On a Stochastic Model in Insurance

S Ramasubramanian

Basic aspects of the classical Cramer–Lundberg insurance model are described.

Introduction

Recent cataclysmic events like the tsunami, torrential downpour, floods, cyclones, earthquakes, etc. underscore the fact that 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 a 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.

The author is with the Statistics and Mathematics Unit, Imdian Statistical Institute, Bangalore. His research interests centre around stochastic processes.

A reasonable mathematical theory of insurance can possibly provide a scientific basis for the trust between the insured and the company.

Keywords
Poisson process, exponential distribution, arrival times, claim size distribution, independent increments, order statistics, uniform distribution, discounted sum, premium, risk/ surplus process, ruin problem, ruin probability, random walk, subexponential distribution.

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

Insurance mathematics today is considered a part of applied probability theory.

various results and the heuristics can be appreciated. While proofs of some of the basic results are given, for some others only partial proofs or heuristic arguments are indicated; of course, in a few cases we are content with just citing an appropriate reference. [1,2] 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 [3–6] for elucidation of terms like random variable, distribution, density, expectation, independence, independent identically distributed (i.i.d.) random variables, etc. (Some of the earlier articles in *Resonance* compiled in [7] 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.

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.

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 $\lim_{h \to 0} \frac{f(h)}{h} = 0$; that is, if f decays at a faster rate than h.

Poisson Processes

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,

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

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

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

  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.

In formulating a model it is desirable that the hypotheses are realistic and simple.

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

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

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 $0 \leq P_0(t) \leq 1$),

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

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

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

Once a model is chosen, theoretical properties and their implications should be considerably rich and obtainable with reasonable ease.

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

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

The assumptions (N1)–(N6) are qualitative, whereas the conclusion is quantitative.

where

$$I_1 = P(N(t) = n, N(t+h) - N(t) = 0),$$
$$I_2 = P(N(t) = n-1, N(t+h) - N(t) = 1),$$
$$I_3 = P(N(t) \leq n-2, N(t+h) - N(t) \geq 2),$$

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

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

Using the initial values $P_n(0) = P(N(0) = n) = 0$, $n \geq 1$ it is fairly easy to inductively solve (8) and get

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

Thus we have proved the following theorem.

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 [3-5]. □

Poisson distribution and Poisson process do come up in diverse fields.

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 Exp(λ) distribution (= exponential distribution with parameter $\lambda > 0$); that is,

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

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, $F_{(T_1,T_2)}(t_1, t_2) = P(T_1 \leq t_1, T_2 \leq t_2)$. As $0 \leq$

$T_1 \leq T_2$ it is enough to look at $F_{(T_1,T_2)}(t_1, t_2)$ for $0 \leq t_1 \leq t_2$. It is clear that for $0 \leq t_1 \leq t_2$,

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

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

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

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 the following theorem.

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.* □

The interarrival times are independent random variables having exponential distribution.

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

Order statistics property is crucial in explicit computations involving Poisson processes.

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

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

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* $\{N(t) : t \geq 0\}, T_1, T_2, \cdots$ *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*

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

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

$$\begin{aligned} P(s_1 < T_1 < s_1 + h_1, s_2 < T_2 < s_2 + h_2 \mid N(t) = 2) \\ = \frac{1}{P(N(t)=2)}. \end{aligned}$$

$$P\left(\begin{array}{c} N(s_1) = 0, N(s_1+h_1) - N(s_1) = 1, \\ N(s_2) - N(s_1+h_1) = 0, \\ N(s_2+h_2) - N(s_2) = 1, N(t) - N(s_2+h_2) = 0 \end{array}\right)$$

$$
\begin{aligned}
&= \frac{1}{e^{-\lambda t}(\lambda t)^2/2!} \\
&\quad \cdot\{e^{-\lambda s_1}\lambda h_1 e^{-\lambda(s_1+h_1-s_1)}e^{-\lambda(s_2-(s_1+h_1))} \\
&\quad \cdot\lambda h_2 e^{-\lambda(s_2+h_2-s_2)}e^{-\lambda(t-(s_2+h_2))}\} \\
&= \frac{2!}{t^2}h_1 h_2.
\end{aligned}
$$

Dividing by h_1h_2 and letting $h_1, h_2 \downarrow 0$ we get the desired result. □

