POISSON PROCESS AND INSURANCE : AN INTRODUCTION[1]

S.RAMASUBRAMANIAN
Statistics and Mathematics Unit
Indian Statistical Institute
8th Mile, Mysore Road
Bangalore - 560059.

Abstract: Basic aspects of the classical Cramer-Lundberg insurance model are discussed.

Introduction

Recent cataclysmic events like tsunami, torrential downpour, floods, cyclones, earthquakes, etc. underscore the fact that practically everyone would like to be assured that there is some (non-supernatural) agency to bank upon in times of grave need. If the affected parties are too poor, then it is the responsibility of the government and the "haves" to come to the rescue. However, there are also sizeable sections of the population who are willing to pay regular premium to suitable agencies during normal times to be assured of insurance cover to tide over crises. Insurance has thus become an important aspect of modern society. In such a set up, a significant proportion of the financial risk is shifted to the insurance company. The implicit trust between the insured and the insurance company is at the core of the interaction. A reasonable mathematical theory of insurance can possibly provide a scientific basis for this trust.

Certain types of insurance policies have been prevalent in Europe since the latter half of the 17th century. But the foundations of modern actuarial mathematics were laid only in 1903 by the Swedish mathematician Filip Lundberg, and later in the 1930's by the famous Swedish probabilist Harald Cramer. Insurance mathematics today is considered a part of applied probability theory, and a major portion of it is described in terms of continuous time stochastic processes.

This article should be accessible to anyone who has taken a course in probability theory. At least statements of the various results and the heuristics can be appreciated. While proofs of some of the basic results are given, for some others only a partial proof or heuristic arguments are indicated; of course, in a few cases we are content with just citing an appropriate reference. [Mi], [RSST]

[1] These notes form part of a series of lectures given at a Refresher course on Applied Stochastic Processes, sponsored by the Indian Academy of Sciences, Bangalore, held at the Indian Statistical Institute, Delhi in December 2005.

are very good books where an interested reader can find more information. It is inevitable that a bit of jargon of basic probability theory is assumed. One may look up [Fe], [HPS], [Ro1], [Ro2] for elucidation of terms like random variable, distribution, density, expectation, independence, independent identically distributed (i.i.d.) random variables, etc. Some of the articles compiled in [DKR] also contain a few introductory accounts.

Collective risk model

We shall mainly look at one model, known as the *Cramer-Lundberg model*; it is the oldest and the most important model in actuarial mathematics. This model is a particular type of a collective risk model. In a collective risk model there are a number of anonymous but very similar contracts or policies for similar risks, like insurance against fire, theft, accidents, floods or crop damage/ failure, etc. The main objectives are modelling of claims that arrive in an insurance business, and decide how premiums are to be charged to avoid ruin of the insurance company. Study of probability of ruin and obtaining estimates for such probabilities are also some of the interesting aspects of the model.

There are three main assumptions in a collective risk model:

1. The total number of claims, say N, occurring in a given period is random. Claims happen at times $\{T_i\}$ satisfying $0 \leq T_1 \leq T_2 \leq \cdots$ We call them *claim arrival times* (or just *arrival times*).

2. The i-th claim arriving at time T_i causes a payment X_i. The sequence $\{X_i\}$ is assumed to be an i.i.d. sequence of nonnegative random variables. These random variables are called *claim sizes.*

3. The claim size process $\{X_i\}$ and the claim arrival times $\{T_j\}$ are assumed to be independent. So $\{X_i\}$ and N are independent.

The first two assumptions are fairly natural, whereas the third one is more of a mathematical convenience.

Take $T_0 = 0$. Define the *claim number process* by

$$\begin{aligned} N(t) &= \max\{i \geq 0 : T_i \leq t\} \\ &= \text{number of claims occurring by time } t, \quad t \geq 0 \end{aligned} \tag{1}$$

Also define the *total claim amount process* by

$$S(t) = X_1 + X_2 + \cdots + X_{N(t)} = \sum_{i=1}^{N(t)} X_i, t \geq 0. \tag{2}$$

These two stochastic processes will be central to our discussions. Note that a sample path of N and the corresponding sample path of S have jumps at the same times T_i, by 1 for N and by X_i for S.

A function $f(\cdot)$ is said to be $o(h)$ if $\lim\limits_{h \to 0} \frac{f(h)}{h} = 0$; that is, if f decays at a faster rate than h.

Poisson process

We first consider the claim number process $\{N(t) : t \geq 0\}$. For each $t \geq 0$, note that $N(t, \cdot)$ is a random variable on the same probability space $(\Omega, \mathcal{F}, P)$. We list below some of the obvious/ desired properties of N (rather postulates for N), which may be taken into account in formulating a model for the claim number process.

- **(N1):** $N(0) \equiv 0$. For each $t \geq 0$, $N(t)$ is a nonnegative integer valued random variable.
- **(N2):** If $0 \leq s < t$ then $N(s) \leq N(t)$. Note that $N(t) - N(s)$ denotes the number of claims in the time interval $(s, t]$. So N is a nondecreasing process.
- **(N3):** The process $\{N(t) : t \geq 0\}$ has *independent increments*; that is, if $0 < t_1 < t_2 < \cdots < t_n < \infty$ then $N(t_1), N(t_2) - N(t_1), \cdots, N(t_n) - N(t_{n-1})$ are independent random variables, for any $n = 1, 2, \cdots$. In other words, claims that arrive in disjoint time intervals are independent.
- **(N4):** The process $\{N(t)\}$ has *stationary increments*; that is, if $0 \leq s < t, h > 0$ then the random variables $N(t) - N(s)$ and $N(t+h) - N(s+h)$ have the same distribution (probability law). This means that the probability law of the number of claim arrivals in any interval of time depends only on the length of the interval.
- **(N5):** Probability of two or more claim arrivals in a very short span of time is negligible; that is,

$$P(N(h) \geq 2) = o(h), \quad \text{as} \quad h \downarrow 0. \tag{3}$$

- **(N6):** There exists $\lambda > 0$ such that

$$P(N(h) = 1) = \lambda h + o(h), \quad \text{as} \quad h \downarrow 0. \tag{4}$$

  The number λ is called the *claim arrival rate.* That is, in very short time interval the probability of exactly one claim arrival is roughly proportional to the length of the interval.

Remark 1: The first two postulates are self evident. The hypothesis (N3) is quite intuitive; it is very reasonable at least as a first stage approximation to many real situations. (N5),(N6) are indicative of the fact that between two arrivals there will be a gap, but may be very small; (note that bulk arrivals are not considered here). (N4) is a time homogeneity assumption; it is not very crucial.

