Insurance Risk Theory

Leda D.Minkova

Lecture Notes

2010

Contents

Chapter 1

Basic Risk Model

The foundation of the modern risk theory goes back to the works of Filip Lundberg and Harald Cramér. The Poisson process was proposed by Filip Lundberg in 1903 as a simple process in solving the problem of the first passage time. In 1930 Harald Cramér extended the Lundberg's work for modeling the ruin of an insurance company as a first passage time problem. The basic model is called a Cramér - Lundberg model or classical risk model. Insurance Risk Theory is a synonym of non-life insurance mathematics.

The basic process of the general risk model is given by

$$X(t) = \Pi(t) - S(t) \tag{1.1}$$

and is called **a risk process**. Here $\Pi(t)$ is the total amount of the premiums to the insurance company up to time t. $S(t)$ is the accumulated sum of claims up to time t. The risk reserve of the insurance company with initial capital u is given by

$$U(t) = u + \Pi(t) - S(t), \quad t \geq 0. \tag{1.2}$$

The stochastic process in (1.2) is $S(t)$ and it can be described by the following elements:

(i) The times $0 \leq \sigma_1 \leq \sigma_2 \leq \dots,$ of claim arrivals. Suppose that $\sigma_0 = 0$. The random variables $T_n = \sigma_n - \sigma_{n-1},\ n = 1, 2, \dots,$ called *inter - occurrence* or *inter - arrival times* are nonnegative.

(ii) $N(t) = \sup\{n : \sigma_n \le t\},\ t \ge 0$ is the number of claims up to time t. The relations between the times $\{\sigma_0, \sigma_1, \ldots\}$ and the counting process $\{N(t),\ t \ge 0\}$ are given by

$$\{N(t) = n\} = \{\sigma_n \le t < \sigma_{n+1}\},\ n = 0, 1, \ldots .$$

(iii) The sequence $\{Z_n,\ n = 1, 2, \ldots\}$ of independent identically distributed random variables represents the amounts of the successful claims to the insurance company. Suppose that the sequence $\{Z_n\}$ is independent of the counting process $N(t)$.

The accumulated sum of claims up to time t is given by

$$S(t) = \sum_{i=1}^{N(t)} Z_i, \quad t \ge 0.$$

The process $S = (S(t))_{t \ge 0}$ is defined by the sum $S_n = Z_1 + \ldots + Z_n$, where n is a realization of the random variable $N(t)$:

$$S(t) = Z_1 + \ldots + Z_{N(t)} = S_{N(t)}, \quad t \ge 0,$$

or a random sum of random variables. Suppose that $S(t) = 0$, if $N(t) = 0$.

Chapter 2

Random Variables

All the random variables are defined on a complete probability space $(\Omega, \mathcal{F}, P)$.

Definition 2.1 *The function*

$$F_X(x) = P(X \leq x) = P(X \in (-\infty, x]). \tag{2.1}$$

is called a distribution function of the random variable X.

The distribution function $F_X(x)$ satisfies the following conditions.

a) $F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0$.

b) $F_X(\infty) = \lim_{x \to \infty} F_X(x) = 1$.

c) $F_X(x)$ is nondecreasing, i. e., if $x < y$, then

$$F_X(x) \leq F_X(y).$$

d) $F_X(x)$ is a right continuous, i. e. for any x,

$$F_X(x + 0) = F_X(x).$$

The function $\overline{F}_X = 1 - F_X(x) = P(X > x)$ is called a **survival function** for X and represents the tail distribution.

The random variables are characterized by the properties of the distribution function in three different types.

The random variable X is called *discrete*, if the distribution function is a stepwise and has countable many jumps.

For the discrete random variable there exists a countable subset $E = \{x_0, x_1, \ldots\}$ of $\mathbb{R}$ such that $P(X \in E) = 1$. In this case the probability function is given by

$$p_k = P(X = x_k), \quad k = 0, 1, \ldots$$

and the distribution function

$$F_X(x) = \sum_{x_k \le x} P(X = x_k),$$

where the sum is over all jump points of $F_X(x)$.

A discrete r. v. X is called *arithmetic with a step h*, if the values of X belong to the set $\{x = ih : i = 0, \pm 1, \pm 2, \ldots,\}$. We use *arithmetic distribution* in the case of $h = 1$. An arithmetic distribution over the nonnegative values is called a *counting distribution*. So the counting random variables are discrete with integer nonnegative values, i. e.

$$p_k = P(X = k), \quad k = 0, 1, \ldots.$$

The r. v. X is called *continuous*, if the distribution function $F_X(x)$ is continuous. The continuous r. v. X is called *absolutely continuous*, if $F_X(x)$ is differentiable over all $\mathbb{R}$.

For the continuous r. v., the function

$$f_X(x) = \frac{d}{dx} F_X(x),$$

if it exists, is called a *probability density function* of X. The function $f_X(x)$ is nonnegative. It is easy to see that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1 \quad \text{and} \quad F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$

In a more general case for any event A,

$$P(A) = \int_A f(t)dt.$$

Intuitively, we could write that $P(X = x) = f(x)dx$ and consider $P(X = x)$ like the probability $P(x < X \le x + dx)$ over an infinitesimal interval.

The r. v. X is of a *mixed type*, if it is neither discrete nor continuous. If x is a jump point of $F(x)$, then we say that X has a mass $P(X = x)$ at the point x. In the points where $F(x)$ is absolutely continuous we consider the r. v. like continuous.

2.1 Hazard rate function

Let X be a continuous r. v. with distribution function F and density function f.

Definition 2.2 *The function*

$$\lambda(t) = \frac{f(t)}{\overline{F}(t)}$$

is called **hazard rate function** *for the r. v. X.*

Depending of the applications, the hazard rate function is called also *failure rate function, force of default, intensity rate function.* Let X be the life distribution. Then the hazard rate function is equal to the probability that the individual, survived the time t dies in the additional time dt, i. e.

$$\begin{aligned} P(X \in (t, t+dt)|X > t) &= \frac{P(X \in (t, t+dt), X > t)}{P(X > t)} = \frac{P(X \in (t, t+dt))}{P(X > t)} \\ &\approx \frac{f(t)dt}{\overline{F}(t)} = \lambda(t)dt. \end{aligned}$$

The function $\lambda(t)$ is interpreted like the intensity of deathes in the set of individuals at age t and characterizes the distribution of the r. v. Let $\Lambda(t) = \int_0^t \frac{f(s)}{\overline{F}(s)} ds$. Since $\lambda(t) = \frac{-\frac{d}{dt}\overline{F}(t)}{\overline{F}(t)}$ it follows that $\Lambda(t) = -\log \overline{F}(t)$. Consequently

$$\overline{F}(t) = \exp\left(-\int_0^t \lambda(s)ds\right) = \exp(-\Lambda(t)). \tag{2.2}$$

The function

$$\Lambda(t) = \int_0^t \lambda(s)ds$$

is called a **hazard function**.

For a discrete random variable with distribution $\{p_k\}_{k=0}^{\infty}$, the hazard rate function is given by

$$\lambda(k) = \frac{p_k}{\sum_{j=k}^{\infty} p_j}, \quad k = 0, 1, \dots .$$

In this case $\lambda(k) \leq 1$.

2.2 Moments

Mathematical expectation $\mu = EX$ of the r. v. X is defined by

$$EX = \sum_k x_k p(x_k),$$

if X is a discrete and

$$EX = \int_{-\infty}^{\infty} x f(x) dx,$$

if X is continuous, conditionally on $\sum_k |x_k| p(x_k) < \infty$ and $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$. The mathematical expectation is called also *mean value* or simply *a mean.*

In the more general case, for any measurable function $g : \mathbb{R} \longrightarrow \mathbb{R}$, the mathematical expectation $Eg(X)$ is defined by the Lebesgue - Stieltjes integral $Eg(X) = \int_{-\infty}^{\infty} g(x) dF(x)$, providing that $\int_{-\infty}^{\infty} |g(x)| dF(x) < \infty$.

The mathematical expectation $\mu_k = E(X^k)$, for any integer k, is called a $k-$ th moment of the r. v. X. $E(X - \mu)^k$ is called a $k-$th central moment. The second central moment $Var(X) = \sigma^2 = E(X - \mu)^2$ is called a *dispersion* or *variance.* The variance is often used as a measure of the risk, modeled by the random variable. Another risk measure is the *standard deviation* $\sigma = \sqrt{Var(X)}$. Essentials in applications are the coefficient of variation $CV = \frac{\sigma}{\mu}$ and the index of dispersion $I = \frac{\sigma^2}{\mu}$, known also as a *Fisher index.* The *skewness* coefficient is defined by the third moment: $\gamma_1 = \frac{E(X-\mu)^3}{\sigma^3}$. The *kurtosis* coefficient is $\gamma_2 = \frac{E(X-\mu)^4}{\sigma^4}$. The kth factorial moment is $\mu_{(k)} = E[X(X-1)\dots(X-k+1)]$.

2.3 Transforms

2.3.1 Probability Generating Functions

Let X be a nonnegative integer - valued random variable with PMF

$$p_k = P(X = k), \quad k = 0, 1, 2, \ldots.$$

Definition 2.3 *The* **Probability generating function (PGF)** *is defined as*

$$P_X(s) = Es^X = \sum_{k=0}^{\infty} s^k p_k, \tag{2.3}$$

provided the expectation is finite.

It is easy to verify that

$$p_k = \frac{1}{k!}\frac{d^k}{dt^k}P_X(s)|_{s=0} = \frac{P_X^{(k)}(0)}{k!}$$

and the factorial moments

$$\mu_{(k)} = E[X(X-1)\ldots(X-k+1)] = \frac{d^k}{dt^k}P_X(s)|_{s=1} = P_X^{(k)}(1).$$

2.3.2 Moment Generating Functions

Let X be a real - valued random variable.

Definition 2.4 *The function*

$$M_X(s) = Ee^{sX} = \int e^{sx}dF_X(x), \tag{2.4}$$

is called a **Moment generating function (MGF)** *of the random variable* X, *whenever the expectation is finite.*

Note that the MGF is finite if $s = 0$ and $M(0) = 1$. The power series of the exponential function implies that

$$M_X(s) = \int \sum_{i=0}^{\infty} \frac{(sx)^i}{i!} dF_X(x) = 1 + \mu_1 s + \mu_2 \frac{s^2}{2!} + \mu_3 \frac{s^3}{3!} + \ldots,$$

where

$$\mu_k = EX^k = \frac{d^k}{dx^k}M_X(s)|_{s=0}, \ k = 1, 2, \ldots$$

2.3.3 Characteristic Functions

Definition 2.5 *Let X be an $\mathbb{R}-$ valued random variable. Then*

$$\varphi(s) = Ee^{isX} = \int_{-\infty}^{\infty} e^{sx} dF_X(x), \;\; s \in \mathbb{R}$$

is called **a characteristic function** *of X.*

Apart from a minus sign in the exponent and the factor $\frac{1}{\sqrt{2\pi}}$, the characteristic functions coincide with Fourier transforms in the absolutely continuous case and with Fourier series in the lattice case.

2.3.4 Laplace Transform

Suppose that the random variable X is nonnegative.

Definition 2.6 *The function*

$$LT_X(s) = Ee^{-sX} = \int_{0}^{\infty} e^{-sx} dF_X(x), \tag{2.5}$$

if it exists is called Laplace transform (LT) of the random variable X or its distribution.

The Laplace transform is defined for complex values of s with positive real part. For our purposes it is sufficient to assume that s is a positive real number.

The LT (2.5) of the random variable is called also *Laplace - Stijeltes transform (LST)* of the distribution function $F_X(x)$ and actually is the LT of the density function $f_X(x)$, if it exists.

Properties of the Laplace transform

Suppose that the random variable X is nonnegative with density function $f(x)$. Then

$$LT_X(s) = LT_f(s) = \int_{0}^{\infty} e^{-sx} f(x) dx, \tag{2.6}$$

where s is a positive real number. It is easy to show that if $f(x)$ is continuous, then the integral in (2.6) exists, and there exist numbers $\alpha_1, \;\alpha_2$ and $\beta > 0$, such that, for all $x > \beta$,

$$|f(x)| \leq \alpha_1 e^{\alpha_2 x}.$$

Example 2.1 *Let us find the Laplace transform of* $f(x) = e^{\mu x}$, $x > 0$. *According to the definition*

$$LT_f(s) = \int_0^\infty e^{-sx} e^{\mu x} dx = \int_0^\infty e^{-(s-\mu)x} dx = \frac{1}{s-\mu}, \ s > \mu.$$

Suppose that the $LT_f(s)$ in (2.6) exists. Then

1. $LT_f(s-a) = \int_0^\infty e^{-sx}[e^{ax} f(x)]dx;$

2. $e^{-as} LT_f(s) = \int_a^\infty e^{-sx} f(x-a)dx;$

3. $\frac{1}{a} LT_f(\frac{s}{a}) = \int_0^\infty e^{-sx} f(ax)dx;$

4. $sLT_f(s) - f(0) = \int_0^\infty e^{-sx} f'(x)dx;$

5. $\frac{1}{s} LT_f(s) = \int_0^\infty e^{-sx} \left[\int_0^x f(y)dy\right] dx;$

6. $\frac{d}{ds} LT_f(s) = \int_0^\infty e^{-sx}[-xf(x)]dx;$

7. $\frac{d^n}{ds^n} LT_f(s) = \int_0^\infty e^{-sx}[(-1)^n x^n f(x)]dx;$

8. $\int_s^\infty LT_f(v)dv = \int_0^\infty e^{-sx} \left[\frac{f(x)}{x}\right] dx;$

9. If $LT_i(s) = \int_0^\infty e^{-sx} f_i(x)dx, \quad i = 1, 2,$ then

$$LT_1(s) LT_2(s) = \int_0^\infty e^{-sx} \left[\int_0^x f_1(y) f_2(x-y)dy\right] dx.$$

10. For a distribution function F with density f

$$LT_f(s) = LST_F(s) = sLT_F(s).$$

2.4 Counting Random Variables

1. Bernoulli random variables. The random variable X has a Bernoulli distribution with parameter p, if $P(X = 1) = 1 - P(X = 0) = p$. It is easy to check that

$$EX = EX^2 = p \text{ and } Var(X) = p(1-p).$$

2. Geometric distribution. The random variable N has a geometric distribution with parameter p, $(N \sim Ge_1(p))$ if

$$P(N = k) = p(1-p)^{k-1}, \quad k = 1, 2, \ldots .$$

The mean and the distribution function are given by

$$EN = \frac{1}{p}, \quad P(N \le n) = 1 - (1-p)^n.$$

The random variable N is interpreted as the number of trials up to the first success in a sequence of independent Bernoulli trials. The moment generating function is

$$M_N(s) = \frac{pe^s}{1-(1-p)e^s}, \quad s < -\log(1-p)$$

and the PGF

$$P_N(s) = \frac{ps}{1-(1-p)s}.$$

A characterization property of the geometric distribution is a lack of memory:

$$P(N = n+k | N > k) = \frac{P(N = n+k)}{P(N > k)} = \frac{(1-p)^{n+k-1}p}{(1-p)^k} = P(N = n),$$

for any $n > 0$ and $k > 0$.

3. Binomial distribution. The random variable N has binomial distribution with parameters n and $p \in (0,1)$, $(N \sim Bi(n,p))$ if

$$P(N = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad n = 0, 1, \ldots n.$$

The mean and the variance are given by

$$EN = np \quad \text{and} \quad Var(N) = np(1-p).$$

The MGF is

$$M_N(s) = [1 - p(1-e^s)]^n$$

and the PGF

$$P_N(s) = [1 - p(1-s)]^n.$$

The binomial distributed random variable N is interpreted as the number of successes in a sequence of n independent Bernoulli trials.

4. Poisson distribution. The random variable N has a Poisson distribution with parameter λ, $(N \sim Po(\lambda))$ if

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

The mean and the variance of the Poisson distribution are equal:

$$EN = Var(N) = \lambda.$$

The MGF is

$$M_N(s) = e^{-\lambda(1-e^s)}$$

and the PGF

$$P_N(s) = e^{-\lambda(1-s)}.$$

2.5 Continuous random variables

2.5.1 Exponential distribution and lack of memory property

The random variable X is exponentially distributed with parameter $\lambda > 0$, $(X \sim \exp(\lambda))$, if

$$F(x) = 1 - e^{-\lambda x} \quad \text{and} \quad f(x) = \lambda e^{-\lambda x}, \; x \geq 0.$$

The MGF of the exponential distribution is

$$M_X(s) = Ee^{sX} = \frac{\lambda}{\lambda - s}, \quad t < \lambda \tag{2.7}$$

and the Laplace transform

$$LT_X(s) = \frac{\lambda}{\lambda + s}.$$

All the moments of the random variable could be determine by differentiation of (2.7).

$$EX = \frac{1}{\lambda}, \quad \text{and} \quad Var(X) = \frac{1}{\lambda^2}$$

and

$$EX^n = \frac{n!}{\lambda^n}.$$

If $\lambda = 1$, the r. v. $X \sim \exp(1)$ is called a standard exponentially distributed. Let $X \sim \exp(1)$ and $Y = a + \frac{X}{\lambda}, \quad \lambda > 0, \quad -\infty < a < \infty$. Then the distribution of Y is determined by the following distribution function and probability density function

$$F(x) = 1 - e^{-\lambda(x-a)} \quad \text{and} \quad f(x) = \lambda e^{-\lambda(x-a)}, \ x \geq a$$

and is noted by $Y \sim \exp(a, \lambda)$.

The advantages in applications of the exponential distribution is the lack of memory property.

The random variable X satisfies the **lack of memory property**, if for any $t, s \geq 0$ such that $P(X \geq t) > 0$

$$P(X \geq t + s | X \geq t) = P(X \geq s). \tag{2.8}$$

Let X be the lifetime distribution. Then (2.8) means the probability that the individual will survive another s years, after having attained the age t is equal to the probability that the individual will survive age s. If the distribution function is not degenerate at zero, the condition (2.8) is equivalent to

$$\frac{P(X \geq t + s, X \geq t)}{P(X \geq t)} = P(X \geq s)$$

or

$$P(X \geq t + s) = P(X \geq s)P(X \geq t). \tag{2.9}$$

The equation (2.9) contains also the degenerated case. The decision is given by the following theorem [5].

Theorem 2.1 *There are only two solutions of the equation (2.9) among distribution functions. Either $F(x)$ is degenerate at zero, or, with some constant $\lambda > 0$, $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$.*

Since $e^{-\lambda(t+s)} = e^{-\lambda t}e^{-\lambda s}$, the only exponentially distributed random variables are memoryless. If the lifetime distribution is exponential **the hazard rate function is constant:**

$$\lambda(t) = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda.$$

Let X and Y be independent exponentially distributed random variables with respective parameters λ and μ. Then the distribution of $Z = \min(X, Y)$ is given by

$$\begin{aligned} P(Z \leq z) &= 1 - P(Z > z) = 1 - P(X > z, Y > z) = \\ &= 1 - P(X > z)P(Y > z) = 1 - e^{-\lambda z}e^{-\mu z} = 1 - e^{-(\lambda+\mu)z}. \end{aligned}$$

Consequently, $Z = \min(X, Y)$ is again exponentially distributed with parameter $\lambda + \mu$.

2.5.2 Gamma distribution

Let $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx,\ \alpha > 0$ be the Gamma function.

Properties of the Gamma function

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1),$$

$$\Gamma(n) = (n - 1)!, \quad n \geq 1$$

and

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

The random variable X is Gamma distributed with parameters $\alpha > 0$ and $\beta > 0$, $(X \sim \Gamma(\alpha, \beta))$, if the density function is given by

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

The parameter α is called a shape parameter. Small value of α results in a long tail to the right, or right skewed distribution.

In the special case α integer, it is an Erlang distribution.

The MGF is

$$M(s) = \left(1 - \frac{s}{\beta}\right)^{-\alpha}, \quad t > \beta$$

and the n th moment

$$EX^n = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\beta^n}.$$

In particular

$$EX = \frac{\alpha}{\beta} \quad \text{and} \quad Var(X) = \frac{\alpha}{\beta^2}.$$

2.5.3 Beta distribution

For $\alpha > 0$ and $\beta > 0$, the function

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

is called a Beta function. The most important property of the Beta function:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

The random variable X is Beta distributed ($X \sim B(\alpha, \beta)$) with parameters α and β, if the density function is

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in (0, 1).$$

The moments of the Beta distributed random variable:

$$EX^n = \frac{B(n+\alpha, \beta)}{B(\alpha, \beta)} = \frac{\Gamma(n+\alpha)\Gamma(\alpha+\beta)}{\Gamma(n+\alpha+\beta)\Gamma(\alpha)}.$$

2.5.4 Weibull distribution

The random variable X has Weibull distribution with parameters $\beta > 0$ and $\sigma > 0$, ($X \sim W(\beta, \sigma)$) if the distribution function and density function are

$$F(x) = 1 - e^{-(\frac{x}{\sigma})^\beta}, \quad x \geq 0$$

and

$$f(x) = \frac{\beta}{\sigma}\left(\frac{x}{\sigma}\right)^{\beta-1} e^{-(\frac{x}{\sigma})^\beta}, \quad x \geq 0.$$

The mean and the variance:

$$EX = \sigma\Gamma(1+\frac{1}{\beta}), \quad Var(X) = \sigma^2\left[\Gamma(1+\frac{2}{\beta}) - \left(\Gamma(1+\frac{1}{\beta})\right)^2\right].$$

If $\beta = 1$, $W(1,\sigma) = exp(\sigma)$.

