UNIVERSITY OF TARTU

FACULTY OF SCIENCE AND TECHNOLOGY

INSTITUTE OF MATHEMATICS AND STATISTICS

KIRILL SMIRNOV

Approximation of Ruin Probability using Phase-Type Distributions

ACTUARIAL AND FINANCIAL ENGINEERING

MASTER'S THESIS (30 ECTS)

SUPERVISOR: PROF. KALEV PÄRNA

TARTU 2020

Laostumistõenäosuse lähendamine faas-tüüpi jaotuste abil

Magistritöö

Kirill Smirnov

Lühikokkuvõte. Käesoleva magistritöö eesmärk on leida laostumistõenäosuse lähend, mis on täpsem, kui hästituntud De Vylderi meetod, aga samal ajal on matemaatiliselt lihtne. Töös esitatud uus lähendusmeetod kaustab ära De Vylderi meetodi idee, kuid eksponetjaotuse asemel kasutab faas-tüüpi jotusi. Töö teoreetilises osas antakse ülevaade riskiprotsessidest, faas-tüüpi jaotustest ja De Vylderi meetodist ning tuletatakse valemid faasi-tüüpi lähendjaotuse parameetrite arvutamiseks. Töö praktilises osas võrreldakse kuue uudse lähendusmeetodi täpsust De Vylderi meetodi täpsusega. Võrdlemine toimub numbriliselt nelja erinevate riskiprotsessi põhjal ning tulemused näitavad uute meetodite suuremat täpsust võrreldes De Vylderi meetodiga.

CERCS teaduseriala: P160 Statistika, operatsioonanalüüs, programmeerimine, finants- ja kindlustusmatemaatika.

Märksõnad: Riskiteooria, laostumistõenäosus, faas-tüübi jaotus, risksprotsees, eksponentjaotus, De Vylderi meetod.

Approximation of Ruin Probability using Phase-Type Distributions

Master's thesis

Kirill Smirnov

Abstract. The purpose of this master's thesis is to find an approximation of ruin probabilities that is more accurate than well-known De Vylder's method, but at the same time is mathematically simple enough. This new approximation method is based on the idea of De Vylder's approximation, but instead of exponential distribution of claims some more complicated phase-type distributions are used. In theoretical part of the thesis an overview of main concepts of risk theory, the notion of phase-type distribution and De Vylder's approximation is given. In practical part accuracy of six approximations of ruin probability based on phase-type distributions are compared with De Vylder's method. The comparison is based on numerical examples of four different risk processes. According to the results, new methods are more accurate than De Vylder's approximation.

CERCS research specialisation: P160 Statistics, operation research, programming, actuarial mathematics.

Keyword: Risk theory, ruin probability, phase-type distribution, risk process, exponential distribution, De Vylder's approximation.

Contents

Introduction

The main question of classical risk theory is the calculation of the ruin probability of an insurance company. One of the factors affecting the probability of ruin is the claim distribution. There is a whole list of distributions fitting claim sizes depending on the type of insurance. Unfortunately, it is not possible to evaluate the exact formula of the ruin probability for most of claim distributions. One possibility to estimate the ruin probability, if the exact formula is not available, is using of approximations.

One of the most famous and successful approximations is De Vylder's approximation. The idea of this method is very simple. Assume that it is not possible to calculate the exact ruin probability for some distribution of claims. Than it is needed to replace initial distribution with well-fitting exponential one and calculate the ruin probability using the exact formula for exponentially distributed claims.

Exponential distribution is the simplest case of phase-type distributions and exact formula of the ruin probability for exponentially distributed claims can be expanded for all phase-type distributions. Hence, an idea has sparked to modify De Vylder's approximation by using phase-type distributions instead of exponential distribution. Intuitively, more complicated phase-type distribution will give more accurate estimation of the ruin probability. From here follows the main goal of the thesis: to find more accurate but at the same time mathematically simple enough approximation of the ruin probability based on phase-type distribution using the idea of De Vylder's method.

The main part of this thesis is divided into four chapters. In Chapter 1 we study the main basic concepts of the classical risk theory that are needed in the future research. Chapter 2 is devoted to De Vylder's approximation. Here is given a brief explanation of the method's idea and shown its application based on three numerical examples. In Chapter 3 we meet up with the notion of phase-type distributions, consider main special cases of this class of distributions and modify De Vylder's method using six different phase-type distributions. In the last chapter we apply received modifications of De Vylder's methods on the examples from the Chapter 2.

Most of the calculations in the thesis are done using software: RStudio 3.3.1; Maxima 5.42.2.

1 Main concepts and results of classical risk theory

In this chapter we will give a brief explanation of main concepts of classical risk theory, such as stochastic processes, ruin probability and classical risk model. This chapter is mainly based on [1] and [2].

1.1 The classical risk model

Let's consider the main cash-flows of an insurance company. The finance operations of insurer can be presented as a series of inflows and outflows (Figure 1). The main source of income for this business segment is selling of premiums. Also, insurer receive money through reinsurance recoveries, investments, etc. The main outflows are claims payout, reinsurance premiums, dividends paid to shareholders and bonuses paid to policyholders. The most important component of an insurance company's expenses is usually payout of claims.[5]

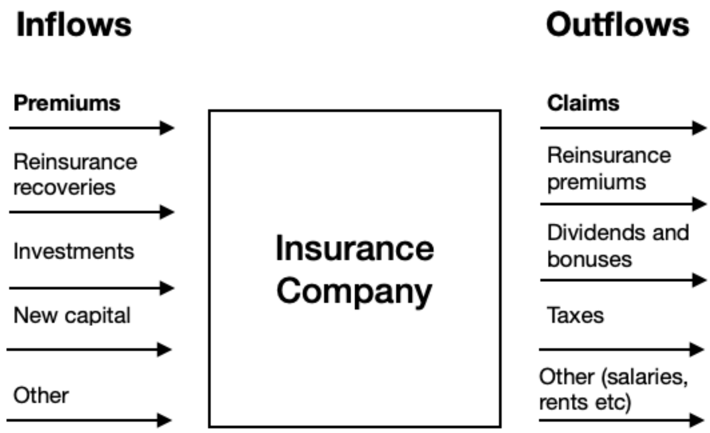

Figure 1. Main cash-flows of an insurance company

The number of claims arriving during the time interval and their sizes are usually unknown and can take different values i.e. they are stochastic. That is why the amount of outflows is changeable and at different moments of time can exceed the inflows amount or not.

