated tail distribution (2.1) is equivalent to

$$(2.2) \quad \varphi(u) = \varphi(0) + \frac{1}{1+\rho} \varphi * F_I(u),$$

where  $*$  denotes the convolution. By monotone convergence it follows from (2.1) that

$$(2.3) \quad \varphi(\infty) = \varphi(0) + \frac{1}{1+\rho} \varphi(\infty),$$

which gives

$$(2.4) \quad \varphi(0) = \frac{\rho}{1+\rho}.$$

Using Laplace transforms and (2.4), an analytical solution of (2.2) can be obtained in the form

$$\varphi(u) = \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} \left( \frac{1}{1+\rho} \right)^n F_I^{n*}(u).$$

It is easy to check that this is the unique solution of (2.2). Only for special choices of the claim size distribution  $F$  is it possible to get an explicit expression.

*Example.* For exponentially distributed claims, differentiation of (2.1) leads to

$$\psi(u) = 1 - \varphi(u) = \frac{1}{1+\rho} \exp \left\{ -\frac{\rho}{\mu(1+\rho)} u \right\},$$

where we also have used (2.3) and (2.4).

The asymptotic behavior of the ruin probability  $\psi(u)$  for an initial reserve  $u \rightarrow \infty$  can be obtained by renewal limit theorems. To this end we rewrite (2.1) as follows:

$$(2.5) \quad \psi(u) = \frac{\kappa}{c} \int_u^\infty \bar{F}(z) dz + \frac{\kappa}{c} \int_0^u \psi(u-z) \bar{F}(z) dz.$$

Since

$$\frac{\kappa}{c} \int_0^\infty \bar{F}(z) dz = \frac{\kappa\mu}{c} < 1,$$

(2.5) is a defective renewal equation. The defect can be removed by an Esscher transformation (if the appropriate exponential moment exists). If there exists a constant  $R$  such that

$$(2.6) \quad \frac{\kappa}{c} \int_0^\infty e^{Rz} \bar{F}(z) dz = 1,$$

then  $(\kappa/c)e^{Rz}\bar{F}(z)$  is the density of a proper distribution function and multiplication of (2.5) by  $e^{Ru}$  yields

$$e^{Ru}\psi(u) = \frac{\kappa}{c} e^{Ru} \int_u^\infty \bar{F}(z) dz + \frac{\kappa}{c} \int_0^u e^{R(u-z)} \psi(u-z) e^{Rz} \bar{F}(z) dz.$$

This is a proper renewal equation and from Smith's key renewal theorem (see [39]) the following famous result follows.

**THEOREM 2.1.1** (Cramér–Lundberg theorem). *Suppose there exists a constant  $R$  such that*

$$\int_0^\infty e^{Rz} \bar{F}(z) dz = \frac{c}{\kappa}.$$

If

$$\mu^* = \frac{\kappa}{c} \int_0^\infty z e^{Rz} \bar{F}(z) dz < \infty,$$

then

$$(2.7) \quad \psi(u) \sim \frac{\rho}{(1+\rho)R\mu^*} e^{-Ru}, \quad u \rightarrow \infty,$$

and if  $\mu^* = \infty$ , then

$$(2.8) \quad \psi(u) = o(e^{-Ru}), \quad u \rightarrow \infty.$$

For the above example of exponentially distributed claims  $R$  can easily be calculated, and in this case the Cramér–Lundberg approximation is exact.

The number  $R$  in (2.6) is called *Lundberg coefficient* or *adjustment coefficient*. An upper bound for the ruin probability is given by the *Lundberg inequality*, which states that

$$(2.9) \quad \psi(u) \leq e^{-Ru}, \quad \forall u \in [0, \infty),$$

where  $R$  is again the Lundberg coefficient given as solution of (2.6). This can be proved in different ways. If we define  $\psi_n(u)$  as the probability of ruin up to the  $(n+1)$ st claim, then  $\psi(u) = \lim_{n \rightarrow \infty} \psi_n(u)$  and it suffices to show that  $\psi_n(u) \leq e^{-Ru}$  for all  $n \in \mathbf{N}$ . This can be done by induction. The most elegant way to prove (2.9) is, however, via a martingale approach that goes back to Gerber [40]. Martingale methods also allow one to consider various generalizations concerning the underlying claim number

process as well as to take inflation and interest into account. An interesting recent summary is to be found in [46].

Definitions and results on martingale theory can be found in any standard textbook (see, e.g., [24], [78]). We consider martingales (respectively, super- or submartingales)  $M = (M(t))_{t \geq 0}$  with respect to a given filtration  $\mathbf{F} = (\mathcal{F}(t))_{t \geq 0}$ . In most cases,  $\mathbf{F}$  will be the natural filtration  $(\mathcal{F}^U(t))_{t \geq 0}$  of the underlying risk process. Moreover, all processes  $M$  we consider will be càdlàg.

The basic tool to prove Lundberg type inequalities is the optional stopping theorem which states that, for any right-continuous  $\mathbf{F}$ -martingale  $M$  ( $\mathbf{F}$ -supermartingale,  $\mathbf{F}$ -submartingale) and any bounded stopping time  $T$ ,

$$\mathbf{E}[M_T | \mathcal{F}_s] = M_s \quad \forall T: T \geq s \quad \text{a.s.}$$

$$(\mathbf{E}[M_T | \mathcal{F}_s] \leq M_s, \quad \mathbf{E}[M_T | \mathcal{F}_s] \geq M_s).$$

Denote by

$$\hat{f}(s) = \int_0^\infty e^{-sx} dF(x)$$

the Laplace–Stieltjes transform of the distribution function  $F$ , then

$$(2.10) \quad \mathbf{E}[e^{-sQ(t)}] = e^{-sct} \mathbf{E}[e^{sS(t)}] = e^{tg(s)},$$

where  $g(s) = \kappa(\hat{f}(-s) - 1) - sc$ .

Now let  $\mathbf{F}^Q$  be the natural filtration for  $(Q(t))_{t \geq 0}$ , i.e.,  $\mathcal{F}_t^Q = \sigma\{Q(s); s \leq t\}$  and set

$$T_u = \inf \{t \geq 0; u + Q(t) < 0\},$$

then  $T_u$  is a stopping time for  $(Q(t))_{t \geq 0}$  and  $\psi(u) = \mathbf{P}(T_u < \infty)$ . Set, for given  $r$  in the domain of convergence of  $\hat{f}$ ,

$$(2.11) \quad M_u(t) = \frac{e^{-r(u+Q(t))}}{e^{tg(r)}}.$$

Then, for all  $t \geq s$ ,

$$\begin{aligned} \mathbf{E}[M_u(t) | \mathcal{F}_s^Q] &= \mathbf{E}\left[\frac{e^{-r(u+Q(s))}}{e^{sg(r)}} \frac{e^{-r(Q(t)-Q(s))}}{e^{(t-s)g(r)}} \middle| \mathcal{F}_s^Q\right] \\ &= M_u(s) \mathbf{E}\left[\frac{e^{-r(Q(t)-Q(s))}}{e^{(t-s)g(r)}} \middle| \mathcal{F}_s^Q\right] = M_u(s) \end{aligned}$$

by (2.10) and the fact that  $Q$  has independent increments; hence  $(M_u(t))_{t \geq 0}$  is an  $\mathbf{F}^Q$ -martingale.

For any  $t_0 < \infty$ ,  $T_u \wedge t_0$  is a bounded stopping time and by the optional stopping theorem we obtain

$$\begin{aligned} e^{-ru} &= M_u(0) = \mathbf{E}[M_u(t_0 \wedge T_u) | \mathcal{F}_0^Q] = \mathbf{E}[M_u(t_0 \wedge T_u)] \\ &= \mathbf{E}[M_u(t_0 \wedge T_u) | T_u \leq t_0] \mathbf{P}(T_u \leq t_0) \\ &\quad + \mathbf{E}[M_u(t_0 \wedge T_u) | T_u > t_0] \mathbf{P}(T_u > t_0) \\ (2.12) \quad &\geq \mathbf{E}[M_u(T_u) | T_u \leq t_0] \mathbf{P}(T_u \leq t_0). \end{aligned}$$

Since  $T_u \leq t_0$  implies that  $U(T_u) = u + Q(T_u) \leq 0$ , it follows that

$$\mathbf{P}(T_u \leq t_0) \leq \frac{e^{-ru}}{\mathbf{E}[M_u(T_u) | T_u \leq t_0]} \leq \frac{e^{-ru}}{\mathbf{E}[e^{-T_u g(r)} | T_u \leq t_0]} \leq e^{-ru} \sup_{0 \leq t \leq t_0} e^{tg(r)},$$

and for  $t_0 \rightarrow \infty$  we obtain

$$\psi(u) \leq e^{-ru} \sup_{t \geq 0} e^{tg(r)}.$$

To obtain this inequality as sharp as possible we would like to take  $r$  large, but such that  $\sup_{t \geq 0} e^{tg(r)} < \infty$ . If  $R$  denotes this optimal value for  $r$ , then

$$R = \sup \{r; g(r) \leq 0\}$$

and hence  $R$  is the Lundberg coefficient defined in (2.6) and we obtain (2.9).

It is clear that the above calculation can be made for more general (so-called additive) models for  $Q$  (see [46]).

For a more refined treatment we again consider (2.12). With  $r = R$  we have

$$(2.13) \quad \begin{aligned} e^{-Ru} &= \mathbf{E} \left[ e^{-R(u+Q(T_u))} | T_u \leq t_0 \right] \mathbf{P}(T_u \leq t_0) \\ &+ \mathbf{E} \left[ e^{-R(u+Q(t_0))} | T_u > t_0 \right] \mathbf{P}(T_u > t_0). \end{aligned}$$

From this we obtain

$$\begin{aligned} 0 &\leq \mathbf{E} \left[ e^{-R(u+Q(t_0))} | T_u > t_0 \right] \mathbf{P}(T_u > t_0) = \mathbf{E} \left[ e^{-R(u+Q(t_0))} I(T_u > t_0) \right] \\ &\leq \mathbf{E} \left[ e^{-R(u+Q(t_0))} I(u + Q(t_0) \geq 0) \right]. \end{aligned}$$

Since

$$0 \leq \exp \left\{ -R(u + Q(t_0)) \right\} I(u + Q(t_0) \geq 0) \leq 1$$

and  $(Q(t))_{t \geq 0}$  almost surely drifts to  $+\infty$ , it follows by dominated convergence that

$$\lim_{t_0 \rightarrow \infty} \mathbf{E} \left[ e^{-R(u+Q(t_0))} | T_u > t_0 \right] \mathbf{P}(T_u > t_0) = 0,$$

and thus by (2.13)

$$(2.14) \quad \psi(u) = \frac{e^{-R(u)}}{\mathbf{E} \left[ e^{-R(u+Q(T_u))} | T_u < \infty \right]}.$$

Unfortunately, the denominator is only tractable in very few cases as, e.g., for exponentially distributed claims because of the lack of memory property. The above presented general consideration presents only one aspect of applications of martingale techniques to insurance. The method is very powerful in obtaining inequalities but less so for approximations of the Cramér–Lundberg type (2.7). Moreover, this whole theory rests on the ability to spot the relevant martingale (as  $M_u$  above). In §3 we shall present a general method for solving the latter problem.

**2.2. Generalizations of the claim number process.** The Poisson process assumption imposes certain constraints on the model. For example, stationarity implies that the size of the portfolio cannot increase (or decrease). Note that few managers would believe in a model that does not allow for a growth of their business. We refer to this situation as *size fluctuations*. Furthermore, there may be fluctuations in the

underlying risk. Typical examples are automobile and fire insurance (see, e.g., [9, Fig. 2.7.2]). We shall refer to this phenomenon as *risk fluctuation* ([45, Chap. 2]). In this subsection we generalize the claim number process  $(N(t))_{t \geq 0}$  and consider more general point processes.

Traditionally, the mixed Poisson model is considered if short-time risk fluctuations are taken into account. For each point  $t$  the claim numbers  $N(t)$  up to time  $t$  are Poisson with parameter  $\Lambda t$  for some positive random variable  $\Lambda$ .  $\Lambda$  is called a *structure variable* and its distribution function  $H(\lambda) = \mathbf{P}(\Lambda \leq \lambda)$  a *structure distribution*. Then

$$\mathbf{P}(N(t) = k) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^k}{k!} dH(\lambda),$$

and  $(N(t))_{t \geq 0}$  is called a *mixed Poisson process*.

The most important case from a practical point of view is the mixed Poisson process where  $H$  is  $\Gamma(q, \kappa q)$  for some  $q > 0$ ,  $\kappa > 0$ . Then  $N(t)$  is negative binomial with parameters  $\kappa q$  and  $q/(q+t)$  and  $(N(t))_{t \geq 0}$  is called a *Pólya process*. The corresponding risk model is also known as the Pólya–Eggenberger model and has first been mentioned in 1923 as contagion model ([38, § V. 2]). If one compares the total claim amount up to time  $t$  for the Poisson model, then for equal means the variance of the Pólya model is of course bigger than in the Poisson model.

We have taken the following considerations from [16]. They clearly show that fluctuations in risk have to be compensated for by stochastic premiums. Suppose  $(N(t))_{t \geq 0}$  is mixed Poisson with structure variable  $\Lambda$  with unbounded support such that  $\mathbf{E}\Lambda < \infty$  and structure distribution  $H$ . Then the distribution function of the total claims up to time  $t$  is given by

$$G_t(x) = \sum_{k=0}^{\infty} p_k(t) F^{n*}(x),$$

where

$$p_k(t) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^k}{k!} dH(\lambda).$$

The ruin probability can be calculated by conditioning and we obtain

$$\psi(u) = \int_0^\infty \psi(u | \Lambda = \lambda) dH(\lambda).$$

Now suppose the premium rate  $c$  is constant, then for  $\lambda \geq c/\mu$  we have  $\psi(u | \Lambda = \lambda) = 1$ , and hence

$$\psi(u) \geq \mathbf{P}(\Lambda \geq c/\mu) > 0 \quad \text{for all } u.$$

This implies that any insurer who does not constantly adjust his premium rate  $c$  according to the risk fluctuations deserves to be ruined.