2.5.5 Pareto distribution

The random variable X has a Pareto distribution with parameters $\alpha > 0$ and $\lambda > 0$, $(X \sim Par(\alpha,\lambda))$, if the distribution function is given by

$$F(x) = 1 - \left(\frac{\lambda}{\lambda+x}\right)^\alpha, \quad x > 0.$$

The density function is

$$f(x) = \frac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}, \quad x > 0.$$

The distribution defined by the density

$$f_1(x) = f(x-\lambda) = \frac{\alpha\lambda^\alpha}{x^{\alpha+1}}, \quad x > \lambda$$

is called a shifted Pareto distribution.

The moments of the random variable with density $f(x)$ are

$$m_k = \frac{\alpha}{\lambda}\int_0^\infty x^k\left(\frac{\lambda}{\lambda+x}\right)^{\alpha+1}dx.$$

The change of variables $y = \frac{\lambda}{\lambda+x}$ leads to

$$m_k = \int_0^1 \frac{\alpha}{\lambda}\left[\frac{\lambda(1-y)}{y}\right]^k y^{\alpha+1}\lambda y^{-2}dy = \alpha\lambda^k\int_0^1 y^{\alpha-k-1}(1-y)^k dy,$$

which is a Beta function and

$$m_k = \alpha\lambda^k\frac{\Gamma(\alpha-k)\Gamma(k+1)}{\Gamma(\alpha+1)} = \lambda^k k!\frac{\Gamma(\alpha-k)}{\Gamma(\alpha)}, \quad \alpha > k.$$

The mean and the variance are

$$EX = m_1 = \frac{\lambda}{\alpha - 1}, \quad \alpha > 1$$

and

$$Var(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2.$$

It is easy to see that only finite number of moments exist. Even the mean and the variance not always exist. This means that the Pareto distribution is a heavy tailed.

2.5.6 The normal distribution

The normal distribution is important in insurance and finance since it appears like limiting distribution in many cases. The normal density function is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad -\infty < x < \infty.$$

The notation $X \sim N(\mu, \sigma)$ means that X has a normal distribution with parameters μ and σ.

The mean value and the variance are $EX = \mu$ and $Var(x) = \sigma^2$. The most important property is that the random variable $Z = \frac{X-\mu}{\sigma}$ is normally distributed with parameters 0 and 1 ($Z \sim N(0,1)$). The distribution function is denoted Φ and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx.$$

The characteristic function of X is

$$\varphi_X(s) = e^{i\mu s - \frac{s^2\sigma^2}{2}}, \ s \in \mathbb{R}$$

and the MGF

$$M_X(s) = e^{\mu s - \frac{s^2\sigma^2}{2}}.$$

2.5.7 Log-normal distribution

The random variable X has a log - normal distribution, if $Y = \log X$ has a normal distribution. If $Y \sim N(\mu, \sigma)$, the density function of X is

$$f_X(x) = f_Y(\log x)\frac{1}{x} = \frac{1}{\sigma x\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\log x-\mu}{\sigma}\right)^2}, \quad x > 0,$$

where $\sigma > 0, \quad -\infty < \mu < \infty$. The distribution function of X is defined by

$$F(x) = \Phi\left(\frac{\log x - \mu}{\sigma}\right), \quad x > 0,$$

where $\Phi(.)$ is the standard normal distribution function. The moments of the lo g- normal distribution can be calculated from the standard normal distribution function.

$$m_k = EX^k = E(e^{kY}) = M_Y(k) = e^{\mu k + \frac{\sigma^2 k^2}{2}},$$

where $M_Y(k)$ is the MGF of the normal distribution. Particular, the mean and variance are

$$EX = m_1 = e^{\mu + \frac{\sigma^2}{2}}$$

and

$$Var(X) = e^{2\mu+\sigma^2}[e^{\sigma^2} - 1].$$

2.5.8 Inverse Gaussian distribution

If the random variable X is defined by the density

$$f(x) = \frac{\mu}{\sqrt{2\pi\beta x^3}}e^{-\frac{(x-\mu)^2}{2\beta x}}, \; x \geq 0, \tag{2.10}$$

where $\mu > 0$, then X has Inverse Gaussian distribution ($X \sim IG(\mu, \beta)$). In the special case of $\mu = 1$, (2.10) is called Wald distribution. The Inverse Gaussian distribution has greater skewness and a sharper peak then Gaussian.

The distribution function is

$$F(x) = \Phi\left(\frac{x-\mu}{\sqrt{\beta x}}\right) + e^{\frac{2\mu}{\beta}}\Phi\left(-\frac{x+\mu}{\sqrt{\beta x}}\right), \quad x > 0.$$

The Laplace transform:

$$LT(s) = \int_0^\infty e^{-sx} \frac{\mu}{\sqrt{2\pi\beta x^3}} e^{-\frac{(x-\mu)^2}{2\beta x}} dx = e^{\frac{\mu}{\beta}[1-\sqrt{1+2\beta s}]}, \quad s \geq -\frac{1}{2\beta}.$$

The mean and the variance:

$$EX = \mu \text{ and } Var(X) = \mu\beta.$$

2.6 Functions of random variables

Theorem 2.2 *Let X be a continuous random variable and $\varphi(x)$ is a strictly increasing function, defined on the set of values of X. Let $Y = \varphi(X)$ and F_X and F_Y are the distribution functions of X and Y. Then*

$$F_Y(y) = F_X(\varphi^{-1}(y)).$$

If $\varphi(x)$ is strictly decreasing over the set of values of X, then

$$F_Y(y) = 1 - F_X(\varphi^{-1}(y)).$$

Proof. Since $\varphi(x)$ is strictly increasing over the set of values of X, the events $\{X \leq \varphi^{-1}(y)\}$ and $\{\varphi(X) \leq y\}$ are equivalent. Consequently

$$F_Y(y) = P(Y \leq y) = P(\varphi(X) \leq y) = P(X \leq \varphi^{-1}(y)) = F_X(\varphi^{-1}(y)).$$

If $\varphi(x)$ is strictly decreasing, then

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(\varphi(X) \leq y) = \\ &= P(X > \varphi^{-1}(y)) = 1 - P(X \leq \varphi^{-1}(y)) = 1 - F_X(\varphi^{-1}(y)). \end{aligned}$$

□

If f_X and f_Y are the density functions of X and Y and $\varphi(x)$ is strictly increasing, then

$$f_Y(y) = f_X(\varphi^{-1}(y)) \frac{d}{dy} \varphi^{-1}(y).$$

If $\varphi(x)$ is strictly decreasing, then

$$f_Y(y) = -f_X(\varphi^{-1}(y)) \frac{d}{dy} \varphi^{-1}(y).$$

The method is applicable in same cases even if the function $\varphi(x)$ is neither increasing nor decreasing. For example, if $Y = X^2$, the function is $\varphi(x) = x^2$. Then

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(-\sqrt{y} < X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

For the density we obtain

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(F_X(\sqrt{y}) - F_X(-\sqrt{y})\right) = \left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right)\frac{1}{2\sqrt{y}}.$$

2.7 Joint density and distribution function

Let $\mathbf{X} = (X, Y)$ be a random vector.

Definition 2.7 *The joint distribution function of* $\mathbf{X}$ *is the function* $H : \mathbb{R}^2 \to [0, 1]$, *defined by*

$$H(x, y) = P(X \leq x, Y \leq y), \quad (x, y) \in \mathbb{R}^2.$$

If X and Y absolutely are continuous random variables, the joint density function satisfies the equation

$$H(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(t_1, t_2)dt_2dt_1.$$

Consequently

$$f(x, y) = \frac{\partial^2 H(x, y)}{\partial x \partial y}.$$

If the random variables X and Y are mutually independent

$$H(x, y) = F_X(x)F_Y(y).$$

The random variables X and Y are called marginal variables and the distribution functions F_X and F_Y - marginal distributions of (X, Y). The joint distribution function is right continuous and satisfies the properties:

1) $\lim_{y\to\infty} H(x, y) = F_X(x)$ and $\lim_{x\to\infty} H(x, y) = F_Y(y)$,

2) $\lim_{(x,y)\to(\infty,\infty)} H(x,y) = 1$, where $(x,y) \to (\infty,\infty)$ means that both variables x and y tend to infinite,

3) $\lim_{x\to-\infty} H(x,y) = \lim_{y\to-\infty} H(x,y) = 0$,

4) for all (x_1, x_2) and (y_1, y_2), such that $x_1 < x_2$ and $y_1 < y_2$,

$$H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1) + H(x_1, y_1) \geq 0.$$

For the random variables X and Y, the second moment

$$Cov(X,Y) = E[(X - EX)(Y - EY)] = E(XY) - (EX)(EY),$$

is called a *covariance.* Note that the variance $Var(X)$ of a square integrable random variable is $Cov(X,X)$. The Pearson correlation coefficient is one of the most useful measures of dependence and is defined by

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

The random variables X and Y are positively correlated if $Cov(X,Y) > 0$, and negatively correlated if $Cov(X,Y) < 0$.

2.8 Conditional distributions

If X and Y are defined over the same probability space, then the **conditional distribution function** of X given $Y = y$ is defined by

$$F_{X|Y}(x|y) = P(X \leq x | Y = y). \tag{2.11}$$

The following equality gives the relation between the conditional distribution function and the distribution function of the random variable X

$$F_X(x) = \int F_{X|Y}(x|y) dF_Y(y).$$

The similar relation exists between the density of X and the conditional density

$$f_X(x) = \int f_{X|Y}(x|y) f_Y(y) dy.$$

For discrete random variables:

$$P(X = x_k) = \sum_{i=0}^{\infty} P(X = x_k|Y = y_i)P(Y = y_i), \quad k = 0, 1, 2, \ldots.$$

If X and Y are independent random variables:

$$F_{X|Y}(x|y) = F_X(x).$$

According (2.11) the conditional mean of X given $Y = y$:

$$E(X|Y = y) = \int_0^{\infty} x dF_{X|Y}(x|y),$$

provided that the random variables are nonnegative. Consequently $E(X|Y)$ is a random variable. For independent random variables

$$E(X|Y = y) = EX.$$

For any function $h(x)$, the conditional mean of $h(X)$ given $Y = y$ is:

$$E[h(X)|Y = y] = \int h(x) dF_{X|Y}(x|y). \tag{2.12}$$

From (2.12) it follows that

$$E[h(X)] = E_Y E_X[h(X)|Y],$$

where on right is taken mathematical expectation of X given fix Y and after that mathematical expectation of Y.

Exercise 2.1 *Prove that for the nonnegative random variable X :*

a) $E[E(X|Y)] = EX$;

b) $Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$.

2.9 Sum of random variables

Let X and Y be independent random variables. The distribution function of the sum $Z = X + Y$, noted

$$F_Z(z) = F_X * F_Y(z)$$

is called a **convolution** of F_X and F_Y and is defined by

$$\begin{aligned} F_Z(z) &= \int P(Z \leq z | Y = y) dF_Y(y) \\ &= \int P(X \leq z - y | Y = y) dF_Y(y) \qquad (2.13) \\ &= \int F_{X|Y}(z - y|y) dF_Y(y) = \int F_X(z - y) dF_Y(y). \end{aligned}$$

If X and Y are continuous random variables, the density of Z is obtained by differentiation of (2.13) relative to z:

$$f_Z(z) = f_X * f_Y(z) = \int f_X(z - y) f_Y(y) dy.$$

If X and Y are discrete random variables

$$P(Z = k) = \sum_{i=0}^{k} P(X = k - i) P(Y = i), \quad k = 0, 1, 2, \ldots.$$

In the multivariate case, if $X_1, X_2, \ldots X_n$ are random variables, the distribution of the sum $S_n = X_1 + X_2 + \ldots + X_n$ is obtained by recursions. Note $S_1 = X_1$ and $S_j = S_{j-1} + X_j, \quad j = 2, 3, \ldots n$. The Laplace transform of the sum is

$$LT_{S_n}(t) = \prod_{j=1}^{n} LT_{X_j}(t), \qquad (2.14)$$

the distribution function is noted by

$$F_{S_n}(x) = F_{X_1} * F_{X_2} * \ldots * F_{X_n}(x)$$

and the density

$$f_{S_n}(x) = f_{X_1} * f_{X_2} * \ldots * f_{X_n}(x).$$

If $X_1, X_2, \ldots X_n$ are identically distributed with distribution function $F_X(x)$ and density function $f_X(x)$, then

$$F_{S_n}(x) = F_X^{*n}(x)$$

and

$$f_{S_n}(x) = f_X^{*n}(x).$$

Here $*n$ denotes $n-$ th convolution of X. In that case (2.14) implies that the transforms of $S_n = X_1 + X_2 + \ldots + X_n$ are $n-$th power of the corresponding transform of X. For example, the PGF is

$$P_{S_n}(t) = [P_X(t)]^n.$$

2.9.1 Negative binomial distribution

Let $X_1, X_2, \dots X_r$ be independent $Ge_1(p)$ distributed random variables. The random variable $N = X_1 + X_2 + \dots X_r$ has probability mass function

$$p_k = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \dots$$

and is called negative binomial distributed ($N \sim NB(r,p)$). The interpretation of the NB distributed random variable is the number of trials up to the rth success in a sequence of Bernoulli trials. The PGF is given by

$$M_N(t) = \left(\frac{pe^t}{1-(1-p)e^t} \right)^r, \quad t < -\ln(1-p).$$

The mean value and the variance:

$$EN = \frac{r}{p} \quad Var(N) = \frac{r(1-p)}{p^2}.$$

An alternative representation of the NBD is by the random variable Y - the number of failures in Bernoulli trials until r successes. In this case $Y = N - r$ and the distribution function is

$$p_k = \binom{r+k-1}{k} p^r (1-p)^k, \quad k = 0, 1, 2, \dots$$

The MGF is

$$M_Y(t) = \left(\frac{p}{1-(1-p)e^t} \right)^r, \quad t < -\ln(1-p).$$

The mean and the variance are

$$EY = \frac{r(1-p)}{p} \quad Var(Y) = \frac{r(1-p)}{p^2}.$$

2.9.2 Erlang distribution

Let $X_1, X_2, \dots X_n$ be independent, exponentially distributed with parameter λ. Let us find the distribution of the random variable $S_n = X_1 + \dots + X_n$. It is known that the exponential distribution is $\Gamma(1, \lambda)$.

Suppose that

$$f_{X_1+\ldots+X_{n-1}}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-2}}{(n-2)!}.$$

For S_n we obtain

$$\begin{aligned} f_{X_1+\ldots+X_n}(t) &= \int_0^\infty f_{X_n}(t-s) f_{X_1+\ldots+X_{n-1}}(s)ds \\ &= \int_0^\infty \lambda e^{-\lambda(t-s)}\lambda e^{-\lambda s}\frac{(\lambda s)^{n-2}}{(n-2)!}ds = \lambda\frac{(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t}. \end{aligned}$$

Definition 2.8 *The random variable with density function*

$$f_X(t) = \lambda\frac{(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t}, \quad t \geq 0, \tag{2.15}$$

is called Erlang distributed random variable with parameters n and λ, $X \sim Erl(n,\lambda)$, $n = 1, 2, \ldots$.

The distribution function is

$$F(t) = 1 - \sum_{j=0}^{k-1}\frac{(\lambda t)^j}{j!}e^{-\lambda t}, \quad t \geq 0. \tag{2.16}$$

The λ is a location parameter, n a shape parameter.

The mean and the variance are given by

$$EX = \frac{n}{\lambda} \quad \text{and} \quad Var(X) = \frac{n}{\lambda^2}.$$

The Laplace - Stieltjes transform:

$$LS_X(t) = \left(\frac{\lambda}{\lambda+t}\right)^n.$$

It is known that (2.15), with real n is the density of Gamma distributed random variable ($X \sim \Gamma(n,\lambda)$.)

2.10 Mixed distributions

Let X be a random variable with distribution function $F_X(x,\theta)$ and density function $f_X(x,\theta)$, where θ is a parameter.

Suppose that θ is a realization of a continuous, positive random variable Θ. For fixed θ we use the notation $F_X(x|\theta)$, i.e. conditional distribution function. Let $U(\theta)$ and $u(\theta)$ be the distribution function and density function of the random variable Θ. The distribution function of X is defined by

$$F_X(x) = \int_\theta F_X(x|\theta) dU(\theta) = E_\theta F_X(x|\theta), \tag{2.17}$$

where E_θ denotes the mathematical expectation relative θ.

Definition 2.9 *The random variable X with a distribution (2.17) is called mixed by mixing distribution $U(\theta)$.*

The equation (2.17) gives the unconditional distribution of X. Unconditional density function of X is defined by

$$f_X(x) = \int_\theta f_X(x|\theta) dU(\theta) = E_\theta f_X(x|\theta).$$

The moments:

$$EX^k = \int_\theta E(X^k|\theta) dU(\theta) = E_\theta E(X^k|\theta), \quad k = 1, 2, \ldots$$

The unconditional transforms are obtained by the mathematical expectation over the set of all possible values of θ :

$$M_X(z) = E_\theta M_X(z|\theta)$$

$$P_X(z) = E_\theta P_X(z|\theta)$$

$$\varphi_X(z) = E_\theta \varphi_X(z|\theta)$$

$$LT_X(z) = E_\theta LT_X(z|\theta).$$

Example 2.2 *(Poisson mixture) Let X be a Poisson distributed random variable with MGF*

$$M_X(s|\lambda) = e^{\lambda(e^s-1)}.$$

Then, for any distribution $\Lambda(\lambda)$, the unconditional MGF of X is

$$M_X(s) = E_\lambda M_X(s|\lambda) = E_\lambda[e^{\lambda(e^s-1)}] = M_\Lambda(e^s-1),$$

wich is the MGF of Λ with the argument e^s-1.

Example 2.3 *(NBD) Let X be a Poisson distributed random variable with probability mass function*

$$P(X=k|\lambda) = \frac{\lambda^k}{k!}e^{-\lambda},\ \ k=0,1,\ldots.$$

The parameter λ is a realization of Gamma distributed random variable with density function $f(\lambda) = \frac{\beta^r}{\Gamma(r)}\lambda^{r-1}e^{-\beta\lambda},\ \ \beta>0,\ \lambda>0$, where Γ is the Gamma function, r is the shape parameter and β the scale parameter. Prove that

$$P(X=k) = \left(\frac{\beta}{1+\beta}\right)^r \binom{r+k-1}{k} \left(\frac{1}{1+\beta}\right)^k,\quad k=0,1,\ldots.$$

The following construction is again a mixture. Let $p_1, p_2, \ldots$ be a sequence of nonnegative numbers, such that $\sum_{i=1}^{\infty} p_i = 1$. Let $F_1(x), F_2(x), \ldots$ be a sequence of distribution functions. Then

$$F_X(x) = \sum_{i=1}^{\infty} p_i F_i(x) \tag{2.18}$$

is again a distribution function and is called a *mixture.*

Here the weights $\{p_i\}$ are not related to any parameter.

Example 2.4 *Let us construct the mixture between the degenerate at zero distribution*

$$F_1(x) = \begin{cases} 0, & x<0 \\ 1, & x\geq 0 \end{cases}$$

and the exponential distribution

$$F_2(x) = 1-e^{-\lambda x},\ \ x>0$$

by weights ρ and $1-\rho$. The distribution function is given by

$$F_X(x) = \rho.1 + (1-\rho)(1-e^{-\lambda x}) = 1-(1-\rho)e^{-\lambda x}, \quad x \geq 0.$$

The Laplace transform of X is

$$LT_X(s) = \rho LT_1(s) + (1-\rho)LT_2(s) = \rho + (1-\rho)\frac{\lambda}{\lambda+s}.$$

If all $F_i(x)$ in (2.18) are the same, the mixture coincides with (2.17) in the case of discrete parameter θ.