Remark 2: In formulating a model it is desirable that the hypotheses are realistic and simple. Here 'realistic' means that the postulates should capture the essential features of the phenomenon/ problem under study. And 'simple' refers to the mathematical amenability of the assumptions; once a model is chosen, theoretical properties and their implications should be considerably rich and obtainable with reasonable ease. These two aspects can be somewhat conflicting; so success of a mathematical model depends very much on the optimal balance between the two. □

To see what our postulates (N1)–(N6) lead to, put

$$P_n(t) = P(N(t) = n), \quad t \geq 0, n = 0, 1, 2, \cdots \tag{5}$$

Observe that

$$\begin{aligned} P_0(t+h) &= P(N(t) = 0, N(t+h) - N(t) = 0) && \text{(by (N1),(N2))} \\ &= P(N(t) = 0) \cdot P(N(t+h) - N(t) = 0) && \text{(by (N3))} \\ &= P_0(t) \cdot P(N(h) = 0) && \text{(by (N4),(N1))} \\ &= P_0(t) \cdot [1 - \lambda h + o(h)] && \text{(by (N5),(N6))} \end{aligned}$$

whence we get (as $0 \leq P_0(t) \leq 1$),

$$\frac{d}{dt} P_0(t) = -\lambda P_0(t), \qquad t > 0. \tag{6}$$

By (N1), note that $P_0(0) = P(N(0) = 0) = 1$. So the differential equation (6) and the above initial value give

$$\begin{aligned} P_0(t) &= P(N(t) = 0) = P(N(t+s) - N(s) = 0) \\ &= \exp(-\lambda t), \quad t \geq 0, s \geq 0. \end{aligned} \tag{7}$$

Similarly for $n \geq 1$, using (N3)–(N6), we get

$$P_n(t+h) = P(N(t+h) = n) = I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &= P(N(t) = n, N(t+h) - N(t) = 0), \\ I_2 &= P(N(t) = n-1, N(t+h) - N(t) = 1), \\ I_3 &= P(N(t) \leq n-2, N(t+h) - N(t) \geq 2), \end{aligned}$$

and hence

$$P_n(t+h) = P_n(t)[1 - \lambda h + o(h)] + P_{n-1}(t)[\lambda h + o(h)] + o(h).$$

We now get as before

$$\frac{d}{dt} P_n(t) = -\lambda P_n(t) + \lambda P_{n-1}(t), \qquad t > 0. \tag{8}$$

To solve (8) inductively, set $Q_n(t) = e^{\lambda t} P_n(t), t \geq 0, n = 0, 1, 2, \cdots$ Then using (8) it is easy to get

$$\frac{d}{dt} Q_n(t) = \lambda Q_{n-1}(t), t > 0, n = 1, 2, \cdots$$

By (7) note that $Q_0(\cdot) \equiv 1$. So from the equation above it follows that $Q_1(t) = \lambda t + c$ where c is a constant. As $Q_1(0) = P_1(0) = P(N(0) = 1) = 0$, clearly $c = 0$; hence $Q_1(t) = \lambda t, t \geq 0$. Using the above equation for $Q_n(\cdot)$ and the initial value $Q_n(0) = P_n(0) = P(N(0) = n) = 0, n \geq 2$ it is now simple to obtain inductively $Q_n(t) = \frac{1}{n!}\lambda^n t^n$. Hence

$$P_n(t) = e^{-\lambda t} Q_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \qquad n = 0, 1, 2, \cdots, t \geq 0.$$

Thus we have proved

Theorem 1 *Let the stochastic process* $\{N(t) : t \geq 0\}$ *satisfy the postulates (N1)–(N6). Then for any* $t \geq 0, s \geq 0, k = 0, 1, 2, \cdots$

$$P(N(t+s) - N(s) = k) = P(N(t) = k) = \frac{(\lambda t)^k}{k!} \exp(-\lambda t). \tag{9}$$

The stochastic process $\{N(t)\}$ is called a *time homogeneous Poisson process with arrival rate* $\lambda > 0$. □

Remark 3: The assumptions (N1)–(N6) are qualitative, whereas the conclusion is quantitative. Such a result is usually indicative of a facet of nature; that is, phenomena observed in different disciplines, in unrelated contexts may exhibit the same law/ pattern. In fact, Poisson distribution and Poisson process do come up in diverse fields like physics, biology, engineering, and economics. See [Fe], [KT], [Ro1]. □

The Poisson arrival model owes its versatility to the fact that many natural (and, of course, useful) quantities connected with the model can be explicitly determined. We give a few examples which are relevant in the context of insurance as well.

Interarrival and waiting time distributions

Let $\{N(t) : t \geq 0\}$ be a Poisson process with arrival rate $\lambda > 0$. Set $T_0 \equiv 0$. For $n = 1, 2, \cdots$ define $T_n = \inf\{t \geq 0 : N(t) = n\}$ = time of arrival of n-th claim (or waiting time until the n-th claim arrival). Put $A_n = T_n - T_{n-1}, n = 1, 2, \cdots$ so that A_n = time between $(n-1)$-th and n-th claim arrivals. Recall from our initial comments that we had in fact defined the process $\{N(t)\}$ starting from $\{T_i\}$. The random variables $T_0, T_1, T_2, \cdots$ are called *claim arrival times* (or *waiting times*); the sequence $\{A_n : n = 1, 2, \cdots\}$ is called the sequence of *interarrival times.*

For any $s > 0$ note that $\{T_1 > s\} = \{N(s) = 0\}$; hence by (9)

$$P(A_1 > s) = P(T_1 > s) = P(N(s) = 0) = \exp(-\lambda s). \tag{10}$$

So $P(A_1 \leq s) = 1 - e^{-\lambda s}, s \geq 0$. Therefore the random variable A_1 has an $\text{Exp}(\lambda)$ distribution (= exponential distribution with parameter $\lambda > 0$); that is,

$$P(A_1 \in (a,b)) = \int_a^b \lambda e^{-\lambda s} ds, \qquad 0 \le a \le b < \infty. \tag{11}$$