From the changeability of the outflows follows the main classical question of the risk theory - to study the ruin probability of a company, i.e. the probability that company's balance will become negative at some point of time.

Definition 1.1. **Stochastic process** *is defined as a family of random variables* $\{X(t) : t \in T\}$*, where t is time parameter and* T *is the set of possible values of t. The set* T *can be discrete (*$T = \{1, 2, \dots\}$*) or continuous (*$T = [0, \infty)$*).*

Counting process is a special case of stochastic processes. Let us consider an event A that happens at random time points $S_1, S_2, \dots$ The number of occurrences of A within the time interval $[0, t]$ is called a counting process:

$$N(t) = \#\{i : S_i \in [0, t]\}.$$

So the number of claims $N(t)$ arriving within the time interval $[0, t]$ is a counting process. Now we can formulate the definition of the standard risk model.

Definition 1.2. **Risk process** *is a stochastic process defined as*

$$X(t) = ct - \sum_{k=1}^{N(t)} Z_k, \tag{1}$$

where

- c *- positive real constant meaning gross premium rate i.e. company receives c money units per time unit;*
- $N(t)$ *- counting process with* $N(0) = 0$*, interpreted as the number of claims arrived within the time interval* $(0, t]$*;*
- $\{Z_k\}_{k=1}^{\infty}$ *- sequence of independent and identically distributed random variables with mean value* μ*, and variance* σ^2*.* Z_k *means the size of* $k-$*th claim.*

This is the standard risk model of an insurance company, which is interpreted as follows. An insurer receives c money units per time unit and loses random amounts of money Z_1, $Z_2, \dots, Z_{N(t)}$ at time points $S_1, S_2, \dots, S_{N(t)} \in (0, t]$. Hence, risk process $X(t)$ means the profit of a company within the time period $(0, t]$.

The most famous special case of counting processes is Poisson process. Let's define waiting times T_i of the events as time difference between current and previous occurrence of the event:

$$T_i = S_i - S_{i-1}.$$

Definition 1.3. *A counting process $N(t)$ is called* **Poisson process** *if its waiting times T_1, T_2, $\ldots$ are independent random variables from the same exponential distribution with rate parameter α. The parameter α is called the intensity of the Poisson process.*

The classical risk theory is based on the Poisson process.

Definition 1.4. *Risk process $X(t)$ is called* **classical risk process** *if counting process $N(t)$ is Poisson process.*

Further in this thesis we consider only classical risk processes.

Assume that N has intensity α. It means, $E\left(N(t)\right) = \alpha t$. Hence, expected profit of the company within time period $(0, t]$ is

$$E\left(X(t)\right) = E(ct) - E\left(N(t)\right) \cdot E\left(Z_k\right)\right) = ct - \alpha\mu t = \left(c - \alpha\mu\right) t.$$

The ratio $\frac{c-\alpha\mu}{\alpha\mu}$ is called relative safety loading, denoted by ρ. If relative safety loading is positive ($\rho > 0$), then the risk process $X(t)$ has a drift to $+\infty$ and it is said that the company is profitable.

Usually, the company starts its activities having some starting capital u. Now we can define the main concept of classical risk theory - ruin probability - for an insurance company with the risk process $X(t)$ described by equation (1) and starting capital u.

Definition 1.5. **The ruin probability** *of an insurance company with initial capital u and risk process $X(t)$ is a probability, that at some time point $t > 0$ company's balance $u + X(t)$ will be negative.*

$$\Psi(u) = P\left\{u + X(t) < 0 \, \textit{for some} \; t > 0\right\}.$$

From here follows the definition of non-ruin probability $\Phi(u)$ which is defined as $1 - \Psi(u)$.

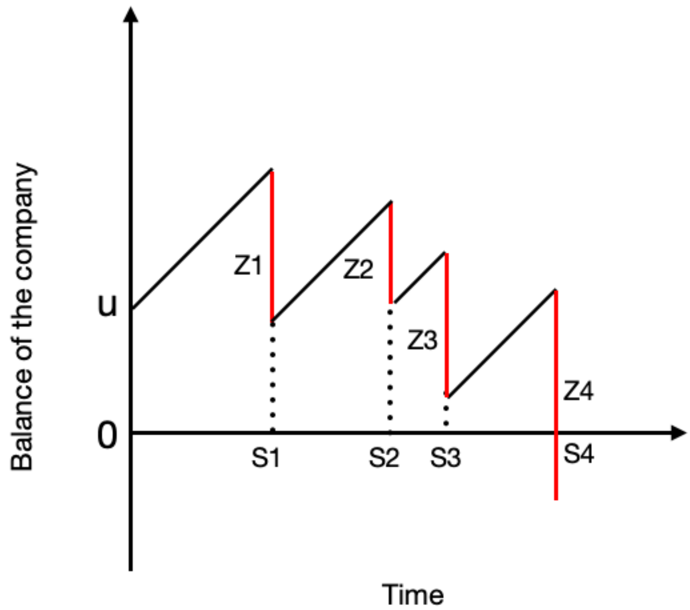

Figure 2. Illustration of a risk process's trajectory

The concept of the risk process is clearly illustrated in Figure 2. Starting with initial capital u at time $t = 0$ the company receives c money units per time unit. In this way company's balance is equal to $u + c \cdot S_1$ at the time point S_1, before the first claim's arrival. Paying out the claim company's balance decreases by Z_1. So the profit of the company within the time interval $(0, S_1]$ is equal to $c \cdot S_1 - Z_1$. After that the process repeats. At the time point S_4 the fourth claim Z_4 arrives. The size of claim is greater than the current balance of the company. Hence, paying out of this claim leads to reserve's drop below zero. It means that at the time S_4 accrues ruin of observed company.

1.2 Exact formula of the ruin probability for exponentially distributed claims

As it was mentioned in the previous section, the main classical question of the theory of risks is the calculation of the ruin probability. The ruin probability depends on the starting capital, the intensity of Poisson process and the distribution of claims. In this section we consider the risk process with exponentially distributed claims and derive the exact formula of the ruin probability for this case.