2.11 Compound distributions

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with distribution function $F_X(x)$, characteristic function $\varphi_X(z)$, mean μ and variance σ_X^2. Then the sum $X_1 + X_2 + \ldots + X_n, \quad n \geq 1$ has a distribution function $F_X^{*n}(x)$, characteristic function $[\varphi_X(z)]^n$, mean $n\mu_X$ and variance $n\sigma_X^2$.

Consider the sum

$$S_N = X_1 + X_2 + \ldots X_N, \tag{2.19}$$

where N is a discrete random variable and $S_N = 0$ if $N = 0$. Then the distribution function of S_N is

$$F_{S_N}(x) = P(S_N \leq x) = \sum_{n=0}^{\infty} P(S_N \leq x | N = n)P(N=n) = \sum_{n=0}^{\infty} P(N=n)F_X^{*n}(x),$$

where

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

Definition 2.10 *The distribution of the random sum (2.19) is called compound distribution.*

The MGF of the compound distribution is given by

$$\begin{aligned} M_{S_N}(s) &= Ee^{sS_N} = \textstyle\sum_{n=0}^{\infty} E[e^{s(X_1+\ldots X_n)}|N=n]P(N=n) \\ &= \textstyle\sum_{n=0}^{\infty} P(N=n)[M_X(s)]^n = E[[M_X(s)]^N], \end{aligned}$$

or

$$M_{S_N}(s) = P_N(M_X(s)).$$

Analogously

$$P_{S_N}(s) = P_N(P_X(s))$$

$$\varphi_{S_N}(s) = P_N(\varphi_X(s))$$

$$LT_{S_N}(s) = P_N(LT_X(s)).$$

Exercise 2.2 *Show that*

$$ES_N = E(N)E(X) \tag{2.20}$$

and

$$Var(S_N) = Var(N)(EX)^2 + E(N)Var(X). \tag{2.21}$$

If N has a Poisson distribution, (2.19) is called compound Poisson distribution.

Theorem 2.3 *Let $S_i \sim CPo(\lambda_i, F_i(x)),\ i = 1, \ldots n$ be independent Compound Poisson random variables with parameters λ_i and $Z \sim F_i(x),\ \ x > 0$. Show that $S_n = S_1 + \ldots S_n$ is also Compound Poisson with parameters $\lambda = \sum_{i=1}^n \lambda_i$ and $F(x) = \sum_{i=1}^n \frac{\lambda_i}{\lambda} F_i(x)$.*

Example 2.5 *(Compound geometric - exponential distribution) Let $X_1, X_2, \ldots$ be independent, identically exponentially distributed with parameter $\lambda > 0$, density function*

$$f_X(x) = \lambda e^{-\lambda x}\ x > 0$$

and

$$LT_X(s) = \frac{\lambda}{\lambda + s}.$$

The Laplace transform of S_N is

$$LT_S(s) = \sum_{n=0}^{\infty} P(N = n) \left(\frac{\lambda}{\lambda + s}\right)^n.$$

Suppose that N has a geometric distribution with PMF

$$P(N = n) = p(1-p)^n, \quad n = 0, 1, 2, \ldots$$

and PGF

$$P_N(s) = \frac{p}{1-(1-p)s}.$$

In that case the LT of S_N is

$$LT_S(s) = p\left[1-(1-p)\frac{\lambda}{\lambda+s}\right]^{-1} = p\frac{\lambda+s}{p\lambda+s} = p + (1-p)\frac{p\lambda}{p\lambda+s}.$$

The LT of S_N is a mixture of the Laplace transform of degenerate at zero random variable and the Laplace transform of $\exp(p\lambda)$ *random variable.*

According the properties of the LT, the density function of S_N is

$$f_{S_N}(x) = \begin{cases} p, & x = 0 \\ (1-p)p\lambda e^{-p\lambda x}, & x > 0 \end{cases}$$

and the distribution function

$$F_{S_N}(x) = 1-(1-p)e^{-p\lambda x}, \quad x \geq 0.$$

Example 2.6 *(Pólya - Aeppli distribution). Let $X_1, X_2, \ldots$ be independent identically $Ge_1(1-\rho)$ distributed random variables with parameter $\rho \in [0,1)$ and probability mass function*

$$P(X_1 = i) = \rho^{i-1}(1-\rho), \quad i = 1, 2, \ldots.$$

The random variable $N \sim Po(\lambda)$ is independent of $X_i, i = 1, 2, \ldots$. The probability mass function of $S_N = X_1 + X_2 + \ldots + X_N$ is given by

$$P(S_N = k) = \begin{cases} e^{-\lambda}, & k = 0, \\ & \\ e^{-\lambda}\sum_{i=1}^{k}\binom{k-1}{i-1}\frac{[\lambda(1-\rho)]^i}{i!}\rho^{k-i}, & i = 1, 2, \ldots \end{cases} \tag{2.22}$$

Definition 2.11 *The distribution, defined by the PMF (2.22) is called a Pólya - Aeppli distribution.*

2.11.1 Hiperexponential distribution

Let $X_i \sim \exp(\lambda_i),\ i = 1, 2,\ \lambda_1 \neq \lambda_2$ be independent. The distribution of the sum $X_1 + X_2$ is given by

$$\begin{aligned} f_{X_1+X_2} &= \textstyle\int_0^t f_{X_1}(s) f_{X_2}(t-s) ds = \int_0^t \lambda_1 e^{-\lambda_1 s} \lambda_2 e^{-\lambda_2 (t-s)} ds \\ &= \lambda_1 \lambda_2 e^{-\lambda_2 t} \textstyle\int_0^t e^{-(\lambda_1 - \lambda_2)s} ds = \frac{\lambda_2}{\lambda_2 - \lambda_1} \lambda_1 e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_2 t}. \end{aligned}$$

It can be proved by induction that the density function of $S_n = X_1 + \ldots + X_n$, where $X_i \sim \exp(\lambda_i),\ i = 1, \ldots, n,\ \lambda_i \neq \lambda_j$ for $i \neq j$, is

$$f_{S_n}(t) = \sum_{i=1}^{n} p_i \lambda_i e^{-\lambda_i t}, \tag{2.23}$$

$$p_i = \Pi_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i}.$$

Definition 2.12 *The random variable defined by density function (2.23) for some weights p_i, $\sum_{i=1}^n p_i = 1$ is called hiperexponential distributed random variable.*

The notation is $H_n(p_1, \ldots p_n; \lambda_1, \ldots \lambda_n)$ or H_n. That is a random variable equal to X_i with probability p_i, $i = 1, \ldots n$. The distribution (2.23) is a mixture of exponentially distributed random variables. The mean value is

$$ES_n = \sum_{i=1}^{n} \frac{p_i}{\lambda_i}$$

and the Laplace transform

$$LT_{S_n}(s) = \sum_{i=1}^{n} p_i \frac{\lambda_i}{\lambda_i + s}.$$

Chapter 3

Counting processes

The stochastic process $\{N(t),\ t \geq 0\}$ is called a counting process, if $N(t)$ is equal to the number of events occurred up to time t.

The counting process satisfies the following:

1. $N(t) \geq 0$;

2. $N(t)$ has integer values;

3. If $s < t$, then $N(s) \leq N(t)$;

4. For $s < t$, $N(t) - N(s)$ is the number of events in the time interval (s, t).

The counting process is called a process with independent increments if the number of events in disjoint intervals are independent random variables.

The counting process has stationary increments if the distribution of the number of events occurred in a given time interval depends only on the length of the interval. This means that for $t > 0$ and $h > 0$ the distribution of $N(t+h) - N(t)$ coincides with the distribution of $N(h)$.

3.1 Poisson process

One of the basic counting processes is the Poisson process.

Definition 3.1 *The counting process $\{N(t),\ t \geq 0\}$ is called a Poisson process with intensity rate $\lambda > 0$, if*

1. $N(0) = 0$;

2. *The process has stationary independent increments;*
3. $P(N(h) = 1) = \lambda h + \circ(h)$;
4. $P(N(h) \geq 2) = \circ(h)$.

(The function f is called $\circ(h)$, if $\lim_{h\to 0} \frac{f(h)}{h} = 0$.) It is clear that the process is integer valued and nondecreasing.

Notation: $P_n(t) = P(N(t) = n)$. We will show that the following system of differential equations follows from the postulates in the definition.

$$\begin{aligned} P_0'(t) &= -\lambda P_0(t) \\ P_n'(t) &= -\lambda P_n(t) + \lambda P_{n-1}(t), \quad n = 1, 2, \ldots \end{aligned} \tag{3.1}$$

For $n = 0$

$$\begin{aligned} P_0(t+h) &= P(N(t) = 0, N(t+h) - N(t) = 0) = \\ &= P(N(t) = 0)P(N(t+h) - N(t) = 0) = P_0(t)[1 - \lambda h + \circ(h)] \end{aligned}$$

and

$$\frac{P_0(t+h) - P_0(t)}{h} = -\lambda P_0(t) + \frac{\circ(h)}{h}.$$

By $h \to 0$ we obtain the first equation of (3.1).

Let $n \geq 1$

$$\begin{aligned} P_n(t+h) &= P(N(t) = n, N(t+h) - N(t) = 0) \\ &+ P(N(t) = n-1, N(t+h) - N(t) = 1) + \\ &+ \textstyle\sum_{k=2}^{\infty} P(N(t) = n-k, N(t+h) - N(t) = k). \end{aligned}$$

According to the fourth postulate, the sum is $\circ(h)$. Consequently

$$P_n(t+h) = P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + \circ(h) = (1-\lambda h)P_n(t) + \lambda h P_{n-1}(t) + \circ(h),$$

or

$$\frac{P_n(t+h) - P_n(t)}{h} = -\lambda P_n(t) + \lambda P_{n-1}(t) + \frac{\circ(h)}{h}.$$

For $h \to 0$ we obtain the second equation of (3.1).

To solve (3.1), we use the method of integrating factor. The condition $N(0) = 0$ implies the initial conditions

$$P_0(0) = 1 \quad \text{and} \quad P_n(0) = 0, \quad n = 1, 2, \ldots. \tag{3.2}$$

The solution of the first equation of (3.1) is given by

$$P_0(t) = C_0 e^{-\lambda t}.$$

Together with (3.2) this leads to

$$P_0(t) = e^{-\lambda t}. \tag{3.3}$$

Inserting (3.3) in the second equation of (3.1) for $n = 1$ yield

$$P_1'(t) + \lambda P_1(t) = \lambda e^{\lambda t}$$

with the solution

$$P_1(t) = (\lambda t + C_1)e^{-\lambda t}.$$

By the initial condition (3.2) for $n = 1$, the solution is

$$P_1(t) = \lambda t e^{-\lambda t}.$$

In the general case we suppose that

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

By the second equation of (3.1)

$$P_n'(t) + \lambda P_n(t) = \lambda P_{n-1}(t) = \lambda \frac{(\lambda t)^{n-1}}{(n-1)!} e^{\lambda t}$$

with the solution

$$P_n(t) = \left(\frac{(\lambda t)^n}{n!} + C_n \right) e^{-\lambda t}.$$

Together with the initial condition this gives the solution

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \; n = 1, 2, \ldots \tag{3.4}$$

The Poisson process: the number of events in $(0, t]$ has a Poisson distribution with parameter λt, i.e. $EN(t) = \lambda t$.

This leads to the second definition.

Definition 3.2 *The counting process* $\{N(t),\ t \geq 0\}$ *is called a Poisson process with intensity rate* $\lambda > 0$, *if*

1. $N(0) = 0$;

2. The process has independent increments;

3. For $s < t$, *the number of claims in the interval* $(s, t]$ *has a Poisson distribution with parameter* $\lambda(t-s)$:

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

By the law of large number or by the Chebyshev's inequality, it follows that $\frac{N(t)}{t} \xrightarrow{\mathrm{P}} \lambda$, as $t \to \infty$. In fact, almost sure convergence holds. This means that the intensity measures the average frequency or density of claim arrivals.

According to the second definition, if $h > 0$ is small

$$P(N(t+h) - N(t) = 1) = \lambda h e^{-\lambda h} = \lambda h - \lambda h(1 - e^{-\lambda h}) = \lambda h + o(h),$$

as $h \to 0$. The last equality follows from the inequality $0 < 1 - e^{-x} < x$, for $x > 0$.

Furthermore, for $\lambda h \in (0, \frac{1}{2})$

$$\sum_{n=2}^{\infty} \frac{(\lambda h)^n}{n!} \leq \frac{1}{2} \sum_{n=2}^{\infty} (\lambda h)^n = \frac{1}{2} \frac{(\lambda h)^2}{1 - \lambda h} \leq (\lambda h)^2.$$

This implies that, as $h \to 0$

$$P(\text{at least 2 claims arrive in}(t, t+h]) = \sum_{n=2}^{\infty} \frac{(\lambda h)^n}{n!} e^{-\lambda h} \leq (\lambda h)^2 = o(h),$$

which is the fourth postulate of the first definition.

In order to prove that the definitions 3.1 and 3.2 are equivalent, we have to show that the increments $N(t) - N(s) \sim Po(\lambda(t-s)),\ 0 \leq s < t$. It follows from the stationarity condition.

The interarrival times

Let T_1 be the time until the first claim. For $n \geq 2,\ T_n$ is the time between the $(n-1)$th and the nth claim. $T_2, T_3, \ldots$ is a sequence of the interarrival times.

The distribution of T_n : Note that the event $\{T_1 > t\}$ occurs if and only if, no one event occurs up to time t and

$$P(T_1 > t) = P(N(t) = 0) = e^{-\lambda t}.$$

Hence T_1 is exponentially distributed with parameter λ. Given T_1, the distribution of T_2 is

$$\begin{aligned} &P(T_2 > t) = P(T_2 > t | T_1 = s) = \\ &= P(0 \text{ claims in}(s, s+t] | T_1 = s) = P(0 \text{ claims in}(s, s+t]) = e^{-\lambda t}, \end{aligned} \tag{3.5}$$

where the last two equalities follow from the conditions of independent, stationary increments. From (3.4) it follows that T_2 is independent of T_1 and is exponentially distributed with parameter λ. By the same arguments and by induction it could be proved the following.

Theorem 3.1 *The interarrival times* $T_n,\ n = 1, 2, \ldots$ *are independent, identically* $\exp(\lambda)$ *distributed random variables.*

Exercise 3.1 *Prove that the waiting time* $\sigma_k = T_1 + \ldots + T_k \sim \Gamma(k, \lambda),\ k = 1, 2, \ldots,$ *i.e.*

$$f_{\sigma_k}(t) = \frac{\lambda}{\Gamma(k)} (\lambda t)^{k-1} e^{-\lambda t},\ t > 0$$

and

$$F_{\sigma_k}(t) = 1 - \sum_{j=0}^{k-1} \frac{(\lambda t)^j}{j!} e^{-\lambda t}.$$

The properties of Theorem 3.1 and Exercise 3.1 characterize the Poisson process. This states in the following

Theorem 3.2 *Let* $N(t),\ t \geq 0$ *be a stochastic process with* $N(0) = 0$. T_1 *is the time until the first claim,* $T_2, T_3, \ldots$ *the interarrival times. Let* $T_k,\ k = 1, 2, \ldots$ *be independent, identically* $\exp(\lambda)$ *distributed random variables and* $N(t) =$ *the number of claims up to time* t*. Then* $N(t)$ *is a Poisson process.*

Proof. Follows from

$$P(\sigma_k \leq t) = P(N(t) \geq k), \quad k = 0, 1, \ldots$$

and the lack of memory property of the exponential distribution.

□

The Fisher index of dispersion for the Poisson process is:

$$FI(t) = \frac{Var(N(t))}{EN(t)} = 1.$$

3.1.1 Order statistics property

Let $U_1, \dots U_n$ be independent uniformly distributed random variables ($U_i \sim U([a, b])$. The order statistics of this sample are

$$U_{(1)} \leq U_2 \leq \dots \leq U_{(n)},$$

where $U_{(n)} = \max\{U_i,\ i = 1, \dots, n\}$, $U_{(1)} = \min\{U_i, i = 1, \dots n\}$. The joint density function of $(U_{(1)}, \dots U_{(n)})$ is given by

$$f(u_1, \dots u_n) = \frac{n!}{(b-a)^n} I_{\{a<u_1 \dots <u_n<b\}}$$

and is called the Dirichlet distribution ($D_n([a, b])$.

Theorem 3.3 *Let $T_1, \dots, T_n$ be independent $\exp(\lambda)$ distributed random variables and $\sigma_k = T_1 + \dots T_k,\ k = 1, \dots n$. Then*

a) $(\sigma_1, \dots \sigma_n)$ has a probability density in $\mathbb{R}^n$ given by

$$f_{\sigma_1,\dots,\sigma_n}(t_1, \dots t_n) = \lambda^n e^{-\lambda t_n}(t_1, \dots, t_n), \quad t_1 < t_2 < \dots < t_n. \tag{3.6}$$

b) The joint conditional density of $(\sigma_1, \dots \sigma_n)$, given that $N(t) = n$ is

$$f_{\sigma_1,\dots,\sigma_n|N(t)=n}(t_1, \dots t_n) = \begin{cases} \frac{n!}{t^n}, & t_1 < t_2 < \dots < t_n \\ \\ 0, & \textit{otherwise.} \end{cases} \tag{3.7}$$

Proof. a) The joint distribution of $T_1, \dots, T_n$ is given by

$$f_{T_1,\dots,T_n}(x_1, \dots x_n) = \lambda^n \prod_{i=1}^{n} e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^n x_i},\ x_i \geq 0.$$

Change of variables

$$\begin{array}{lcl} t_1 = x_1 & & x_1 = t_1 \\ t_2 = x_1 + x_2 & & x_2 = t_2 - t_1 \\ t_3 = x_1 + x_2 + x_3 & \text{and} & x_3 = t_3 - t_2 \\ \ldots & & \ldots \\ t_n = x_1 + \ldots x_n & & x_n = t_n - t_{n-1} \end{array}$$

yields (3.6).

b)

$$P(\sigma_1 \le t_1, \ldots \sigma_n \le t_n | N(t) = n) = \frac{P(\sigma_1 \le t_1, \ldots \sigma_n \le t_n, N(t) = n)}{P(N(t) = n)}$$

$$= \frac{\int \ldots \int P(N(t) - N(s_n) = 0) f_{\sigma_1, \ldots \sigma_n}(s_1, \ldots s_n) ds_1 \ldots ds_n}{P(N(t) = n)}$$

$$= \frac{\int_0^{t_1} \int_{s_1}^{t_2} \ldots \int_{s_{n-1}}^{t_n} e^{-\lambda(t-s_n)} \lambda^n e^{-\lambda s_n} ds_n ds_{n-1} \ldots ds_1}{\frac{(\lambda t)^n}{n!} e^{-\lambda t}}$$

$$= \int_0^{t_1} \int_{s_1}^{t_2} \ldots \int_{s_{n-1}}^{t_n} \frac{n!}{t^n} ds_n ds_{n-1} \ldots ds_1.$$

□

3.2 Renewal process

In terms of point processes: Suppose that $0 = \sigma_0 \le \sigma_1 \le \ldots$ are random points in $\mathbb{R}^+$, at which a certain event occurs. The counting process is defined by the number of points in the interval $(0, t]$ and is given by

$$N(t) = \sum_{n=1}^{\infty} I_{\{\sigma_n \le t\}}, \quad t \ge 0.$$

Assume this counting process has finite values for each t. This is equivalent to $\sigma_n \to \infty$ a.s. as $n \to \infty$. The point counting process $N(t)$ is *simple* if the occurrence times are distinct: $0 < \sigma_1 < \sigma_2 < \ldots$ a. s.

Definition 3.3 *A simple point process $\{\sigma_n,\ n \ge 1\}$ is a renewal process if the inter - arrival times $T_n = \sigma_n - \sigma_{n-1}, \quad n \ge 1$ are independent, identically* *distributed with common distribution function F_T, $F_T(0) = 0$ and $\sigma_0 = 0$. The points $\{\sigma_n\}$ are called renewal times.*

Suppose that the time between claims are not concentrated at zero, i. e. $P(T = 0) < 1$. Let

$$N(t) = \max\{n, \sigma_n \leq t\} = \min\{n, \sigma_{n+1} > t\}, \tag{3.8}$$

or $N(t)$ is equal to the number of renewals in the interval $(0, t]$.

Definition 3.4 *The process (3.8) is called a renewal counting process.*

The equivalence of the processes $\{\sigma_n\}$ and $\{N(t)\}$ follows from the equivalence of the events

$$\{N(t) = n\} \text{ and } \{\sigma_n \leq t < \sigma_{n+1}\} \tag{3.9}$$

and

$$\{N(t) \geq n\} = \{\sigma_n \leq t\}, \quad n = 1, 2, \ldots. \tag{3.10}$$

Note that $\{N(t),\ t \geq 0\}$ is defined in continuous time and the sample paths are continuous on the right.