As a consequence we consider the premium process

$$c(t) = (1 + \rho)\lambda(t)\mu,$$

where we estimate the process  $(\lambda(t))_{t \geq 0}$  by

$$(2.15) \quad \hat{\lambda}(t) = \mathbf{E}[\Lambda | N(s); s \leq t] = \frac{\int_0^\infty \lambda \frac{e^{-\lambda t} (\lambda t)^{N(t)}}{N(t)!} dH(\lambda)}{\int_0^\infty \frac{e^{-\lambda t} (\lambda t)^{N(t)}}{N(t)!} dH(\lambda)}.$$



Then the surplus process is

$$U(t) = u + c(t) - S(t) = u + (1 + \rho)\mu\lambda(t) - S(t).$$

In the important Pólya model we obtain

$$\begin{aligned}\hat{\lambda}(t) &= \frac{q + N(t)}{\kappa q + t} = \frac{\kappa q}{\kappa q + t} \frac{1}{\kappa} + \frac{t}{\kappa q + t} \frac{N(t)}{t} \\ &= (1 - R(t)) \mathbf{E} \Lambda + R(t) \frac{N(t)}{t}.\end{aligned}$$

The last formula can be interpreted as the weighted mean of expectation and of the observed mean; it gives one instance of the celebrated credibility formula [14].

The simplest way to take size fluctuations of the portfolio into account is to let  $(N(t))_{t \geq 0}$  be an inhomogeneous Poisson process. So suppose  $A(t)$  is a continuous nondecreasing function with  $A(0) = 0$  and  $A(t) < \infty$  for each  $t < \infty$  and  $N(t)$  is Poisson distributed with mean  $A(t)$ , then  $(N(t))_{t \geq 0}$  is called *inhomogeneous Poisson process with intensity measure A*.

By a fundamental theorem on random time change (see, e.g., [13, p. 41]), the Poisson process  $\tilde{N} = N \circ A^{-1}$ , where  $A^{-1}$  is the inverse function of  $A$ , is a standard Poisson process, i.e., a Poisson process with intensity  $\kappa = 1$ . The function  $A^{-1}$  is called *operational time scale*. On the other hand, starting with a standard Poisson process  $\tilde{N}$ , we can construct an inhomogeneous Poisson process  $N$  with given intensity measure  $A$  by defining  $N = \tilde{N} \circ A$ . If the function  $A$  is absolutely continuous with Lebesgue density  $\alpha$ , then  $\alpha$  is called *an intensity function*.

To model size fluctuations, one takes, e.g.,  $\alpha(t)$  proportional to the number of policies at time  $t$ . By considerations similar to those for the mixed Poisson process, we choose for a fixed safety loading  $\rho$  also the premium rate proportional to  $\alpha$ , i.e.,  $dc(t) = (1 + \rho)\mu\alpha(t) dt$ . The surplus process is then of the form

$$(2.16) \quad U(t) = u + (1 + \rho)\mu \int_0^t \alpha(s) ds - S(t) = u + (1 + \rho)\mu A(t) - S(t).$$

A further generalization of the inhomogeneous Poisson process is given by the so-called Cox models. They also include mixed Poisson processes and we shall calculate ruin probabilities for this more general class of risk processes in the next subsection.

**2.3. Cox processes.** The time scale  $A^{-1}$  of the inhomogeneous Poisson process can be defined for a very wide class of point processes. We shall apply it for Cox processes, which seem to form a natural class to model risk and size fluctuations. As became obvious in § 2.2 it is mathematically irrelevant for the intensity function  $\alpha$  to fluctuate as long as these fluctuations are compensated by the premiums. We first start with the definition of Cox processes.

Suppose  $\Lambda = (\Lambda(t))_{t \geq 0}$  is a random measure; that is a.s.  $\Lambda(0) = 0$ ,  $\Lambda(t) < \infty$ , for  $t < \infty$ , and  $\Lambda$  has nondecreasing trajectories. Furthermore, let  $\tilde{N}$  be a standard Poisson process, independent of  $\Lambda$ . Then the point process  $N = \tilde{N} \circ \Lambda$  is called a *Cox process*.

In a rigorous mathematical sense it is only required that  $N$  and  $\tilde{N} \circ \Lambda$  are equal in distribution. These questions and related measurability problems are, for instance, discussed in [44, pp. 9–16].

If  $\Lambda$  has continuous trajectories a.s. then  $\Lambda$  is called *diffuse*. In that case, for any realization  $A$  of  $\Lambda$  the Cox process  $N$  is an inhomogeneous Poisson process with intensity measure  $A$ .

Now let  $\Lambda$  be a diffuse random measure with  $\Lambda(\infty) = \infty$  a.s. and  $N$  be the corresponding Cox process. As a generalization of (2.16) we obtain the surplus process

$$(2.17) \quad U(t) = u + (1 + \rho)\mu\Lambda(t) - \sum_{k=1}^{N(t)} Y_k.$$

Define  $\mathcal{F}_\infty^\Lambda := \sigma\{\Lambda(s); s < \infty\}$  and

$$\tilde{U}(t) = U \circ \Lambda^{-1}(t) = u + (1 + \rho)\mu t - \sum_{k=1}^{\tilde{N}(t)} Y_k,$$

where  $\tilde{N} = N \circ \Lambda^{-1}$  for each realization of  $\Lambda$  is a standard Poisson process. This allows us to calculate the ruin probability, where we denote by  $\tilde{\psi}$  the ruin probability corresponding to a standard Poisson process. We obtain

$$(2.18) \quad \begin{aligned} \psi(u) &= \mathbf{P} \left( \inf_{t \geq 0} U(t) < 0 \right) = \mathbf{E} \left[ \mathbf{P} \left( \inf_{t \geq 0} U(t) < 0 \mid \mathcal{F}_\infty^\Lambda \right) \right] \\ &= \mathbf{E} \left[ \mathbf{P} \left( \inf_{t \geq 0} \tilde{U}(t) < 0 \mid \mathcal{F}_\infty^\Lambda \right) \right] = \mathbf{E} \tilde{\psi}(u) = \tilde{\psi}(u). \end{aligned}$$

This is not surprising since it basically means that something that works for every realization also works for a randomly chosen one.

If  $\Lambda$  has the representation  $\Lambda(t) = \int_0^t \lambda(s) ds$ , then  $\lambda = (\lambda(t))_{t \geq 0}$  is called the *intensity process*. If  $\lambda$  has right-continuous and Riemann integrable trajectories, then the corresponding Cox process is well defined ([44, p. 14]). The premium rate is then  $c(t) = (1 + \rho)\mu\lambda(t)$ , i.e., it is a stochastic process.

If  $\lambda(t) = \lambda$  for some positive random variable  $\lambda$ , then for every realization  $\kappa$  of  $\lambda$  the process  $N$  is a Poisson process with intensity  $\kappa$ , i.e.,  $N$  is a mixed Poisson process, and if  $\lambda$  is  $\Gamma(q, \kappa q)$ , then  $N$  is a Pólya process.

To calculate the ruin probability it must be possible to observe  $\lambda$ . Gerber ([41, pp. 25–31 and 142–143]) suggests taking an estimator based on  $N(s)$  for  $s \leq t$ , e.g.,

$$\lambda^*(t) = \mathbf{E} \left[ \lambda(t) \mid \mathcal{F}_t^N \right]$$

(compare with (2.15)), or, if some additional information is used,

$$\hat{\lambda}(t) = \mathbf{E} \left[ \lambda(t) \mid \mathcal{F}_t \right],$$

where  $\mathcal{F}_t^N \subseteq \mathcal{F}_t$  for all  $t$ . Then to use (2.18) we must ensure that  $N \circ \hat{\Lambda}^{-1}$  is a standard Poisson process.

To this end we link point processes and martingales as follows. Let  $N$  be a simple point process and  $\mathbf{F} = (\mathcal{F}_t)_{t \geq 0}$  a filtration such that  $N$  is adapted to  $\mathbf{F}$ . Furthermore, let  $\hat{\Lambda}$  be a diffuse random measure.  $\hat{\Lambda}$  is called the  *$\mathbf{F}$ -compensator of  $N$*  if  $\hat{\Lambda}$  is adapted to  $\mathbf{F}$  and  $N - \hat{\Lambda}$  is an  $\mathbf{F}$ -martingale. Now we go back to Cox processes. A point process  $N$  with  $\mathbf{F}$ -compensator  $\hat{\Lambda}$  is called an  *$\mathbf{F}$ -Cox process* if  $\hat{\Lambda}$  is  $\mathcal{F}_0$ -measurable and, conditioned on  $\mathcal{F}_0$ ,  $N(t) - N(s)$  is Poisson distributed with mean  $\hat{\Lambda}(t) - \hat{\Lambda}(s)$  for  $s < t$ . Proposition 18 of [45] states that a Cox process corresponding to  $\Lambda$ , where  $\Lambda$  is a diffuse random measure with  $\mathbf{E} \Lambda(t) < \infty$  for all  $t < \infty$ , and an  $\mathbf{F}$ -Cox process are equivalent for  $\mathbf{F} = (\mathcal{F}_t)_{t \geq 0}$  and  $\mathcal{F}_t = \mathcal{F}_\infty^\Lambda \vee \mathcal{F}_t^N$ . This means that whether the underlying measure is denoted by  $\hat{\Lambda}$  or by  $\Lambda$  is indeed only a matter of notation.

Hence for an  $\mathbf{F}$ -Cox process  $N$  with observed compensator  $\widehat{\Lambda}$  for the corresponding surplus process  $U$  defined in (2.17), the above arguments imply that  $\psi(u) = \widetilde{\psi}(u)$ , where  $\widetilde{\psi}$  is the ruin probability corresponding to a standard Poisson process.

In a certain sense Cox processes seem to be natural point processes for modeling claim arrivals since they appear as limiting processes of certain thinning procedures. A rigorous mathematical treatment of this idea is given in [58]; in insurance mathematics the idea goes back to [4]. Consider claims which are caused by “risk situations” or *incidents*. Each incident becomes a claim with probability  $p$  independent of all the other incidents. Under these assumptions the claim number process is the result of a thinning procedure of the incident number process. Below this idea is made mathematically precise.

For any point process  $N$ , let  $N^{(p)}$  be the point process resulting from a  $p$ -thinning of  $N$ ; i.e., each point of  $N$  is kept with probability  $p$  and deleted with probability  $1 - p$  (independently for every point of  $N$ ). Let  $\mathcal{P}$  be the set of point processes and  $\mathcal{C}$  the set of Cox processes. Furthermore, let  $D_p: \mathcal{P} \rightarrow \mathcal{P}$  denote the thinning operator and  $\mathcal{D}_p = \{D_p N; N \in \mathcal{P}\}$ ; i.e., the set of all point processes which can be obtained by  $p$ -thinning. Note that the operator  $D_p$  is one-to-one; i.e., the inverse operator  $D_p^{-1}: \mathcal{D}_p \rightarrow \mathcal{P}$  is unique.

For all Poisson processes  $N$  with intensity  $\kappa$ ,  $N^{(p)} = D_p N$  is of course a Poisson process with intensity  $p\kappa$  for all  $p \in (0, 1]$ ; furthermore,  $N \in \mathcal{D}_p$  and  $D_p^{-1} N$  is a Poisson process with intensity  $\kappa/p$  for all  $p \in (0, 1]$ . Also Cox processes are closed with respect to  $p$ -thinning, i.e.,  $D_p N \in \mathcal{C}$  for  $N \in \mathcal{C}$  and  $\mathcal{C} \subset \mathcal{D}_p$  and  $D_p^{-1} N \in \mathcal{C}$  for  $N \in \mathcal{C}$  for all  $p \in (0, 1]$ .

The following theorem due to [58] explains the importance of Cox processes in connection with  $p$ -thinning, and consequently explains their importance as claim arrival models in insurance.

**THEOREM 2.3.1.** *Let  $(N_k)_{k \in \mathbf{N}}$  be a sequence of point processes and let  $(p_k)_{k \in \mathbf{N}} \subset (0, 1)$  with  $\lim_{k \rightarrow \infty} p_k = 0$ . Then there exists a point process  $N$  such that*

$$D_{p_k} N_k \Longrightarrow N, \quad k \rightarrow \infty,$$

*if and only if there exists a random measure  $\Lambda$  such that*

$$p_k N_k \Longrightarrow \Lambda, \quad k \rightarrow \infty,$$

(“ $\Longrightarrow$ ” means convergence in distribution). In that case  $N$  is a Cox process with random measure  $\Lambda$ .

The following corollary due to Mecke [65] gives a characterization of Cox processes.

**COROLLARY 2.3.2.** *A point process  $N$  can be obtained by  $p$ -thinning for every  $p \in (0, 1)$  if and only if it is a Cox process; i.e.,  $\mathcal{C} = \bigcap_{0 < p < 1} \mathcal{D}_p$ .*

Finally, a further natural way in which thinning occurs in insurance can be found in reinsurance policies where exceedance of claims over an increasing sequence of levels (retentions) is considered.

**2.4. Renewal processes.** An important point in the generalizations of the Poisson model in §§ 2.2 and 2.3 was the possibility to compensate risk and size fluctuations by the premiums. Thus the premium rate has to be constantly adapted to the development of the total claims. We shall show in this subsection that for renewal claim number processes a constant premium rate allows for a constant safety loading. Furthermore, since Cox models seem to be a natural class of risk models we are

particularly interested in point processes that are Cox processes as well as renewal processes.