Theorem 1.1. *If claim sizes Z_k of the classical risk process $X(t)$ are exponentially distributed, i.e. $Z_k \sim Exp(\sigma)$, then the ruin probability $\Psi(u)$ can be calculated by the following formula:*

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

where $\mu = \frac{1}{\sigma}$ is mean value of exponential distribution and $\rho = \frac{c-\alpha\mu}{\alpha\mu}$ is relative safety loading.

In order to prove this theorem, we need to derive some other important results of the classical theory of risks.

Let's consider the classical risk process X for a company with an initial capital equal to u. Suppose that the first claim Z_1 arrives at the time moment S_1, then $X(S_1) = c \cdot S_1 - Z_1$. In this way, at time S_1 starts, so say, a "new" risk process X with an initial capital equal to $u + c \cdot S_1 - Z_1$.
Since we assume, that ruin can not happen within time interval $(0, S_1)$:

$$\Phi(u) = E\left(\Phi(u + c \cdot S_1 - Z_1)\right) = \int_0^\infty \int_0^\infty \Phi(u + cs - z) dF(z) dF_{S_1}(s),$$

where $F_{S_1}(s)$ and $F(z)$ are distribution functions of S_1 and Z_1 respectively.

As we consider classical risk process, S_1 is exponentially distributed with intensity α. Hence, $dF_{S_1}(s) = \alpha \cdot e^{-\alpha s} ds$.

$$\Phi(u) = \int_0^\infty \alpha \cdot e^{-\alpha s} \int_0^\infty \Phi(u + cs - z) dF(z) ds.$$

If the claim size is greater or equal to $u + cs$, occurs ruin of the company at time S_1. Assuming that, we get

$$\Phi(u) = \int_0^\infty \alpha \cdot e^{-\alpha s} \int_0^{u+cs} \Phi(u + cs - z) dF(z) ds.$$

Let's apply the change variable $x := u + cs$. It follows, that $ds = \frac{ds}{c}$.

$$\Phi(u) = \frac{\alpha}{c} \cdot e^{\frac{\alpha u}{c}} \int_u^\infty \alpha \cdot e^{-\alpha \cdot \frac{x}{c}} \int_0^x \Phi(x - z) dF(z) dx.$$

Differentiating of Φ by u we get

$$\Phi'(u) = \frac{\alpha^2}{c^2} \cdot e^{\frac{\alpha u}{c}} \int_u^{\infty} \alpha \cdot e^{-\alpha \cdot \frac{x}{c}} \int_0^x \Phi(x-z) dF(z) dx - \frac{\alpha}{c} \cdot e^{\frac{\alpha \cdot u}{c}} \cdot e^{-\frac{\alpha \cdot u}{c}} \int_0^u \Phi(u-z) dF(z).$$

The first term of the right side is equal to $\frac{\alpha}{c}\Phi(u)$. Then we get

$$\Phi'(u) = \frac{\alpha}{c}\Phi(u) - \frac{\alpha}{c}\int_0^u \Phi(u-z)dF(z). \tag{2}$$

Integration of the received equation over $(0, \hat{u})$ leads to

$$\begin{aligned}
\Phi(\hat{u}) - \Phi(0) = \frac{\alpha}{c}\int_0^{\hat{u}} \Phi(u)du + \frac{\alpha}{c}\int_0^{\hat{u}}\int_0^u \Phi(u-z)\cdot(1-F(z))du = \\
\frac{\alpha}{c}\int_0^{\hat{u}} \Phi(u)du + \\
\frac{\alpha}{c}\int_0^{\hat{u}} \left(\Phi(0)\cdot(1-F(u)) - \Phi(u) + \int_0^u \Phi'(u-z)d(1-F(z))dz\right) du = \\
\frac{\alpha}{c}\Phi(0)\int_0^{\hat{u}}(1-F(u))du + \frac{\alpha}{c}\int_0^{\hat{u}}(1-F(z))dz\int_z^{\hat{u}} \Phi'(u-z)du = \\
\frac{\alpha}{c}\Phi(0)\int_0^{\hat{u}}(1-F(u))du + \frac{\alpha}{c}\int_0^{\hat{u}}(1-F(z))dz(\Phi(\hat{u}-z) - \Phi(0))dz.
\end{aligned}$$

From here we get the final integral equation for non-ruin probability:

$$\Phi(u) = \Phi(0) + \frac{\alpha}{c}\int_0^u \Phi(u-z)\cdot(1-F(z))dz. \tag{3}$$

Assume now that initial capital tends to infinity ($u \to \infty$) and the company is profitable ($\rho > 0$). It is possible to show using Monotone Convergence Theorem that in this case equation 3 leads to

$$\Phi(\infty) = \Phi(0) + \frac{\alpha\mu}{c}\Phi(\infty). \tag{4}$$

If ρ is positive, then $\lim_{t\to\infty} X(t) = +\infty$ $a.s.$ Hence there is time T, such that for all time moments $t > T$ profit of the company is positive ($X(t) > 0$). It means that ruining of the company can not occur after the time T. Consider the time period $[0, T]$. Within this time interval arrives finite number of claims ($N(T)$ is Poisson process) and the size of each claim is finite too. These facts lead to conclusion that total sum of expenses

$(\sum_{k=1}^{N(T)} Z_k)$ is finite. Therefore, having infinite initial capital, company can not ruin in period $[0, T]$. We have shown that the ruin probability of the company in this case is equal to zero in time interval $[0, T]$ and for all $t > T$. From here follows that $\Phi(\infty) = 1$. Substituting this result into equation (4) we get

$$1 = \Phi(0) + \frac{\alpha\mu}{c}.$$

As $\Psi(u) = 1 - \Phi(u)$ and using the definition of relative safety loading ρ we get

$$\Psi(0) = \frac{\alpha\mu}{c} = \frac{1}{1+\rho}.$$

Now we have got all needed results to prove the theorem 1.1.

<u>Proof of the Theorem 1.1</u>:
Let's find the ruin probability of a company with an initial capital u, assuming that claims are exponentially distributed with the mean value μ.

Distribution function of exponential distribution $F(z)$ is $1 - e^{-\frac{z}{\mu}}$. It follows that $dF(z) = \frac{1}{\mu}e^{-\frac{z}{\mu}}dz$. Substituting this result into the equation 2 we get

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

Let's apply change variable $v := u - z$. It follows $dz = -dv$.