For the risk model, again (σ_n) is a sequence of arrival times and (T_n) is the sequence of inter - arrival times.

Example 3.1 (Homogeneous Poisson process) *Let $T_n \sim \exp(\lambda)$. The renewal process is called a homogeneous Poisson process with intensity λ. In this case, by the convolution properties*

$$P(\sigma_n \leq t) = F_T^{*n}(t) = \int_0^t \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} dx.$$

This is gamma distribution with parameters n and λ and

$$P(\sigma_n \leq t) = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Using the relation (3.10) we arrive at

$$P(N(t) \leq n) = 1 - P(N(t) \geq n+1) = 1 - P(\sigma_{n+1} \leq t) = \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

The motivation for introducing the renewal counting process is that the homogeneous Poisson process does not describe in an adequate way the claim arrivals. In many cases it is quite natural to model the inter - arrival times by log - normal, Pareto or another distribution.

In general, the Poisson process is a special case of renewal process, but many of the asymptotic properties are the same.

Suppose that the inter - arrival times T are defined by the probability distribution function $F_T(x)$ and $ET = \frac{1}{\lambda}$. In some cases the distribution F_T is defective. This means that $\lim_{x\to\infty} F_T(x) < 1$, or the random variable T can be equal to ∞ with positive probability $1 - F(\infty)$. Here $F(\infty) = \lim_{x\to\infty} F(x)$. We will call in this case the renewal process terminating. We will show that the limit $N(\infty) = \lim_{t\to\infty} N(t) < \infty$ with probability 1 and $\sigma_{N(\infty)}$ has a compound geometric distribution.

Theorem 3.4 (Strong law of large numbers) *If $F(\infty) = 1$, then the samples of $\{N(t),\ t \geq 0\}$ are increasing with probability 1 and*

$$\lim_{t\to\infty} \frac{N(t)}{t} = \lambda = (ET)^{-1} \quad a.\ s. \tag{3.11}$$

Proof. The limit $\lim_{t\to\infty} N(t)$ exists, since the sample paths of the process are increasing. From (3.10) it follows that $\lim_{t\to\infty} N(t) = \infty$ with probability 1. The strong law of large numbers for (T_n) yields that $\lim_{n\to\infty} \frac{\sigma_n}{n} = \frac{1}{\lambda} > 0$ with probability 1. Consequently, $\lim_{n\to\infty} \frac{\sigma_{N(t)}}{N(t)} = \frac{1}{\lambda}$ with probability 1. Again from (3.10), it follows that $\sigma_{N(t)} \leq t < \sigma_{N(t)+1}$ and

$$\frac{\sigma_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{\sigma_{N(t)+1}}{N(t)+1} \frac{N(t)+1}{N(t)}.$$

Let $t \to \infty$ and obtain (3.11) in the case of $\lambda > 0$. When $\lambda = 0,\ ET = \infty$. Note that

$$N(\infty) = \lim_{t\to\infty} N(t) = \min\{n : T_n = \infty\} - 1,$$

i. e. $N(\infty)$ is geometrically distributed with parameter $P(N(\infty) = 0) = 1 - F(\infty)$. From the total probability law it follows that the random variable $\sigma_{N(\infty)}$ has a geometric distribution with parameter $F(\infty)$.

□

The strong law of large numbers for the renewal process shows that the mean value $EN(t)$ is approximated by λt and plays a key role in the asymptotic analysis. Note that in the case of homogeneous Poisson process the exact mean value $EN(t) = \lambda t$ is known.

Theorem 3.5 (Elementary Renewal Theorem) *Suppose that* $F_T(\infty) = 1$. *Then*

$$\lim_{t\to\infty} \frac{EN(t)}{t} = \lambda.$$

Proof. Since $N(t)$ is not always a stopping time we define the first passage time

$$\nu(t) = \min\{n : \sigma_n > t\}.$$

Note that $\nu(t) = N(t) + 1$. It follows that

$$EN(t) = E\nu(t) - 1 = \lambda E\sigma_{\nu(t)} - 1 = \lambda t + \lambda E(\sigma_{\nu(t)} - t) - 1,$$

and

$$\frac{N(t)}{t} = \lambda + \lambda \frac{E(\sigma_{\nu(t)} - t)}{t} - \frac{1}{t}. \tag{3.12}$$

Since $\sigma_{\nu(t)} - t \geq 0$, we obtain that

$$\liminf_{t\to\infty} \frac{EN(t)}{t} \geq \lambda. \tag{3.13}$$

Note that for some $M \geq 0$,

$$P(T_k \leq M) = 1, \ k = 1, 2, \dots .$$

Together with (3.12) this implies that

$$\frac{EN(t)}{t} \leq \lambda + \frac{M}{t}$$

and

$$\limsup_{t\to\infty} \frac{EN(t)}{t} \leq \lambda,$$

which together with (3.13) proves the theorem.

□

The next theorem states the asymptotic behavior of the variance of the renewal process.

Theorem 3.6 *Suppose that $Var(T) < \infty$. Then*

$$\lim_{t\to\infty} \frac{Var(N(t))}{t} = \frac{Var(T_1)}{(ET_1)^3}.$$

Theorem 3.7 (Central Limit Theorem) *Let $Var(T_1) < \infty$. Then for $t \to \infty$*

$$\frac{N(t) - \lambda t}{\sqrt{\frac{Var(T_1)}{(ET_1)^3} t}} \longrightarrow Z \sim N(0, 1).$$

3.2.1 Renewal function

Definition 3.5 *The mean value*

$$m(t) = 1 + EN(t) = E\nu(t), \ \ t \geq 0$$

is called a renewal function.

Since $\{N(t) \geq k\} = \{\sigma_k \leq t\}, \ k = 1, 2, \ldots$ we have

$$m(t) = 1 + \sum_{k=1}^{\infty} P(N(t) \geq k) = 1 + \sum_{k=1}^{\infty} P(\sigma_k \leq t) = \sum_{k=0}^{\infty} F_T^{*k}(t), \qquad (3.14)$$

that is $m(t)$ is the expected number of renewals in $[0, t]$ and is called also a *renewal measure.*

Note that $m(t)$ is similar to a distribution function. It is nondecreasing and right - continuous on $\mathbb{R}$, but has a unit jump at $t = 0$ and $m(t) \uparrow \infty$ as $t \to \infty$.

An important property of the renewal function is that it determines uniquely the distribution F_T. Taking the Laplace transform of both sides in (3.14) we have

$$LT_{m(t)}(s) = \sum_{k=0}^{\infty} LT_{F_T^{*k}}(s) = \sum_{k=0}^{\infty} (LT_{F_T}(s))^k = \frac{1}{1 - LT_{F_T}(s)}.$$

This yields the following theorem.

Theorem 3.8 *The Laplace transforms of $m(t)$ and F_T are determine each other uniquely by the relation*

$$LT_{m(t)}(s) = \frac{1}{1 - LT_{F_T}(s)}.$$

Hence $m(t)$ and F_T uniquely determine each other.

This result can be used for identification of the renewal process. For example, the renewal function of the Poisson process is $m(t) = \lambda t + 1$ and any renewal process with this type of renewal function is a Poisson process.

Theorem 3.9 (The Integral Equation for Renewal process) *The renewal function $m(t)$ satisfies the integral equation*

$$m(t) = I_{[0,\infty)}(t) + F_T(t) + \int_0^t m(t-s)dF_T(s). \tag{3.15}$$

Moreover, $m(t)$ is the unique solution of (3.15), which is bounded on finite intervals.

Proof. According to (3.14), it follows that

$$\begin{aligned} m(t) &= 1 + F_T(t) + \textstyle\sum_{k=2}^{\infty} F_T^{*k}(t) = I_{[0,\infty)}(t) + F_T(t) + \sum_{k=1}^{\infty} F_T^{*(k+1)}(t) \\ &= 1 + F_T(t) + \sum_{k=1}^{\infty} (F_T^{*k} * F_T)(t), \end{aligned}$$

or

$$m(t) = I_{[0,\infty)}(t) + F_T(t) + (m * F_T)(t),$$

which is equivalent to (3.15).

□

The equation (3.15) is called a *renewal equation*. In the general case the renewal equation is given by

$$U(t) = u(t) + \int_0^t U(t-y)dF(y), \tag{3.16}$$

where all functions are defined on $[0,\infty)$. The function $U(t)$ is unknown, $u(t)$ is given and $F(y)$ is a distribution function. If F is a defective distribution, then (3.16) is called a *defective renewal equation*.

The solution of the equation (3.16) is defined by the next theorem.

Theorem 3.10 *If $u(t)$ is bounded on finite intervals, then*

$$\begin{aligned} U(t) &= \sum_{k=0}^{\infty} \int_0^t u(t-s)dF^{*k}(s) \\ &= \int_0^{\infty} u(t-s)dm(s), \quad t \geq 0, \end{aligned}$$

is the only solution of the renewal equation (3.16).

Theorem 3.11 (Key renewal theorem) *If in addition $u(t)$ is directly Riemann integrable, then*

$$\lim_{t\to\infty} U(t) = \lambda \int_0^{\infty} u(s)ds.$$

3.2.2 Recurrence times of a renewal process

Consider the renewal sequence $\{\sigma_n, \quad n = 1, 2, \ldots \quad \sigma_0 = 0\}$ and $T_n > 0$. Note that

$$\{N(t) = n\} = \{\sigma_n \leq t < \sigma_{n+1}\}.$$

In particular,

$$\sigma_{N(t)} \leq t < \sigma_{N(t)+1}.$$

For $t \geq 0$, $B(t) = t - \sigma_{N(t)}$ is the time since the last renewal prior to t and is called a *backward recurrence time* of the renewal process (or the age process).

$F(t) = \sigma_{N(t)+1} - t$ is the time to the next renewal after t and is called *forward time* (or excess life or residual life).

We will show that for fixed $0 \leq x < t$, the distribution function $P(B(t) \leq x)$ satisfies the renewal equation. Since $B(t) \leq t$ a.s. it is sufficient to consider $x < t$. Hence for $x \geq t, \quad P(B(t) \leq x) = 1$.

Let us start with the identity

$$P(B(t) \leq x) = P(B(t) \leq x, T_1 \leq t) + P(B(t) \leq x, T_1 > t), \; x > 0. \tag{3.17}$$

If $T_1 > t$, no jumps occur up to t and $N(t) = 0$. Consequently $B(t) = t$, and hence

$$P(B(t) \leq x, T_1 > t) = (1 - F_{T_1}(t)), \quad x \leq t.$$

For $T_1 \le t$ we will show that

$$P(B(t) \le x, T_1 \le t) = \int_0^t P(B(t-y) \le x) dF_{T_1}(y). \tag{3.18}$$

According to the properties of the renewal process

$$\begin{aligned} P(B(t) \le x, T_1 \le t) &= P(t - \sigma_{N(t)} \le x, N(t) \ge 1) \\ &= \textstyle\sum_{n=1}^{\infty} P(t - \sigma_{N(t)} \le x, N(t) = n) \\ &= \textstyle\sum_{n=1}^{\infty} P(t - \sigma_n \le x, \sigma_n \le t < \sigma_{n+1}). \end{aligned}$$

For every one of the summands, given $\{T_1 = y\}$ for $y \le t$ we have

$$\begin{aligned} &P(t - \sigma_n \le x, \sigma_n \le t < \sigma_{n+1} | T_1 = y) \\ &= P(t - [y + \textstyle\sum_{i=2}^{n} T_i] \le x, y + \sum_{i=2}^{n} T_i \le t < y + \sum_{i=2}^{n+1} T_i) \\ &= P(t - y - \sigma_{n-1} \le x, \sigma_{n-1} \le t - y \le \sigma_n) \\ &= P(t - y - \sigma_{N(t-y)} \le x, N(t-y) = n-1), \end{aligned}$$

and hence

$$\begin{aligned} P(B(t) \le x, T_1 \le t) &= \textstyle\sum_{n=1}^{\infty} \int_0^t P(t - y - \sigma_{N(t-y)} \le x, N(t-y) = n) dF_{T_1}(y) \\ &= \textstyle\int_0^t P(B(t-y) \le x) dF_{T_1}(y), \end{aligned}$$

which proves (3.18).

Combining (3.17) and (3.18) we get

$$P(B(t) \le x) = (1 - F_{T_1}(t)) I_{[0,x]}(t) + \int_0^t P(B(t-y) \le x) dF_{T_1}(y),$$

which is the renewal equation for $u(t) = (1 - F_{T_1}(t)) I_{[0,x]}(t)$ and $U(t) = P(B(t) \le x)$.

Similarly

$$P(F(t) > x) = \int_0^t P(F(t-y) > x) dF_{T_1}(y) + (1 - F_{T_1}(t+x)).$$

It is known that the only solution of the renewal equation is given by

$$U(t) = P(B(t) \le x) = \int_0^t (1 - F_{T_1}(t-y)) I_{[0,x]}(t-y) dm(y).$$

Consider the case of homogeneous Poisson process with intensity λ. In this case $m(t) = \lambda t + 1, \;\; 1 - F_{T_1}(x) = e^{-\lambda x}$ and

$$P(B(t) \le x) = P(t - \sigma_{N(t)} \le x) = \begin{cases} 1 - e^{-\lambda x}, & x < t \\ 1, & x \ge t. \end{cases}$$

Analogously,

$$P(F(t) \le x) = P(\sigma_{N(t)+1} - t \le x) = 1 - e^{-\lambda x}, \; x > 0.$$

3.2.3 Delayed Renewal process

In many cases the renewal process starts at random point. In this case the time T_1 has a different distribution. The process $N(t)$ is called a *delayed renewal process.* The distribution of T_1 is called the *delayed distribution.*

Definition 3.6 *A continuous time stochastic process $N(t), \; t \ge 0$ is stationary or has stationary increments, if for every points $0 = t_1 < \ldots < t_k$ and $h > 0$*

$$(N(t_1 + h), \ldots, N(t_k + h)) \stackrel{d}{=} (N(t_1), \ldots, N(t_k)).$$

A basic property of the stationary point process is that the mean value function is linear.

Theorem 3.12 *Let $N(t)$ be a stationary point process and $EN(1)$ is finite. Then $EN(t) = EN(1)t$.*

Proof. Consider

$$\begin{aligned} EN(s+t) &= EN(s) + [EN(s+t) - EN(s)] \\ &= EN(s) + EN(t). \end{aligned}$$

This is a functional equation $f(s+t) = f(s) + f(t), \quad s, t \ge 0$. The only nondecreasing function that satisfies this equation is $f(ct)$ for some constant c. In our case $c = f(1) = EN(1)$, and hence $EN(t) = EN(1)t$.

$\square$

We are ready to characterize the stationary renewal process.

Theorem 3.13 *The delayed renewal process $N(t)$ is stationary if and only if the forward recurrence time process $F(t) = \sigma_{N(t)+1} - t$ is stationary.*

Proof. Using $\sigma_n = \inf\{s, N(s) = n\}$, we have

$$\begin{aligned} F(t) &= \sigma_{N(t)+1} - t = \inf\{s - t : N(s) = N(t) + 1\} \\ &\stackrel{d}{=} \inf\{s : N(t, s] = 1\} = \{N(0, s - t] = 1\}. \end{aligned}$$

Consequently, the stationarity property of N implies $F(t) \stackrel{d}{=} F(0), \ t \geq 0$. Then F is stationary process (it is a Markov process). Conversely, since N counts the number of times $F(t)$ jumps upward,

$$N(A + t) = \sum_{s \in A} I_{\{F(u+s) > F((u+t)-)}.$$

Therefore the stationarity of F implies N is stationary.

□

Theorem 3.14 *Let $N(t)$ be a stationary delayed renewal process. Then*

a) $EN(t) = \lambda t, \quad t \geq 0$.

b) $F_{T_1}(t) = \lambda \int_0^t [1 - F_{T_2}(s)]ds = \frac{1}{ET_2} \int_0^t [1 - F_{T_2}(s)]ds$.

Proof. a) Let $N(t)$ be stationary. Theorem 3.12 ensure that $EN(t) = EN(1)t$. Also $EN(1) = \lambda$, since $\frac{EN(t)}{t} \to \lambda, \quad t \to \infty$. Therefore $EN(t) = \lambda t$.

b) Please, see the proof in [19].

□

3.3 Mixed Poisson process

The modeling by homogeneous Poisson process is not realistic.

Suppose that the parameter λ is a realization of the random variable Λ with distribution function F_Λ. Then

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

where $F_\Lambda(\lambda) = P(\Lambda \le \lambda)$ is the distribution function of the mixing distribution Λ.

For the mixed Poisson process Fisher index of dispersion is

$$FI(t) = 1 + \frac{Var(\Lambda)}{E\Lambda} > 1,$$

i. e. it is over - dispersed related to the Poisson process.

Example 3.2 *Let $\Lambda \sim \Gamma(\alpha, \beta)$ with density function*

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

Then

$$p_k(t) = P(N(t) = k) = \binom{\alpha + k - 1}{k} \left(\frac{\beta}{\beta + t}\right)^\alpha \left(\frac{t}{\beta + t}\right)^k, \quad k = 0, 1, 2, \dots, \tag{3.19}$$

i. e. the number of claims has negative binomial distribution with parameters α and $\frac{\beta}{\beta+t}$ ($N(t) \sim NBD\ (\alpha, \frac{\beta}{\beta+t})$). The counting process, defined by (3.18) is called a **Pólya process.**

3.4 Compound Poisson process

Let $N_1(t)$ be a homogeneous Poisson process and $X_1, X_2, \dots$ a sequence of independent, identically distributed random variables, independent of $N_1(t)$. Define the process

$$N(t) = \sum_{i=1}^{N(t)} X_i. \tag{3.20}$$

Definition 3.7 *The process (3.20) is called a compound Poisson process.*

The distribution of X is called a compounding distribution.

3.4.1 Pólya - Aeppli process

The Pólya - Aeppli process is defined in [16]. Suppose that the compounding random variable X has a geometic distribution with parameter $1 - \rho$, i. e. $P(X = i) = \rho^{i-1}(1 - \rho), \quad i = 1, 2, \ldots$. Then, the compound Poisson process is called a Pólya - Aeppli process.

Definition 3.8 *A counting process $\{N(t), t \geq 0\}$ is said to be a Pólya - Aeppli process if*

a) $N(0) = 0$;

b) $N(t)$ has independent, stationary increments;

c) for each $t > 0$, $N(t)$ is Pólya - Aeppli distributed.

Theorem 3.15 *Suppose that the inter-arrival times $\{T_k\}_{k\geq 2}$ of the stationary renewal process are equal to zero with probability ρ and with probability $1 - \rho$ exponentially distributed with parameter λ. Then the number of renewals up to time t, has the Pólya - Aeppli distribution with parameters λ and ρ.*

Chapter 4

Claim Size Models

The claim amounts to the insurance company can be described by discrete and by continuous random variables. In the case of continuous distributed claims, the basic is to find an adequate model for the claim amount. The probability distributions are separated in two families - light tailed and heavy tailed distributions. The exponential distribution is the border one.

Definition 4.1 *The distribution F is called light tailed, if for $\lambda > 0$,*

$$\limsup_{x\to\infty} \frac{\overline{F}(x)}{e^{-\lambda x}} < \infty.$$

For the light tailed distribution there are constants $a > 0$ and $\lambda > 0$, such that $\overline{F}(x) \leq ae^{-\lambda x}$ and there is $z > 0$, such that $M_X(z) < \infty$.

According to the definition, the exponential distribution is light tailed for every $\lambda > 0$.

Definition 4.2 *If for every $\lambda > 0$,*

$$\liminf_{x\to\infty} \frac{\overline{F}(x)}{e^{-\lambda x}} > 0,$$

the distribution F has a heavy tail.

For the heavy tailed distributions, for every $a > 0$ and $\lambda > 0$, $\overline{F}(x) > ae^{-\lambda x}$ and for every $z > 0$, $\quad M_X(z) = \infty$.

The truncated normal distribution is called a standard distribution. This is a random variable $Z = |Y|$, where Y is a normal distributed random variable with distribution function $F(z) = P(|Y| \leq z)$. If Y has a standard normal distribution, $F(z) = 2(\Phi(z) - \frac{1}{2}), \quad x > 0$, where $\Phi(z)$ is a standard normal distribution function.

It is easy to show that if $\varphi(z)$ is a standard normal density, then

$$\lim_{z\to\infty} \frac{z\Phi(z)}{\varphi(z)} = 1,$$

and hence the truncated normal distribution is a light tailed.