We first introduce some notation. Let  $N$  be a point process with independent interarrival times  $(T_i)_{i \in \mathbf{N}}$ . If the random variables  $T_i$ , for  $i \geq 2$ , have the same distribution function  $H$ , then  $N$  is a *renewal process*.  $N$  is an *ordinary renewal process* if also  $T_1$  has distribution function  $H$ ; and  $N$  is called a *stationary renewal process* if  $T_1$  has the distribution  $H_1$  given by

$$H_1(t) = \kappa \int_0^t \bar{H}(s) ds.$$

The only renewal process that is ordinary and stationary is the Poisson process. Also note that a stationary renewal process is simple if and only if  $H(0) = 0$ .

For a fixed safety loading  $\rho$  it is required that for any time interval the expected income is proportional to the expected expenses. Define  $Q_0 = 0$  and for  $V_k = \sum_{i=1}^k T_i$ ,  $Q_k = -(Q(V_k) - Q(V_{k-1})) = Y_k - cT_k$ ,  $k \in \mathbf{N}$ , then  $(Q_k)_{k \in \mathbf{N}}$  is a sequence of independent and identically distributed random variables and the expected loss between two claims is

$$\mathbf{E} Q_k = \mathbf{E} Q_1 = \mu - c/\kappa \quad \forall k \in \mathbf{N}.$$

This expected loss can be compensated by a constant premium rate  $c$  as in the Poisson model and we choose

$$c = (1 + \rho)\kappa\mu.$$

Now suppose  $N$  is a Cox process. We want to derive conditions such that  $N$  is also a renewal process. Suppose  $N$  has associated random measure  $\Lambda$  ( $\Lambda$  not necessarily diffuse,  $\Lambda(0) = 0$  and  $\Lambda(\infty) = \infty$ ). Then  $\tilde{N} = N \circ \Lambda^{-1}$ , where  $\Lambda^{-1}$  is path-wise the generalized inverse of  $\Lambda$ , i.e.,  $\Lambda^{-1}(t) = \sup\{s; \Lambda(s) \leq t\}$ . This means that  $\tilde{N}(t) = \sup\{k \in \mathbf{N}; \Lambda(V_k) \leq t\}$ . Hence  $\tilde{N}$  is defined by  $(\tilde{V}_k)_{k \in \mathbf{N}} = (\Lambda(V_k))_{k \in \mathbf{N}}$  and  $N$  is a Cox process if and only if  $\tilde{V}_1, (\tilde{V}_{k+1} - \tilde{V}_k)_{k \in \mathbf{N}}$  are independent and have the same standard exponential distribution. The following characterization is due to Kingman [60].

**THEOREM 2.4.1.** *Suppose  $N$  is a Cox process with random measure  $\Lambda$ . Then  $N$  is a renewal process if and only if  $\Lambda^{-1}$  has stationary and independent increments.*

Using the theory of infinitely divisible distributions, one can obtain characterizations for ordinary and stationary renewal processes. We refer to [44], [45] for a more refined treatment.

**THEOREM 2.4.2.** *Let  $N$  be a renewal process with  $\hat{h}(s) = \int_0^\infty e^{-sx} dH(x)$ .*

(i) *If  $N$  is ordinary, then  $N$  is a Cox process if and only if*

$$\hat{h}(s) = \frac{1}{1 - \log \hat{g}(s)},$$

where  $\hat{g}(s) = \int_0^\infty e^{-sx} dG(x)$  for some infinitely divisible distribution function  $G$  such that  $G(0) < 1$ . Furthermore,  $\hat{g}(s) = \mathbf{E} \exp\{s\Lambda^{-1}(1)\}$  and  $\Lambda^{-1}(0) = 0$ , where  $\Lambda$  is the random measure corresponding to  $N$ .

(ii) *If  $N$  is stationary, then  $N$  is a Cox process if and only if*

$$(2.19) \quad \hat{h}(s) = \left(1 + bs + \int_0^\infty (1 - e^{-sx}) dB(x)\right)^{-1}, \quad s \geq 0,$$

where  $b \geq 0$  and  $B$  is a measure on  $(0, \infty)$ , such that  $\int_0^\infty x dB(x) < \infty$ .

We discuss the stationary case in more detail. Consider (2.19) and set  $d = \int_0^\infty dB(x)$  and  $D(x) = B(x)/d$  if  $d < \infty$ . We know that  $N$  is simple if and only if  $H(0) = 0$ . Since  $\lim_{s \rightarrow \infty} \hat{h}(s) = H(0)$ ,  $N$  is simple except for  $b = 0$  and  $d < \infty$ .

(a)  $b = 0$ ,  $d < \infty$ . Then  $\lim_{s \rightarrow \infty} \hat{h}(s) \neq 0$  and  $H(0) \neq 0$ , i.e.,  $N$  is not simple or  $\Lambda$  is not diffuse, but has the representation

$$\Lambda(t) = \sum_{k=1}^{\tilde{N}(t)} E_k,$$

where  $\tilde{N}$  is a stationary renewal process with inter renewal time distribution  $D$  and  $(E_k)_{k \in \mathbf{N}}$  are independent identically distributed exponential with parameter  $d$ , independent of  $\tilde{N}$ .

(b)  $b = 0$ ,  $d = \infty$ . In that case,  $\Lambda$  is diffuse but the right derivatives are with probability one equal to zero for almost all  $t \geq 0$ . We shall come back to this case later. If  $b > 0$ , then  $\Lambda(t) = \int_0^t \lambda(s) ds$ , where  $(\lambda(t))_{t \geq 0}$  is a stochastic process, not identically zero.

(c)  $b > 0$ ,  $d = 0$ . Then  $N$  is a Poisson process with intensity  $1/b$ .

(d)  $b > 0$ ,  $0 < d \leq \infty$ . Then  $\Lambda(t) = \int_0^t \lambda(s) ds$ , where  $(\lambda(t))_{t \geq 0}$  has trajectories that alternate between the values 0 and  $1/b$ . For  $d < \infty$  the intervals where  $\lambda(t)$  is 0 or  $1/b$ , respectively, are independent random variables and the lengths of the intervals where  $\lambda(t) = 1/b$  are exponential with mean  $b/d$  and the lengths of the intervals where  $\lambda(t) = 0$  have distribution function  $D$ .

Note that the case (d) where  $D$  is also exponential is the most interesting one from the point of view of Cox processes. In this case the intensity process  $(\lambda(t))_{t \geq 0}$  is a two-state Markov process. For solvency simulations of the latter process see [36]. Markovian intensities have been investigated in general and again we refer to [45]. A detailed analysis of the general two-state case has been carried out in full detail by Reinhard [69]. In the above two-state Markov situation with one state zero the inter-renewal times distribution  $H$  is a mixture of two exponentials (see [45, Ex. 2.37]). This example can be generalized to the following result of Thorin (for a proof see [45]).

**THEOREM 2.4.3.** *Suppose  $N$  is a stationary renewal process with*

$$H(t) = \int_0^\infty (1 - e^{-t\theta}) dV(\theta),$$

where  $V$  is a distribution function with  $\int_0^\infty (1/\theta) dV(\theta) < \infty$ , then  $N$  is a Cox process.

On the other hand, suppose that  $N$  is a stationary renewal process with  $\Gamma(\gamma)$  inter renewal times. Then it can be shown for  $0 < \gamma < 1$  that  $H$  is a mixture of exponentials as in the theorem above; i.e.,  $N$  is a Cox process. For  $\gamma = 1$ , of course,  $N$  is a Poisson process and, for  $\gamma > 1$ , it can be shown that  $N$  is not a Cox process. For  $\gamma < 1$ , we use Theorem 2.4.2.(b) and the adjoining discussion. We obtain  $b = 0$  and since  $N$  is simple we must have  $d = \infty$ ; i.e.,  $\Lambda$  is a.s. singular with respect to the Lebesgue measure.

Renewal processes as claim arrival processes allow for a constant premium rate. This simplifies the calculation of the ruin probability considerably. The first treatment of the ruin problem for renewal models is due to Sparre Anderson [5]; this is the reason why the ordinary renewal model is sometimes called the *Sparre Anderson model*. A systematic study based on Wiener–Hopf methods has been carried out in a series of papers by Thorin (see, e.g., [73], [74]).

**2.4.1. The ruin probability for ordinary renewal models.** Let the claim number process  $N$  be an ordinary renewal process with interclaim distribution  $H$  with finite mean  $1/\kappa$ . Define  $Z_0 = 0$ ,  $Z_n = \sum_{k=1}^n Q_k$ ,  $n \in \mathbf{N}$ , then note that  $Z_n = -Q(V_n)$  is the loss immediately after the  $n$ th claim. We always denote by  $\psi^0$  the ruin probability for the ordinary renewal case. Since ruin can only occur at renewal points we have

$$\psi^0(u) = \mathbf{P}(u + Q(V_n) < 0 \text{ for some } n \in \mathbf{N}) = \mathbf{P}\left(\max_{n \geq 1} Z_n > u\right).$$

Let  $K$  denote the distribution function of  $Q_1$  and  $\gamma = \mathbf{E}Q_1 = -\mu\rho < 0$ . Then the Laplace–Stieltjes transform of  $K$  is

$$\widehat{k}(r) = \mathbf{E}e^{-rQ_1} = \mathbf{E}\left[e^{-r(Y_1 - cT_1)}\right] = \widehat{f}(r)\widehat{h}(-rc).$$

We assume that  $K(0) < 1$ , the case  $K(0) = 1$  is formally possible (for example, see [45, § 3.1, Rem. 2]) but implies that  $\mathbf{P}(Q_1 \leq 0) = \mathbf{P}(Y_1 - cT_1 \leq 0) = 1$  and hence  $\psi^0(u) \equiv 0$ .

Furthermore, we assume that  $\widehat{f}(r) < \infty$  for  $r < 0$  and, to make matters not too complicated, we assume that  $\widehat{f}(r) \rightarrow \infty$  as  $r \rightarrow r_\infty$ , where  $r_\infty \geq -\infty$  is the left abscissa of convergence of  $\widehat{f}$ . If  $r_\infty > -\infty$ , we have

$$\widehat{h}(-r_\infty c) = \int_0^\infty e^{r_\infty cx} dH(x) > 0.$$

This implies that also  $\widehat{k}(r) \rightarrow \infty$  as  $r \rightarrow r_\infty$ . If  $r_\infty = -\infty$ , then  $K(x_0) < 1$  for some  $x_0 \in (0, \infty)$  and hence, for  $r < 0$ ,

$$\widehat{k}(r) \geq \int_{x_0}^\infty e^{-rx} dK(x) \geq e^{-rx_0} \overline{K}(x_0) \rightarrow \infty \quad \text{as } r \rightarrow -\infty.$$

Thus  $\widehat{f}(r) \rightarrow \infty$  implies  $\widehat{k}(r) \rightarrow \infty$  as  $r \rightarrow r_\infty$ . Furthermore,  $\widehat{k}(0) = 1$  and  $\widehat{k}'(0) = \mathbf{E}Q_1 = -\mu\rho < 0$ . From this it follows that there exists some  $R$  such that  $\widehat{k}(R) = 1$ ;  $R$  is again called *Lundberg coefficient* or *adjustment coefficient*. Indeed if  $T_1$  is exponential, then  $R$  is the Lundberg coefficient from the classical model.

The process  $(Z_k)_{k \in \mathbf{N}_0}$  is a random walk and hence has stationary and independent increments. This is exactly the property we needed in the classical case for the martingale approach that proved the Lundberg inequality. Hence the same argument can be used in this more general setting.

Consider the filtration  $\mathbf{F}^Z = (\mathcal{F}_k^Z)_{k \in \mathbf{N}_0}$ , where

$$\mathcal{F}_k^Z = \sigma\{Z_i; i = 0, \dots, k\}.$$

Denote by  $N_u = \min\{k, Z_k > u\}$  the number of the claim causing ruin, then  $N_u$  is a stopping time and

$$\psi^0(u) = \mathbf{P}(N_u < \infty).$$

Set

$$M_u(n) = \frac{e^{-r(u - Z_n)}}{\widehat{k}^n(r)}.$$

Then up to a change of sign, which is made for future convenience,  $(M_u(n))_{n \in \mathbf{N}}$  is equivalent to (2.11). The same argument now shows that  $(M_u(n))_{n \in \mathbf{N}}$  is an  $\mathbf{F}^Z$ -martingale. Again  $N_u \wedge n_0$  is a bounded  $\mathbf{F}^Z$ -stopping time for  $n_0 < \infty$  and by the optional stopping theorem we obtain as before that

$$\psi^0(u) \leq e^{-ru} \sup_{n \geq 0} \widehat{k}^n(r).$$

The best choice of  $r$  is the Lundberg coefficient  $R$ . Thus we have Lundberg's inequality

$$(2.20) \quad \psi^0(u) \leq e^{-Ru} \quad \forall u \in [0, \infty).$$

As in the derivation of (2.14) we obtain

$$\psi^0(u) = \frac{e^{-Ru}}{\mathbf{E} \exp \{ -R(u - Z_{N_u}) \mid N_u < \infty \}}.$$

For exponential claims this can be calculated in view of the lack of memory and gives

$$\psi^0(u) = (1 - \mu R)e^{-Ru}.$$

To derive an asymptotic result as in the Cramér–Lundberg theorem we use a renewal argument as before. To this end we define a random variable  $A_1 = Z_{N_0}$  on  $\{N_0 < \infty\}$ , where  $N_0 = \min\{k; Z_k > 0\}$ ,  $A(y) = \mathbf{P}(A_1 \leq y, N_0 < \infty)$ , and note that  $A(\infty) = \psi^0(0)$ . Thus  $A$  is a defective distribution function where the defect  $1 - A(\infty)$  is the probability that  $A_1$  is undefined. By separating the cases  $A_1 > u$  and  $A_1 \leq u$ , we obtain

$$(2.21) \quad \psi^0(u) = A(\infty) - A(u) + \int_0^u \psi^0(u - y) dA(y), \quad u \geq 0,$$

which is as (2.5) a defective renewal equation. Again the defect can be removed by an Esscher transformation (if the appropriate exponential moment exists). So assume that there exists a constant  $\beta$  such that

$$\int_0^\infty e^{\beta y} dA(y) = 1;$$

then we multiply (2.21) by  $e^{\beta u}$  which gives a proper renewal equation and Smith's key renewal theorem again yields, for nonarithmetic  $A$ ,

$$\lim_{u \rightarrow \infty} e^{\beta u} \psi^0(u) = \frac{D_1}{D_2},$$

where

$$D_1 = \int_0^\infty e^{\beta y} (A(\infty) - A(y)) dy = (1 - A(\infty)) / \beta,$$

$$D_2 = \int_0^\infty y e^{\beta y} dA(y).$$

Using random walk theory one can show that  $\beta = R$ , where  $R$  is the Lundberg coefficient, i.e., the solution of  $\widehat{k}(R) = 1$ . Unfortunately, the constant  $D_1/D_2$  cannot be explicitly calculated since  $A$  is unknown; it is only known to satisfy  $0 < D_1/D_2 < \infty$ .