Next let us consider the joint distribution of (T_1, T_2). Let $F_{(T_1,T_2)}$ denote the joint distribution function of (T_1, T_2); that is, $F_{(T_1,T_2)}(t_1, t_2) = P(T_1 \le t_1, T_2 \le t_2)$. As $0 \le T_1 \le T_2$ it is enough to look at $F_{(T_1,T_2)}(t_1, t_2)$ for $0 \le t_1 \le t_2$. It is clear that for $0 \le t_1 \le t_2$,

$$\begin{aligned} \{T_1 &\le t_1, T_2 \le t_2\} = \{N(t_1) \ge 1, N(t_2) \ge 2\} \\ &= \{N(t_1) = 1, N(t_2) - N(t_1) \ge 1\} \cup \{N(t_1) \ge 2\}, \end{aligned}$$

where the r.h.s. is a disjoint union. So using properties (N3),(N4) and equation (9) we get

$$\begin{aligned} F_{(T_1,T_2)}(t_1,t_2) &= P(N(t_1) = 1, N(t_2) - N(t_1) \ge 1) + P(N(t_1) \ge 2) \\ &= \lambda t_1 e^{-\lambda t_1}(1 - e^{-\lambda(t_2 - t_1)}) + [1 - (e^{-\lambda t_1} + \lambda t_1 e^{-\lambda t_1})] \\ &= -\lambda t_1 e^{-\lambda t_2} + H(t_1) \end{aligned}$$

where H is a function depending only on t_1. Consequently the joint probability density function $f_{(T_1,T_2)}$ of (T_1, T_2) is given by

$$\begin{aligned} f_{(T_1,T_2)}(t_1,t_2) &\triangleq \frac{\partial^2}{\partial t_2 \partial t_1} F_{(T_1,T_2)}(t_1,t_2) \\ &= \left\{ \begin{array}{ll} \lambda^2 e^{-\lambda t_2}, & \text{if} \quad 0 < t_1 < t_2 < \infty \\ 0, & \text{otherwise} \end{array} \right\} \end{aligned} \tag{12}$$

To find the joint distribution of (A_1, A_2) from the above, note that

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} T_1 \\ T_2 - T_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}. \tag{13}$$

The linear transformation given by the (2×2) matrix in (13) has determinant 1, and transforms the region $\{(t_1, t_2) : 0 < t_1 < t_2 < \infty\}$ in $1-1$ fashion onto $\{(a_1, a_2) : a_1 > 0, a_2 > 0\}$. So the joint probability density function $f_{(A_1,A_2)}$ of (A_1, A_2) is given by

$$\begin{aligned} f_{(A_1,A_2)}(a_1,a_2) &= f_{(T_1,T_2)}(a_1, a_1 + a_2) \\ &= \left\{ \begin{array}{ll} (\lambda e^{-\lambda a_1})(\lambda e^{-\lambda a_2}), & \text{if} \quad a_1 > 0, a_2 > 0 \\ 0, & \text{otherwise} \end{array} \right\} \end{aligned} \tag{14}$$

Thus A_1, A_2 are independent random variables each having an exponential distribution with parameter λ.

With more effort one can prove

Theorem 2 *Let* $\{N(t) : t \geq 0\}$ *be a time homogeneous Poisson process with arrival rate* $\lambda > 0$*. Let* $A_1, A_2, \cdots$ *denote the interarrival times. Then* $\{A_n : n = 1, 2, \cdots\}$ *is a sequence of independent, identically distributed random variables (or in other words an i.i.d. sequence) having Exp(λ) distribution.* $\Box$

In view of the argument above for the case $n = 2$, the general idea of the proof is clear. One proves first that the joint distribution function of $T_1, T_2, \cdots, T_n$ is given by

$$F_{(T_1,T_2,\cdots,T_n)}(t_1, t_2, \cdots, t_n) = -\lambda^{n-1}\left(\prod_{i=1}^{n-1} t_i\right) e^{-\lambda t_n} + H(\underline{t}),$$

if $0 \leq t_1 < t_2 < \cdots < t_n < \infty$, where $H(\cdot)$ is a function such that $\partial^n H/(\partial t_1 \partial t_2 \cdots \partial t_n) = 0$. In fact $H(\cdot)$ is a sum of a finite number of terms; each term is a product of powers of t_i and $e^{-\lambda t_j}$ with atleast one $t_k, k \geq 2$ missing! Establishing this is the tedious part of the proof. Once this is done the joint probability density function of $T_1, T_2, \cdots, T_n$ is given by

$$f_{(T_1,T_2,\cdots,T_n)}(t_1, t_2, \cdots t_n) = \left\{ \begin{array}{cc} \lambda^n \exp(-\lambda t_n), & \text{if } 0 < t_1 < t_2 < \cdots < t_n < \infty \\ 0, & \text{otherwise} \end{array} \right\}$$

Note that the analogue of (13) is

$$\begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_n \end{pmatrix} = \begin{pmatrix} T_1 \\ T_2 - T_1 \\ T_3 - T_2 \\ \vdots \\ T_n - T_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \\ T_3 \\ \vdots \\ T_n \end{pmatrix}$$

One can now proceed exactly as in the earlier case to obtain the theorem. The reader is invited to work out the details at least when $n = 3, 4$. $\Box$

Note: As A_1 has Exp(λ) distribution, its expectation is given by $E(A_1) = \frac{1}{\lambda}$; so $\frac{1}{\lambda}$ is the mean arrival time. Thus the arrival rate being λ is consistent with this conclusion.

Note: It is an easy corollary of the theorem that $T_n = A_1 + A_2 + \cdots + A_n$ has the gamma distribution $\Gamma(n, \lambda)$.

Remark 4: One can also go in the other direction. That is, let $0 = T_0 \leq T_1 \leq T_2 \leq \cdots$ be the claim arrival times; let $A_n = T_n - T_{n-1}, n \geq 1$. Suppose $\{A_n\}$ is an i.i.d. sequence having Exp(λ) distribution. Define $\{N(t)\}$ by (1). Then the stochastic process $\{N(t) : t \geq 0\}$ can be shown to be time homogeneous Poisson process with rate λ. In the jargon of the theory of stochastic processes, Poisson process is the renewal process with i.i.d. exponential arrival rates.

Order statistics property

This is another important property of the Poisson process. Recall that for events G, H, the conditional probability of G given H is defined by $P(G \mid H) \triangleq \frac{P(G\cap H)}{P(H)}$. We first prove

Theorem 3 *Notation as earlier. For* $0 \le s \le t$,

$$P(A_1 < s \mid N(t) = 1) = \frac{s}{t}; \tag{15}$$

that is, given that exactly one arrival has taken place in $[0,t]$*, the time of the arrival is uniformly distributed over* $(0,t)$.

Proof: As the Poisson process has independent increments,

$$\begin{aligned} P(A_1 < s \mid N(t) = 1) &= \frac{P(A_1 < s, N(t) = 1)}{P(N(t) = 1)} \\ &= \frac{P(N(s) = 1, N(t) - N(s) = 0)}{P(N(t) = 1)} \\ &= \frac{P(N(s) = 1) \cdot P(N(t) - N(s) = 0)}{P(N(t) = 1)} \\ &= \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} = \frac{s}{t}, \end{aligned}$$

completing the proof. □

A natural question is: If $N(t) = n$, what can one say about the conditional distribution of $T_1, T_2, \cdots, T_n$?