The Pareto distribution with parameters $\alpha > 0$ and $\lambda > 0$

$$1 - F(x) = \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}, \quad x \ > 0$$

has a heavy tail.

The Weibull distribution

$$1 - F(x) = e^{-\left(\frac{x}{\sigma}\right)^{\beta}}, \quad x > 0, \quad \lambda > 0$$

is heavy tailed for $\beta < 1$ and light tailed for $\beta \geq 1$.

The most useful light tailed and heavy tailed distributions are given in the next tables.

Light tailed distributions		
Name	Parameters	Density
Exponential	$\lambda > 0$	$f_X(x) = \lambda e^{-\lambda x}$
Gamma	$\alpha > 0,\ \beta > 0$	$f_X(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$
Weibull	$\beta > 0, \quad \tau \geq 1$	$f_X(x) = \beta\tau x^{\tau-1} e^{-\beta x^{\tau}}$
Hyperexponential	$\lambda_i > 0, \quad \sum_{i=1}^{n} p_i = 1$	$f_X(x) = \sum_{i=1}^{n} p_i \lambda_i e^{-\lambda_i x}$

Heavy tailed distributions		
Name	Parameters	Density
Weibull	$\beta > 0,\ 0 < \tau < 1$	$f_X(x) = \beta\tau x^{\tau-1} e^{-\beta x^{\tau}}$
Lognormal	$\mu \in \mathbb{R}, \quad \sigma > 0$	$f_X(x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$
Loggamma	$\alpha > 0, \quad \beta > 0$	$f_X(x) = \frac{\beta^{\alpha} (\ln x)^{\alpha-1}}{x^{\beta+1}\Gamma(\alpha)}$
Pareto	$\alpha > 0, \quad \lambda > 0$	$f_X(x) = \frac{\alpha}{\lambda+x} \left(\frac{\lambda}{\lambda+x}\right)^{\alpha}$
Burr	$\alpha > 0, \quad \lambda > 0, \quad \tau > 0$	$f_X(x) = \frac{\alpha\tau\lambda^{\alpha} x^{\tau-1}}{(\lambda+x^{\tau})^{\alpha+1}}$

4.1 Heavy tailed distributions

Let $\alpha_F = \limsup_{x\to\infty} \frac{\Lambda(x)}{x}$, where $\Lambda(x) = -\log \overline{F}(x)$ is the hazard function of F. If F is continuously differentiable, then $\Lambda(x)$ is also differentiable and $\frac{d\Lambda(x)}{dx} = \lambda(x)$, where $\lambda(x)$ is the intensity rate function. Suppose that $F(0-) = 0$.

Theorem 4.1 *If $\alpha_F = 0$, then F is heavy tailed distribution function.*

Proof. Suppose that $\alpha_F = 0$, i. e. $\limsup_{x\to\infty} \frac{\Lambda(x)}{x} = 0$. Then, for every $\varepsilon > 0$, there is $x' > 0$, such that for $x > x'$, $\Lambda(x) \le \varepsilon x$. Consequently, there is a constant $c > 0$, such that for every $x \ge 0$, $\overline{F}(x) \ge ce^{-\varepsilon x}$, and therefore for every $t \ge \varepsilon$,

$$\int_0^\infty e^{tx}\overline{F}(x)dx = \infty. \tag{4.1}$$

Letting $\varepsilon \to 0$ yields that (4.1) holds for every $t > 0$.

□

Remark 4.1 *Suppose that F is heavy tailed distribution function. Then for every $t > 0$,*

$$\lim_{x\to\infty} e^{tx}\overline{F}(x) = \infty. \tag{4.2}$$

The most popular heavy tailed distributions are the distributions with regularly varying tails and subexponential distributions.

4.2 Regularly varying functions

Definition 4.3 *The positive, measurable function f is called* **regularly varying at** ∞ *with index $\alpha \in \mathbb{R}$, if*

$$\lim_{x\to\infty} \frac{f(tx)}{f(x)} = t^\alpha \quad \textit{for every } t > 0. \tag{4.3}$$

If $\alpha = 0$, f is called **slowly varying function**.

The family of regularly varying functions with index α is denoted by $\mathbf{RV}(\alpha)$.

Suppose that in (4.3) the limit, if it exists, is finite and positive for every $t > 0$. Then the limiting function satisfies the equation $K(ts) = K(t)K(s)$. The only solution of this equation is the power function.

Remark 4.2 *If $f \in RV(\alpha)$ then*

$$f(x) = x^{\alpha} L(x),$$

where $L(x)$ is a slowly varying function ($L(x) \in RV(0)$).

Remark 4.3 *The function $f(x)$ is called regularly varying at 0, if $f(\frac{1}{x})$ is regularly varying at ∞.*

Examples: Slowly varying functions are the positive constants and the functions $\log(1+x)$, $\log\log(e+x)$.

According the Remark 4.1, the following functions:

$$x^{\alpha}, \quad x^{\alpha}\ln(1+x), \quad [x\ln(1+x)]^{\alpha}, \quad x^{\alpha}\ln(\ln(e+x))$$

are in $RV(\alpha)$. Probability distributions whose tails are regularly varying are

$$1 - F(x) = x^{-\alpha}, \; x \geq 1, \; \alpha > 0$$

and the extreme value distribution

$$\Phi_{\alpha}(x) = \exp(-x^{-\alpha}), \quad x \geq 0.$$

4.2.1 Properties

Theorem 4.2 (Karamata's theorem) *Let L be a slowly varying function. Then there exists $t_0 > 0$ such that L is locally bounded over $[t_0, \infty)$ and*

a) For $\alpha > -1$

$$\int_{t_0}^{t} s^{\alpha} L(s) ds \sim (\alpha+1)^{-1} t^{\alpha+1} L(t), \; t \to \infty. \tag{4.4}$$

b) For $\alpha < -1$ or $\alpha = -1$ and $\int_0^{\infty} \frac{L(s)}{s} ds < \infty$,

$$\int_t^\infty s^\alpha L(s)ds \sim -(\alpha+1)^{-1}t^{\alpha+1}L(t), \; t\to\infty. \tag{4.5}$$

Conversely, if (4.4) holds with $\alpha > -1$, then $L \in RV(0)$. If (4.5) holds with $\alpha > -1$, then $L \in RV(0)$.

Remark 4.4 *Let $f \in RV(\alpha)$ and local bounded over $[t_0,\infty)$ for some $t_0 \geq 0$. Then*

a) For $\alpha > -1$

$$\lim_{t\to\infty} \frac{\int_{t_0}^t f(s)ds}{tf(t)} = \frac{1}{\alpha+1}.$$

b) For $\alpha < -1$

$$\lim_{t\to\infty} \frac{\int_t^\infty f(s)ds}{tf(t)} = -\frac{1}{\alpha+1}.$$

If $\alpha = -1$ and for some positive function f, local bounded over $[t_0,\infty)$ one of the conditions a) or b) satisfies, then $f \in RV(\alpha)$.

Theorem 4.3 (Representation theorem) *Let $f \in RV(\alpha)$. Then there exist measurable functions $\varepsilon : \mathbb{R}^+ \to \mathbb{R}$ and $c : \mathbb{R}^+ \to \mathbb{R}$, such that*

$$\lim_{t\to\infty} \varepsilon(t) = \alpha \quad \textit{and} \quad \lim_{t\to\infty} c(t) = c_0 > 0 \tag{4.6}$$

and $t_0 \in \mathbb{R}^+$, such that for $t > t_0$

$$f(t) = c(t)\exp\left\{\int_{t_0}^t \frac{\varepsilon(s)}{s}ds\right\}. \tag{4.7}$$

Conversely, if (4.7) holds with ε and c satisfying (4.6), then $f \in RV(\alpha)$.

For example, the function $L(t) = \ln t$ is slowly varying and the representation (4.7) holds with $t_0 = e,\; c(t) = 1$ and $\varepsilon(t) = (\ln t)^{-1}$.

Remark 4.5 *From the representation theorem, it follows that the functions of $RV(\alpha)$ satisfy*

$$\lim_{t\to\infty} f(t) = \begin{cases} \infty, & \alpha > 0 \\ 0, & \alpha < 0. \end{cases}$$

If L is slowly varying, then for every $\alpha > 0$,

$$\lim_{t\to\infty} t^{-\alpha}L(t) = 0 \;\textit{and}\; \lim_{t\to\infty} t^\alpha L(t) = \infty.$$

4.2.2 Regularly varying random variables

Definition 4.4 *The nonnegative random variable X and its distribution are called regularly varying with index $\alpha \geq 0$, if the right tail distribution $\overline{F}_X(x) \in RV(-\alpha)$.*

Theorem 4.4 (Regularly varying distributions) *Let F be a distribution function and $F(x) < 1$ for every $x \geq 0$.*

a) If the sequences (a_n) and (x_n) satisfy $\frac{a_n}{a_{n+1}} \to 1, \quad x_n \to \infty$ and for some real function g and all values λ of some subset of $(0, \infty)$,

$$\lim_{n\to\infty} a_n \overline{F}(\lambda x_n) = g(\lambda) \in (0, \infty),$$

then $g(\lambda) = \lambda^{-\alpha}$ for some $\alpha \geq 0$ and $\overline{F} \in RV$.

b) Let F be absolutely continuous with density f, such that $\lim_{x\to\infty} \frac{xf(x)}{\overline{F}(x)} = \alpha$ for some $\alpha > 0$. Then $f \in RV(-(1+\alpha))$ and $\overline{F} \in RV(-\alpha)$.

c) Let $f \in RV(-(1+\alpha))$ for $\alpha > 0$. Then $\lim_{x\to\infty} \frac{xf(x)}{\overline{F}(x)} = \alpha$. This means that if for $\alpha > 0, \quad \overline{F} \in RV(-\alpha)$, then the density f is monotone function.

d) Let $X \in RV(\alpha)$ be a nonnegative random variable and $\alpha > 0$. Then

$$EX^\beta < \infty, \quad \text{if} \quad \beta < \alpha;$$

$$EX^\beta = \infty, \quad \text{if} \quad \beta > \alpha.$$

e) Suppose that $\overline{F} \in RV(-\alpha), \quad \alpha > 0, \quad \beta \geq \alpha$. Then

$$\lim_{x\to\infty} \frac{x^\beta \overline{F}(x)}{\int_0^x y^\beta dF(y)} = \frac{\beta - \alpha}{\alpha}.$$

The opposite holds if $\beta > \alpha$. If $\beta = \alpha$, the only that we can say is that $\overline{F}(x) = \circ(x^{-\alpha}L(x))$ for some slowly varying function L.

f) The following are equivalent:

$$(1) \quad \int_0^x y^2 dF(y) \in RV(0),$$

$$(2) \quad \overline{F}(x) = \circ(x^{-2} \int_0^x y^2 dF(y)), \quad x \to \infty.$$

Example 4.1 Distributions similar to Pareto *These are the Pareto distribution, Cauchy distribution, Burr distribution, stable distribution with index $\alpha < 2$. The right tails of all these distributions are given by*

$$\overline{F}(x) \sim Kx^{-\alpha}, \quad x \to \infty$$

for some constants $\alpha > 0$ and K. It is clear that $\overline{F}(x) \in RV(\alpha)$.

The regularly varying distributions are widely used in practice. The motivation is given in the next

Lemma 4.1 *Let X and Y be independent nonnegative random variables from $RV(\alpha)$, $\alpha \geq 0$. Then $X + Y \in RV(\alpha)$ and*

$$P(X + Y > x) \sim P(X > x) + P(Y > x), \quad x \to \infty.$$

4.3 Subexponential distributions

Let $F(x)$ be a regularly varying distribution function, defined on $(0, \infty)$, such that for every $n \geq 2$,

$$\lim_{x\to\infty} \frac{\overline{F^{*n}}(x)}{\overline{F}(x)} = n, \tag{4.8}$$

where $\overline{F^{*n}}(x)$ means the nth convolution of $\overline{F(x)} = 1 - F(x)$.

The condition (4.8) means that for $x \to \infty$

$$P(X_1 + \ldots + X_n) > x) \sim nP(X_1 > x).$$

It is easy to show that (4.7) for $n = 2$, implies $n \geq 2$.

Let $X_1, X_2, \ldots$ be nonnegative identically distributed random variables. Then for $x \to \infty$,

$$\begin{aligned} P(\max(X_1, \ldots X_n) > x) &= 1 - [F(x)]^n = \\ &= [1 - F(x)] \textstyle\sum_{k=0}^{n-1} [F(x)]^k \sim n[1 - F(x)], \end{aligned},$$

and hence (4.8) is equivalent to

$$\lim_{x\to\infty} \frac{P(X_1 + \ldots + X_n) > x)}{P(\max(X_1, \ldots X_n) > x)} = 1. \tag{4.9}$$

This means that the largest claims have the main contribution to the sum $X_1 + \ldots + X_n$. This property defines a large family of probability distributions.

Definition 4.5 Subexponential distribution *The nonnegative random variable X and the distribution of X are called subexponential if the independent copies $X_1, X_2, \ldots, X_n$ satisfy (4.8) and (4.9).*

The family of subexponential distributions (SE) is introduced by V. Chistyakov in 1964.

4.3.1 Properties

The main properties are given in the next lemma

Lemma 4.2 *Let $F \in SE$. Then*

a) For all $y \in \mathbb{R}$

$$\lim_{x\to\infty} \frac{\overline{F}(x-y)}{\overline{F}(x)} = 1. \tag{4.10}$$

b) If (4.10) holds, then for every $r > 0$

$$\lim_{x\to\infty} e^{rx}(1 - F(x)) = \infty$$

and

$$\int_{0-}^{\infty} e^{rx} dF(x) = \infty.$$

c) For given $\varepsilon > 0$, there is a finite constant K, such that for $n \geq 2$

$$\frac{1 - F^{*n}(x)}{1 - F(x)} \leq K(1+\varepsilon)^n, \quad x \geq 0. \tag{4.11}$$

Proof. a) For $0 \leq y \leq x$ we have

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

Thus for large x, such that $F(x) - F(y) \neq 0$, we have

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

The assertion (4.10) follows by letting $x \to \infty$.

If $y < 0$ then

$$\lim_{x\to\infty} \frac{1-F(x-y)}{1-F(x)} = \lim_{x\to\infty} \frac{1-F(x-y)}{1-F((x-y)-(-y))} = \lim_{v\to\infty} \left[\frac{1-F(v-(-y))}{1-F(v)}\right]^{-1} = 1,$$

where $v = x - y$.

b) For the MGF we have

$$\begin{aligned}\int_{0-}^{\infty} e^{rx} dF(x) &= 1 + \int_0^\infty \int_0^x re^{ry} dy dF(x) \\ &= 1 + r\int_0^\infty \int_y^\infty e^{ry} dF(x) dy = 1 + r\int_0^\infty e^{ry}(1-F(y))dy = \infty.\end{aligned}$$

c) Let

$$C_n = \sup_{x\geq 0} \frac{1-F^{*n}(x)}{1-F(x)}.$$

Then

$$\begin{aligned}\frac{1-F^{*(n+1)}(x)}{1-F(x)} &= 1 + \frac{F(x)-F^{*(n+1)}(x)}{1-F(x)} = 1 + \int_0^x \frac{1-F^{*n}(x-t)}{1-F(x)} dF(t) \\ &= 1 + \int_0^x \left(\frac{1-F^{*n}(x-t)}{1-F(x-t)}\frac{1-F(x-t)}{1-F(x)}\right) dF(t).\end{aligned}$$

For $T < \infty$ and $n \geq 1$

$$C_{n+1} \leq 1 + \sup_{0\leq x\leq T} \int_0^x \frac{\overline{F^{*n}}(x-y)}{\overline{F}(x)} dF(y) + \sup_{x\geq T} \int_0^x \frac{\overline{F^{*n}}(x-y)}{\overline{F}(x-y)} \frac{\overline{F}(x-y)}{\overline{F}(x)} dF(y) \leq$$

$$\leq 1 + \frac{1}{\overline{F}(T)} + C_n \sup_{x\geq T} \frac{F(x)-F^{*2}(x)}{\overline{F}(x)},$$

where $\frac{1}{\overline{F}(T)} < \infty$.

Since $F \in SE$, then for a given $\varepsilon > 0$, one can choose $T = T(\varepsilon)$ such that

$$\sup_{x\geq T} \frac{F(x)-F^{*2}(x)}{\overline{F}(x)} < 1 + \varepsilon,$$

and hence

$$C_{n+1} \leq 1 + \frac{1}{\overline{F}(T)} + C_n(1+\varepsilon).$$

Note that $C_1 = 1$. We obtain the assertion recursively

$$\begin{aligned} C_n &\leq 1 + \tfrac{1}{\overline{F}(T)} + C_{n-1}(1+\varepsilon) \leq 1 + \tfrac{1}{\overline{F}(T)} + (1 + \tfrac{1}{\overline{F}(T)})(1+\varepsilon) + C_{n-2}(1+\varepsilon)^2 \leq \ldots \\ &\leq (1 + \tfrac{1}{\overline{F}(T)}) \textstyle\sum_{i=0}^{n-1} (1+\varepsilon)^i \leq \frac{(1+\frac{1}{\overline{F}(T)})}{\varepsilon}((1+\varepsilon)^n. \end{aligned}$$

□

Remark 4.6 *The assertion b) explains the name of the sub - exponential distributions $F \in SE$. Since for every $\varepsilon > 0$*

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

MGF of the distributions $F \in SE$, doesn't exist. Consequently, the Laplace - Stieltjes also doesn't exist.

The following lemma gives a sufficient condition for subexponentiality.

Lemma 4.3 *Let for $z \in (0,1]$ the limit*

$$\gamma(z) = \lim_{x\to\infty} \frac{1 - F(zx)}{1 - F(x)}$$

exists and is left - continuous at 1. Then F is a subexponential distribution function.

Proof. Note that

$$F^{*2}(x) = P(X_1 + X_2 \leq x) \leq P(X_1 \leq x, X_2 \leq x) = [F(x)]^2.$$

Assume that $F(0) = 0$. Hence

$$\liminf_{x\to\infty} \frac{1 - F^{*2}(x)}{1 - F(x)} \geq \liminf_{x\to\infty} \frac{1 - [F(x)]^2}{1 - F(x)} = \liminf_{x\to\infty} [1 + F(x)] = 2.$$

For fixed $n \geq 1$

$$\begin{aligned}
\limsup_{x\to\infty} \frac{1-F^{*2}(x)}{1-F(x)} &= 1 + \limsup_{x\to\infty} \int_0^x \frac{1-F(x-y)}{1-F(x)} dF(y) \\
&\leq 1 + \limsup_{x\to\infty} \sum_{k=1}^{n} \frac{1-F(x-\frac{kx}{n})}{1-F(x)} \left[F\left(\frac{kx}{n}\right) - F\left(\frac{(k-1)x}{n}\right) \right] \\
&= 1 + \gamma(1-\frac{1}{n}).
\end{aligned}$$

Since n is arbitrary, γ is left continuous at 1.

$\square$

Example 4.2 *Consider the Pareto(α, β) distribution.*

$$\frac{1-F(zx)}{1-F(x)} = \frac{\left(\frac{\beta}{\beta+zx}\right)^{\alpha}}{\left(\frac{\beta}{\beta+x}\right)^{\alpha}} = \left(\frac{\beta+x}{\beta+zx}\right)^{\alpha} \longrightarrow z^{-\alpha},$$

as $x \to \infty$. It follows that Pareto distribution is subexponential.

Remark 4.7 *The condition (4.10) gives another definition of the heavy tailed distributions. For the random variable X with distribution F, the condition (4.10) can be written as*

$$\lim_{x\to\infty} P(X > x+y | X > x) = \lim_{x\to\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = 1, \quad y > 0.$$

The light tail of the exponential distribution $\overline{F}(x) = e^{-\lambda x},\ x > 0$, for $\lambda > 0$, satisfies the condition

$$\lim_{x\to\infty} \frac{\overline{F}(x+y)}{\overline{F}(x)} = e^{-\lambda y}, \quad y > 0.$$

Chapter 5

Cramér - Lundberg model

5.1 Ruin probability

The ruin probability is a measure for the risk related to the some company. For convenience we consider an insurance company.

Consider the usual risk model. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with

a) Counting process $N(t), \quad N(0) = 0$;

b) A sequence $\{Z_k\}_1^\infty$ of independent identically distributed random variables with distribution function F such that $F(0) = 0$, mean μ and variance σ^2.