**2.4.2. The ruin probability for stationary renewal models.** Let  $\psi^S$  be the ruin probability for the stationary renewal model and  $\psi^0$  for the associated ordinary renewal model. Then exactly as one derives (2.5) one obtains

$$(2.22) \quad \psi^S(u) = \frac{\kappa}{c} \int_u^\infty \bar{F}(z) dz + \frac{\kappa}{c} \int_0^u \psi^0(u-z) \bar{F}(z) dz.$$

A Lundberg inequality can be derived using the result for the ordinary case. By (2.20),

$$\psi^S(u) \leq \frac{\kappa}{c} \int_0^\infty e^{-R(u-z)} \bar{F}(z) dz = \frac{\kappa}{cR} (\hat{f}(R) - 1) e^{-Ru},$$

and hence Lundberg's inequality holds, but the constant may be greater than one.

To establish a Cramér–Lundberg approximation we use dominated convergence and obtain

$$\begin{aligned} \lim_{u \rightarrow \infty} e^{Ru} \psi^S(u) &= \lim_{u \rightarrow \infty} \frac{\kappa}{c} e^{Ru} \int_u^\infty \bar{F}(z) dz \\ &\quad + \lim_{u \rightarrow \infty} \frac{\kappa}{c} \int_0^u e^{R(u-z)} \psi^0(u-z) e^{Ru} \bar{F}(z) dz \\ &= 0 + \frac{\kappa}{cR} (\hat{f}(R) - 1) \frac{D_1}{D_2} := D. \end{aligned}$$

It follows immediately that  $0 < D < \infty$ . This result is due to Thorin [73, p. 97].

The value of the Lundberg coefficient is a measure of the dangerousness of the risk business. We denote by  $R_P$  the Lundberg coefficient for the Poisson model and by  $R$  for the renewal model, where  $\kappa$ ,  $c$ , and  $F$  are equal. It can be shown that for Cox processes  $R \leq R_P$ ; i.e., Cox processes with  $\mathbf{E}T_1 = 1/\kappa$  are more dangerous than the Poisson process with intensity  $\kappa$  ([45, § 3.3]).

**2.5. Large claims problems.** It has been well known for a long time that, for models where large claims may occur with high probability, the Cramér–Lundberg theory is not applicable since the Lundberg coefficient does not exist. Typical examples are Pareto or lognormal claim sizes. The ruin probability for these two examples have been derived by von Bahr [76] and Thorin and Wilkstad [75], respectively. A general theory has been developed by Embrechts and Veraverbeke [34] (see also [28]). They derived an asymptotic expression for the ordinary renewal process setup whenever the Lundberg coefficient does not exist.

In this paper we restrict ourselves to the classical Poisson model, and for ease of presentation we only consider claim size distributions whose Laplace–Stieltjes transform has zero as an essential singularity. For the so-called intermediate case where the left abscissa of convergence is negative, say  $-\gamma$ , but  $\int_0^\infty e^{\gamma y} \bar{F}(y) dy < c/\kappa$ , we refer to [34] (see also [61, Thm. 7] and [25]).

The unique solution of (2.2) together with (2.4) can be represented as

$$\varphi(u) = \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} \left( \frac{1}{1+\rho} \right)^n F_I^{n*}(u).$$

The main result on asymptotic ruin estimates when the Lundberg coefficient does not exist is based on subexponentiality of  $F_I$ .



In general, a distribution function ( $G \in \mathcal{S}$ ), i.e.,  $G$  is a *subexponential distribution*, if and only if

$$(2.23) \quad \lim_{x \rightarrow \infty} \frac{\overline{G}^{n*}(x)}{\overline{G}(x)} = n \quad \forall n \in \mathbf{N}.$$

This equation holds for all  $n$  if and only if it holds for  $n = 2$ . Furthermore, every  $G \in \mathcal{S}$  has the property

$$(2.24) \quad \lim_{x \rightarrow \infty} \frac{\overline{G}(x - y)}{\overline{G}(x)} = 1 \quad \forall y \in \mathbf{R},$$

which implies that  $\widehat{g}(s) = \infty$  for all  $s > 0$ . To explain why  $\mathcal{S}$  can be used to model large claims we reformulate (2.23) as follows: If  $Z_1, Z_2, \dots, Z_n$  are independent and identically distributed random variables with distribution function  $G \in \mathcal{S}$ , then

$$(2.25) \quad \mathbf{P} \left( \sum_{i=1}^n Z_i > x \right) \sim \mathbf{P} \left( \max_{1 \leq i \leq n} Z_i > x \right) \quad \text{as } x \rightarrow \infty.$$

Asymptotic ruin estimates involving the class  $\mathcal{S}$  were proved in [34, Thm. 4.6].

THEOREM 2.5.1. *If  $F_I \in \mathcal{S}$ , then*

$$(2.26) \quad \psi(x) \sim \frac{1}{c/\kappa - \mu} \int_x^\infty \overline{F}(y) dy, \quad x \rightarrow \infty.$$

The question arises whether the condition  $F_I \in \mathcal{S}$  can be replaced by a simple requirement on the right tail  $\overline{F}$ . To this end a new class of distribution functions was introduced in [61]. We say  $F \in \mathcal{S}^*$  if and only if

$$\lim_{x \rightarrow \infty} \frac{\overline{F}^{n\otimes}(x)}{\overline{F}(x)} = n\mu^{n-1} < \infty \quad \forall n \in \mathbf{N},$$

where

$$\overline{F}^{(n+1)\otimes}(x) = \int_0^x \overline{F}^{n\otimes}(x - y) \overline{F}(y) dy \quad \forall n \in \mathbf{N}.$$

Tails of distribution function in  $\mathcal{S}^*$  are subexponential densities; i.e.,  $F \in \mathcal{S}^*$  implies  $F_I \in \mathcal{S}$  and hence Theorem 2.5.1 applies. Furthermore,  $F \in \mathcal{S}^*$  implies that (2.25) holds for the claim sizes  $(Y_i)_{i \in \mathbf{N}}$ ; so, asymptotically, the accumulation of  $n$  successive claims is governed by one very big claim.

Before giving the main examples for which (2.26) holds, we formulate some simple sufficient conditions to conclude that  $F \in \mathcal{S}^*$  (see [61, Thm. 3 and its corollary]). Denote by  $Q = -\log \overline{F}$  the *hazard function* of  $F$ . Since  $\mathcal{S}^*$  contains essentially absolutely continuous distribution functions, we can assume that  $Q$  has a density  $q$  that is called *hazard rate* of  $F$ . Furthermore,  $q(x) \rightarrow \gamma$  as  $x \rightarrow \infty$  implies that  $\widehat{f}$  has its singular point in  $\gamma$ . Hence  $\mathcal{S}^*$  contains essentially distribution functions whose hazard rate  $q$  tends to 0. It can be shown that the rate of convergence of  $q$  decides whether  $F \in \mathcal{S}^*$ .

PROPOSITION 2.5.2. *Each of the following conditions imply  $F_I \in \mathcal{S}$ :*

- a)  $\limsup_{x \rightarrow \infty} xq(x) < \infty$ ,
- b)  $\lim_{x \rightarrow \infty} q(x) = 0$ ,  $\lim_{x \rightarrow \infty} xq(x) = \infty$  and one of the following conditions is satisfied:
  - (i)  $\limsup_{x \rightarrow \infty} xq(x)/Q(x) < 1$ ,

- (ii)  $q$  is regularly varying with index  $\delta \in [-1, 0)$ ,  
 (iii)  $Q$  is regularly varying with index  $\delta \in (0, 1)$ , and  $q$  is eventually decreasing,  
 (iv)  $q$  is eventually decreasing, slowly varying, and  $Q(x) - xq(x)$  is regularly varying with index 1.

Note that for regularly varying  $q$  such that  $\lim_{x \rightarrow \infty} q(x) = 0$  the index of regular variation  $\delta$  satisfies  $\delta \in [-1, 0]$ . The case  $\delta < -1$  is impossible since  $\int_0^\infty q(x) dx = \infty$ , on the other hand, for positive  $\delta$ ,  $q$  would be asymptotically equivalent to an increasing function.

These conditions can be applied to check the following examples (for more details and more examples see [61]).

(1) *Pareto model:*

$$\begin{aligned} \bar{F}(x) &= (a/x)^b I_{[a, \infty)}(x) \quad \text{with } a > 0, b > 1, \\ \psi(x) &\sim \frac{a}{(b-1)c/\kappa - a} (a/x)^{b-1}, \quad x \rightarrow \infty. \end{aligned}$$

(2) *Lognormal model:*

$$\begin{aligned} \bar{F}(x) &\sim \frac{a}{\sqrt{2\pi}(\log x - \log b)} \exp\left\{-\frac{(\log x - \log b)^2}{2a^2}\right\}, \quad x \rightarrow \infty, \\ &\quad \text{with } a > 0, b \geq 1, \\ \psi(x) &\sim \frac{a^3}{\sqrt{2\pi}(c/\kappa - b \exp\{a^2/2\})} \frac{x}{(\log x - \log b)^2} \exp\left\{-\frac{(\log x - \log b)^2}{2a^2}\right\}, \\ &\quad x \rightarrow \infty. \end{aligned}$$

(3) *Weibull model:*

$$\begin{aligned} \bar{F}(x) &= \exp\{-x^a\} \quad \text{with } 0 < a < 1, \\ \psi(x) &\sim \frac{x^{1-a} \exp\{-x^a\}}{ac/\kappa - \Gamma(1/a)}, \quad x \rightarrow \infty. \end{aligned}$$

Further results can also be found in [35]. For asymptotic ruin estimates in a more general model see [7].

**3. Martingales and insurance risk.** As already indicated in § 2.1, martingales play a fundamental role in insurance mathematics. For a review of some of the recent results, together with a list of further references, see [26] and the already mentioned [46]. The key idea for the fact that martingales are inherent in general risk models was already hinted at in § 2.1. The general structure of a risk process consists of:

$$\text{“initial capital”} + \text{“premium income”} - \text{“claims.”}$$

If for the moment we forget about the initial capital and assume that our liabilities (claims) follow a general stochastic process, still denoted by  $S(t)$ ; then to find the fair premium for  $S(t)$  it seems natural to construct a predictable random process  $P(t)$ , the premium process, which is defined to make the difference

$$M(t) = P(t) - S(t), \quad t \geq 0,$$

a fair game (i.e., a martingale) between insured and insurer. The construction of generalizations of the classical Cramér–Lundberg model using this approach is worked

out in [23]. In this section, we shall present an alternative approach based on the theory of piecewise deterministic Markov processes as introduced by Davis [21]. In §3.1 we recall the basic definition of a piecewise-deterministic Markov process. As an application in §3.2 we discuss a risk model taking borrowing and investment into account. The use of diffusion approximations will be highlighted in §3.3.

**3.1. Piecewise-deterministic Markov processes.** Once notions like borrowing, investment, and inflation are introduced in the classical Cramér–Lundberg model, it is clear that the interclaim evolution of the risk process will not be linear anymore. However, in many cases we still encounter a more general deterministic behavior. In a general Markov process set-up, Davis [21] let the deterministic paths between jumps follow the integral curves  $\varphi(t, z)$  of a vector field  $\chi$ , i.e., the curve  $\varphi(0, z) = z$  and  $(d/dt)f(\varphi(t, z)) = (\chi f)(\varphi(t, z))$  for all differentiable functions  $f$ . For the classical risk process this amounts to  $\varphi(t, z) = z + ct$ , i.e.,  $\chi = c(d/dz)$ . For the definition given below, we follow [70]. Let  $I$  be a countable set endowed with the discrete topology and let  $(M_i: i \in I)$  be given open subsets from  $\mathbf{R}^{d_i}$ , for some  $d_i \in \mathbf{N}$ . Set

$$E = \{(i, \xi): i \in I, \xi \in M_i\}$$

and denote by  $\mathcal{E}$  the Borel sets of  $E$ . For every  $i \in I$ , let  $\chi_i$  be a vector field on  $M_i$  such that, for every point  $z \in M_i$  there exists exactly one integral curve  $\varphi_i(t, z)$  through  $z = \varphi_i(0, z)$ . Denote by  $\partial M_i$  the boundary of  $M_i$  and let

$$\begin{aligned} \partial^* M_i &= \{z \in \partial M_i: \exists(t, \xi) \in \mathbf{R}^+ \times M_i, z = \varphi_i(t, \xi)\}, \\ \Gamma &= \{(i, z) \in \partial E: i \in I, z \in \partial^* M_i\}, \\ t^*(i, z) &= \inf \{t > 0: \varphi_i(t, z) \notin M_i\}. \end{aligned}$$

We shall assume that  $\varphi_i(t^*(i, z), z) \in \Gamma$  if  $t^*(i, z) < \infty$ . Hence,  $\Gamma$  denotes the set of boundary points of  $E$  that can be reached from  $E$  via integral curves within finite time and  $t^*(i, z)$  is the time needed to reach the boundary from the point  $(i, z)$  ( $\infty$  if  $\varphi_i(t, z) \in M_i$  for all  $t \geq 0$ ). The condition  $\varphi_i(t^*(i, z), z) \in \Gamma$  assures that the integral curves cannot ramify.