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

Differentiation and simplification of this equation leads to

$$\begin{aligned}\Phi''(u) = \frac{\alpha}{c}\Phi'(u) + \frac{1}{\mu}\left(\frac{\alpha}{\mu}\Phi(u) - \Phi'(u)\right) - \frac{\alpha}{c\mu}\Phi(u) = \\ = \left(\frac{\alpha}{c} - \frac{1}{\mu}\right) \cdot \Phi'(u) = -\frac{\rho}{\mu(1+\rho)} \cdot \Phi'(u).\end{aligned}$$

Now we need to solve received second order differential equation. Note that $\ln(\Phi'(u))' =$

$\frac{\Phi''(u)}{\Phi'(u)} = -\frac{\rho}{\mu(1+\rho)}$. Hence,

$$\begin{aligned}
\ln \Phi'(u) &= -\frac{\rho}{\mu(1+\rho)} \cdot u + C_1, \\
\Phi'(u) &= e^{C_1} \cdot \exp\left(-\frac{\rho}{\mu(1+\rho)} \cdot u\right) = C_2 \cdot \exp\left(-\frac{\rho}{\mu(1+\rho)} \cdot u\right), \\
\Phi(u) &= C_3 \cdot \exp\left(-\frac{\rho}{\mu(1+\rho)} \cdot u\right) + C_4.
\end{aligned}$$

We have proved above that $\Phi(\infty) = 1$ and $\Phi(0) = 1 - \frac{1}{1+\rho}$. Using these results we can find C_3 and C_4 values:

$$\begin{cases} 1 - \frac{1}{1+\rho} = C_3 \cdot \exp\left(-\frac{\rho}{\mu(1+\rho)} \cdot 0\right) + C_4 \\ 1 = C_3 \cdot \exp\left(-\frac{\rho}{\mu(1+\rho)} \cdot \infty\right) + C_4 \end{cases}.$$

From the second equation of the system follows:

$$\begin{aligned}
&1 = C_3 \cdot 0 + C_4, \\
&C_4 = 1.
\end{aligned}$$

Substituting C_4 into first equation of the system we get:

$$\begin{aligned}
1 - \frac{1}{1+\rho} &= C_3 \cdot 1 + 1, \\
C_3 &= -\frac{1}{1+\rho}.
\end{aligned}$$

Therefore,

$$\Psi(u) = 1 - \Phi(u) = \frac{1}{1+\rho} \exp\left(-\frac{\rho u}{\mu(1+\rho)}\right).$$

□

In many scientific articles it is customary to assign gross premium rate c equal to one for simplicity. It means that c is taken as money unit and all other quantities (including claims sizes) are measured in this units.[3] Assume that we have a risk process $X(t) = c \cdot t - \sum_{k=1}^{N(t)} Z_k$ with $c \neq 1$. We can define a new risk process $\hat{X}(t) = \frac{X(t)}{c}$ which has gross premium rate $\hat{c} = 1$:

$$\hat{X}(t) = \frac{ct}{c} - \sum_{k=1}^{N(t)} \frac{Z_k}{c} = 1 \cdot t - \sum_{k=1}^{N(t)} \frac{Z_k}{c}.$$

According to the definition of ruin probability is easy to show that $\psi(u) = \hat{\psi}(u)$. Thus, we can present the risk process with any gross premium rate, as a new risk process with $c = 1$, which has the same ruining probabilities as initial process. Using this property we can simplify the result of Theorem 1.1.

$$\Psi(u) = \frac{1}{1 + \frac{1-\alpha\mu}{\alpha\mu}} \cdot \exp\left(-\frac{\frac{1-\alpha\mu}{\alpha\mu} \cdot u}{\mu(1+\rho)}\right) = \alpha\mu \cdot \exp\left(-\left(\frac{1}{\mu} - \alpha\right) \cdot u\right). \tag{5}$$

Further in this thesis we assume that gross premium rate is equal to one, $c = 1$.

Example 1.1. *Let's consider an insurance company with initial capital $u = 0.1$ and gross premium rate $c = 1$. Assume that claims are exponentially distributed with mean value $\mu = 0.5$ and intensity of Poisson process $\alpha = 1$.*

In this case we can calculate the probability of ruining using equation (5).

$$\Psi(0.1) = 1 \cdot 0.5 \cdot \exp\left(-\left(\frac{1}{0.5} - 1\right) \cdot 0.1\right) = 0.5 \cdot \exp(-0.1) \approx 0.45242$$

Thus, starting with 0.1 money units on its balance the company will ruin with probability 0.45. Obviously, increasing initial capital of the company the ruin becomes less likely. This process is shown in Figure 3. For example, if the company increases its starting capital up to 1.9 money units, the ruining probability will be 0.0748.