Theorem 4 *Let* $\{N(t) : t \ge 0\}, T_1, T_2, \cdots$ *be as before. For any* $t > 0$, *and any* $n = 1, 2, \cdots$ *the conditional density of* $(T_1, T_2, \cdots, T_n)$ *given* $N(t) = n$ *is*

$$f_{T_1,T_2,\cdots,T_n}((s_1, s_2, \cdots, s_n) \mid N(t) = n) = n! \cdot \frac{1}{t^n}, \tag{16}$$

for $0 < s_1 < s_2 < \cdots < s_n < t$.

Proof: For notational simplicity we take $n = 2$; the general case is similar. Let $0 < s_1 < s_2 < t$; take $h_1, h_2 > 0$ small enough that $0 < s_1 < s_1 + h_1 < s_2 < s_2 + h_2 < t$. Then again using the independent increment property and (9), we get

$$\begin{aligned} P(s_1 &< T_1 < s_1 + h_1, s_2 < T_2 < s_2 + h_2 \mid N(t) = 2) \\ &= \frac{1}{P(N(t) = 2)} P \begin{pmatrix} N(s_1) = 0, N(s_1 + h_1) - N(s_1) = 1, \\ N(s_2) - N(s_1 + h_1) = 0, \\ N(s_2 + h_2) - N(s_2) = 1, N(t) - N(s_2 + h_2) = 0 \end{pmatrix} \\ &= \frac{e^{-\lambda s_1} \lambda h_1 e^{-\lambda(s_1 + h_1 - s_1)} e^{-\lambda(s_2 - (s_1 + h_1))} \lambda h_2 e^{-\lambda(s_2 + h_2 - s_2)} e^{-\lambda(t - (s_2 + h_2))}}{e^{-\lambda t}(\lambda t)^2/2!} \\ &= \frac{2!}{t^2} h_1 h_2. \end{aligned}$$

Dividing by $h_1 h_2$ and letting $h_1, h_2 \downarrow 0$ we get the desired result. $\square$

Remark 5: Let $Y_1, Y_2, \cdots, Y_n$ be i.i.d. random variables with probability density function $f(\cdot)$; that is, $P(Y_i \in G) = \int\limits_G f(x)dx$, for any reasonable subset G of $\mathbb{R}$. Let $Y_{(1)} \leq Y_{(2)} \leq \cdots Y_{(n)}$ denote the order statistics of $Y_1, Y_2, \cdots, Y_n$; (that is, $Y_{(k)}(\omega)$ is the k-th smallest value among $Y_1(\omega), Y_2(\omega), \cdots, Y_n(\omega), k = 1, 2, \cdots, n$ for any $\omega \in \Omega$). Clearly the joint probability density function of $Y_1, Y_2, \cdots, Y_n$ is $f_{(Y_1, Y_2, \cdots, Y_n)}(y_1, y_2, \cdots, y_n) = f(y_1)f(y_2)\cdots f(y_n)$. Note that $(Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)})$ takes values in the set $\Delta = \{y_1 < y_2 < \cdots < y_n\}$. (Why?). Let $B \subset \Delta$ be a reasonable set. Let $B_i, i = 1, 2, \cdots, (n!)$ correspond to disjoint sets obtained by permutation of coordinates in B. Observe that

$$\begin{aligned}
P((Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)}) \in B) &= P((Y_1, Y_2, \cdots, Y_n) \in \bigcup_{i=1}^{n!} B_i) \\
&= \sum_{i=1}^{n!} P((Y_1, Y_2, \cdots, Y_n) \in B_i) \\
&= \sum_{i=1}^{n!} \int\limits_{B_i} f(z_1)f(z_2)\cdots f(z_n)dz_1 dz_2 \cdots dz_n \\
&= (n!) \int\limits_B f(z_1)f(z_2)\cdots f(z_n)dz_1 dz_2 \cdots dz_n.
\end{aligned}$$

So the joint probability density function of $Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)}$ is given by

$$g(y_1, y_2, \cdots, y_n) = \left\{ \begin{array}{cc} (n!)f(y_1)f(y_2)\cdots f(y_n), & \text{if } y_1 < y_2 < \cdots < y_n \\ 0 & \text{otherwise} \end{array} \right\} \tag{17}$$

Now let $V_1, V_2, \cdots, V_n$ be i.i.d. random variables each having a uniform distribution over $(0, t)$, where $t > 0$ is fixed. Note that the probability density function of each V_i is given by $f_{V_i}(s) = \frac{1}{t}$, if $0 < s < t$ (and it is 0 otherwise). Let $V_{(1)} \leq V_{(2)} \leq \cdots \leq V_{(n)}$ denote the order statistics of $V_1, V_2, \cdots, V_n$. That is, $V_{(1)}(\omega), V_{(2)}(\omega), \cdots, V_{(n)}(\omega)$ denotes $V_1(\omega), V_2(\omega), \cdots, V_n(\omega)$ arranged in increasing order for any $\omega \in \Omega$. By (17) it is clear that the joint probability density function of $V_{(1)}, V_{(2)}, \cdots, V_{(n)}$ is given by the r.h.s. of (16). So the preceding theorem means that $((T_1, T_2, \cdots, T_n) \mid N(t) = n) \stackrel{d}{=} (V_{(1)}, V_{(2)}, \cdots, V_{(n)})$, where $\stackrel{d}{=}$ denotes that two sides have the same probability distribution. If $U_1, U_2, \cdots, U_n$ are i.i.d. U(0,1) random variables (that is, having uniform distribution over $(0, 1)$), then the above can be expressed as

$$((T_1, T_2, \cdots, T_n) \mid N(t) = n) \stackrel{d}{=} (tU_{(1)}, tU_{(2)}, \cdots, tU_{(n)}). \tag{18}$$

$\square$

An important consequence of Theorem 4 and Remark 5 is the following result whose proof is quite involved; see [Mi].