The risk process is given by

$$X(t) = ct - \sum_{k=1}^{N(t)} Z_k \quad (\sum_{k=1}^{0} \stackrel{def}{=} 0),$$

where c is a positive constant, representing the premium incomes per unit time. The constant c is called *gross risk premium rate.*

Suppose that $N(t)$ has an intensity λ and $EN(t) = \lambda t$. The profit of this risky business in $(0, t]$ is $X(t)$. The expected profit is

$$EX(t) = ct - EN(t)EZ_k = (c - \lambda\mu)t.$$

Define the *safety loading* coefficient θ by

$$\theta = \frac{c - \lambda\mu}{\lambda\mu} = \frac{c}{\lambda\mu} - 1.$$

The risk process X(t) has a positive safety loading, if $\theta > 0$, i. e. $c > \lambda\mu$. In this case X(t) has a trend to $+\infty$ and we say that there is a *net profit condition,* (NPC).

The ruin probability $\Psi(u)$ for the insurance company with initial capital u is defined by

Definition 5.1

$$\Psi(u) = P\{u + X(t) < 0 \text{ for some } t > 0\}.$$

Sometimes it is more convenient to use the probability of non ruin $\Phi(u) = 1 - \Psi(u)$.

From the definition it follows that $\Psi(u) = 1$ for $u < 0$.

Definition 5.2 *If $N(t)$ is a homogeneous Poisson process with intensity λ, i. e. $N(t) \sim Po(\lambda t)$, $X(t)$ is called a* **classical risk model**, *or* **Cramér - Lundberg model.**

5.2 Integral equation of ruin probability

Consider the Poisson process as a renewal process. At first we will derive the equation for the non-ruin probability $\Phi(u)$. Let T_1 be the time to the first claim. Then $X(T_1) = cT_1 - Z_1$. Conditioning on the no claim in $(0, T_1)$, we obtain

$$\begin{aligned} \Phi(u) &= E[\Phi(u + cT_1 - Z_1)] = \\ &= \int_0^\infty \lambda e^{-\lambda s} \int_{0-}^{u+cs} \Phi(u + cs - z) dF(z) ds. \end{aligned}$$

The change of variables $x = u + cs$ leads to

$$\Phi(u) = \frac{\lambda}{c} e^{\lambda \frac{u}{c}} \int_u^\infty e^{-\lambda \frac{x}{c}} \int_{0-}^{x} \Phi(x - z) dF(z) dx,$$

and since Φ is differentiable we have

$$\Phi'(u) = \frac{\lambda}{c}\Phi(u) - \frac{\lambda}{c} \int_{0-}^{u} \Phi(u - z) dF(z). \tag{5.1}$$

Integrating from 0 to t leads to

$$\begin{aligned}
&\Phi(t) - \Phi(0) \\
&= \frac{\lambda}{c}\int_0^t \Phi(u)du + \frac{\lambda}{c}\int_0^t \int_0^u \Phi(u-z)d(1-F(z))du \\
&= \frac{\lambda}{c}\int_0^t \Phi(u)du + \frac{\lambda}{c}\int_0^t \left[\Phi(0)(1-F(u)-\Phi(0)+\int_0^u \Phi'(u-z)(1-F(z))dz\right]du \\
&= \frac{\lambda}{c}\Phi(0)\int_0^t (1-F(u))du + \frac{\lambda}{c}\int_0^t (1-F(z))dz\int_z^t \Phi'(u-z)du \\
&= \frac{\lambda}{c}\Phi(0)\int_0^t (1-F(u))du + \frac{\lambda}{c}\int_0^t (1-F(z))(\Phi(t-z)-\Phi(0))dz
\end{aligned}$$

and to the equation

$$\Phi(u) = \Phi(0) + \frac{\lambda}{c}\int_0^u \Phi(u-z)(1-F(z))dz. \tag{5.2}$$

It follows that for $u \to \infty$, $\Phi(\infty) = \Phi(0) + \frac{\lambda\mu}{c}\Phi(\infty)$. According to the law of large numbers we have

$$\lim_{t\to\infty}\frac{X(t)}{t} = c - \lambda\mu \text{ with probability 1.}$$

The NPC says that $\theta > 0$ and $c > \lambda\mu$. Consequently there is a dependent of N and $\{Z_k\}$ random variable T, such that $X(t) > 0$ for every $t > T$. Since only a finite number of claims can arrive up to T,

$$\inf_{t>0} X(t) < \infty \text{ with probability 1}$$

and $\Phi(\infty) = 1$. Consequently

$$1 = 1 - \Psi(0) + \frac{\lambda\mu}{c},$$

i. e.

$$\Psi(0) = \frac{\lambda\mu}{c} = \frac{1}{1+\theta}.$$

Example 5.1 *Suppose that the claims are exponentially distributed with mean* μ, *i. e.* $F(x) = 1 - e^{-\frac{x}{\mu}}, \quad x \geq 0$. *The equation (5.1) is given by*

$$\Phi'(u) = \frac{\lambda}{c}\Phi(u) - \frac{\lambda}{c\mu}\int_0^u \Phi(u-z)e^{-\frac{z}{\mu}}dz = \frac{\lambda}{c}\Phi(u) - \frac{\lambda}{c\mu}\int_0^u \Phi(z)e^{-\frac{u-z}{\mu}}dz.$$

The second derivative

$$\Phi''(u) + \frac{1}{\mu}\frac{\theta}{1+\theta}\Phi'(u) = 0.$$

and the initial conditions

$$\Phi(\infty) = 1 \quad and \quad \Phi(0) = \frac{\theta}{1+\theta}$$

give the solution

$$\Phi(u) = 1 - \frac{1}{1+\theta}e^{-\frac{1}{\mu}\frac{\theta}{1+\theta}u}.$$

From the equation (5.2) and $\Psi(u) = 1 - \Phi(u)$ it follows that

$$\Psi(u) = \frac{\lambda}{c}\int_u^\infty (1-F(z))dz + \frac{\lambda}{c}\int_0^u \Psi(u-z)(1-F(z))dz. \tag{5.3}$$

5.3 Cramér - Lundberg approximation

From the definition of the safety loading coefficient it follows that $\frac{\lambda}{c} = \frac{1}{\mu(1+\theta)}$. Denote by

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

the **integrating tail distribution** related to $F(x)$. In terms of the safety loading, the equation (5.2) for the nonruin probability has the form

$$\Phi(u) = \frac{\theta}{1+\theta} + \frac{1}{1+\theta}\int_0^u \Phi(u-z)dF_I(z).$$

Denote $q = \frac{1}{1+\theta}$ and rewrite the equation for $\Psi(u)$.

$$\Psi(u) = q\overline{F}_I(u) + \int_0^u \Psi(u-x)d(qF_I(x)). \tag{5.4}$$

This is a renewal equation relative the measure $qF_I(x)$. It is easy to see that $\lim_{x\to\infty} qF_I(x) = q < 1$, i. e. $qF_I(x)$ is not a probability measure. The equation (5.4) is a defective renewal equation.

For some r and $x > 0$, we define

$$F^{(r)}(x) = \int_0^x e^{rz} d(qF_I(z)) = \frac{q}{\mu} \int_0^x e^{rz}(1 - F(z))dz,$$

which is a distribution function. It is nondecreasing and $\lim_{x\to\infty} F^{(r)}(x) = 1$, i. e.

$$\frac{q}{\mu} \int_0^\infty e^{rz}(1 - F(z))dz = 1. \tag{5.5}$$

The distribution, generated by $F^{(r)}(x)$ is called **Esscher transform** of F. The equation (5.5) is called a **Cramér condition.**

From (5.5) it follows that

$$f^{(r)}(x) = \frac{q}{\mu} e^{rx}(1 - F(x)), \quad x > 0.$$

is a density of a proper probability distribution.

Multiplying the equation (5.4) by e^{ru} :

$$\begin{aligned} e^{ru}\Psi(u) &= qe^{ru}\overline{F}_I(u) + \textstyle\int_0^u e^{r(u-x)}\Psi(u-x)e^{rx}d(qF_I(x)) \\ &= qe^{ru}\overline{F}_I(u) + \int_0^u e^{r(u-x)}\Psi(u-x)dF^{(r)}(x)). \end{aligned} \tag{5.6}$$

This is a renewal equation

$$U(t) = u(t) + \int_0^t U(t-y)dF(y),$$

for $u(t) = qe^{ru}\overline{F}_I(u)$, $\quad F = F^{(r)}$ and $U(t) = e^{rt}\Psi(t)$. The function $U(t)$ is bounded on finite intervals. By the Theorem 3.10, the solution of the equation (5.6) is given by

$$U(t) = e^{rt}\Psi(t) = \int_0^t u(t-y)dm^{(r)}(y) = q\int_0^t e^{r(t-y)}\overline{F}_I(t-y)dm^{(r)}(y),$$

where $m^{(r)}$ is the renewal function, corresponding to the renewal process with $F^{(r)}$ distributed interarrival times. In general, the function $m^{(r)}$ is unknown. The Key renewal theorem gives the asymptotic solution of the equation for $u \to \infty$.

If the integrals $C_1 = q\int_0^\infty e^{rz}\overline{F}_I(z)dz$ and $C_2 = q\int_0^\infty ze^{rz}\overline{F}_I(z)dz$ exist, then for the solution of (5.6) we obtain

$$\lim_{u\to\infty} e^{ru}\Psi(u) = \frac{C_1}{C_2}. \tag{5.7}$$

From (5.5) it follows that the Cramér condition is given by

$$\frac{M_Z(r)-1}{r} = \mu(1+\theta),$$

where $M_Z(r)$ is the MGF of Z. We can see that the adjustment coefficient is independent of the Poisson parameter λ.

It can be shown that the function $M_Z(r) - 1 - r\mu(1+\theta)$ is convex. The equation

$$M_Z(r) - 1 = r\mu(1+\theta) \tag{5.8}$$

has a solution equal to zero. If no zero solution exists, it is positive. Let R be the positive solution of (5.8). Then for C_1 and C_2 we get

$$C_1 = \frac{1}{R}\frac{\theta}{1+\theta}$$

and

$$C_2 = \frac{1}{R}\frac{1}{1+\theta}\frac{1}{\mu}\left(M_Z'(R) - \mu(1+\theta)\right)$$

and the limit (5.7) is

$$\lim_{u\to\infty} e^{Ru}\Psi(u) = \frac{\theta\mu}{M_Z'(R) - \mu(1+\theta)}. \tag{5.9}$$

Definition 5.3 *(5.9) is called* **Cramér - Lundberg approximation**. *The constant R, the non negative solution of the equation (5.5) is called a* **Lundberg exponent** *or* **adjustment coefficient.**

Example 5.2 *We continue the Example 5.1 for exponentially distributed claims. The MGF is given by $M_Z(r) = \frac{1}{\mu}\int_0^\infty e^{rz}e^{-\frac{z}{\mu}}dz = \frac{1}{1-\mu r}$. The constant R is the positive solution of the equation*

$$\frac{\mu r}{1-\mu r} = \mu r(1+\theta),$$

and hence

$$R = \frac{1}{\mu}\frac{\theta}{1+\theta}.$$

The derivative of the MGF is $M'_Z(r) = \mu(1+\theta)^2$ *and (5.9) is given by*

$$\lim_{u\to\infty} e^{Ru}\Psi(u) = \frac{1}{1+\theta}.$$

This result and the ruin probability of Example 5.1 show that the Cramér - Lundberg approximation for exponentially distributed claims is exact.

5.4 Martingale approximation

Remember two theorems, important for the approximation.

Theorem 5.1 (Stopping Time Theorem) *Let* τ *be a finite stoping time (markov moment), i. e.* $\tau \le t_0 < \infty$ *and* M *is a right continuous* $F-$*martingale (supermartingale). Then*

$$E[M(\tau)|\mathcal{F}_0] = (\le)M(0), \quad P-a.\ s.$$

Theorem 5.2 *Let* $X(t)$ *be a continuous process such that*

1. $X(0) = 0 \quad P-a.\ s.$;
2. X *has stationary independent increments;*
3. $EX(t) = \beta t, \quad \beta > 0$;
4. $Ee^{-rX(t)} < \infty$ *for* $r > 0$.

Then $Ee^{-rX(t)} = e^{g(r)t}$ *for some function* $g(\cdot)$.

If $X(t)$ is a classical risk model with NPC, then $X(t)$ has stationary independent increments with $\beta = c - \lambda\mu$ and

$$\begin{aligned} Ee^{-rX(t)} &= e^{-rct}Ee^{r(Z_1+\ldots+Z_{N(t)})} = e^{-rct}\textstyle\sum_{k=0}^{\infty}\frac{(\lambda t)^k}{k!}e^{-\lambda t}[M_Z(r)]^k \\ &= e^{-rct}e^{\lambda M_Z(r)t}e^{-\lambda t} = e^{[\lambda(M_Z(r)-1)-cr]t}. \end{aligned}$$

In this case

$$g(r) = \lambda(M_Z(r)-1) - cr.$$

It is easy to show that

$$M_u(t) = \frac{e^{-r(u+X(t))}}{e^{g(r)t}}$$

is a martingale relative the $\sigma-$ algebra, generated by the process X.

Let $T_u = \inf\{t \geq 0 : u + X(t) < 0\}$ be the time to ruin for a company with initial capital u. T_u is a stoping time relative the filtration $\mathcal{F}^X$ and the ruin probability is

$$\Psi(u) = P(T_u < \infty).$$

For $t_0 < \infty$, $t_0 \wedge T_u$ is finite $\mathcal{F}^X$ - stoping time. $\mathcal{F}_0^X$ is a trivial $\sigma-$ algebra and $M_u > 0$. According to the Theorem 5.1

$$e^{-ru} = M_u(0) = E[M_u(t_0 \wedge T_u)] =$$

$$= E[M_u(t_0 \wedge T_u)|T_u \leq t_0]P(T_u \leq t_0) + E[M_u(t_0 \wedge T_u)|T_u > t_0]P(T_u > t_0) \geq$$

$$\geq E[M_u(t_0 \wedge T_u)|T_u \leq t_0]P(T_u \leq t_0) = E[M_u(T_u)|T_u \leq t_0]P(T_u \leq t_0).$$

Since $u + X(T_u) \leq 0$ for $T_u < \infty$, then $e^{-r(u+X(T_u))} \geq 1$. Consequently

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

Let $t_0 \to \infty$ and obtain

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

Let $R = \sup\{r : g(r) \leq 0\}$, i. e. R is the positive solution of the equation $g(r) = 0$. This is just the equation (5.8) and R is the Lundberg exponent. We obtain the inequality

$$\Psi(u) \leq e^{-Ru}, \tag{5.10}$$

named a **Lundberg inequality.**

This approximation gives an interpretation of the Lundberg coefficient. Consider the surplus process $U(t) = u + X(t)$. We can show that the process $e^{-rU(t)}$ is a martingale:

$$Ee^{-rU(t)} = Ee^{-r(u+ct-S(t))} = e^{-ru}e^{-rct}e^{\lambda[M_Z(r)-1]t} = e^{-ru}.$$

So, R is the unique positive number, such that $e^{-rU(t)}$ is a martingale. It is known that the martingale property is related to the fair game.

We are ready to give the basic theorem.

Theorem 5.3 (Cramér - Lundberg)

1. *If the constant $R > 0$ exists and*

$$\int_0^\infty e^{Rx} dF_I(x) = 1 + \theta,$$

then

$$\Psi(u) \leq e^{-Ru},$$

where R is the Lundberg exponent.

2. *Suppose that the integral $\int_0^\infty x e^{Rx} \overline{F}_I(x) dx = C$ is finite. Then*

$$\lim_{u \to \infty} \Psi(u) = \frac{\theta \mu}{RC} e^{-ru}.$$

Chapter 6

Renewal Risk Model

Consider the surplus process

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

where $S(t) = \sum_{k=1}^{N(t)} Z_k$, the counting process $N(t)$ is a renewal process, $0 = \sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \dots$ are the claim arrival times. The inter-arrival times $T_k = \sigma_k - \sigma_{k-1},\ k = 1, 2, \dots$ are independent identically distributed random variables with mean value $ET_k = \frac{1}{\lambda}$. The claims Z_k are independent identically distributed random variables, independent of $N(t)$.

Let $F_{T_1}(t)$ be the distribution function of the time to the first claim and $F_T(t)$ be the distribution function of $T_2, T_3, \dots$. If $F_{T_1}(t) = F_T(t)$, the counting process is called an *ordinary renewal process* and the risk model is called a *Sparre Andersen model.*

If the stationarity condition

$$F_{T_1}(t) = \lambda \int_0^t [1 - F_T(x)]dx \tag{6.1}$$

holds, the counting process is stationary and the risk model is called a *stationary renewal risk model.*

6.1 Ordinary renewal risk model

In the ordinary case we use $\Psi^0(u)$ and $\Phi^0(u) = 1 - \Psi^0(u)$ for ruin and non ruin probability. According to the argument that ruin can occur only at claim times, the ruin probability for this model $\Psi^0(u) = P(U(t) \leq 0, \;\; t \geq 0)$ can be given by

$$\begin{aligned}\Psi^0(u) &= P(u + c\sigma_n - S(\sigma_n) \leq 0, \;\; n \geq 1) \\ &= P\left(u + \textstyle\sum_{k=1}^n (cT_k - Z_k) \leq 0, \;\; n \geq 1\right) \\ &= P\left(\sup_{n\geq 1} \textstyle\sum_{k=1}^n (Z_k - cT_k) \geq u\right).\end{aligned}$$

Notation: $X_k = Z_k - cT_k, \; W_n = \sum_{k=1}^n X_k$ and

$$M_n = \sup_{n\geq 1} \sum_{k=1}^n X_k = \sup_{n\geq 1} W_n. \tag{6.2}$$

Then $\Phi^0(u) = 1 - \Psi^0(u) = P(M_n \leq u)$. The defined random variables $X_k, k = 1, 2, \ldots$ are independent identically distributed. A sum of independent identically distributed random variables is called a *random walk.* This proves the following

Proposition 1 *The ruin probability for the zero-delayed case can be represented as $\Psi^0(u) = P(M_n > u)$, where M_n is given by (6.2) with W_n a discrete time random walk with increments distributed as the difference $Z - cT$ between claims Z and the interarrival time cT.*

The relative safety loading $\theta = \frac{cET_1}{EZ_1} - 1 = \frac{c}{\lambda\mu} - 1 > 0$, i.e. the premium received per unit time exceeds the expected claim payments per unit time. The NPC implies that

$$EX_1 = E(Z_1 - cT_1) = \mu - \frac{c}{\lambda} < 0$$

and that $W_n \to -\infty$, a.s. It is easy to proof that $0 \leq M_n < \infty$.

The safety loading coefficient θ is the same like in the Cramér - Lundberg model. The difference is that here $EN(t) \neq \lambda t$, and hence excluding the case $T \sim \exp(\lambda)$,

$$E\left(U(t) - u\right) \neq (c - \lambda\mu)t.$$

According to the Law of large numbers

$$\lim_{t\to\infty} \frac{E(U(t) - u)}{t} = c - \lambda\mu,$$

and $\Psi(u) \to 0$ for $u \to \infty$.

6.1.1 Lundberg exponent

Let $U(t)$ be an ordinary renewal risk model.

Lemma 6.1 *Suppose that $M_Z(r) < \infty$ and for $r \geq 0$, $g(r)$ is the unique solution of the equation*

$$M_Z(r)M_T(-g(r) - cr) = 1. \tag{6.3}$$

Then the discrete time process $e^{-rU_{\sigma_k} - g(r)\sigma_k}$ is a martingale.