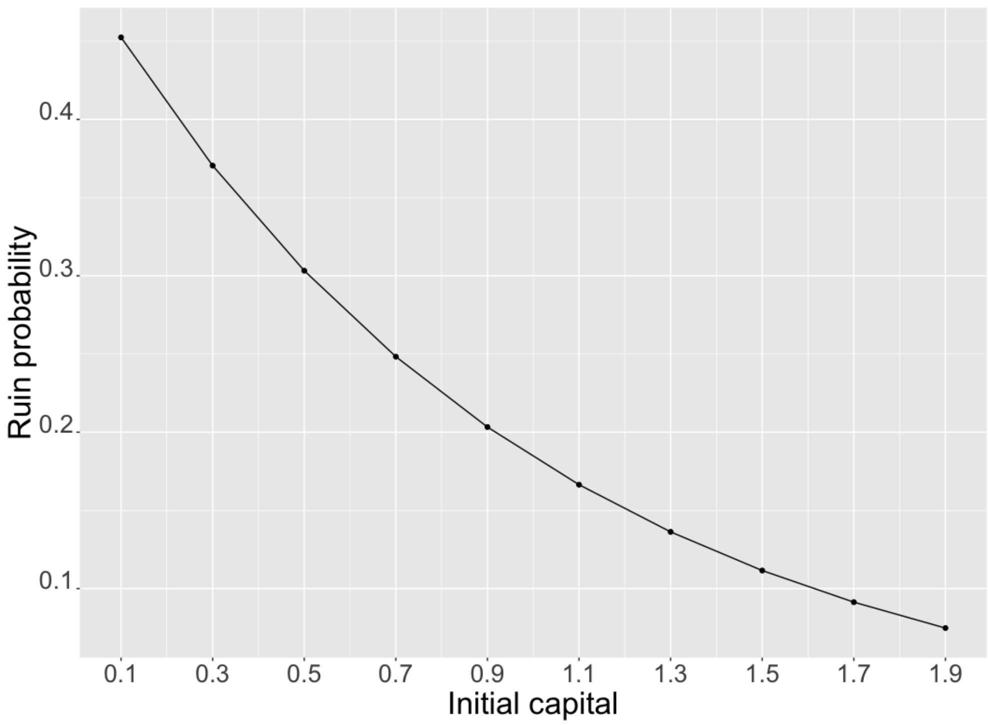

Figure 3. Ruin probability for different values of initial capital

The exact formula of the ruin probability in case of exponentially distributed claims is very simple and convenient in application. But the risk process with exponentially distributed claims is just one rare example of risk processes where an exact formula exists. For example, Cramer-Lundberg approximation (6) does not work in case of heavy-tailed distributions.

$$\lim_{u\to\infty} \Psi(u) \cdot e^{-Ru} = \frac{\rho u}{h'(R) - \frac{c}{\alpha}}, \tag{6}$$

where R is positive constant (Lundberg exponent) and $h(r) = \int_0^\infty e^{rz} dF(z) - 1$.

2 De Vylder's approximation for the ruin probabilities

As it was shown in the previous chapter, it is possible to calculate exact ruin probabilities of risk processes with exponentially distributed claims, occurring according to the Poisson process. But exponential distribution is not the only possible distribution used to describe sizes of claims. There is a whole list of distributions - that fit the data much better in different situations. For example, in non-life insurance exponential distribution is considered to be not enough realistic for using in real life models and preferences are given to other distributions, such as Gamma, Weibull and Lognormal.[5]

There are some other claim's distributions besides exponential one, where exact formula of the ruin probability can be derived, for example, class of phase-type distributions. But in most case of claim distributions the ruin probability can be calculated only via simulation process or using some approximations.

Several approximations to the ruin probability have been proposed. The most famous of them are Cramer-Lundberg's, Beekman-Bowers's and De Vylder's approximations. Futher we will focus on the De Vylder's method of finding of ruin probabilities. This chapter is mostly based on [1], [2] and [6].

2.1 Concept of De Vylder's method

De Vylder's approximation (1978), is the most successful and mathematically simple approximation of the ruin probability. Consider the risk process $X(t)$ having such distribution of claims, that it is not possible to calculate exact ruin probability of it. Assume that the intensity of Poisson process is α, and the gross premium rate is c. De Vylder's method is based on the idea to replace $X(t)$ with a new risk process $\hat{X}(t)$ having exponentially distributed claims, such that the first three moments of $X(t)$ match with the first three moments of $\hat{X}(t)$. It means

$$E[X^k(t)] = E[\hat{X}^k(t)] \quad \text{for } k = 1, 2, 3.$$

The number three comes from the fact that the risk process with exponentially distributed claims has three parameters: the premium gross rate $\hat{c}$, the rate of exponential distribution $\hat{\sigma}$ and the intensity of Poisson distribution $\hat{\alpha}$.

When parameters of $\hat{X}(t)$ are estimated, it is possible to calculate the ruin probability of this process, which is approximately equal to the ruin probability of $X(t)$

$$\Psi_{\hat{X}(t)}(u) \approx \Psi_{X(t)}(u).$$

To derive the first three moments of $X(t)$ the characteristic function is used.

Definition 2.1. [8] *For a scalar random variable X* **the characteristic function** *is defined as the expected value of e^{ivX}, where i is the imaginary unit, and $v \in \mathbb{R}$ is the argument of the characteristic function:*

$$\varphi(v) = E\left[e^{ivX}\right].$$

For simplicity of the calculations we take the logarithm of characteristic function (or characteristic exponent) of $X(t)$:

$$\begin{aligned}\log\left(E\left[e^{ivX(t)}\right]\right) &= \log\left(E\left[e^{iv\left(ct-\sum_{k=1}^{N(t)} Z_k\right)}\right]\right)\\ &= t\left(icv + \alpha\left(E\left[e^{-ivZ_k}\right] - 1\right)\right).\end{aligned}$$

According to the Taylor series $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Hence,

$$\begin{aligned}e^{-ivZ_k} &= \sum_{n=0}^{3} \frac{(-ivZ_k)^n}{n!} + o(v)^3\\ &= 1 + \frac{(-ivZ_k)}{1} + \frac{(-ivZ_k)^2}{2} + \frac{(-ivZ_k)^3}{6} + o(v)^3\\ &= 1 - ivZ_k - \frac{v^2 Z_k^2}{2} + \frac{iv^3 Z_k^3}{6} + o(v)^3.\end{aligned}$$

Hence,

$$E\left[e^{-ivZ_k}\right] = 1 - iv\zeta_1 - \frac{v^2\zeta_2}{2} + \frac{iv^3\zeta_3}{6} + o(v)^3,$$

where $\zeta_1, \zeta_2, \zeta_3$ are the first three moments of claim distribution Z_k respectively.

Then

$$\begin{aligned}&\log\left(E\left[e^{ivX(t)}\right]\right) =\\ &\quad = t\left(icv + \alpha\left(1 - iv\zeta_1 - \frac{v^2\zeta_2}{2} + \frac{iv^3\zeta_3}{6} + o(v)^3 - 1\right)\right) =\\ &\qquad = t\left(iv(c - \alpha\zeta_1) - \frac{v^2\zeta_2\alpha}{2} + \frac{v^3 i\alpha\zeta_3}{6} + o(v)^3\right).\end{aligned}$$

Taking exponent of both sizes of the last equation we get characteristic function of the risk process $X(t)$.

$$\begin{aligned}\varphi_{X(t)}(v) = E\left[e^{ivX(t)}\right] = \\ = \exp\left(t\left(iv(c-\alpha\zeta_1) - \frac{v^2\zeta_2\alpha}{2} + \frac{v^3 i\alpha\zeta_3}{6} + o(v)^3\right)\right).\end{aligned}$$