Theorem 5 *Let* $\{N(t) : t \geq 0\}$ *be a time homogeneous Poisson process with rate* $\lambda > 0$; *let* $0 < T_1 < T_2 < \cdots$ *denote the claim arrival times corresponding to* $N(\cdot)$. *Let* $\{X_i : i = 1, 2, \cdots\}$ *be an i.i.d. sequence independent of the process* $\{N(t)\}$. *Then there exists a sequence* $\{U_j : j = 1, 2, \cdots\}$ *such that (i)* $\{U_j\}$ *is a sequence of i.i.d. random variables having U(0,1) distribution, (ii) the families* $\{U_j\}, \{X_i\}, \{N(t)\}$ *are independent of each other, (iii) for any reasonable function* g *of two variables*

$$\sum_{i=1}^{N(t)} g(T_i, X_i) \stackrel{d}{=} \sum_{i=1}^{N(t)} g(tU_i, X_i), \quad t \geq 0. \tag{19}$$

$\square$

The basic strategy for proving Theorem 5 can be easily stated. Conditioning the l.h.s. of (19) by $\{N(t) = n\}$, we use Theorem 4 to replace T_i by $tU_{(i)}$. Then invoking independence of the families $\{U_i\}, \{X_j\}$ and the fact that X_j's are i.i.d.'s, we permute $X_1, X_2, \cdots, X_n$ suitably to facilitate the desired conclusion. Mathematical justification requires measure theoretic machinery.

Cramer-Lundberg model

This is the classical and very versatile model in insurance. The claim arrivals $\{T_i\}$ happen as in a time homogeneous Poisson process with rate $\lambda > 0$. The claim sizes $\{X_i\}$ are i.i.d. nonnegative random variables. The sequences $\{X_i\}, \{T_j\}$ are independent of each other. The *total claim amount* upto time t in this model is given by

$$S(t) = X_1 + X_2 + \cdots + X_{N(t)} = \sum_{i=1}^{N(t)} X_i, \quad t \geq 0.$$

which is the same as (2). Note that $\{S(t) : t \geq 0\}$ is an example of a *compound Poisson process.*

We now look at the *discounted sum* corresponding to the above model. Let $r > 0$ denote the *interest rate.* Define

$$S_0(t) = \sum_{i=1}^{N(t)} e^{-rT_i} X_i, \quad t \geq 0. \tag{20}$$

This is the "present value" (at time 0) of the cumulative claim amount over the time horizon $[0, t]$. By Theorem 5 for any $t \geq 0$

$$S_0(t) \stackrel{d}{=} \sum_{i=1}^{N(t)} e^{-rtU_i} X_i, \tag{21}$$

where $\{U_i\}$ is an i.i.d. U(0,1) sequence as in the theorem. Therefore using the independence of the three families of random variables we get

$$\begin{aligned}
E(S_0(t)) &= E\left(\sum_{i=1}^{N(t)} e^{-rtU_i} X_i\right) \\
&= \sum_{n=0}^{\infty} E\left[\sum_{i=1}^{n} e^{-rtU_i} X_i \mid N(t)=n\right] \cdot P(N(t)=n) \\
&= \sum_{n=0}^{\infty} n \cdot E\left[e^{-rtU_1}\right] \cdot E(X_1) \cdot P(N(t)=n) \\
&= E(N(t)) \cdot E(X_1) \cdot \left(\int_0^1 e^{-rty} dy\right) = \lambda \frac{1}{r}(1-e^{-rt}) \cdot E(X_1).
\end{aligned}$$

So we have proved

Theorem 6 *With the notation as above*

$$E\left(\sum_{i=1}^{N(t)} e^{-rT_i} X_i\right) = \lambda \frac{1}{r}(1-e^{-rt}) \cdot E(X_1). \tag{22}$$

That is, in the Cramer-Lundberg model, the average/ expected amount needed to take care of claims over $[0,t]$ *is given by (22).* □

Next let $p(t)$ denote the *premium income* in the time interval $[0,t]$. In the Cramer-Lundberg model it is assumed that $p(\cdot)$ is a deterministic linear function; that is, $p(t)=ct, t \geq 0$ where $c>0$ is a constant called the *premium rate.* With the total claim amount $S(\cdot)$ defined by (2), put for $t \geq 0$,

$$U(t) = u + p(t) - S(t) = u + ct - \sum_{i=1}^{N(t)} X_i. \tag{23}$$

The process $\{U(t) : t \geq 0\}$ is called the *risk process* (or *surplus process*) of the model; here u is the *initial capital.* Note that $U(t)$ is the insurance company's capital balance at time t. Letting $r \downarrow 0$ in (22) or otherwise, note that $E(S(t)) = \lambda t E(X_1)$ and hence

$$E(U(t)) = u + ct - E(S(t)) = u + ct - \lambda t E(X_1) \tag{24}$$

By (24), a minimal requirement in choosing the premium rate may be taken to be

$$c > \lambda E(X_1) \tag{25}$$

so that on the average, claim payments are taken care of by premium income. This somewhat simple criterion can be justified by other considerations also, as we shall see later. A more prudent condition is to require that $c > (1+\rho)\lambda E(X_1)$, where $\rho > 0$ is called a *safety loading factor.*

Ruin problem in the Cramer-Lundberg model

As mentioned earlier, in an insurance set-up the financial risk is shifted to the insurance company to a large extent. There have been many instances when insurance companies have gone bankrupt unable to cope up with claims during major catastrophes. So a theoretical understanding of conditions leading to ruin of the company, probability of ruin, severity of ruin, etc. will help at least in avoiding certain pitfalls. Study of ruin problems has, therefore, a central place in insurance mathematics.

The event that the surplus $U(\cdot)$ falls below zero is called *ruin.* Set

$$T = \inf\{t > 0 : U(t) < 0\}; \tag{26}$$

T is called the *ruin time*; it is the first time the surplus falls below zero. The *probability of ruin* is then

$$\psi(u) = P(T < \infty \ \mid \ U(0) = u) \tag{27}$$

for $u > 0$; it is considered as a function of the initial capital u. Note that $\psi(\cdot)$ depends on the premium rate c as well. A very natural question is: For what premium rates c and initial capital u can it happen that $\psi(u) = 1$? That is, when is ruin certain?

By the definition of $U(\cdot)$, note that $U(\cdot)$ increases in the intervals $[T_n, T_{n+1}), n \geq 0$. Therefore ruin can occur only at some T_n. Now for $n \geq 1$,

$$\begin{aligned} U(T_n) &= u + cT_n - \sum_{i=1}^{n} X_i \qquad (\text{because } \ N(T_n) = n) \\ &= u + \sum_{i=1}^{n} (cA_i - X_i) \qquad (\text{because } \ T_n = \sum_{i=1}^{n} A_i) \end{aligned} \tag{28}$$

Put $Z_i = X_i - cA_i, i \geq 1, S_0 = 0, S_n = \sum\limits_{i=1}^{n} Z_i, n \geq 1$. Then (28) is just $U(T_n) = u - S_n, n \geq 1$. Since "ruin" $=$ $\{U(T_n) < 0$ for some $n\}$, it is now easy to see that

$$\psi(u) = P(\sup_{n \geq 1} S_n > u). \tag{29}$$

Since the families $\{A_i\}$ and $\{X_j\}$ are mutually independent, and each is a sequence of i.i.d.'s, note that $\{Z_i\}$ is a sequence of i.i.d.'s and hence $\{S_n : n \geq 0\}$ is a *random walk* on the real line $\mathbb{R}$. The following result concerning random walks on $\mathbb{R}$ is known.