Proof. If $r \geq 0$, then $M_Z(r) \geq 1$. $M_T(r)$ is increasing continuous function, defined for $r \leq 0$. Consequently $M_T(r) \to 0$ for $r \to -\infty$ and there is an unique solution $g(r)$ of the equation (6.3). Then

$$\begin{aligned} E\left[e^{-rU_{\sigma_{k+1}} - g(r)\sigma_{k+1}}|\mathcal{F}_{\sigma_k}\right] &= E\left[e^{-r[c(\sigma_{k+1}-\sigma_k)Z_{k+1}] - g(r)(\sigma_{k+1}-\sigma_k)}|\mathcal{F}_{\sigma_k}\right] e^{-rU_{\sigma_k} - g(r)\sigma_k} \\ &= E\left[e^{rZ_{k+1}} e^{-(cr+g(r))(\sigma_{k+1}-\sigma_k)}|\mathcal{F}_{\sigma_k}\right] e^{-rU_{\sigma_k} - g(r)\sigma_k} \\ &= E\left[M_Z(r)M_T(-cr - g(r))\right] e^{-rU_{\sigma_k} - g(r)\sigma_k} = e^{-rU_{\sigma_k} - g(r)\sigma_k}. \end{aligned}$$

□

Example 6.1 *For exponentially distributed interarrival times $F(t) = 1 - e^{-\lambda t}$. The equation (6.3) is given by*

$$M_Z(r)\frac{\lambda}{\lambda + g(r) + cr} = 1.$$

As in the classical case, the function $g(r)$ is convex and $g(0) = 0$. There exists a nonnegative solution R to the equation $g(R) = 0$. This solution is called again *adjustment coefficient* or *Lundberg exponent.*

Recall that R is the only positive solution of the equation

$$M_Z(r)M_T(-cr) = 1. \tag{6.4}$$

Example 6.2 *Let $U(t)$ be a renewal risk model with* $\exp(1)$ *distributed claims, premium rate $c = 4$ and iterarrival time distribution $F(t) = 1 - \frac{1}{2}\left(e^{-2t} + e^{-3t}\right)$.*

It follows that $M_Z((r)$ exists for $r < 1$, $M_T((r)$ exists for $r < 2$ and $\lambda = 2.4$. The NPC $4 > 2.4$ is fulfilled. The equation to solve is

$$\frac{1}{1-r}\frac{1}{2}\left(\frac{2}{2+4r}+\frac{3}{3+4r}\right)=1.$$

Thus

$$2(3+4r)+3(2+4r)=2(1-r)(2+4r)(3+4r)$$

or equivalently

$$4r^3+r^2-r=0.$$

We find the solutions $r = 0$ and

$$r_{1,2}=\frac{-1\pm\sqrt{17}}{8},$$

such that $r_1 = \frac{-1+\sqrt{17}}{8} > 0$ and $r_2 = \frac{-1-\sqrt{17}}{8} < 0$. We proved that there is only no negative solution. Why do we get $r_2 < 0$. Obviously $M_Z(r_2) < 1 < \infty$. But $-cr = \frac{1+\sqrt{17}}{2} > 2$ and thus $M_T(-cr_2) = \infty$ and r_2 is not a solution of (6.4).

6.1.2 Pollaczeck - Khinchine formula (Ruin probability as a compound geometric probability)

The Poisson process is an ordinary renewal process. By the renewals arguments we can proof the following. If $\theta > 0$, then

$$\Phi^0(u)=\frac{\theta}{1+\theta}\sum_{n=0}^{\infty}\left(\frac{1}{1+\theta}\right)^n F_I^{*n}(u), \tag{6.5}$$

where F_I^{*n} is the $n-$ th convolution of F_I. From the equation of the non ruin probability

$$\Phi^0(u)=\Phi^0(0)+\frac{1}{1+\theta}\int_0^u \Phi^0(u-z)dF_I(z) \tag{6.6}$$

by the Laplace transform follows (6.5). Denote by

$$LT_{\Phi^0}(s)=\int_0^{\infty} e^{-sz}\Phi^0(z)dz$$

the Laplace transform and

$$LST_{\Phi^0}(s) = \int_0^\infty e^{-sz} d\Phi^0(z)$$

the Laplace - Stieltjes transform of Φ^0. Recall the relation $LT_{\Phi^0}(s) = \frac{1}{s} LST_{\Phi^0}(s)$. Taking the Laplace transform in both sides of the equation (6.6) gives

$$\begin{aligned}
LT_{\Phi^0}(s) &= \frac{\Phi^0(0)}{s} + \frac{1}{1+\theta} \int_0^\infty e^{-sz} \int_0^z \Phi^0(z-t) dF_I(t) dz \\
&= \frac{\Phi^0(0)}{s} + \frac{1}{1+\theta} \int_0^\infty \int_t^\infty e^{-sz} \Phi^0(z-t) dz dF_I(t) \\
&= \frac{\Phi^0(0)}{s} + \frac{1}{1+\theta} \int_0^\infty \int_0^\infty e^{-s(x+t)} \Phi^0(x) dx dF_I(t) \\
&= \frac{\Phi^0(0)}{s} + \frac{1}{1+\theta} \int_0^\infty e^{-st} \left[\int_0^\infty e^{-sx} \Phi^0(x) dx \right] dF_I(t) \\
&= \frac{\Phi^0(0)}{s} + \frac{1}{1+\theta} LT_{\Phi^0}(s) \int_0^\infty e^{-st} dF_I(t) \\
&= \frac{\Phi^0(0)}{s} + \frac{1}{1+\theta} LT_{\Phi^0}(s) LST_{F_I}(s).
\end{aligned}$$

Hence, for the Laplace transform we have

$$LT_{\Phi^0}(s) = \frac{\Phi^0(0)}{s\left[1 - \frac{1}{1+\theta} LST_{F_I}(s)\right]}. \tag{6.7}$$

The Laplace - Stieltjes transform is

$$LST_{\Phi^0}(s) = \frac{\Phi^0(0)}{1 - \frac{1}{1+\theta} LST_{F_I}(s)} = \Phi^0(0) \sum_{n=0}^\infty \left[\frac{1}{1+\theta} LST_{F_I}(s) \right]^n.$$

According to the inversion formula with initial condition $\Phi^0(0) = \frac{\theta}{1+\theta}$, the probability of non ruin $\Phi^0(u)$ in the ordinary case is given by

$$\Phi^0(u) = \frac{\theta}{1+\theta} \sum_{n=0}^\infty \left(\frac{1}{1+\theta} \right)^n F_I^{*n}(u), \tag{6.8}$$

called **Pollaczeck - Khinchine formula**.

It is easy to see that this formula is a compound geometric sum, i. e. a geometric sum of independent, identically distributed random variables with distribution function $F_I(z)$.

Exercise 6.1 *Show that the ruin probability for the ordinary renewal risk model is given by*

$$\Psi^0(u) = \frac{\theta}{1+\theta}\sum_{n=1}^{\infty}\left(\frac{1}{1+\theta}\right)^n \overline{F}_I^{*n}(u),$$

6.2 Stationary case

Note that the ruin probability for the delayed case with $T_1 = s$ can be expressed in terms of the zero - delayed case as

$$\Psi(u) = 1 - F(u+cs) + \int_0^{u+cs} \Psi^0(u+cs-y)dF(y).$$

Indeed, the first term represents the probability $P(U_1 - cs > u)$ of ruin at the time s of the first claim. The second term is $P(\tau(u) < \infty, U_1 - cs \leq u)$, as follows easily by noting that the evolution of the risk process after time s is that of a renewal risk model with initial reserve $U_1 - cs$.

Proposition 2 *The non - ruin probability $\Phi(u)$ and the ruin probability $\Psi(u)$ in the stationary case satisfy the integral representations*

$$\Phi(u) = \Phi(0) + \frac{\lambda}{c}\int_0^u \Phi^0(u-z)(1-F(z))dz \tag{6.9}$$

and

$$\Psi(u) = \frac{\lambda}{c}\left[\int_u^{\infty}(1-F(z))dz + \int_0^u \Psi^0(u-z)(1-F(z))dz\right]. \tag{6.10}$$

Since $\Phi(\infty) = \Phi^0(\infty) = 1$ when $c > \lambda\mu$, we have

$$\Phi(0) = 1 - \frac{\lambda\mu}{c}.$$

Taking the Laplace transform of (6.9) and applying (6.7) we have

$$LT_\Phi(s) = \frac{\Phi(0)}{s} + \frac{\lambda\mu}{c} LST_{F_I}(s) \frac{\Phi^0(0)}{s[1 - LST_{F_I}(s)]}.$$

Again, the standard properties of the transforms lead to

$$LST_\Phi(s) = 1 - \frac{\lambda\mu}{c} + \frac{\lambda\mu}{c}\Phi^0(0) LST_{F_I}(s) \sum_{n=0}^{\infty} \left(\frac{1}{1+\theta}\right)^n [LST_{F_I}(s)]^n.$$

So, the ruin probability in the stationary case is given by

$$\Psi(u) = \frac{\lambda\mu}{c}\left[\overline{F}_I(u) + [1 - \frac{\lambda\mu}{c}] * F_I(u) \sum_{n=1}^{\infty} \left(\frac{\lambda\mu}{c}\right)^n \overline{F}_I^{*n}(u)\right], \qquad (6.11)$$

where $\overline{F}_I(u) = 1 - F_I(u)$.

In terms of the relative safety loading the ruin probability is given by

$$\Psi(u) = \frac{1}{1+\theta}\left[\overline{F}_I(u) + \frac{\theta}{1+\theta} F_I(u) * \sum_{n=1}^{\infty} \left(\frac{1}{1+\theta}\right)^n \overline{F}_I^{*n}(u)\right].$$

Example 6.3 *Consider again the case in which the claim amount distribution is exponential with mean value μ. Applying the argument of the ordinary case we obtain the ruin probability*

$$\Psi(u) = \frac{1}{1+\theta} \ \exp\left\{-\frac{1}{\mu}\frac{\theta}{1+\theta}u\right\}. \qquad (6.12)$$

6.3 Ruin probability for heavy tailed distributions

Recall that in the Cramér - Lundberg model the following relation for $\Psi(u)$ holds

$$\Psi(u) = \frac{\theta}{1+\theta}\sum_{n=1}^{\infty}\left(\frac{1}{1+\theta}\right)^n \overline{F}_I^{*n}(u),$$

where $F_I(u) = \frac{1}{\mu}\int_0^u [1 - F(x)]dx$ is the integrated tail distribution. Under the condition that $\overline{F}_I \in RV(-\alpha)$ for some $\alpha \geq 0$ we might hope that the following asymptotic estimate holds

$$\frac{\Psi(u)}{\overline{F}_I(u)} = \frac{\theta}{1+\theta}\sum_{n=1}^{\infty}\left(\frac{1}{1+\theta}\right)^n \frac{\overline{F_I^{*n}}(u)}{\overline{F}_I(u)} \tag{6.13}$$

$$\to \frac{\theta}{1+\theta}\sum_{n=1}^{\infty}\left(\frac{1}{1+\theta}\right)^n n = \frac{1}{\theta}, \quad u \to \infty. \tag{6.14}$$

(6.13) is a natural estimate of ruin probability whenever $\overline{F}_I$ is regularly varying. We shall show that a similar estimate holds true for much wider class of distribution functions. (6.13) can be reformulated as follows.

Proposition 3 *For claim size distributions with regularly varying tails, the ruin probability $\Psi(u)$ for large initial capital u is essentially determined by the tail $\overline{F}(z)$ of the claim size distribution for large values of z, i. e.*

$$\Psi(u) \sim \frac{1}{\mu\theta}\int_u^{\infty}\overline{F}(z)dz, \quad u \to \infty.$$

The main step in obtaining (6.14) is the property of the subexponential distributions

$$\overline{F_I^{*n}}(u) \sim n\overline{F}_I(u) \;\text{ for any arbitrary } n \geq 2 \;\text{ and }\; u \to \infty.$$

This leads to

Theorem 6.1 (Cramér - Lundberg theorem for large claims, I) *Consider the Cramér - Lundberg model with NPC and $F_I(z) \in SE$. Then*

$$\Psi(u) \sim \frac{1}{\theta}\overline{F}_I(u), \quad u \to \infty. \tag{6.15}$$

Proof. Since $\frac{1}{1+\theta} < 1$, there exists an $\varepsilon > 0$ such that $\frac{1}{1+\theta}(1+\varepsilon) < 1$. Together with the basic property

$$\frac{\overline{F^{*n}}(u)}{\overline{F}(u)} \leq K(1+\varepsilon)^n, \quad u \geq 0$$

it follows that

$$\frac{1}{(1+\theta)^n}\frac{\overline{F^{*n}}(u)}{\overline{F}(u)} \leq \frac{1}{(1+\theta)^n}K(1+\varepsilon)^n, \quad u \geq 0,$$

which allows by dominated convergence the interchange of limit and sum in (6.14), yielding the result.

□

For claim size distributions with subexponential integrated tail distribution, ultimate ruin probability $\Psi(u)$ is given by (6.15).

From mathematical point of view this result can be substantially improved.

Theorem 6.2 (Cramér - Lundberg theorem for large claims, II) *Consider the Cramér - Lundberg model with NPC. Then the following assertions are equivalent.*

a) $\overline{F}_I(u) \in SE$.

b) $1 - \Psi(u) \in SE$.

c) $\dfrac{\Psi(u)}{\overline{F}_I(u)} \longrightarrow \dfrac{1}{\theta}, \quad u \to \infty$.

Consequently, the estimate (6.15) is possible only under the condition $\overline{F}_I(u) \in SE$.

Chapter 7

Premium Calculation Principles

Denote by Π_X the premium that an insurer charges to cover the risk X. The risk X means that claims from this risk are described by the random variable X and the distribution of X. The premium Π_X is a function X, for example $\Pi_X = \phi(X)$. The rule that assigns the numerical value of Π_X is referred to as a premium calculation principle.

Properties of premium principles:

1. $\Pi_X \geq EX$ (nonnegative loading);

2. If X_1 and X_2 are independent, then $\Pi_{X_1+X_2} = \Pi_{X_1} + \Pi_{X_2}$ (additivity);

3. If $Z = aX$, where $a > 0$, then $\Pi_Z = a\Pi_X$ (scale invariance);

4. If $Y = X + c$, where $c > 0$, then $\Pi_Y = \Pi_X + c$ (consistency);

5. If there is a finite maximum value of the claim amount x_m then $\Pi_X \leq x_m$.

7.1 Premium calculation principles

7.1.1 Pure premium principle

$$\Pi_X = EX.$$

The pure premium is not very attractive.

7.1.2 Expected value principle

$$\Pi_X = (1+\theta)EX,$$

where $\theta > 0$ is the safety loading factor. θEX is the loading in the premium. The premium is easy to calculate. It assigns the same premium to all risks with the same mean value and is not sensible to heavy tailed distributions.

7.1.3 The variance principle

$$\Pi_X = EX + \alpha Var(X),$$

where $\alpha > 0$. The loading is proportional to $Var(X)$. This principle counts two characteristics of the risk - the mean value and the variance and is more sensible to higher risks.

7.1.4 Standard deviation principle

$$\Pi_X = EX + \alpha\sqrt{Var(X)},$$

where $\alpha > 0$. The loading is proportional to the standard deviation of X. The loss can be written as

$$\Pi_X - X = \sqrt{Var(X)}\left(\alpha - \frac{X - EX}{\sqrt{Var(X)}}\right),$$

or the loss is equal to the loading parameter minus a random variable with mean value 0 and variance 1.

7.1.5 Modified Variance Principle

In the Variance Principle, the changing in monetary unite changes the security loading. The following modification

$$\Pi_X = \begin{cases} EX + \alpha\frac{Var(X)}{EX}, & EX > 0 \\ 0, & EX = 0 \end{cases}$$

for $\alpha > 0$ can change this.

7.1.6 The Principle of Zero Utility

The worst thing that may happen for the company is a very high accumulated sum of claims. Therefore high losses should be weighted stronger then small losses. Hence, the company chooses an *utility function* v, which should have the following properties:

1) $v(0) = 0$.

2) $v(x)$ is strictly increasing.

3) $v(x)$ is strictly concave.

The first property is for convenience. The second means that less losses are preferred. The last condition gives stronger weights for higher losses. The premium is defined by the equation

$$v(u) = E[v(u + \Pi_X - X)], \tag{7.1}$$

where u is the insurer's surplus. This means that the expected utility is the same whether the insurance contract is taken or not. In general, the premium depends on the surplus.

Lemma 7.1 *1. If the solution of (7.1) exists, it is unique.*

2. If for every $x < \Pi$, $v''(x) < 0$, then $\Pi > EX$;

3. The premium is independent of u if and only if $v''(x) = 0$ or $v(x) = A(1 - e^{-\alpha x})$, $A > 0$, $\alpha > 0$.

Proof.

1. Let $\Pi_1 > \Pi_2$ be two solutions of (7.1). Since $v'(x) > 0$, then

$$v(u) = E[v(u + \Pi_1 - X)] > E[v(u + \Pi_2 - X)] = v(u),$$

which is a contradiction.

2. The Jensen's inequality is fulfilled

$$v(u) = E[v(u + \Pi - X)] < v(u + \Pi - EX)$$

and since $v'(x) > 0$, we obtain $u + \Pi - EX > u$.

3. It is easy to see that if $v''(x) = 0$ or $v(x) = A(1 - e^{-\alpha x})$, then the premium principle is independent of u.

Suppose that the premium is independent of u. Let $P(X = 1) = 1 - P(X = 0) = q$ and $\Pi(q)$ is the premium. Then

$$qv(u + \Pi(q) - 1) + (1 - q)v(u + \Pi(q)) = v(u). \tag{7.2}$$

Differentiation in (7.2) relative to u leads to

$$qv'(u + \Pi(q) - 1) + (1 - q)v'(u + \Pi(q)) = v'(u). \tag{7.3}$$

The derivative of (7.2) relative to q is

$$v(u+\Pi(q)-1)-v(u+\Pi(q))+\Pi'(q)(qv'(u+\Pi(q)-1)+(1-q)v'(u+\Pi(q)) = 0. \tag{7.4}$$

Insert (7.4) in (7.2) and obtain

$$v(u + \Pi(q) - 1) - v(u + \Pi(q)) + \Pi'(q)v'(u) = 0. \tag{7.5}$$

Note that $\Pi'(q) > 0$. The derivative of (7.5) relative to u is

$$v'(u + \Pi(q) - 1) - v'(u + \Pi(q)) + \Pi'(q)v''(u) = 0,$$

and relative to q

$$\begin{aligned} \Pi'(q) \quad & (v'(u + \Pi(q) - 1) - v'(u + \Pi(q))) + \Pi''(q)v'(u) \\ & = -[\Pi'(q)]^2 v''(u) + \Pi''(q)))v'(u) = 0. \end{aligned}$$

Consequently $\Pi''(q) \leq 0$ and for $\alpha \geq 0$,

$$\frac{v''(u)}{v'(u)} = -\alpha$$

which proves the statement.

$\square$

Remark 7.1 *In the case of exponential utility function from the equation (7.1) we obtain*

$$\Pi_X = \alpha^{-1} \log Ee^{\alpha X} \tag{7.6}$$

and it is called an exponential principle.

7.1.7 The Esscher Principle

$$\Pi_X = \frac{E[Xe^{hX}]}{E[e^{hX}]},$$

where $h > 0$.

The Esscher premium can be interpreted as a pure premium for Y, related to X as follows. Let X be a continuous nonnegative random variable with density function f. Define the function g, such that

$$g(x) = \frac{e^{hX} f(x)}{\int_0^\infty e^{hx} f(x)dx}. \tag{7.7}$$

The function g, defined by (7.7) is the density of the random variable Y. The distribution function of Y is given by

$$G(x) = \frac{\int_0^t e^{hy} f(y)dy}{M_X(h)}$$

and is called *Esscher transform* for the function F with parameter h. $G(x)$ is related to the distribution function of the risk X, but does give more weight to larger losses. From the equation (7.7), the MGF of Y is

$$M_Y(t) = \frac{M_X(t+h)}{M_X(h)}.$$

Example 7.1 *Let $F(x) = 1 - \exp(-\lambda x)$, $x \geq 0$ be the distribution function of the random variable X. Find the Esscher transform with parameter $h < \lambda$. The MGF is given by $M_X(t) = \frac{\lambda}{\lambda - t}$, and hence*

$$M_Y(t) = \frac{M_X(t+h)}{M_X(h)} = \frac{\lambda - h}{\lambda - h - t}.$$

The Esscher transform of the function F is

$$G(x) = 1 - e^{-(\lambda - h)x}.$$

The density g is an weighted version of f. From (7.7) it follows that $g(x) = w(x)f(x)$, where $w(x) = \frac{e^{hx}}{M_X(h)}$. Since $h > 0$, then the weights increase when x increases. This means that the transforms are useful for the heavy tailed distributions. The expected value of Y is

$$EY = \frac{\int_0^\infty xe^{hx} f(x)dx}{\int_0^\infty e^{hx} f(x)dx} = \frac{E[Xe^{hX}]}{E[e^{hX}]} = \Pi_X.$$

Example 7.2 *Let* $X \sim \exp(1)$. *Find the premium by the Esscher principle with parameter* $h < 1$.

7.1.8 Risk adjusted premium principle

$$\Pi_X = \int_0^\infty [1 - F(x)]^{\frac{1}{\rho}} dx,$$

where $\rho \geq 1$ is called a *risk index.*

This principle is defined for nonnegative random variable X with distribution function F.

Let Z be a random variable with distribution function H, defined by the equation

$$1 - H(x) = [1 - F(x)]^{\frac{1}{\rho}}.$$

The expected value of Z :

$$EZ = \int_0^\infty [1 - H(x)] dx,$$

and hence $\Pi_X = EZ$.

Example 7.3 *Let* $X \sim \exp(\frac{1}{\lambda})$. *Find the risk adjusted premium* Π_X.
Here

$$1 - F(x) = e^{-\lambda x}$$

and

$$1 - H(x) = e^{-\frac{\lambda x}{\rho}}.$$

Consequently $Z \sim \exp(\frac{\rho}{\lambda})$ *and* $\Pi_X = \frac{\rho}{\lambda}$.