Property 2.1. [8] *If a random variable X has moments up to k-th order, then the characteristic function φ_X is k times continuously differentiable on the entire real line. In this case*

$$E\left(X^k\right) = i^{-k} \cdot \varphi^{(k)}(0).$$

Using Property 2.1 we can derive expressions of the first moments of $X(t)$:

$$\begin{aligned}E\left[X(t)\right] = i^{-1}\varphi'(0) = i^{-1}t\left(i(c-\alpha\zeta_1) - v\zeta_2\alpha + \frac{v^2 i\alpha\zeta_3}{2}\right)\varphi_{X(t)}(v)\bigg|_{v=0} = \\ = i^{-1} \cdot i \cdot t(c-\alpha\zeta_1)e^0 = t(c-\alpha\zeta_1).\end{aligned}$$

Same result was obtained in the section 1.1. The second and the third moments of $X(t)$ can be derived in the same way. As a result, we get

$$\begin{aligned}E\left[X(t)\right] =& t(c-\alpha\zeta_1),\\ E\left[X^2(t)\right] =& \alpha\zeta_2 t + (c-\alpha\zeta_1)^2 t^2,\\ E\left[X^3(t)\right] =& -\alpha\zeta_3 t + 3(c-\alpha\zeta_1)(\alpha\zeta_2)t^2 + (c-\alpha\zeta_1)^3 t^3.\end{aligned}$$

Since the claims of $\hat{X}(t)$ are exponentially distributed, the n-th moment of Z_k can be found as follow:

$$E\left[Z_k^n\right] = \frac{n\,!}{\hat{\sigma}^n}.$$

Hence,

$$\begin{aligned}\hat{\zeta}_1 =& E\left[Z_k\right] = \frac{1}{\hat{\sigma}},\\ \hat{\zeta}_2 =& E\left[Z_k^2\right] = \frac{2}{\hat{\sigma}^2},\\ \hat{\zeta}_3 =& E\left[Z_k^3\right] = \frac{6}{\hat{\sigma}^3}.\end{aligned}$$

Thus, to match the first three moments of $X(t)$ and $\hat{X}(t)$ parameters $\hat{\sigma}$, $\hat{c}$, $\hat{\alpha}$ must satisfy the next system of equations.

$$\begin{cases} c - \alpha\zeta_1 = \hat{c} - \hat{\alpha}\frac{1}{\hat{\sigma}} \\ \alpha\zeta_2 = 2\hat{\alpha}\frac{1}{\hat{\sigma}^2} \\ \alpha\zeta_3 = 6\hat{\alpha}\frac{1}{\hat{\sigma}^3} \end{cases} . \tag{7}$$

Dividing the third equation by the second one we can find the estimation of the parameter $\hat{\sigma}$:

$$\frac{\alpha\zeta_3}{\alpha\zeta_2} = \frac{6\hat{\alpha}\frac{1}{\hat{\sigma}^3}}{2\hat{\alpha}\frac{1}{\hat{\sigma}^2}},$$
$$\frac{\zeta_3}{\zeta_2} = \frac{3}{\hat{\sigma}},$$
$$\hat{\sigma} = \frac{3\zeta_2}{\zeta_3}.$$

Substitution received $\hat{\sigma}$ into the second equation of the system leads to the estimation of $\hat{\alpha}$:

$$\alpha\zeta_2 = 2\hat{\alpha}\frac{\zeta_3^2}{9\zeta_2^2},$$
$$\hat{\alpha} = \frac{9\zeta_2^3}{2\zeta_3^2}\alpha.$$

Assuming that $c = 1$ and substituting $\hat{\sigma}$ and $\hat{\alpha}$ into the first equation of the system we get the estimation of $\hat{c}$:

$$1 - \alpha\zeta_1 = \hat{c} - \frac{9\zeta_2^3}{2\zeta_3^2} \cdot \frac{\zeta_3}{3\zeta_2}\alpha,$$
$$1 - \alpha\zeta_1 = \hat{c} - \frac{3\zeta_2^2}{2\zeta_3}\alpha,$$
$$\hat{c} = \frac{3\zeta_2^2}{2\zeta_3}\alpha - \alpha\zeta_1 + 1.$$

Letting $\alpha^* := \frac{\hat{\alpha}}{\hat{c}}$ we can calculate an approximate ruin probability of a company with the risk process $X(t)$ by De Vylder's approximation using formula 5.

$$\Psi(u) \approx \frac{\alpha^*}{\hat{\sigma}}e^{-(\hat{\sigma}-\alpha^*)\cdot u}.$$

2.2 Application of De Vylder's approximation

De Vylder's method for estimation of the ruin probability is considered to be one of the most accurate approximations. Numerical calculations demonstrate that De Vylder's approximation outperforms other theoretically justified approximations of ruin probability like so called "diffusion approximation" and Beekman-Bower's approximation [2] (however, we do not discuss these methods in details in this thesis). In this section De Vylder's method is applied on three examples.

Example 2.1. [2] *Consider risk process $X(t)$ with gross premium rate $c = 1$, claim sizes are from Gamma distribution, $Z_k \sim Gamma\left(\alpha' = \frac{1}{100}, \beta' = \frac{1}{100}\right)$, and an intensity of the Poisson process is $\alpha = \frac{10}{11}$. Using De Vylder's approximation we will find the ruin probability of $X(t)$ in case of starting capital $u = 300, 600, \ldots, 3000$.*

First of all, let's the first three moments of observed Gamma distribution. The n-th moment of random variable $Y \sim Gamma(\alpha', \beta')$ can be found as follow

$$E\left(Y\right) = \frac{(\alpha' + n - 1) \cdot \dots \cdot \alpha'}{(\beta')^n}. \qquad [9]$$

Hence in our case

$$\begin{aligned}
\zeta_1 &= E\left(Z_k\right) = \frac{\alpha'}{\beta'} = 1, \\
\zeta_2 &= E\left(Z_k^2\right) = \frac{(\alpha' + 1) \cdot \alpha'}{(\beta')^2} = 101, \\
\zeta_3 &= E\left(Z_k^3\right) = \frac{(\alpha' + 2) \cdot (\alpha' + 1) \cdot \alpha'}{(\beta')^3} = 20301.
\end{aligned}$$

Now using relations derived above we can calculate estimations of parameters $\hat{c}$, $\hat{\alpha}$, $\hat{\sigma}$.

$$\begin{aligned}
\hat{c} &= \frac{3\zeta_2^2}{2\zeta_3} \cdot \alpha - \alpha\zeta_1 + 1 = 0.7761194, \\
\hat{\sigma} &= \frac{3\zeta_2}{\zeta_3} = 0.01492537, \\
\hat{\alpha} &= \frac{9\zeta_2^3}{2\zeta_3^2} \cdot \alpha = 0.01022702, \\
\alpha^* &:= \frac{\hat{\alpha}}{\hat{c}} = 0.01317712.
\end{aligned}$$

According to formula 5 the ruin probability of a company with the risk process $X(t)$ by De Vylder's approximation is

$$\Psi(u) \approx 0.8828671 \cdot e^{-0.001748252 \cdot u}.$$

Comparison of results obtained by De Vylder's approximation with exact ruin probability of a company with the risk process $X(t)$ is presented in Table 1.