Theorem 7 *Let* $\{Z_i\}, \{S_n\}$ *be as above; assume that* Z_i *is not identically zero, and* $E(Z_i)$ *exists.*

(a) If $E(Z_1) > 0$, *then* $P(\lim_{n\to\infty} S_n = +\infty) = 1$.

(b) If $E(Z_1) < 0$, *then* $P(\lim_{n\to\infty} S_n = -\infty) = 1$.

(c) If $E(Z_1) = 0$, *then* $P(\limsup_{n\to\infty} S_n = +\infty, \liminf_{n\to\infty} S_n = -\infty) = 1$.

□

Note: While (a),(b) above are immediate consequences of the strong law of large numbers, assertion (c) requires a lengthy proof; see [RSST]. The special case when Z_i can take only the values ± 1 (so that at each stage the transition is only to one of the nearest neighbours) is well known; see [KT], [Ro1], [Ro2]. The intuition behind the general case is similar. □

From (29) and the above theorem it follows that $\psi(u) = 1$ for all $u > 0$, if $E(X_1) - cE(A_1) \geq 0$; note that $E(A_1) = \frac{1}{\lambda}$ as A_1 has an exponential distribution with parameter λ; so in the Cramer-Lundberg model ruin is certain if (25) is not satisfied. The condition (25) is called the *net profit condition,* which is generally assumed to be satisfied.

If (25) holds, the above does not imply that ruin is avoided; it only means that one may hope to have $\psi(u) < 1, u > 0$.

Now let $u > 0$. Suppose there exists $r > 0$ such that

$$E[\exp(rZ_1)] = E[\exp(r(X_1 - cA_1))] = 1. \tag{30}$$

Such an $r > 0$ is called the *adjustment or Lundberg coefficient*. This leads to a useful martingale. Recall that a sequence $\{Y_n : n = 0, 1, 2, \cdots\}$ of integrable random variables is said to be a *martingale* if the conditional expectation $E(Y_{n+1}|Y_0 = y_0, Y_1 = y_1, \cdots, Y_n = y_n) = y_n$ for any $y_0, y_1, \cdots, y_n$ with $P(Y_i = y_i, i = 0, 1, 2, \cdots, n) > 0$, for $n = 0, 1, 2, \cdots$. Martingale theory is a powerful tool in a probabilist's kit. We illustrate this in the following discussion. One may see [KT], [Ro2] for some of the elementary aspects and applications of martingales, and [RSST] for applications to insurance.

Theorem 8 *Let* $\{Z_i\}, \{S_n\}$ *be as above. Suppose there exists an adjustment coefficient* $r > 0$*. Then* $\{\exp(rS_n) : n = 0, 1, 2, \cdots\}$ *is a martingale.*

Proof: Integrability of e^{rS_n} is left as an exercise. (Use the fact that if W_1, W_2 are independent integrable random variables then $W_1 W_2$ also has finite expectation and $E(W_1 W_2) = E(W_1)E(W_2)$). To prove the theorem we just need to prove for $n = 0, 1, 2, \cdots$

$$E[\exp(rS_{n+1}) \quad | \quad S_0 = 0, S_1 = s_1, \cdots, S_n = s_n] = \exp(rs_n).$$

Indeed, by independence of Z_{n+1} and $\{S_i : i \leq n\}$, using (30) we get

$$\begin{aligned} \text{l.h.s. of the above} &= E[e^{rS_n} e^{rZ_{n+1}} \mid S_0 = 0, S_1 = s_1, \cdots, S_n = s_n] \\ &= E[e^{rs_n} e^{rZ_{n+1}} \mid S_0 = 0, S_1 = s_1, \cdots, S_n = s_n] \\ &= e^{rs_n} E[e^{rZ_{n+1}} \mid S_0 = 0, S_1 = s_1, \cdots, S_n = s_n] \\ &= e^{rs_n} E(e^{rZ_{n+1}}) \\ &= e^{rs_n} \ , \end{aligned}$$

completing the proof. □

As $E(e^{rZ}) = 1$ if $r = 0$ for any random variable Z, no useful purpose will be served if the adjustment coefficient is allowed to be zero. So we seek a unique $r > 0$ satisfying (30). In this direction we have

Theorem 9 *Notation as above. Suppose* $E(Z_1) < 0$; *this is the net profit condition. Assume moreover that there is* $h_1 > 0$ *such that the moment generating function* $m(h) \triangleq E(e^{hZ_1}) < \infty$ *for all* $-h_1 < h < h_1$, *and that* $\lim_{h \uparrow h_1} m(h) = \infty$. *Then there is a unique adjustment coefficient* $r > 0$.

Proof: As the moment generating function exists in a neighbourhood of zero, note that $m(\cdot)$ has derivatives of all orders in $(-h_1, h_1)$. It can be shown that $m'(h) = E(Z_1 e^{hZ_1}), m''(h) = E(Z_1^2 e^{hZ_1})$ for $-h_1 < h < h_1$; these can be justified using the dominated convergence theorem. As $E(Z_1) < 0$ it follows that $P(Z_1 \neq 0) > 0$, and hence $m''(h) > 0$. So m is convex on $(-h_1, h_1)$. Again by hypothesis $m'(0) = E(Z_1) < 0$; therefore m is decreasing in a neighbourhood of 0. Since $m(h) \uparrow \infty$ as $h \uparrow h_1$, and $m(0) = 1$, it now follows that there exists a unique $s \in (0, h_1)$ such that $m(s) < 1, m'(s) = 0$, and that on (s, h_1) the function $m(\cdot)$ is strictly increasing to $+\infty$. Consequently there is a unique $r \in (0, h_1)$ such that $m(r) = 1$. As the moment generating function of Z_1 does not exist on $[h_1, \infty)$ (Why?), uniqueness on $(0, \infty)$ follows. □