Remark 5. 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

$$
\begin{aligned}
f_{V_i}(s) &= \frac{1}{t}, \quad \text{if } \ 0 < s < t \\
&= 0, \qquad \text{otherwise.}
\end{aligned}
$$

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 $\omega \in \Omega$. It is not difficult to show 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

$$
((T_1, T_2, \cdots, T_n) \ \mid \ N(t) = n) \stackrel{d}{=} (V_{(1)}, V_{(2)}, \cdots, V_{(n)}), \tag{17}
$$

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 (17) can be expressed as

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

□

An important consequence of Theorem 4 and Remark 5 is the following result whose proof is quite involved; see [1].

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*

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

□

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 the following theorem.

Equation (22) gives the average amount needed to take care of claims over an initial period, when premium income might not be sizable.

Theorem 6. *With the notation as above*

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

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 $t \geq 0$,

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

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

$$E(U(t)) = u + ct - E(S(t)) = u + ct - \lambda t E(X_1). \quad (24)$$

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

$$c > \lambda E(X_1) \qquad (25)$$

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.

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.

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 $n \geq 1$,

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

Put $Z_i = X_i - cA_i, i \geq 1, S_0 = 0, S_n = \sum_{i=1}^{n} Z_i, n \geq 1$. Then (28) is just $U(T_n) = u - S_n, n \geq 1$. Since "ruin"

$= \{U(T_n) < 0 \text{ for some } n\}$, it is now easy to see that

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

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 [2,3]. □

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.

In the Cramer–Lundberg model, ruin is certain if equation (25) is not satisfied. The condition (25) is called the net profit condition.

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

Theorem 8. *In the Cramer–Lundberg model assume that the net profit condition (25) holds. Assume also that there exists $r > 0$ such that*

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

Then

$$\psi(u) \leq \exp(-ru), \quad \textit{for} \quad u > 0. \tag{31}$$

□

The constant r in (30) is called the *adjustment coefficient*; in the Cramer–Lundberg model, this coefficient exists if (25) holds and X_1 has moment generating function (or Laplace transform) in a neighbourhood of 0. The inequality (31) is known as *Lundberg inequality.* See [2] for proof and extensions.

Note: An elegant way of proving Theorem 8 is through *martingale* methods. Martingale theory is a powerful tool in a probabilist's kit. One may see [6,8] for some of the elementary aspects and applications of martingales, and [2] for applications to insurance.

Remark 6. 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

$$\psi(t) = \frac{\lambda E(X_1)}{c}[1 - G(t)] + \frac{\lambda E(X_1)}{c} \int_0^t \psi(t-y)dG(y). \tag{32}$$

Equation (32) 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 (32). See [1,2] 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

Risks regarding insurance of airplanes, skyscrapers, dams, bridges, etc. are very high. Companies have also faced ruin due to a very small number of very huge claims.

of Theorem 8, 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

$$\lim_{x\to\infty} \frac{P(S_n > x)}{P(M_n > x)} = 1, \text{ for } n \geq 2, \qquad (33)$$

then F is said to be *subexponential.* 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 [1,2].

(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 8 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 8. So ruin is much more formidable if the claim size distribution is heavy tailed. See [1,2].

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.

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

Suggested Reading

[1] **T Mikosch,** ***Non-life Insurance Mathematics: an Introduction with Stochastic Processes*****, Springer, New Delhi, 2004.**

[2] **T Rolski, H Schmidli, V Schmidt and J Teugels,** ***Stochastic Processes for Insurance and Finance,*** **Wiley, New York, 2001.**

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

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

[5] **S M Ross,** ***Introduction to Probability Models*****, 8th edition. Elsevier India, New Delhi, 2005.**

[6] **S M Ross,** ***Stochastic Processes*****, 2nd edition, (Wiley Student Edition), John Wiley, Singapore, 2004.**

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

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

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.

[1,2] and the references therein deal with the above issues and more.

Acknowledgement

This is an expanded version of the Professor Wazir Hasan Abdi memorial lecture given on October 1, 2005 at the Department of Mathematics, Cochin University of Science and Technology, Kochi.

Address for Correspondence
S Ramasubramanian
Statistics and Mathematics Unit
Indian Statistical Institute
8th Mile, Mysore Road
Bangalore 560059.
Email:ram@isibang.ac.in