To define the relevant Markov process on  $(E, \mathcal{E})$  one further needs the jump intensity  $\lambda: E \rightarrow \mathbf{R}^+$  and the Markov transition measure

$$Q: \mathcal{E} \times (E \cup \Gamma) \rightarrow [0, 1].$$

The piecewise deterministic Markov process (PDMP)  $(X_t)_{t \geq 0}$  with starting point  $x_0 = (n, z)$  is now constructed as follows. Let

$$1 - F_1(t) = \begin{cases} \exp\left(-\int_0^t \lambda(n, \varphi_n(s, z)) ds\right), & 0 \leq t < t^*(x_0), \\ 0, & t \geq t^*(x_0) \end{cases}$$

and define the random variable  $T_1$  on  $\mathbf{R}^+$  with distribution function  $\mathbf{P}(T_1 \leq t) = F_1(t)$ . Moreover, define the random vector  $(N_1, Z_1)$  on  $E$  with distribution function  $\mathbf{P}((N_1, Z_1) \in \cdot | T_1) = Q(\cdot, \varphi_n(T_1, z))$ . Set

$$X_t = \begin{cases} (n, \varphi_n(t, z)), & 0 \leq t < T_1, \\ (N_1, Z_1), & t = T_1. \end{cases}$$

Assume now that the process  $(X_t)_{t \geq 0}$  is constructed up to time  $T_{k-1}$ ,  $k > 1$ . Let

$$1 - F_k(t) = \begin{cases} \exp\left(-\int_0^t \lambda(N_{k-1}, \varphi_{N_{k-1}}(s, Z_{k-1})) ds\right), & 0 \leq t < t^*(X_{k-1}), \\ 0, & t \geq t^*(X_{k-1}) \end{cases}$$

and define the random variables  $T_k \geq T_{k-1}$  on  $\mathbf{R}$  with distribution function

$$\mathbf{P}(T_k \leq T_{k-1} + t \mid \sigma(X_s: s \leq T_{k-1})) = F_k(t)$$

and  $(N_k, Z_k)$  on  $E$  with distribution function

$$\mathbf{P}((N_k, Z_k) \in \cdot \mid \sigma(X_s: s \leq T_{k-1}), T_k) = Q(\cdot, \varphi_{N_{k-1}}(T_k - T_{k-1}, Z_{k-1})).$$

Set

$$X_t = \begin{cases} (N_{k-1}, \varphi_{N_{k-1}}(t - T_{k-1}, Z_{k-1})), & T_{k-1} < t < T_k, \\ (N_k, Z_k), & t = T_k. \end{cases}$$

The number of jumps of the process in  $(0, t]$  is given by  $J_t = \sum_{i \in \mathbf{N}} I_{(t \geq T_i)}$  and we assume that  $\mathbf{E}[J_t] < \infty$  for all  $t \in \mathbf{R}^+$ . Davis [21] showed that the above constructed process is strong Markov. The key steps to be taken in practice now are

- (i) show that the considered generalized risk process is in fact a PDMP;
- (ii) calculate the generator  $A$  of the process together with its domain  $\mathcal{D}(A)$ , and
- (iii) solve  $Af = 0$  for  $f \in \mathcal{D}(A)$  in order to construct the relevant martingale  $f(X_t)$  via Dynkin's theorem.

For more details on this see [20] and [70]. For most applications in insurance, the following result of [20, p. 185] turns out to be useful.

**PROPOSITION 3.1.1.** *Let  $(X_t)_{t \geq 0}$  be a PDMP and let  $f: (E \cup \Gamma) \rightarrow \mathbf{R}$  be a measurable function satisfying*

- (i) *the function  $(0, t^*(i, z)) \rightarrow \mathbf{R}$ ,  $t \mapsto f(i, \varphi_i(t, z))$  is absolutely continuous  $\forall (i, z) \in E$ ,*
- (ii)  *$f(x) = \int_E f(y) Q(dy, x) \forall x \in \Gamma$  (boundary condition), and*
- (iii)  *$\mathbf{E} \left[ \sum_{T_i \leq t} |f(X_{T_i}) - f(X_{T_i^-})| \right] < \infty$ .*

*Then  $f \in \mathcal{D}(A)$  and the generator of  $(X_t)_{t \geq 0}$  is given by*

$$Af(x) = \chi f(x) + \lambda(x) \left[ \int_E (f(y) - f(x)) Q(dy, x) \right].$$

As indicated above, once the risk process is formulated as a PDMP, one solves  $Af = 0$  to obtain that  $(f(X_t): t \geq 0)$  is a martingale with respect to the natural filtration of  $(X_t)_{t \geq 0}$ . In most cases, there are various ways in which the risk process can be made into a PMDP. Considering the Cramér–Lundberg model from § 2.1, where

$$U(t) = u + ct - S(t), \quad S(t) = \sum_{i=1}^{N(t)} Y_i,$$

Dassios and Embrechts [20] used the following constructions.

**MODEL 1.** Consider  $(U(t))_{t \geq 0}$  directly as a PDMP.

**MODEL 2.** Consider the process  $(U(t))_{t \geq 0}$  until the ruin time  $T_u = \inf\{t \geq 0: U(t) < 0\}$ , at which time the process jumps to an absorbing state.

Using MODEL 1, one recovers Gerber’s martingale from § 2.1. MODEL 2 yields expressions for the Laplace transform

$$\int_0^\infty \mathbf{P}(T_u < \infty) e^{-su} du.$$

From the latter expression, using standard Tauberian arguments, ruin estimates for  $u \rightarrow \infty$  or  $u \rightarrow 0$  can be obtained.

**3.2. A general insurance risk model.** For this section, we follow Schmidli [70], see also [33]. As a generalization of the classical Cramér–Lundberg model, we assume that a company can borrow money if needed (i.e., for negative of “low” surplus) and gets interest for capital above a certain (deterministic) level  $\Delta$ , the amount of capital the company retains as a liquid reserve. The assumed constant forces of interest are denoted by  $\beta_1$  for invested money and  $\beta_2$  for borrowed money, i.e., after time  $t$  a capital  $z$  becomes  $ze^{\beta t}$ ,  $\beta \in \{\beta_1, \beta_2\}$ . Using the PDMP-language of the previous subsection, the associated vector field becomes

$$\chi = \begin{cases} (\beta_1(x - \Delta) + c), & \Delta \leq x, \\ c \frac{\partial}{\partial x}, & 0 \leq x < \Delta, \\ (\beta_2 x + c) \frac{\partial}{\partial x}, & x < 0. \end{cases}$$

We denote the corresponding risk process by  $(U_g(t))_{t \geq 0}$ ,  $g$  standing for “general.” The integral curve corresponding to the vector field  $\chi$  is decreasing for  $x < -c/\beta_2$ . Whenever the risk process hits the boundary  $-c/\beta_2$ , the company will a.s. not be able to repay its debts. So

$$T_{g,u} = \inf \{t > 0: U_g(t) < -c/\beta_2\}$$

will be called the ruin time. The model where  $\Delta = \infty$  was studied in [20]. Using the theory from § 3.1, one can show that for  $\Delta \in [0, \infty]$ :

$$\mathbf{P}(T_{g,u} < \infty) = 1 - \frac{f(u)}{f(\infty)},$$

where  $f$  is to be obtained as a solution of some complicated integral-differential (generator) equation. Moreover,  $\mathbf{P}(T_{g,u} < \infty) = 1$  if and only if  $\Delta = \infty$  and  $c \leq \kappa\mu$ . In the case of exponential claims, the function  $f$  above can be calculated explicitly as follows:

$$f(x) = f_1(x)I_{[\Delta, \infty)}(x) + f_2(x)I_{[0, \Delta)}(x) + f_3(x)I_{\mathbf{R}^-}(x),$$

where

$$f_3(x) = K \int_0^{x+c/\beta_2} s^{(\kappa/\beta_2)-1} e^{-s/\mu} ds,$$

$$f_2(x) = f_3(0) + \frac{f'_3(0)}{1/\mu - \kappa/c} (1 - e^{-(1/\mu - \kappa/c)x}),$$

$$f_1(x) = f_2(\Delta) + \left(\frac{\beta_1}{c}\right)^{(\kappa/\beta_1)-1} e^{c/(\beta_1\mu)} f'_2(\Delta) \int_{c/\beta_1}^{x+c/\beta_1-\Delta} s^{(\kappa/\beta_1)-1} e^{-s/\mu} ds,$$

for some constant  $K$  which can be calculated. As a consequence of this result one obtains the following adjustment coefficient estimate:

$$\lim_{u \rightarrow \infty} \mathbf{P}(T_{g,u} < \infty) e^{ru} = \begin{cases} 0, & r < 1/\mu \text{ or } (r = 1/\mu \text{ and } \kappa < \beta_1), \\ c, & r = 1/\mu \text{ and } \kappa = \beta_1, \\ \infty, & \text{otherwise,} \end{cases}$$

where  $c = e^{\Delta\mu(\beta_1/c)^{(\kappa/\beta_1)-1}\mu f_2'(\Delta)/f_1(\infty)}$ .

Further analytic results can be obtained in the case of Erlang claim sizes; it is clear, however, that, for more general models, no explicit analytic results like those above can be obtained. Hence, the need arises for approximate solutions like those based on the theory of phase-type distributions [6] or Monte-Carlo simulation (see, for instance, [9]). For more details and further references, see [70].

**3.3. Diffusion approximations.** Rather than using the nondiffusion type PDMP-theory, one might want to use the usual time-space rescaling and hope for a reasonable diffusion limit on which the ruin times (hitting times) can be calculated more easily. Although this approach is common throughout applied probability, it was first introduced into the insurance framework through a paper by Iglehart [55] (see also [6]). Recently, Schmidli [70] obtained a rather general result that allows immediate calculations for the diffusion limits of risk processes like the one encountered in the previous subsection.

**THEOREM 3.3.1.** *Let  $\delta: \mathbf{R} \rightarrow \mathbf{R}$  be a Lipschitz-continuous function,  $(M^{(n)}: n \in \mathbf{N})$  a sequence of semimartingales, and  $M$  a semimartingale such that  $M(0) = M^{(n)}(0) = 0$ . If  $X^{(n)}$  is a sequence of stochastic processes satisfying the SDE*

$$dX^{(n)}(t) = \delta(X^{(n)}(t)) dt + dM^{(n)}(t), \quad X^{(n)}(0) = u,$$

and  $Z$  is a diffusion satisfying the SDE

$$dZ(t) = \delta(Z(t)) dt + dM(t), \quad Z(0) = u,$$

then  $M^{(n)} \Rightarrow M$  for  $n \rightarrow \infty$  is equivalent to  $X^{(n)} \Rightarrow Z$ .

This result can be applied to the general risk process of § 3.2, where

$$\delta(x) = \begin{cases} \beta_1(x - \Delta), & x \geq \Delta, \\ 0, & 0 \leq x < \Delta, \\ \beta_2x, & x < 0. \end{cases}$$

In this case, the semimartingales  $(M^{(n)}, M)$  are just Cramér–Lundberg  $(U^{(n)}, U)$  risk processes with  $n$ -dependent parameters  $(\kappa_n, \mu_n, c_n)$ .

Under the set of conditions,

(i) for all  $n \in \mathbf{N}$ :  $\rho_n \kappa_n \mu_n = \rho_1 \kappa_1 \mu_1, \kappa_n \sigma_n^2 = \kappa_1 \sigma_1^2,$

(ii) for all  $\varepsilon > 0$ :  $\kappa_n \int_\varepsilon^\infty x_2 dF_n(x) \rightarrow 0$  as  $n \rightarrow \infty$ , and

(iii)  $Q^{(n)}(0) \Rightarrow Q(0),$

it follows that  $Q^{(n)} \Rightarrow Q$  where  $Q(t) = \eta B(t) + dt$  with  $(B(t))_{t \geq 0}$  a standard Wiener process and  $d = \rho_1 \kappa_1 \mu_1, \eta^2 = \lim_{n \rightarrow \infty} \kappa_n (\mu_n^2 + \sigma_n^2)$  (see [70]). The notation above follows that in § 2.1.

Combining these results, one is able to calculate the limiting diffusion  $(Z(t))_{t \geq 0}$  of the general risk processes from § 3.2; indeed one obtains the following result.

**THEOREM 3.3.2.**

(i) The process  $f(Z(t))_{t \geq 0}$  is a martingale for  $f: \mathbf{R} \rightarrow \mathbf{R}$ , with

$$f(x) = \begin{cases} \frac{\eta}{d} \left( 1 - \exp\left(-\frac{2d}{\eta^2} \Delta\right) \right) + 2\sqrt{\frac{\pi}{\beta_1}} \exp\left(\frac{d^2}{\eta^2 \beta_1} - \frac{2d}{\eta^2} \Delta\right) \\ \quad \times \left( \Phi\left(\frac{\sqrt{2\beta_1}}{\eta} \left(x - \Delta + \frac{d}{\beta_1}\right)\right) - \Phi\left(\frac{d}{\eta} \sqrt{\frac{2}{\beta_1}}\right) \right), & \Delta \leq x, \\ \frac{\eta}{d} \left( 1 - \exp\left(-\frac{2d}{\eta^2} x\right) \right), & 0 \leq x < \Delta, \\ 2\sqrt{\frac{\pi}{\beta_2}} \exp\left(\frac{d^2}{\eta^2 \beta_2}\right) \left( \Phi\left(\frac{\sqrt{2\beta_2}}{\eta} \left(x + \frac{d}{\beta_2}\right)\right) - \Phi\left(\frac{d}{\eta} \sqrt{\frac{2}{\beta_2}}\right) \right), & x < 0, \end{cases}$$

where  $\Phi$  stands for the standard normal distribution function.