Table 1. Accuracy of De Vylder's method in case of Gamma distributed claims (excerpt from Table 1 in [2]).

u	Exact $\Psi(u)$	$\Psi_{DV}(u)$	Relative error of DV
300	0.52114	0.52254	0.4%
600	0.30867	0.30927	0.3%
900	0.18287	0.18305	0.1%
1200	0.10834	0.10834	0.0%
1500	0.06418	0.06412	-0.1%
1800	0.03803	0.03795	-0.3%
2100	0.02253	0.02246	-0.4%
2400	0.01335	0.01329	-0.6%
2700	0.00791	0.00787	-0.8%
3000	0.000468	0.00466	-0.9%

From Table 1 we can see that De Vylder's approximation gives very accurate estimation of the probability of ruining of the company with the risk process $X(t)$ with Gamma distributed claims. In case of observed u values absolute relative errors of the estimation do not increase 0.9%.

Example 2.2. [2] *Consider risk process $X(t)$ with gross premium rate $c = 1$, relative safety loading $\rho = 0.05$ and claims' sizes are from mixed exponential distribution with distribution functions $F(z)$.*

$$F(z) = 1 - 0.0039793 \cdot e^{-0.014631 \cdot z} - 0.1078392 \cdot e^{-0.190206 \cdot z} - 0.8881815 \cdot e^{-5.514588 \cdot z}.$$

Using De Vylder's approximation we calculate the ruin probability of $X(t)$ if company's inintal capital $u = 10, 100, 1000$.

The n-th moment of random variable $Y \sim MixedExp$ with distribution function

$$F(z) = 1 - w_1 \cdot e^{-\sigma_1 \cdot z} - w_2 \cdot e^{-\sigma_2 \cdot z} - w_3 \cdot e^{-\sigma_3 \cdot z}$$

can be found as follow

$$E\left[Y^{n}\right]=n!\cdot\sum_{k=1}^{3}\frac{w_{k}}{\sigma_{k}^{n}}. \quad [10]$$

Hence in our case, the first three moments of claims' distribution Z_k are

$$\begin{aligned}
\zeta_1 &= E\left(Z_k\right)=\sum_{k=1}^{3}\frac{w_k}{\sigma_k}=1,\\
\zeta_2 &= E\left(Z_k^2\right)=2\cdot\sum_{k=1}^{3}\frac{w_k}{\sigma_k^2}=43.19817,\\
\zeta_3 &= E\left(Z_k^3\right)=6\cdot\sum_{k=1}^{3}\frac{w_k}{\sigma_k^3}=7717.235.
\end{aligned}$$

From the definition of the relative safety loading we calculate the intensity of Poisson process α of risk process $X(t)$:

$$\begin{aligned}
\rho &= \frac{c-\alpha\zeta_1}{\alpha\zeta_1}=\frac{1}{\alpha\zeta_1}-1,\\
\alpha &= \frac{c}{\zeta_1(\rho+1)}=0.9523831.
\end{aligned}$$

Using relations derived in section 2.1 we can calculate estimations of parameters $\hat{c}$, $\hat{\alpha}$, $\hat{\sigma}$.

$$\begin{aligned}
\hat{c} &= \frac{3\zeta_2^2}{2\zeta_3}\cdot\alpha-\alpha\zeta_1+1=0.3930586,\\
\hat{\sigma} &= \frac{3\zeta_2}{\zeta_3}=0.01679287,\\
\hat{\alpha} &= \frac{9\zeta_2^3}{2\zeta_3^2}\cdot\alpha=0.005800922,\\
\alpha^* &:= \frac{\hat{\alpha}}{\hat{c}}=0.01475841.
\end{aligned}$$

According to formula 5 the ruin probability of a company with the risk process $X(t)$ by De Vylder's approximation is

$$\Psi(u)\approx 0.87885\cdot e^{-0.002034456\cdot u}.$$

Comparison of results obtained by De Vylder's approximation with exact ruin probability of a company with the risk process $X(t)$ is presented in Table 2

Table 2. Accuracy of De Vylder's method in case of Mixed Exponential distribution of claims (excerpt from Table 2 in [2]).

u	Exact $\Psi(u)$	$\Psi_{DV}(u)$	Relative error
10	0.8897	0.86115	-3.21%
100	0.7144	0.71706	0.37%
1000	0.1149	0.11491	0.01%

From Table 2 we can see that De Vylder's approximation works well in case of mixed exponential distribution of claims if initial capital is big enough. In case of small values of u absolute relative error is much bigger.

Example 2.3. [2] *Consider risk process $X(t)$ with gross premium rate $c = 1$, relative safety loading $\rho = 0.05$ and claims' sizes are from lognormal distribution with variance $\sigma_L^2 = 3.24$ and mean value $\mu_L = -1.62$. Using De Vylder's approximation we will calculate the ruin probability of $X(t)$ if company's initial capital $u = 100, 1000$.*

The n-th moment of random variable $Y \sim LN(\mu_L, \sigma_L^2)$ can be found as follow

$$E\left[Y^n\right] = \exp\left(n \cdot \mu_L + \frac{1}{2} \cdot n^2 \cdot \sigma_L^2\right). \quad [2]$$