Remark 6: In the Cramer-Lundberg model $Z_1 = X_1 - cA_1, A_1$ has $\text{Exp}(\lambda)$ distribution, $c > 0, \lambda > 0$. So for $h > 0$, we have $m(h) \triangleq m_{Z_1}(h) = m_{X_1}(h) m_{A_1}(-ch) = m_{X_1}(h) \frac{\lambda}{\lambda + ch}$; (here for a random variable W we write $m_W(h) = E(e^{hW})$). Thus, if the net profit condition (25) holds, and if there exists $h_1 > 0$ such that $m_{X_1}(h) < \infty, h < h_1$ and $\lim_{h \uparrow h_1} m_{X_1}(h) = +\infty$, then by the above theorem there is a unique adjustment coefficient $r > 0$. In addition suppose that $X_1 \sim \text{Exp}(\theta)$; that is, claim sizes are i.i.d. exponentially distributed random variables with parameter $\theta > 0$. Then $m_{X_1}(h) = \frac{\theta}{\theta - h}$ for $h < \theta$ and clearly $\lim_{h \uparrow \theta} m_{X_1}(h) = +\infty$. Observe that (25) holds $\Leftrightarrow \frac{\lambda}{\theta} < c \Leftrightarrow \theta - (\lambda/c) > 0$. It is easily checked that $m_{Z_1}(h) = 1, h > 0 \Leftrightarrow \frac{\theta}{(\theta - h)} \frac{\lambda}{(\lambda + ch)} = 1, h > 0 \Leftrightarrow h = \theta - (\lambda/c) > 0$. Therefore in this case the unique adjustment coefficient is given by $r = \theta - (\lambda/c)$. □

Note: Suppose $X_1 \geq 0, P(X_1 > 0) > 0$. Then there exist $\varepsilon > 0, \delta > 0$ such that $P(X_1 \geq \varepsilon) \geq \delta$. Consequently for $h > 0$, by Chebyshev's inequality

$$E[e^{hX_1}] \geq e^{h\varepsilon} P(e^{hX_1} \geq e^{h\varepsilon}) \geq \delta e^{h\varepsilon} \to \infty, \quad \text{as} \quad h \uparrow \infty.$$

Thus in the Cramer-Lundberg model, if the net profit condition (25) holds and if the moment generating function of X_1 exists in a neighbourhood of 0, then a unique adjustment coefficient $r > 0$ exists. □

Assume now that the net profit condition (25) holds and that an adjustment coefficient $r > 0$ exists. Put

$$\begin{aligned} \tau_u &= \inf\{n \geq 0 : S_n > u\} \\ &= \text{first time the random walk } \{S_n\} \text{ exceeds } u. \end{aligned}$$

Observe that $\{\tau_u \leq k\} = \bigcup_{i=0}^{k} \{S_i > u\}$; so to know if $\tau_u \leq k$ or not it is enough to look at $S_i, 0 \leq i \leq k$, for any $k = 0, 1, 2, \cdots$; in other words τ_u is a *stopping time.* Clearly $\psi(u) = P(\tau_u < \infty)$. Similarly for $a > 0$ put $\sigma_a = \inf\{n \geq 0 : S_n < (-a)\}$; this is also a stopping time. As (25) holds, by Theorem 7(b), $P(\sigma_a < \infty) = 1$. Consequently $(\tau_u \wedge \sigma_a) \triangleq \min\{\tau_u, \sigma_a\}$ is a stopping time with $P((\tau_u \wedge \sigma_a) < \infty) = 1$. Since $\{\exp(rS_n) : n = 0, 1, 2, \cdots\}$ is a martingale by Theorem 8, using the optional sampling theorem we now get

$$\begin{aligned} 1 &= E[e^{rS_0}] = E[\exp(rS_{(\tau_u \wedge \sigma_a)})] \\ &= E[\exp(rS_{\tau_u}) : \{\tau_u < \sigma_a\}] + E[\exp(rS_{\sigma_a}) : \{\sigma_a < \tau_u\}] \\ &\geq E[\exp(rS_{\tau_u}) : \{\tau_u < \sigma_a\}] \\ &\geq e^{ru} P(\tau_u < \sigma_a) \end{aligned} \tag{31}$$

as $S_{\tau_u} > u$. (In the above $E[g : B] \triangleq \int_B g$ denotes expectation of gI_B, for any set B and any random variable g.) It can be shown without much difficulty that $P(\lim_{a\to\infty} \sigma_a = \infty) = 1$; for otherwise, $P(S_k = -\infty) > 0$ for some k, which is not possible as $E(|S_n|) < \infty$ for all n. So letting $a \uparrow \infty$ in (31) we get $P(\tau_u < \infty) = \lim_{a\to\infty} P(\tau_u < \sigma_a) \leq e^{-ru}$. Thus we have proved

Theorem 10 *In the Cramer-Lundberg model let the net profit condition (25) hold, and let an adjustment coefficient* $r > 0$ *exist. Then the ruin probability* $\psi(\cdot)$ *satisfies*

$$\psi(u) = P(\tau_u < \infty) \leq \exp(-ru), \qquad u > 0. \tag{32}$$

(32) is known as Lundberg inequality. □

Note: Note that $|S_n(\omega)| \leq |u| + |a|$, for all $n \leq (\tau_u \wedge \sigma_a)(\omega)$ for all ω. So there is no problem in applying the optional sampling theorem in the above proof.

Remark 7: Consider the Cramer-Lundberg model with $X_1 \sim \text{Exp}(\theta), \theta > 0$. Then Remark 6 and Theorem 10 imply that the ruin probability satisfies

$$\psi(u) \leq \exp\{-(\theta - \frac{\lambda}{c})u\}, \quad u > 0,$$

where $c > 0$ is the premium rate. As expected, larger the initial capital u smaller is the ruin probability $\psi(u)$. Now assume that the premium rate is determined by

$$ct = p(t) = (1+\rho)E(S(t)) = (1+\rho)\frac{\lambda}{\theta}t, \tag{33}$$

where $\rho > 0$ is the safety loading factor; see the comments following (25). Hence $c = (1+\rho)\lambda/\theta$, and consequently the adjustment coefficient is given by $r = \theta - \frac{\lambda}{c} = \theta\rho/(1+\rho)$. Therefore

$$\psi(u) \leq \exp\left\{-\theta\left(\frac{\rho}{1+\rho}\right)u\right\}, \quad u > 0. \tag{34}$$

As $\rho \uparrow \infty$, by (33) $c \uparrow \infty$, while by (34) the bound on $\psi(u)$ does not change significantly; the latter assertion is due to the fact that $\rho/(1+\rho) \uparrow 1$ as $\rho \uparrow \infty$. These observations together with some of the preceding results can be interpreted as follows. In view of the net profit condition (25) and Theorem 7, relatively small values of the safety loading factor ρ afford good cushion against ruin with just marginal increases in premium rate. However, large values of ρ steeply increase the premium rate, without perhaps offering comparable assurances against ruin; in such a case the policy becomes unattractive to the insured without tangible benefits for the company. □