(ii)  $\mathbf{P}(Z(t) \rightarrow -\infty) = \frac{f(\infty) - f(Z(0))}{f(\infty) - f(-\infty)}$ ,

where  $f$  is given in (i).

For a proof of this result, various generalizations and numerical comparisons between the exact ruin estimates given in the previous subsection and the diffusion approximation value of Theorem 3.3.2 (ii) see [70].

**4. Statistical estimation for risk processes.** It is clear that the various parameters in the above discussed risk processes will have to be estimated statistically. The whole branch of empirical studies in insurance is quickly gaining momentum. A prime example on how the transition from data toward model (and backward) is made is the work by Rammlau-Hansen (see [67] and [68]) on solvency. We shall for the present paper content ourselves with the discussion of some statistical problems related to insurance risk. Using the various references given in the text, the reader can get a better view of the totality of statistical problems involved. In §4.1, the important problem of adjustment coefficient estimation will be discussed. We also take the opportunity to highlight the use of bootstrap techniques in insurance. In §4.2 a review will be given of claim size fitting in the context of the numerical estimation of the total claim size distribution. In §4.3, a brief discussion of tail estimation for individual claim size distributions will be given. Finally, in §4.4, a probabilistic definition of a large claim index together with its statistical estimation will be discussed.

**4.1. Estimating the adjustment coefficient.** In §2 (i.e., (2.6)) we have defined the adjustment (or Lundberg) coefficient  $R$  for the classical model via the equation:

$$(4.1) \quad \frac{\kappa}{c} \int_0^\infty e^{Rz} \overline{F}(z) dz = 1,$$

where  $F$  is the claim size distribution. Its main property was that

$$\lim_{u \rightarrow \infty} \mathbf{P}(T_u < \infty) e^{Ru} \in (0, \infty).$$

Thus,  $e^{-Ru}$  gives the asymptotic order of magnitude of ultimate ruin; for a discussion of the existence of  $R$  for more general risk models see also [29]. If we suppose that  $\kappa$  and  $c$  are known, then there are various ways in which  $R$  can be estimated via (4.1). Most of the methods used are based on the link between risk theory and queueing theory or use some empirical version of (4.1). The latter approach was mainly advocated by Csörgö and Teugels [19], Hall, Teugels, and Vanmarcke [47]. See also [50] and [51]. If one

replaces  $F$  in (4.1) by the empirical distribution function (e.d.f.)  $\widehat{F}_n(x) = n^{-1}\#\{i \leq n: Y_i \leq x\}$ , one obtains, by solving (4.1), an estimator  $R_n$  of  $R$ . The problem with the estimator is that its asymptotic behavior strongly depends on conditions relating to the left abscissa of convergence of the Laplace transform of  $F$  (see [19]). In [18] and [22] the following link to queueing theory is exploited. Let  $Q_i = Y_i - cT_i$ ,  $i = 1, 2, \dots$ , for the classical Cramér–Lundberg model, and  $\widehat{q}(s) = \mathbf{E}(e^{sQ_1})$ .

Define  $M_0 = 0$  and  $M_n = \max\{M_{n-1} + Q_n, 0\}$ , for  $n = 1, 2, \dots$ , and  $\nu_0 = 0$  and  $\nu_k = \min\{n \geq \nu_{k-1} + 1: M_n = 0\}$ , for  $k = 1, 2, \dots$ . Deheuvels and Steinebach [22] interpret the stopping times  $\nu_k$  as follows. They first observe that  $U(t) - U(t_{\nu_{i-1}}) > 0$  for  $t_{\nu_{i-1}} < t < t_{\nu_i}$ , while  $U(t_{\nu_i}) - U(t_{\nu_{i-1}}) \leq 0$  for  $i = 1, 2, \dots$ , so that  $U(\cdot)$  has a positive excursion in each interval  $(t_{\nu_{i-1}}, t_{\nu_i})$ . Once  $n \geq 1$ , such positive excursions have been observed, the sequence of random variables  $\{Z_i: 1 \leq i \leq n\}$  is defined by

$$Z_i = \max_{\nu_{i-1} < j \leq \nu_i} M_j, \quad i = 1, 2, \dots$$

It was shown in [8], that the random variables  $Z_1, Z_2, \dots$  are independent and identically distributed and

$$(4.2) \quad \lim_{x \rightarrow \infty} x^{-1} \log\{Z_1 > x\} = -R.$$

(The latter conclusion of course presupposes the existence of  $R$  under the net-profit condition  $\mu - c/\kappa < 0$ .) Using this result, estimators of  $1/R$  based on the order statistics  $\{Z_{n-i+1,n}, 1 \leq i \leq k\}$  of  $Z_1, \dots, Z_n$ , where  $0 \leq Z_{1,n} \leq \dots \leq Z_{m,n}$ , can be constructed. Some examples are

$$\begin{aligned} T_n(k) &= \sum_{i=1}^k Z_{n-i+1,n} && \text{(tail sum estimator),} \\ Q_n(k) &= kZ_{n-k+1,n} && \text{(quantile estimator),} \\ H_n(k) &= T_n(k) - Q_n(k) && \text{(Hill estimator),} \end{aligned}$$

or some convex combination of  $Q_n$  and  $H_n$ , for some sequence  $k = k(n)$  for which  $k(n) \rightarrow \infty, n^{-1}k(n) \rightarrow 0$ .

The construction underlying (4.2) was taken up further by Embrechts and Mikosch [32] as a basis for a bootstrap estimation procedure for estimating  $R$ . Further use of the bootstrap procedure in insurance mathematics is highlighted in [1], where a problem related to stochastic discounting is solved using such a resampling scheme. One final, interesting method which we would like to mention is due to Herkenrath [49]. In the latter paper it is observed that the Lundberg equation (4.1) is equivalent to finding a nontrivial root  $R$  of

$$(4.3) \quad \kappa + Rc = \kappa \int_0^\infty e^{Rx} dF(x).$$

Defining

$$L(R) = \int_0^\infty e^{Rx} dF(x) - \frac{Rc}{\kappa} - 1,$$

one has that

$$L(R) = \mathbf{E} \left( e^{RY_1} - \frac{Rc}{\kappa} - 1 \right);$$

therefore, the problem of estimating  $R$  amounts to the estimation of a root of an unknown regression function by observing corresponding random variables. This is a classical problem of stochastic approximation. See the above mentioned papers for details, numerical comparisons, and further references.



#### 4.2. Numerical approximation of the accumulated claims distribution.

Traditionally, the distribution of the total claims in a fixed time period (e.g., in one year) has been a central topic in risk theory. First there are two independent estimation problems to solve: the choice of the claim number process and the choice of the claim size distribution. We restrict ourselves here to the Poisson model so that the estimation problem for the claim number process reduces to the estimation of  $\kappa$ . Second, for known (or estimated)  $\kappa$  and claim size distribution  $F$  one has to calculate the accumulated (or aggregate) claims distribution

$$G(x) = \sum_{n=0}^{\infty} e^{-\kappa} \frac{\kappa^n}{n!} F^{n*}(x), \quad x \geq 0.$$

Notice that  $F^{n*}$  can be explicitly calculated only for degenerate (deterministic) or exponential claim sizes. In all other cases numerical approximations for  $G$  are required, where the convolutions and the infinite sum cause problems.

Since computers have entered the field the interest in many traditional approximation methods has faded whereas more computer intensive methods like recursions, Fast Fourier transform, and Monte-Carlo methods have gained importance. On the other hand, many of the traditional methods require only a few moments of the claim size distribution and show acceptable accuracy around the mean. Unfortunately, few of them are particularly good in the tail.

All standard textbooks on risk theory contain sections on approximation methods (see, e.g., [9], [41], [48], and [52]). Numerical examples can be found on a diskette included in [52]. It runs on DOS 2.0 or later DOS versions and is written in BASIC or TURBO PASCAL or in [63]. Here we only introduce some basic ideas and refer for a more detailed description and discussion of the approximation methods to the above mentioned textbooks.

For ease of notation we shall always approximate the distribution function  $\tilde{G}$  of the standardized random variable

$$\tilde{S} = \frac{S - \kappa\mu}{\sqrt{\kappa(\sigma^2 + \mu^2)}}, \quad S = S(1),$$

if not stated otherwise. The approximation of  $G$  is then obtained in the obvious way. Furthermore, we shall denote  $\mu_k = \mathbf{E}Y^k$ ,  $k \geq 2$ .

**4.2.1. The normal approximation and related methods.** The central limit theorem applied to  $\tilde{S}$  gives, for large  $\kappa$ ,

$$\tilde{G}(x) \approx \Phi(x).$$

In practical applications, however,  $\kappa$  is often not large enough and the accuracy of this approximation is not satisfactory if the skewness of the claim size distribution is large.

An approximation which takes also higher moments into account is the *Edgeworth expansion*

$$\tilde{G}(x) \approx \Phi(x) - \frac{1}{6} \frac{\mu_3}{\sqrt{\kappa\mu_2^3}} \Phi^{(3)}(x) + \frac{1}{24} \frac{\mu_4}{\kappa\mu_2^2} \Phi^{(4)}(x) + \frac{1}{72} \frac{\mu_3^2}{\kappa\mu_2^3} \Phi^{(6)}(x) + R(x)$$

where  $\Phi^{(k)}$  is the  $k$ th derivative of  $\Phi$ . The first line of the approximation contains terms up to the order  $\kappa^{-1/2}$ , the second line up to  $\kappa^{-1}$ , and  $R(x)$  up to  $\kappa^{-3/2}$ .

This expansion is obtained by means of the moment generating function of  $\tilde{G}$  expanding the exponential in a Mac Laurin series and inverting the terms back to the distribution function. Although the Edgeworth expansion is a divergent series, taking a suitable number of terms, it usually gives acceptable results in the neighbourhood of the mean.

An approximation more reliable also in the tails is given by the so-called *normal power approximation*

$$\tilde{G}(x) \approx \Phi \left( \sqrt{9\kappa\mu_2^3/\mu_3^2 + 6x\sqrt{\kappa\mu_2^3/\mu_3} + 1} - 3\sqrt{\kappa\mu_2^3/\mu_3} \right).$$

The idea is to represent  $\tilde{S}$  as a transformation of a standard normal random variable  $N$ ; i.e.,

$$\tilde{S} \stackrel{d}{=} v(N),$$

for some function  $v$ . Then we approximate

$$\tilde{G}(x) \approx \Phi(v^\leftarrow(x)),$$

where  $v^\leftarrow$  is the inverse of the function  $v$ . A suitable transformation in polynomial form can be obtained by inverting the Edgeworth expansion. Using Newton's approximation one obtains

$$v(y) \approx y + \frac{1}{6} \frac{\mu_3}{\sqrt{\kappa\mu_2^3}}(y^2 - 1) + \frac{1}{24} \frac{\mu_4}{\kappa\mu_2^2}(y^3 - 3y) - \frac{1}{36} \frac{\mu_3^2}{\kappa\mu_2^3}(2y^3 - 5y).$$

The above formula for  $\tilde{G}(x)$  is obtained by taking only the first line of  $v(y)$  into account.

**4.2.2. Approximations using orthogonal polynomials.** This method is based on  $L^2$  approximation theory. Suppose  $I \subset \mathbf{R}$  is an interval and  $w$  a positive continuous (weight) function. Then certain orthogonal polynomials  $(\pi_i)_{i \in \mathbf{N}_0}$  constitute a basis of the Hilbert space  $\mathbf{L}_w^2$ , which denotes the space of all  $L^2$ -integrable functions with respect to the measure  $w(x) dx$ . Then all functions  $f \in \mathbf{L}_w^2$  can be expanded into

$$f(x) = \sum_{i=0}^{\infty} A_i \pi_i(x) w(x),$$

where

$$A_i = \int_I \pi_i(x) f(x) dx / \int_I \pi_i^2(x) w(x) dx, \quad i \in \mathbf{N}_0.$$

Approximations of order  $n$  are obtained by

$$f(x) \approx \sum_{i=0}^n A_i \pi_i(x) w(x).$$

Different intervals  $I$  and weight functions  $w$  yield different approximations.

(i) *The Gamma approximation of Bowers.*

For  $I = \mathbf{R}^+$  and  $w(x) = x^{b-1}e^{-x}/\Gamma(b)$  the Laguerre polynomials constitute a basis of  $\mathbf{L}_w^2$ . In this case the standardized random variable is

$$\tilde{S} = \frac{\mathbf{E} S}{\text{Var } S} S = \frac{\mu}{\mu_2} S$$

and we assume that  $\tilde{S}$  has density  $\tilde{g}$ .

For  $n = 0, 1, 2$  this results in a simple Gamma approximation; i.e.,

$$\tilde{g}(x) \approx w(x) = \frac{1}{\Gamma(b)} x^{b-1} e^{-x}, \quad x \geq 0.$$

For  $n = 3$  one obtains

$$\begin{aligned} \tilde{g}(x) \approx & \frac{1}{\Gamma(b)} x^{b-1} e^{-x} + \frac{1}{6} \left( \frac{\mu^3 \mu_3}{\mu_2^{3/2}} \kappa^{5/2} - b(b+1)(b+2) \right) \\ & \times \left( \frac{x^3}{\Gamma(b+3)} - \frac{3x^2}{\Gamma(b+2)} + \frac{3x}{\Gamma(b+1)} - \frac{1}{\Gamma(b)} \right) e^{-x} x^{b-1}, \quad x \geq 0. \end{aligned}$$

Here  $b$  is chosen in the form  $b = \mathbf{E} \tilde{S} = \kappa \mu^2 / \mu_2$ .