Example 7.4 *Let* $X \sim Par(\alpha, \lambda)$. *Find the risk adjusted premium* Π_X.
In this case

$$1 - F(x) = \left(\frac{\lambda}{\lambda + x} \right)^\alpha$$

and

$$1 - H(x) = \left(\frac{\lambda}{\lambda + x} \right)^{\frac{\alpha}{\rho}}.$$

So $Z \sim Par(\frac{\alpha}{\rho}, \lambda)$ *and* $\Pi_X = \frac{\rho\lambda}{\alpha - \rho}, \quad \rho < \alpha$.

If X is a continuous random variable with density function f, then the density of Z is h, defined by

$$h(x) = \frac{1}{\rho}[1 - F(x)]^{\frac{1}{\rho}-1} f(x). \tag{7.8}$$

This means that the density of Z is an weighted version of f.

Chapter 8

Diffusion Approximation

Consider the surplus process

$$U(t) = u + ct - S_t, \quad t \geq 0,$$

where $S_t = Z_1 + \ldots + Z_{N(t)}$ is the accumulated loss process. $N(t)$ is a homogeneous Poisson process with intensity λ and $S_t = 0$ for $N(t) = 0$. The individual losses $Z_1, Z_2, \ldots$ are positive, independent, identically distributed random variables, independent of $N(t)$. Suppose that $M_Z(r)$ exists.

Here we will give an approximation of the surplus process by a Wiener process with trend, called a generalized Wiener process.

The process $U(t)$ increases continuously with slot c, which is the premium per unit time and in random times $\sigma_1, \sigma_2, \ldots$ has jumps equal to $Z_1, Z_2, \ldots$.

Remember that the risk process $X(t) = ct - S_t, \quad t \geq 0$ satisfies the properties:

1. $X(0) = 0$;
2. $EX(t) = ct - \lambda tEZ$;
3. $Var(X(t)) = \lambda tEZ^2$.

The goal is to construct a continuous time process with the same properties.

Remember the definition of the Wiener process.

Definition 8.1 *Continuous time stochastic process* $\{W_t, \ t \geq 0\}$ *is called a Wiener process, if*

1. $W_0 = 0$;
2. $\{W_t,\ t \geq 0\}$ *has stationary independent increments;*
3. *Foe every* $t > 0,\quad W_t \sim N(0, \sigma^2 t)$, *where* $\sigma > 0$ *is a constant.*

Definition 8.2 *The continuous time stochastic process* $\{W_t,\ t \geq 0\}$ *is called a generalized Wiener process (Wiener process with drift), if it is a Wiener process, with* $EW_t = \mu t$, *i. e.* $W_t \sim N(\mu t, \sigma^2 t)$. *This process is called also a diffusion process.*

We shall prove that the surplus process $\{U(t),\ t \geq 0\}$ can be approximated by a Generalized Wiener process. Consider the limit of the surplus process $U(t)$, conditioning on large expected number of jumps with small sizes. Suppose that the expected values and the variances of the processes are the same. Under these conditions, the surplus process with Poisson counting process can be approximated by a generalized Wiener process.

Let

$$\mu = c - \lambda EZ$$

and

$$\sigma^2 = \lambda EZ^2$$

be the mean value and the variance of the generalized Wiener process. Then

$$c = \mu + \sigma^2 \frac{EZ}{EZ^2}.$$

Suppose that the claim size Z is given by $Z = \alpha Y$, for some random variable Y with arbitrary mean and variance. Then

$$\lambda = \frac{\sigma^2}{EY^2} \frac{1}{\alpha^2}$$

and

$$c = \mu + \sigma^2 \frac{EY}{EY^2} \frac{1}{\alpha}.$$

Let $\alpha \to 0$, then $\lambda \to \infty$. Since the processes $\{S_t,\ t \geq 0\},\ U(t)$ and $X(t),\ t \geq 0$ have stationary independent increments, then the limiting process is the same. $X(0) = 0$, consequently we have to prove only that for every t, the limit of $X(t)$ is normally distributed with parameters μt and $\sigma^2 t$.

Let

$$M_{X(t)}(r) = Ee^{r(ct-S_t)} = e^{[rc+\lambda(M_Z(-r)-1)]t}$$

be the MGF. Then

$$\begin{aligned}
\frac{\log M_{X(t)}(r)}{t} &= rc + \lambda[M_Z(-r) - 1] \\
&= r[\mu + \lambda EZ] + \lambda[1 - rEZ + \tfrac{r^2}{2!}EZ^2 - \tfrac{r^3}{3!}EZ^3 + \ldots - 1] \\
&= r\mu + \tfrac{r^2}{2}\lambda EZ^2 - \lambda[\tfrac{r^3}{3!}EZ^3 - \tfrac{r^4}{4!}EZ^4 + \ldots] \\
&= r\mu + \tfrac{r^2}{2}\sigma^2 - \lambda\alpha^2[\alpha\tfrac{r^3}{3!}EY^3 - \alpha^2\tfrac{r^4}{4!}EY^4 + \ldots].
\end{aligned}$$

For $\alpha \to 0$ we obtain

$$\lim_{\alpha\to 0} \frac{\log M_{X(t)}(r)}{t} = r\mu + \frac{r^2}{2}\sigma^2,$$

and then

$$\lim_{\alpha\to 0} M_{X(t)}(r) = e^{(r\mu+\frac{r^2}{2}\sigma^2)t}.$$

This is the MGF of $N(\mu t, \sigma^2 t)$ - distributed random variable, consequently the limiting process is a generalized Wiener process with mean value μt.

From the definition of the surplus process $U(t)$ it follows that the sample paths are differentiable everywhere except in the jump points. Since in the limiting case the number of points increases, the sample paths of the limiting process are not differentiable. Also, for $\alpha \to 0$, the jump's size tends to zero, consequently the sample paths of the limiting process are continuous with probability 1.

8.1 Ruin Probability for diffusion process

We proved that if W_t is $N(\mu t, \sigma^2 t)$ - distributed Wiener process, $U(t) = u + W_t$ is a risk process with initial capital $U(0) = u$. Consider the ruin probability in a finite time interval $(0, \tau)$ and let $\tau \to \infty$. The ruin probability up to time τ is given by

$$\psi(u,\tau) = 1 - \phi(u,\tau) = P(T_u < \tau) = P(\min_{0<t<\tau} U(t) < 0) = P(\min_{0<t<\tau} W_t < -u).$$

Theorem 8.1 *The ruin probability of the defined diffusion process is given by*

$$\psi(u,\tau) = \Phi\left(-\frac{u+\mu\tau}{\sqrt{\sigma^2\tau}}\right) + e^{-\frac{2\mu}{\sigma^2}u}\Phi\left(-\frac{u-\mu\tau}{\sqrt{\sigma^2\tau}}\right),$$

where Φ is the standard normal distribution function.

From this result, letting $\tau \to \infty$, we obtain

Corrolary 8.1 *The ultimate ruin probability is given by*

$$\psi(u) = 1 - \phi(u) = P(T_u < \infty) = e^{-\frac{2\mu}{\sigma^2}u}.$$

Corrolary 8.2 *The distribution of the time to ruin, given that ruin occurs is*

$$\frac{\psi(u,\tau)}{\psi(u)} = P(T_u < \tau | T_u < \infty) = e^{\frac{2\mu}{\sigma^2}u}\Phi\left(-\frac{u+\mu\tau}{\sqrt{\sigma^2\tau}}\right) + \Phi\left(-\frac{u-\mu\tau}{\sqrt{\sigma^2\tau}}\right), \quad \tau > 0. \tag{8.1}$$

Corrolary 8.3 *Differentiation in (8.1) relative to u gives the probability density function of the time to ruin*

$$f_{T_u}(\tau) = \frac{u}{\sigma\sqrt{2\pi}}\tau^{-\frac{3}{2}}e^{-\frac{(u-\mu\tau)^2}{2\sigma^2\tau}}, \quad \tau > 0. \tag{8.2}$$

Substituting $\frac{u^2}{\sigma^2} = \theta$ and $\frac{u}{\mu} = a$ in the density function we obtain

$$f_{T_u}(\tau) = \left(\frac{\theta}{2\pi\tau^3}\right)^{\frac{1}{2}} e^{-\frac{\theta}{2\tau}\left(\frac{\tau-a}{\sigma}\right)^2}, \quad \tau > 0,$$

which is the standard Inverse Gaussian distribution.

Hence, the time to ruin conditioning that ruin occurs has an Inverse Gaussian distribution with expected value $\frac{u}{\mu}$ and variance $\frac{u\sigma^2}{\mu^3}$. If $\mu = 0$ ruin occurs with probability 1 and the density function is obtain from (8.2) with $\mu = 0$, i. e.

$$f_{T_u}(\tau) = \frac{u}{\sigma\sqrt{2\pi}}\tau^{-\frac{3}{2}}e^{-\frac{u^2}{2\sigma^2\tau}}, \quad \tau > 0.$$

The distribution function is

$$F_{T_u}(\tau) = 2\Phi\left(-\frac{u}{\sigma\sqrt{\tau}}\right), \quad \tau > 0.$$

This is one - ided stable distribution with index $\frac{1}{2}$.

These results could be applied like approximation to the arbitrary risk process $U(t)$, with Poisson counting process. For the approximated risk process we obtain

$$\psi(u,\tau) = \Phi\left[-\frac{u+\theta\lambda\tau EZ}{\sqrt{\lambda\tau EZ^2}}\right] + e^{-\frac{2EZ}{EZ^2}\theta u}\Phi\left[-\frac{u-\theta\lambda\tau EZ}{\sqrt{\lambda\tau EZ^2}}\right], \quad u>0, \tau>0,$$

$$\psi(u) = e^{-\frac{2EZ}{EZ^2}\theta u}, \quad u>0$$

and

$$f_{T_u}(\tau) = \frac{u}{\sqrt{2\pi\lambda EZ^2}}\tau^{-\frac{3}{2}}e^{-\frac{(u-\theta\lambda\tau EZ)^2}{2\lambda\tau EZ^2}}, \quad \tau>0.$$

Here θ is the safety loading factor and $c = (1+\theta)\lambda EZ$.

Similarly, for a given risk process with Poisson counting process there exists a simple numerical approximation.

For example, the expected value of the time to ruin if ruin occurs is given by

$$ET_u = \frac{u}{\mu} = \frac{u}{\theta\lambda EZ}.$$

It is easy to see that it depends on four parameters. If the initial capital is large, the time to ruin increases. The increasing in all the other three parameters causes a decreasing of the time to ruin.

Chapter 9

Reinsurance

In many cases the premiums to the insurance company are not enough to carry the risk. This is the case of large claims. In that cases the insurer shares a part of the risk with other companies. Sharing the risk as well as the premiums is done by reinsurance contracts, which are mutual agreements between insurance companies. Sometimes the insurance companies have agreements about reinsuring certain parts of the portfolios.

We consider reinsurance that applies to the individual claims. If the claim size is z the insurer retains a part $h(z)$, where $h : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing function, such that $h(0) = 0$ and $0 \leq h(z) \leq z$ for all $z \geq 0$. The reinsurer covers the remain part $z - h(z)$. We assume that reinsurance premiums are payable continuously and that the reinsurer pays its share of a claim as soon as that claim occurs. The function $h(x)$ determines the rule of reinsurance. The aggregate sum of claims for the insurer is equal to $S_t^I = \sum_{i=1}^{N(t)} h(Z_i)$. The sum of claims for reinsurer is $S_t^R = S_t - S_t^I$.

9.1 Proportional Reinsurance

Suppose the insurer chooses proportional reinsurance with *retention level* $b \in [0, 1]$. In this case, the function h is $h(z) = bz$. The premium rate for the reinsurance is given by

$$(1 + \eta)(1 - b)\lambda\mu,$$

where $\eta > 0$ is the relative safety loading, defined by the reinsurance company. We consider the case $\eta > \theta$. The premium rate for the insurer is:

$$\lambda\mu[(1+\theta) - (1+\eta)(1-b)] = \lambda\mu[b(1+\eta) - (\eta-\theta)],$$

and the surplus process becomes

$$U(t,b) = u + \lambda\mu[b(1+\eta) - (\eta-\theta)]t - \sum_{k=1}^{N(t)} bZ_k. \tag{9.1}$$

In order that the net profit condition is fulfilled we need

$$\frac{\lambda\mu[b(1+\eta) - (\eta-\theta)]}{\lambda\mu b} > 1,$$

i. e.

$$b > 1 - \frac{\theta}{\eta}.$$

Let $M_Z(r)$ be the moment generating function of the individual claim amount distribution evaluated at r. Then the adjustment coefficient $R(b)$ under proportional reinsurance is the unique positive solution of the equation

$$\lambda[M_Z(br) - 1] - \lambda\mu[b(1+\eta) - (\eta-\theta)]r = 0. \tag{9.2}$$

Let $\Psi(u,b)$ denote the probability of ultimate ruin when the proportional reinsurance is chosen. Then

$$\Psi(u,b) = P(U(t,b) < 0 \text{ for some } \ t > 0).$$

Our objective is to find the retention level that minimizes $\Psi(u,b)$. According to the Lundberg inequality, the retention level will be optimal, if the corresponding Lundberg exponent R is maximal. We know that there is a unique $b \in [0,1]$ where the maximum is attained. If the maximizer $b > 1$, it follows from the uni-modality that the optimal b is 1, i. e. no reinsurance is chosen.

The next result gives the optimal retention level b and maximal adjustment coefficient $R(b)$, see [9].

Lemma 9.1 *The solution of equation (9.2) is given by*

$$R(b(r)) = \frac{(1+\eta)\mu r - [1 - M_Z(r)]}{(\eta - \theta)\mu}, \tag{9.3}$$

where $b \to r(b)$ is invertible.

Proof. Assume that $r(b) = bR((b))$, where $R(b)$ will be the maximal value of the adjustment coefficient and $r(b)$ is invertible. If we consider the function $r \to b(r)$, it follows that

$$b(r) = \frac{(\eta - \theta)\mu r}{(1+\eta)\mu r - [1 - M_Z(r)]}. \tag{9.4}$$

Now $R(b(r)) = \frac{r}{b(r)}$ in details is given by (9.3).

□

Theorem 9.1 *Assume that $M_Z(r) < \infty$. Suppose there is a unique solution r to*

$$M'_Z(r) - (1+\eta)\mu = 0. \tag{9.5}$$

Then $r > 0$, the maximal value of $R(b(r))$ and the retention level $b(r)$ are given by (9.3) and (9.4).

Proof. The necessary condition for maximizing the value of the adjustment coefficient is given by equation (9.5).

Since $R'(b(0)) = \frac{\eta}{\eta - \theta} > 0$, the function $R(b(r))$ is strictly increasing at 0. The second derivative in zero $R''(b(0)) = -\frac{1}{(\eta-\theta)\mu}EZ^2 < 0$ shows that $R(b(r))$ is strictly concave. Consequently, the function $R(b(r))$ has an unique maximum in r, which is the solution of (9.3). The retention level is given by (9.4).

□

Remark 9.1 *Note that the value of the adjustment coefficient does not depend on c but on the relative safety loadings only.*

9.2 Excess - of - Loss Reinsurance(XL)

The Excess - of - Loss reinsurance is nonproportional type of reinsurance. The insurer covers each individual claim up to a certain retention level M, i. e. when a claim of size Z occurs, the insurer pays $Z_M = \min(Z, M) = h(z)$ and the reinsurer $Z_R = Z - Z_M = \max(Z - M, 0) = (Z - M)_+$ so that $Z = Z_M + Z_R$. Suppose the number of claims $N(t)$ follows an ordinary renewal process. Hence the insurer risk process at time t is

$$X_M(t) = (c - c_M)t - \sum_{i=1}^{N(t)} \min(Z_i, M),$$

where c_M is the XL reinsurance premium. For a given M, the adjustment coefficient R_M is the unique positive root of

$$g_M(r) = 1,$$

if it exists with

$$g_M(r) = E[e^{rZ_M}]E[e^{-(c-c_M)rT}].$$

9.3 Stop - Loss Reinsurance

The Stop - Loss reinsurance works similarly to the XL reinsurance, but it covers the total amount of claims. For stop - loss contract with retention level (deductible) d, the amount paid by reinsurer to the insurer is

$$I_d = \begin{cases} 0, & if \quad S \leq d \\ S - d, & if \quad S > d. \end{cases}$$

Sometimes: $I_d = (S - d)_+$. Note that I_d as a function of the aggregate claims S is also a random variable. The amount of claims retained by the insurer is

$$\min(S, d) = S - I_d = \begin{cases} S, & if \quad S \leq d \\ d, & if \quad S > d. \end{cases}$$

Thus, the amount retained is bounded by d, which explains the name *stop - loss contract.*

The expected claims paid by the reinsurer is given by

$$EI_d = \int_d^\infty (x-d) f_S(x) dx \tag{9.6}$$

$$= ES - d + \int_0^d (d-x) f_S(x) dx \tag{9.7}$$

$$= \int_d^\infty [1 - F_S(x)] dx \tag{9.8}$$

$$= ES - \int_0^d [1 - F_S(x)] dx. \tag{9.9}$$

When ES is available, (9.7) and (9.9) are preferable by numerical integration. (9.8) and (9.9) hold for general distribution, including discrete and mixed. If the distribution is given, (9.6) is the most tractable formula.

Example 9.1 *Let $S \sim Gamma(\alpha, \beta)$. Then*

$$\begin{aligned} EI_d &= \int_d^\infty x f(x) dx - d[1 - F_S(d)] \\ &= \int_d^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} x^\alpha e^{-\beta x} dx - d[1 - G(d; \alpha, \beta)] \\ &= \frac{\alpha}{\beta}[1 - G(d; \alpha+1, \beta)] - d[1 - G(d; \alpha, \beta)], \end{aligned}$$

where $G(d; \alpha, \beta) = \dfrac{1}{\Gamma(\alpha)} \displaystyle\int_0^d x^{\alpha-1} e^{-\beta x} dx.$

Bibliography

[1] Asmussen S. (2000). *Ruin Probabilities*, World Scientific, Singapure.

[2] Bowers N.L., Gerber H.U., Hhickman J., Jones D. and Nesbit C. (1997). *Actuarial Mathematics, Society of Actuaries*, Itasca.

[3] Daykin C.D., Pentikämen T. and Pesonen M. (1994). *Practical Risk Theory for Actuaries*, Chapman & Hall, London.

[4] Embrechts P., Kluppelberg C. and Mikosh T. (1997). *Modelling extremal events*, Appl. Math. 33, Springer, Berlin.

[5] Galambos J. and Kotz S. (1978). *Characterizations of Probability Distributions*, Springer - Verlag. 2.5.1

[6] Gerber H.U. (1979). *An Introduction to Mathematical Risk Theory*, Richard D. Irwin, Homewood.

[7] Grandell, J. (1991). *Aspects of Risk Theory*, Springer - Verlag, New York.

[8] Grandell, J. (1997). *Mixed Poisson Processes*, Chapman & Hall, London.

[9] Hald M. and Schmidli H. (2004). On the maximisation of the adjustment coefficient under proportional reinsurance, *ASTIN Bulletin,* **34**, 75 - 83. 9.1

[10] Kaas R., Goovaerts M., Dhaene J. and Denuit M. (2001). *Modern Actuarial Risk Theory*, Klawer, Boston.

[11] Klugman S.A., Panjer H.H. and Willmot G.E. (1998). *Loss Models. From data to decisions*, Wiley.

[12] Johnson N.L., Kemp A. W. and Kotz S. (2005). *Univariate Discrete Distributions*, Third edition, John Wiley & Sons, Hoboken, New Jersey.

[13] Mikosh T. (2004). *Non - life insurance mathematics*, Springer.

[14] Minkova L. D. (2001). A Family of Compound Discrete Distributions, *J. Bulg. Acad. Sci.,* **54**, 11 - 14.

[15] Minkova L.D. (2002). A generalization of the classical discrete distributions, *Comm. Statist. Theory and Methods,* **31**, 871 - 888.

[16] Minkova L.D. (2004). The Polya - Aeppli process and ruin problems, *J. Appl. Math. Stoch. Analysis,* **3**, 221 - 234. 3.4.1

[17] Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*, John Wiley & Sons, Chichester.

[18] Ross S. M. (2007). *Introduction to Probability Models,* 9th edition, Elsevier Inc.

[19] Serfozo R. (2009). *Basic of Applied Stochastic Processes*, Springer. 3.2.3

[20] Willmot G. E. and Lin X. S. (2001). *Lundberg Approximations for Compound Distributions with Insurance Applications*, Springer.