Remark 8: In addition to the above bound one can also derive an integral equation for the ruin probability. Suppose $E(X_1) < \infty$ and that the net profit condition holds. Then one can get a distribution function G (explicitly in terms of the distribution function of X_1) such that

$$\psi(t) = \frac{\lambda E(X_1)}{c}[1 - G(t)] + \frac{\lambda E(X_1)}{c}\int_0^t \psi(t-y)dG(y). \tag{35}$$

The equation (35) is a renewal type equation; however it is a defective renewal equation because $\lambda E(X_1)/c < 1$ (as the net profit condition holds). Still following the methods of renewal theory one can get a series solution to (35). See [Mi], [RSST] for details. □

Claim size distribution

The common distribution of the i.i.d. sequence $\{X_i\}$ is called the *claim size distribution.* With the exception of Theorem 10 and the discussion leading to it, we have not made any specific reference to the claim size distribution so far. A conventional assumption is that X_i have an exponential distribution. In such a case $P(X_i > x) = e^{-\lambda x}, x > 0$, where $\lambda > 0$; that is, the (right) tail of the claim size distribution decays at an exponential rate. Most of the distributions used for modelling in statistics have this property. The ubiquitous normal (or Gaussian) distribution decays even at a faster rate. Such distributions are called *light tailed distributions*; for these distributions the moment generating functions exist in a neighbourhood of 0.

An important development of late is to consider claim sizes that are not necessarily light tailed. Risks regarding insurance of airplanes, skyscrapers, dams, bridges, etc. are very high. In recent years, companies have also faced ruin or near ruin due to a very small number of very huge claims; in some instances, a single massive claim has done the damage. There are quite a few notions of *heavy tailed distributions*; invariably the moment generating function does not exist in any neighbourhood of 0 for these distributions. A versatile notion of heavy-tailedness in the insurance context is given below.

Let F be a distribution function supported on $(0, \infty)$; (this corresponds to a positive random variable). Let $X_1, X_2, \cdots$ be an i.i.d. sequence with common distribution function F. Set $S_n = \sum_{i=1}^{n} X_i, M_n = \max\{X_1, X_2, \cdots, X_n\}$. If

$$\lim_{x \to \infty} \frac{P(S_n > x)}{P(M_n > x)} = 1, \quad \text{for } n \geq 2, \tag{36}$$

then F is said to be *subexponential.* Equation (36) means that the partial sum and the partial maximum have the same tail behaviour; this corresponds to the intuitive notion that largeness of the cumulative claim is basically determined by that of the biggest claim. If F is subexponential then it can be shown that $e^{ax} P(X \geq x) \to \infty$, for any $a > 0$, where X is a random variable with distribution function F. So moment generating function does not exist in any neighbourhood of 0 for such distributions. Two classes of subexponential distributions are given below.

(i) *Weibul distribution:* In this case $\overline{F}(x) \triangleq 1 - F(x) = \exp(-cx^\tau)$, if $x > 0$, and $\overline{F}(x) = 0$, if $x \leq 0$, where $c > 0, \tau > 0$ are constants. This family of distributions has been useful in reliability theory, besides insurance. If $0 < \tau < 1$, then F is subexponential. See [Mi], [RSST].

(ii) *Pareto distribution:* Again it is convenient to define in terms of the right tail of the distribution; here $\overline{F}(x) \triangleq 1 - F(x) = \kappa^\alpha / (\kappa + x)^\alpha, x > 0$, where $\kappa, \alpha > 0$ are constants. This class is subexponential; (even expectation exists only when $\alpha > 1$.) This family has also been used in economics to describe income distributions.

As the moment generating function does not exist for heavy tailed distributions, note that Theorem 10 is not applicable. In fact, when the claim size distribution belongs to an appropriate subclass of subexponential distributions, it can be established that the ruin probability decays only at a power rate, viz. $\psi(u)$ behaves like $Ku^{-\delta}$ for large u, where $K, \delta > 0$ are suitable constants. Contrast this to the exponential rate e^{-ru}, where $r > 0$ in Theorem 10. So ruin is much more formidable if the claim size distribution is heavy tailed. See [Mi], [RSST].

Assorted comments

We have dealt with a few elementary aspects of just one model. Comments below are meant to give a flavour of some other aspects/ models in insurance.

1. A more general model is the *renewal risk model* (also called *Sparre Andersen model*). In this model, the interarrival times $A_1, A_2, \cdots$ are just i.i.d. nonnegative random variables (not necessarily exponentially distributed). The net profit condition is given by an analogue of (25), viz. $c > E(X_1)/E(A_1)$. Lundberg inequality holds provided that the net profit condition is satisfied and that the adjustment coefficient exists. Renewal risk model with subexponential claim sizes continue to be objects of research.

2. Life insurance/ pension insurance models are generally described in terms of continuous time Markov processes with state space having only a finite number of elements; at least one state is absorbing, and certain transitions may be disallowed. For example, in the simplest life insurance model there are only two states, one signifying "alive" and the absorbing state indicating "dead", reflecting the status of the insured.

3. In addition to the basic insurance aspects, more complex models can be considered. For example, an insurance company can invest part of its surplus in bonds giving returns at fixed rates, and another part in stocks which are subject to the volatility of the market. Some problems of interest are how optimally should these investments be made so that the ruin probability is minimized, or so that the dividend payment by the company is maximized.

4. We have not touched upon any statistical aspect like estimation of claim arrival rate, parameters of the claim size distribution, or when does claim size data indicate heavy tailed behaviour, etc.

[Mi], [RSST] and the references therein deal with the above issues and more.

References

[DKR] M.Delampady, T.Krishnan and S.Ramasubramanian (eds.): *Probability and Statistics.* A volume in "Echoes from *Resonance*", Universities Press, Hyderabad, 2001.

[Fe] W.Feller: *An Introduction to Probability Theory and its Applications,* 2 vols. Wiley-Eastern, New Delhi, 1991.

[HPS] P.G.Hoel, S.C.Port and C.J.Stone: *Introduction to Probability Theory.* Universal Book Stall, New Delhi, 1991.

[KT] S.Karlin and H.M.Taylor: *A First Course in Stochastic Processes.* Academic Press, New York, 1975.

[Mi] T.Mikosch: *Non-life insurance mathematics: an introduction with stochastic processes.* Springer, New Delhi, 2004.

[RSST] T.Rolski, H.Schmidli, V.Schmidt and J.Teugels: *Stochastic processes for insurance and finance.* Wiley, New York, 2001.

[Ro1] S.M.Ross: *Introduction to probability models.* 8th edition. Elsevier India, New Delhi, 2005.

[Ro2] S.M.Ross: *Stochastic processes.* 2nd edition. (Wiley student edition). John Wiley, Singapore, 2004.