(ii) *The Gram-Charlier approximation.*

For  $I = \mathbf{R}$  and  $w(x) = \varphi(x) = e^{-x^2/2} / \sqrt{2\pi}$ , the Hermite polynomials constitute a basis of  $\mathbf{L}_w^2$ . For  $\tilde{S} = (S - \kappa \mu) / \sqrt{\kappa \mu_2}$  with density  $\tilde{g}$  we obtain, for  $n = 0, 1, 2$ ,

$$\tilde{g}(x) \approx w(x) = \varphi(x),$$

for  $n = 3$  one obtains

$$\tilde{g}(x) \approx \varphi(x) - \frac{1}{6} \frac{\mu_3}{\sqrt{\kappa \mu_2^3}} \varphi^{(3)}(x),$$

and for  $n = 4$

$$\tilde{g}(x) \approx \varphi(x) - \frac{1}{6} \frac{\mu_3}{\sqrt{\kappa \mu_2^3}} \varphi^{(3)}(x) + \frac{1}{24} \frac{\mu_4}{\kappa \mu_2^2} \varphi^{(4)}(x),$$

where  $\varphi^{(k)}$  is the  $k$ th derivative of the standard normal density  $\varphi$ . Note that for  $n \leq 4$  the Gram-Charlier approximation is exactly the Edgeworth expansion of the corresponding order. Only if one takes terms of higher order into account both expansions differ.

**4.2.3. The Esscher approximation.** Most of the preceding approximations are sufficiently precise around the mean but perform poorly for large  $x$  values. Exponential tilting of the distribution shifts the mean to an arbitrary large  $x$ -value and hence improves the approximation in the tail considerably. For the random variable  $S$  with distribution function  $G$  we define for  $h \in \mathbf{R}$ , whenever  $\hat{g}(-h) < \infty$ , a new random variable  $S_h$  with distribution function  $G_h$  by

$$G_h(x) = \frac{1}{\hat{g}(-h)} \int_0^x e^{hy} dG(y).$$

$G_h$  is called the *Esscher transform* of  $G$ . Then for a given value of  $x$  we define  $h$  so that  $\mathbf{E} S_h = x$  and apply the Edgeworth expansion to  $G_h$ . For  $h > 0$ , the tail  $\bar{G}(x)$  can be approximated for  $x > \mathbf{E} S$  by

$$\bar{G}(x) \approx \exp \left\{ \lambda(\hat{f}(-h) - 1) - hx \right\} \left( E_0(u) - \frac{\hat{f}^{(3)}(-h)}{6\kappa^{1/2}(\hat{f}^{(2)}(-h))^{3/2}} E_3(u) \right),$$

where  $u = h \sqrt{\kappa \hat{f}^{(2)}(-h)}$  and  $E_0$  and  $E_3$  are the so-called Esscher functions

$$E_0(u) = (1 - \Phi(u)) / (\sqrt{2\pi} \varphi(u))$$

and

$$E_3(u) = \left(1 - u^2 + u^3(1 - \Phi(u))/\varphi(u)\right) / \sqrt{2\pi}.$$

The Esscher approximation is similar to the saddle point approximation. For a detailed discussion in the Poisson and the Pólya case see [31]; for more general models see [56].

**4.2.4. Discrete methods.** Note that for all approximation methods introduced in §§ 4.2.1 and 4.2.2 only  $\kappa$  and the first few moments of the claim size distribution have to be known (or estimated); for the Esscher approximation the Laplace transform is needed. For the discrete approximations in this subsection the whole claim size distribution has to be known. A good additional reference for this subsection is [37]. See also [30] for the fast Fourier method. Both approximations introduced in this section work on a finite support and the number of calculations increases with a finer discretization (for absolutely continuous distributions) as well as with a greater support (for heavy tailed distributions).

(i) *The Panjer approximation*

Suppose  $Y$  is concentrated on a lattice, and without loss of generality suppose  $f(i) = \mathbf{P}(Y = i)$ ,  $i \in \mathbf{N}_0$  is given. Set  $g(i) = \mathbf{P}(S = i)$ ,  $i \in \mathbf{N}_0$ , then  $g(i)$  can be calculated recursively by

$$g(0) = \mathbf{P}(N = 0) = e^{-\kappa},$$

$$g(i) = \frac{\kappa}{i} \sum_{j=1}^i j f(j) g(i-j).$$

Similar versions of this recursion hold for more general claim arrival processes; for a characterization theorem see [48]. For an absolutely continuous distribution a discretization is necessary to apply the above recursion; in this case the algorithm is no longer exact.

(ii) *The fast Fourier transform*

Suppose again that  $Y$  is concentrated on a lattice, say  $\mathbf{N}_0$ , i.e.,  $f(k) = \mathbf{P}(Y = k)$ ,  $k \in \mathbf{N}_0$ . Then  $Y$  has the characteristic function

$$\varphi_Y(t_j) = \sum_{k=0}^{\infty} f(k) e^{ik2\pi j/n},$$

for  $t_j = 2\pi j/n \in [0, 2\pi)$ ,  $j = 0, \dots, n-1$ .

If we set

$$\tilde{f}(k) = \sum_{l=-\infty}^{\infty} f(k + ln),$$

where  $f(-k) = 0 \quad \forall k \in \mathbf{N}$ , then by periodicity,

$$\varphi_Y(t_j) = \sum_{k=0}^{n-1} \tilde{f}(k) e^{ik2\pi j/n}.$$

Now the characteristic function of  $S$  is given by

$$\varphi_S(t_j) = \exp \{ \kappa \varphi_Y(t_j) - 1 \}.$$

On the other hand, if  $g(k) = \mathbf{P}(S = k)$ ,  $k \in \mathbf{N}_0$ , then

$$\varphi_S(t_j) = \sum_{k=0}^{\infty} g(k) e^{ik2\pi j/n}.$$

We denote

$$\tilde{g}(k) = \sum_{l=-\infty}^{\infty} g(k + ln)$$

where  $g(-k) = 0 \quad \forall k \in \mathbf{N}$ . Then

$$\varphi_S(t_j) = \sum_{k=0}^{n-1} \tilde{g}(k) e^{ik2\pi j/n},$$

and the problem reduces to determining  $\tilde{g}(k)$ ,  $k \in \mathbf{N}_0$ , from  $\varphi_S(t_j)$ .

Both problems, namely the determination of the characteristic function and its inverse can be considered as follows. Transform a vector  $a = (a_0, \dots, a_{n-1})^T$  into a vector  $b = (b_0, \dots, b_{n-1})^T$  according to the rule

$$b = Wa, \quad W = (e^{kj2\pi i/n})_{k,j=0,\dots,n-1}.$$

Note that  $\overline{W} = (e^{-kj2\pi i/n})_{k,j=0,\dots,n-1}$  and

$$W \cdot \overline{W} = nI,$$

where  $I$  is the unit matrix. Thus, one obtains

$$\begin{aligned} b &= Wa \stackrel{\text{def}}{=} FFT^+(a), \\ a &= \frac{1}{n} \overline{W} b \stackrel{\text{def}}{=} \frac{1}{n} FFT^-(b). \end{aligned}$$

Summarizing the above arguments we obtain

$$\tilde{g} = \frac{1}{n} FFT^-(\exp\{\kappa FFT^+(\tilde{f}) - 1\}).$$

Notice that the error is  $\tilde{g}(k) - g(k) = \sum_{l=1}^{\infty} g_{k+ln} > 0$ . By a special factorization of the matrix  $W$  for  $n = 2^l$ ,  $l \in \mathbf{N}$ , the properties of the unit roots allow a considerable reduction of the number of calculations needed; hence, the name *Fast Fourier Transform* (FFT).

**4.3. Tail estimation.** Tail estimation of the claim size distribution is, particularly in the heavy tailed case, one of the most interesting problems in risk theory. We give some examples: (i) As seen in § 4.2 most of the numerical approximation for the accumulated claims distribution  $G$  do not perform very well in the far end tails. On the other hand, a theoretical result in [28] states that, e.g., for the Poisson model with subexponential claim size distribution  $F$ ,  $\overline{G}(x) \sim \kappa \overline{F}(x)$ . This provides an estimate for  $\overline{G}(x)$  for large  $x$  and the remaining problem is an accurate estimation of the tail of the claim size distribution  $F$ . (ii) Asymptotic estimation of the ruin probability for subexponential claim size distribution is given in (2.26). Here also  $\overline{F}(x)$  is needed for large  $x$ . (iii) By the structure of reinsurance, statistical data are often available only from the upper extreme part of the sample and mainly the fit of the tail is required.

Here we briefly review a semiparametric method which has been introduced in [62], tested by a Monte-Carlo simulation study in [59], and has also been applied to a sample of automobile liability data from an excess of loss reinsurance treaty with limit 100 000 SF [62]. In the same paper some alternative methods are reviewed, in particular, the so-called threshold method and we refer to this article for more details and further references.

To apply this semiparametric tail estimation method one needs at least some information to which class of models the distribution tail belongs. This can very effectively be done using the mean residual life functions. The mean residual life function of a distribution function  $F$  is defined as

$$a_F(x) = \mathbf{E}[Y - x | Y \geq x] = \int_x^\infty \bar{F}(u) du / \bar{F}(x), \quad x > 0.$$

For exponential distributions  $a_F$  is obviously constant, for distributions with tails decreasing faster than exponential it decreases to 0, and for heavy-tailed distributions, like subexponentials, it increases to  $\infty$ ; for a graphical summary of the most important claim size distribution, see, e.g., [54, p. 109].

Moreover, different heavy-tailed distributions show different curves for the mean residual life functions, e.g., for the Pareto it is a straight line with positive slope and for the Weibull (with parameter  $\alpha < 1$ ) it is a concave function. This becomes quite obvious in the tails. Furthermore, to use these functions for the purpose of modeling is relatively easy since their empirical version  $a_{F,n}$  is simply

$$a_{F,n}(x) = \frac{1}{k} \sum_{Y_i > x} Y_i - x,$$

where  $k$  is the number of observations greater than  $x$ .

The function  $a_{F,n}(x)$  allows, for instance, to distinguish between an extended Pareto model

$$\bar{F}_1(x) = l(x)x^{-\alpha}, \quad \alpha > 0,$$

and an extended Weibull model

$$\bar{F}_2(x) = r(x)e^{-x^\alpha}, \quad \alpha \in (0, 1),$$

where  $l$  is a slowly varying function and  $r$  is a regularly varying function (see, e.g., [11]). The advantage of introducing  $l$  or  $r$  is that the model makes a certain departure from the exact Pareto or Weibull model, possible, in particular for the lower and intermediate range of the sample. The parameter  $\alpha$  has to be estimated and this can be done by so-called asymptotic maximum likelihood estimation (AMLE). For the extended Pareto case one obtains Hill's estimator, i.e.,

$$\hat{\alpha}_k^{-1} = \frac{1}{k} \sum_{j=1}^k \log Y_{n-j+1,n} - \log Y_{n-k+1,n}, \quad k \geq 2;$$

for the extended Weibull case the resulting AMLE is the solution of

$$(\alpha^{-1} - 1) \log \log n - \log \alpha = \log \left\{ \frac{1}{k} \sum_{j=1}^k \log Y_{n-j+1,n} - \log Y_{n-k+1,n} \right\}, \quad k \geq 2.$$

For a derivation of the parameters and their asymptotic properties see [62].

**4.4. Closing the gap: Theory and practice.** An important evolution in recent discussions in insurance mathematical concerns the growing need for high-quality empirical work. Much less as in biostatistics or reliability theory, for instance, does the insurance world discuss research problems on the basis of empirical data. As always there are of course happy exceptions, as already indicated in the introduction to

this paper, in such fields as credibility theory, loss distribution modeling, and solvency studies. The main reason, however, for the existing gap between theory and practice is the separation between academic societies and professional ones in many countries. See, for instance, [27]. Recent conferences in the realm of stochastics do include the occasional section (or invited paper) on insurance; if, however, research is really to take off, academic researchers will have to spend more time in participating in the many professional meetings related to insurance to get a better feeling for “what really interests the practitioner.” We would like to close this paper with an example of what sort of research problems may come out of such an encounter. The results to be discussed can be found in [2]. In the latter paper, the problem of large claims is discussed. A practitioner’s view of Pareto claims is described as “those claims for which 20% of the individual claims are responsible for more than 80% of the total claim amount in a particular portfolio.” A way to find a mathematical way in understanding the above rule of thumb goes as follows. Let  $(Y_1, \dots, Y_n)$  be a sample of positive, independent and identically distributed random variables denoting the first  $n$  claims in a portfolio  $\mathbf{Y} = (Y_1, Y_2, \dots)$ , with distribution function  $F$  and finite mean  $\mu$ . Denote the associated order statistics by  $Y_{1,n} \leq \dots \leq Y_{n,n}$ . The total claim amount for the first  $n$  claims is  $S_n = \sum_{i=1}^n Y_i$ . If  $\hat{F}_n$  stands for the empirical distribution function and  $\hat{F}_n^{-1}$  its generalized inverse, then  $\hat{F}_n^{-1}(i/n) = Y_{i,n}$ .  $\hat{F}_n^{-1}$  is known as the quantile function. With this notation, the above rule of thumb can be reformulated in terms of  $T_n(\alpha) = (Y_{[n\alpha],n} + \dots + Y_{n,n})/S_n$ ,  $0 < \alpha < 1$ . Hence  $T_n(\alpha)$  is the proportion of the sum of the  $(n - [n\alpha] + 1)$ st largest claims to the aggregate claim amount  $S_n$  in our portfolio. Using results from the theory of Mallows’ metrics in [2] the following result is proved.

**THEOREM 4.4.1.** *Under the above conditions,*

$$T_n(\alpha) \xrightarrow{\text{a.s.}} \frac{1}{\mu} \int_{\alpha}^1 F^{-1}(x) dx.$$

Consequently one could use the functional  $D_F(\alpha) = \frac{1}{\mu} \int_{\alpha}^1 F^{-1}(x) dx$  as a measure of dangerousness of a claim size distribution function  $F$ . Its value indicates to what extent the  $100(1 - \alpha)\%$  largest claims in a portfolio contribute to the overall portfolio claim amount. Of course,  $D_F = 1 - L_F$ , where  $L_F$  is the Lorentz curve associated with  $F$ . There is a lot of relevant literature that can be brought in at this point. A numerical study now yields the following value for  $D_F$ :

$$F \sim \text{Pareto}(p): p = 1.4, \quad 1 - \alpha = 0.2, \quad D_F(\alpha) = 0.804,$$

indicating that the above rule of thumb may be based on a Pareto distribution behavior for the claims with parameter around 1.5. This observation is substantiated in many recent empirical studies. See [2] for further references, and for numerical results on  $D_F$  across a wide range of claim size distributions. For a recent survey on the stochastic modeling of large claims in nonlife insurance, see, for instance, [10].

**5. Conclusion.** It is clear that the above summary of “some aspects of insurance mathematics” offers only a sketchy view of what the whole field is about. It is, however, to be hoped that enough material and references have been presented so that those interested in getting “more involved” will find the text a useful entrance into the rich world of insurance modeling. In essence, insurance is all about randomness. There is no doubt whatsoever that advanced techniques in probability and mathematical statistics will play an ever increasing role in the further development of the field. It

is therefore our sincerest hope that our paper may contribute to stimulating future students to have a closer look at some of the demanding problems in the area.

## REFERENCES

- [1] M. AEBI, P. EMBRECHTS, AND T. MIKOSCH, *Stochastic discounting and Mallows' Metric*, ETH Preprint, 1992.
- [2] ———, *A large claim index*, Mitteilungen der Schweiz, Vereinigung der Versicherungsmathematiker, to appear.
- [3] P. ALBRECHT, *AFIR-Literatur*, Mitteilungen der Schweiz, Vereinigung der Versicherungsmathematiker, 1992.
- [4] B. ALMER, *Risk analysis in theory and practical statistics*, Transactions XVth International Congress of Actuaries, New York, II (1957), pp. 314–370.
- [5] E. SPARRE ANDERSEN, *On the collective theory of risk in the case of contagion between claims*, Transactions XVth International Congress of Actuaries, New York, II (1957), pp. 219–229.
- [6] S. ASMUSSEN, *Applied Probability and Queues*, John Wiley, Chichester, 1987.
- [7] S. ASMUSSEN, L. FLØEHEMRIKSEN, AND C. KLÜPPELBERG, *Large claims approximations for risk processes in a Markovian environment*, preprint, 1992.
- [8] P. BÁRTFAI, *Limes superior Sätze für die Wartemodelle*, Studia Sci. Math. Hungary, 5 (1970), pp. 317–325.
- [9] R. E. BEARD, T. PENTIKÄINEN, AND R. PESONEN, *Risk Theory*, 3rd ed., Chapman and Hall, London, 1984.
- [10] J. BERLANT AND J. L. TEUGELS, *Modeling large claims in nonlife insurance*, Insurance: Mathematics and Economics, 11 (1992), pp. 17–29.
- [11] N. H. BINGHAM, C. M. GOLDIE, AND J. L. TEUGELS, *Regular Variation*, Cambridge University Press, 1987.
- [12] N. L. BOWERS, H. U. GERBER, J. C. HICKMAN, D. A. JONES, AND C. J. NESBITT, *Actuarial Mathematics*, The Society of Actuaries, Ithaca, Illinois, 1987.
- [13] P. BRÉMAUD, *Point Processes and Queues*, Martingale Dynamics, Springer-Verlag, Berlin, 1980.
- [14] H. BÜHLMANN, *Experience rating and credibility*, ASTIN Bulletin IV, Part III, 1967.
- [15] ———, *Mathematical Methods in Risk Theory*, Springer-Verlag, Berlin, 1970.
- [16] ———, *Tendencies of development in risk theory*, in Centennial Celebration of the Actuarial Profession in North America, 2 (1989), pp. 499–521, Society of Actuaries, Schaumburg, Illinois.
- [17] H. CRAMÉR, *On some questions connected with mathematical risk*, University of California Publication in Statistics, 1954.
- [18] M. CSÖRGÖ AND J. STEINEBACH, *On the estimation of the adjustment coefficient in risk theory via intermediate order statistics*, Insurance: Mathematics and Economics, 10 (1991), pp. 37–50.
- [19] M. CSÖRGÖ AND J. L. TEUGELS, *Empirical Laplace transform and approximation of compound distributions*, J. Appl. Probab., 27 (1990), pp. 88–101.
- [20] A. DASSIOS AND P. EMBRECHTS, *Martingales and insurance risk*, Comm. Statist. Stochastic Models, 5 (1989), pp. 181–217.
- [21] M. H. A. DAVIS, *Piecewise-deterministic Markov processes. A general class of nondiffusion stochastic models*, J. Royal Statist. Soc., B. 46 (1984), pp. 353–388.
- [22] P. DEHEUVELS AND J. STEINEBACH, *On some alternative estimates of the adjustment coefficient in risk theory*, Scand. Actuarial J., (1990), pp. 135–159.
- [23] F. DELBAEN AND J. M. HAEZENDONCK, *Classical risk theory in an economic environment*, Insurance: Mathematics and Economics, 6 (1987), pp. 85–116.
- [24] C. DELLACHERIE AND P. A. MEYER, *Probabilités et Potentiel Théorie des Martingales*, Hermann, Paris, 1980.
- [25] P. EMBRECHTS, *A property of the generalized inverse Gaussian distribution with some applications*, J. Appl. Probab., (1983), pp. 537–544.
- [26] ———, *Martingales in Nonlife Insurance*, Vilnius Conference Proceedings, VNU Press, vol. 1, pp. 314–322, 1990.



- [27] ———, (ed.), *Actuarial Statistics: Its Place in the University Curriculum*, Proceedings of ICOTS III, vol. 2, pp. 404–428, 1991.
- [28] P. EMBRECHTS, C. M. GOLDIE, AND N. VERAVERBEKE, *Subexponentiality and infinite divisibility*, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 49 (1979), pp. 335–347.
- [29] P. EMBRECHTS, J. GRANDELL, AND H. P. SCHMIDL, *Finite-time Lundberg inequalities in the Cox case*, *Scandinavian Actuarial J.*, to appear.
- [30] P. EMBRECHTS, R. GRÜBEL, AND S. PITTS, *Some applications of the fast Fourier transform algorithm in insurance mathematics*, *Statistica Neerlandica*, to appear.
- [31] P. EMBRECHTS, J. L. JENSEN, M. MAEJIMA, AND J. L. TEUGELS, *Approximations for compound Poisson and Pólya processes*, *Adv. Appl. Prob.*, 17 (1985), pp. 623–637.
- [32] P. EMBRECHTS AND T. MIKOSCH, *A bootstrap procedure for estimating the adjustment coefficient*, *Insurance: Mathematics and Economics*, 10 (1991), pp. 181–190.
- [33] P. EMBRECHTS AND H. P. SCHMIDL, *A general insurance risk model*, ETH preprint, 1992.
- [34] P. EMBRECHTS AND N. VERAVERBEKE, *Estimates for the probability of ruin with special emphasis on the possibility of large claims*, *Insurance: Mathematics and Economics*, 1 (1982), pp. 55–72.
- [35] P. EMBRECHTS AND J. VILLASEÑOR, *Ruin estimates for large claims*, *Insurance: Mathematics and Economics*, 7 (1988), pp. 269–274.
- [36] P. EMBRECHTS AND L. WOUTERS, *Simulating risk solvency*, *Insurance: Mathematics and Economics*, 9 (1990), pp. 141–148.
- [37] M. FEILMEIER AND J. BERTRAM, *Anwendung Numerischer Methoden in der Risikotheorie*, Verlag Versicherungswirtschaft e.V., Karlsruhe, 1987.
- [38] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. I, John Wiley, New York, 1957.
- [39] ———, *An Introduction to Probability Theory and Its Applications*, Vol. II, John Wiley, New York, 1971.
- [40] H. U. GERBER, *Martingales in risk theory*, *Mitteilungen der Schweiz, Vereinigung der Versicherungsmathematiker*, 73 (1973), pp. 205–216.
- [41] ———, *An Introduction to Mathematical Risk Theory*, Huebner Foundation Monographs 8, distributed by Richard D. Irwin Inc., Homewood Illinois, 1979.
- [42] ———, *Life Insurance Mathematics*, Springer-Verlag, Berlin, 1990.
- [43] M. J. GOOVAERTS, R. KAAS, A. E. VAN HEERWAARDEN, AND T. BAUWELINCKX, *Effective Actuarial Methods*, North-Holland, Amsterdam, 1990.
- [44] J. GRANDELL, *Doubly Stochastic Poisson Processes*, *Lecture Notes in Mathematics* 529, Springer-Verlag, Berlin, 1976.
- [45] ———, *Aspects of Risk Theory*, Springer-Verlag, Berlin, 1990.
- [46] ———, *Finite time ruin probabilities and martingales*, *Informatica*, 2 (1991), pp. 3–32.
- [47] P. HALL, J. L. TEUGELS, AND A. VANMARCKE, *The abscissa of convergence of the Laplace transform*, *J. Appl. Probab.*, 29 (1992), pp. 353–362.
- [48] W. R. HEILMANN, *Fundamentals of Risk Theory*, Verlag Versicherungswirtschaft e.V., Karlsruhe, 1988.
- [49] U. HERKENRATH, *On the estimation of the adjustment coefficient in risk theory by means of stochastic approximation procedures*, *Insurance: Mathematics and Economics*, 5 (1986), pp. 305–313.
- [50] C. HIPPE, *Efficient estimators for ruin probabilities*, in *Proceedings 4th Prague Symposium. Asymptotics Statistics*, ed. by P. Mandl and M. Hušková, Charles University, Prague, pp. 259–268, 1989.
- [51] ———, *Pollard's CLT and the  $\Delta$ -method*, University of Hamburg, Preprint, 1992.
- [52] C. HIPPE AND R. MICHEL, *Risikotheorie: Stochastische Modelle und Statistische Methoden*, Verlag Versicherungswirtschaft e.V., Karlsruhe, 1990.
- [53] J. M. HOEM, *Markov chain models in life insurance*, *Blätter der Deutschen Gesellschaft für Versicherungsmathematik*, 9 (1969), pp. 91–107.
- [54] R. V. HOGG AND S. A. KLUGMAN, *Loss Distributions*, John Wiley, New York, 1984.
- [55] D. L. IGLEHART, *Diffusion approximation in collective risk theory*, *J. Appl. Probab.*, 6 (1969), pp. 285–292.
- [56] J. L. JENSEN, *Saddle point approximation to the distribution of the total claim amount in some recent risk models*, *Scand. Actuarial J.*, (1991), pp. 154–168.
- [57] W. S. JEWELL, *Up the misty staircase with credibility theory*, *Mitteilungen der Schweiz. Vereinigung der Versicherungsmathematiker*, (1990), pp. 281–311.

- [58] O. KALLENBERG, *Limits of compound and thinned point processes*, J. Appl. Probab., 12 (1975), pp. 269–278.
- [59] B. KELLER AND C. KLÜPPELBERG, *Statistical estimation of large claim distributions*, Mitteilungen der Schweiz. Vereinigung der Versicherungsmathematiker (1991), pp. 203–216.
- [60] J. F. C. KINGMAN, *On doubly stochastic Poisson processes*, Proc. Camb. Phil. Soc., 60 (1964), pp. 923–930.
- [61] C. KLÜPPELBERG, *Estimation of ruin probabilities by means of hazard rates*, Insurance: Mathematics and Economics, 8 (1989), pp. 279–285.
- [62] ———, *Estimation of distribution tails — a semiparametric approach*, Preprint ETH Zürich, 1990.
- [63] J. KUPPER, *Methoden zur Berechnung der Verteilungsfunktion des Totalschadens*, Mittellungen der Schweiz. Vereinigung der Versicherungsmathematiker (1971), pp. 279–315.
- [64] D. MAYERS AND C. W. SMITH JR., *Toward a positive theory of insurance*, Monograph Series in Finance and Economics, New York University, (1982), Monograph 1982-1.
- [65] J. MECKE, *Eine charakteristische Eigenschaft der doppelt stochastischen Poissonschen Prozesse*, Z. Wahr. verw. Gebiete, 11 (1968), pp. 74–81.
- [66] A. PAPATRIANDAFYLOU AND H. R. WATERS, *Martingales in life insurance*, Scand. Actuarial J., (1984), pp. 210–230.
- [67] H. RAMLAU-HANSEN, *A solvency study in nonlife insurance. Part 1: Analysis of fire, windstorm, and glass claims*, Scand. Actuarial J., (1988), pp. 3–34.
- [68] ———, *A solvency study in nonlife insurance. Part 2 : Solvency margin requirements*, Scand. Actuarial J., (1988), pp. 35–60.
- [69] J. M. REINHARD, *On a class of semi-Markov risk models obtained as classical risk models in a Markovian environment*, ASTIN Bulletin, XIV (1984), pp. 23–43.
- [70] H. P. SCHMIDLI, *A General Insurance Risk Model*, Ph. D. Thesis, ETH Zürich, 1992.
- [71] C. W. SMITH, JR., *On the convergence of insurance and finance research*, The Journal of Risk and Insurance, LIII (4) (1986), pp. 693–717.
- [72] E. STRAUB, *Nonlife Insurance Mathematics*, Springer-Verlag, Berlin, 1989.
- [73] O. THORIN, *On the asymptotic behaviour of the ruin probability for an infinite period when the epochs of claims form a renewal process*, Scand. Actuarial J., (1974), pp. 81–99.
- [74] ———, *Probability of ruin*, Scand. Actuarial J., (1982), pp. 65–102.
- [75] O. THORIN AND N. WIKSTAD, *Calculation of ruin probabilities when the claim distribution is lognormal*, ASTIN Bulletin, 9 (1977), pp. 231–246.
- [76] B. VON BAHR, *Asymptotic ruin probability when exponential moments do not exist*, Scand. Actuarial J., (1975), pp. 6–10.
- [77] H. WATERS, *Continuous Mortality Investigation Reports*, Institute of Actuaries, CMR, 12 (1991).
- [78] D. WILLIAMS, *Probability with Martingales*, Cambridge University Press, 1991.