Hence in our case the first three moments of claims' sizes distribution are

$$\begin{aligned}
\zeta_1 &= E\left(Z_k\right) = \exp\left(\mu_L + \frac{1}{2} \cdot \sigma_L^2\right) = 1, \\
\zeta_2 &= E\left(Z_k^2\right) = \exp\left(2 \cdot \mu_L + 2 \cdot \sigma_L^2\right) = 25.53372, \\
\zeta_3 &= E\left(Z_k^3\right) = \left(3 \cdot \mu_L + \frac{9}{2} \cdot \sigma_L^2\right) = 16647.24.
\end{aligned}$$

Similarly to the mixed exponential distribution's example $\alpha = \frac{c}{\zeta_1(\rho+1)} = 0.9523831$. Using relations derived in section 2.1 we can calculate estimations of parameters $\hat{c}$, $\hat{\alpha}$, $\hat{\sigma}$.

$$
\begin{aligned}
\hat{c} &= \frac{3\zeta_2^2}{2\zeta_3} \cdot \alpha - \alpha\zeta_1 + 1 = 0.1035654, \\
\hat{\sigma} &= \frac{3\zeta_2}{\zeta_3} = 0.004601432, \\
\hat{\alpha} &= \frac{9\zeta_2^3}{2\zeta_3^2} \cdot \alpha = 0.0002574435, \\
\alpha^* : &= \frac{\hat{\alpha}}{\hat{c}} = 0.002485806.
\end{aligned}
$$

According to formula 5 the ruin probability of company with risk process $X(t)$ by De Vylder's approximation is

$$\Psi(u) \approx 0.5402243 \cdot e^{-0.002115627 \cdot u}.$$

Comparison of results obtained by De Vylder's approximation with exact ruin probability of company with risk process $X(t)$ is presented in Table 3.

Table 3. Accuracy of De Vylder's method in case of lognormally distributed claims (excerpt from Table 3 in [2]).

u	Exact $\Psi(u)$	$\Psi_{DV}(u)$	Relative error
100	0.55074	0.43721	-20.6%
1000	0.04199	0.06512	55.1%

In case of lognormally distributed claims De Vylder's approximation gives poor results. The reason is that lognormal distribution is heavy tailed distribution and exponentially decreasing approximations (suggested by Cramer-Lundberg approximation formula) can not fit it well. More precisely, in [2] p.23, right asymptotic of ruin probability for lognormal claims is described, which significantly differs from exponential asymptotic.

3 Approximation of the ruin probabilities using phase-type distributions: theoretical aspects

As it was shown in section 1.2, there is exact formula for calculating the ruin probability of a company if its claims are exponentially distributed. Exponential distribution is the simplest non-trivial example of phase-type distributions' class and formula 5 proved for exponential distribution can be extended for all phase-type distributions.[3]

Since that fact, an idea has sparked to modify De Vylder's approximation which is based on exponential distribution and to use instead of it one of phase-type distributions. Intuitively more complicated phase-type distribution will give even more accurate estimation of the ruin probability than usual exponential distribution.

In this chapter we will get to know the concept of phase-type distribution and modify De Vylder's method by using some cases of phase-type distributions. Theoretical background in this chapter is mainly based on [3] and [4].

3.1 Concept of phase-type distribution

The concept of phase-type distribution is based on the notion of Markov process which is a continuous-time version of Markov chain.

Definition 3.1. [12] *Consider continuous-time stochastic process* $\{X(t) : t \geq 0\}$ *on some countable state space* S*. Letting* $\mathcal{F}_{X(s)}$ *denote all the information pertaining to the history of* X *up to time* s *and letting* $j \in S$ *and* $s \leq t$*, we say that* $X(t)$ *satisfies* **Markov property***, if*

$$P\{X(t) = j|\mathcal{F}_{X(s)}\} = P\{X(t) = j|X(s)\}.$$

In other words, Markov property means that future outcome $X(t)$ depends on present outcome $X(s)$ but does not depend on the past path of stochastic process. It is said that continuous-time stochastic process is Markov process if it has Markov property.

Definition 3.2. *Markov process is called* **time homogeneous** *if for any* $s \leq t$ *and any state* $j \in S$

$$P\left\{X(t) = j|X(s)\right\} = P\left\{X(t-s) = j|X(0)\right\}.$$

Conditional probability $p_{i,j} = P\{\text{next state is } j|\text{ current state is } i\}$ is called transition probability from the state i to the state j. If probability that process will remain in the state i is equal to one ($p_{i,i} = 1$), then we say that state i is absorbing state, otherwise it is transient.

Another important notion is the transition matrix $\mathbf{T} = (t_{i,j})$, where $i, j \in S$. The element $t_{i,j}$ of $\mathbf{T}$ is parameter of exponential distribution which determines the time within which the Markov process reaches the state j starting from the state i.

Definition 3.3. *Consider time homogeneous Markov process* $\{\bar{J}\} = \{X(t) : t \geq 0\}$ *with* $n + 1$ *states, such that* n *states are transient and one state is absorbing. The distribution of the time within Markov process reaches its absorbing state* Δ *is called* $\mathbf{phase-type}$ $\mathbf{distribution}$.

The transition matrix Q of this Markov process $\{\bar{J}\}$ can be presented in block-partitioned form.

$$Q = \begin{pmatrix} \mathbf{T} & \mathrm{t} \\ 0 & 0 \end{pmatrix},$$

where $\mathbf{T}$ is transition matrix of n transient states and column vector $\mathbf{t} = -\mathbf{T} \cdot \mathbf{e}$ ($\mathbf{e}$ is $n \times 1$ column vector with all elements equal to one) is exit rate vector, i.e. the i-th component of $\mathbf{t}$ gives the intensity that state i is followed by the absorbing state Δ.

The distribution of probabilities that Markov process with n transient states starts from any concrete state is given by row vector $\mathbf{p} = \{p_1, p_2, \ldots, p_n\}$, such that $\sum_{i=1}^{n} p_i = 1$, where p_i means the probability that the process starts from the i-th state. The vector $\mathbf{p}$ is called the initial distribution.

The simplest special case of phase-type distribution is exponential distribution. Assume that Markov process has one transient state and one absorbing state Δ. Then $\mathbf{p} = \{p_1\} = 1$. Let transition rate from the transient state to the absorbing state is $t_1 := \lambda$. Hence, intensity that process stays in the transient state is $t_{1,1} := -\lambda$. In this case the time within observed Markov process reaches its absorbing state Δ is exponentially distributed with rate parameter λ. Graphically this process is presented in Figure 4.

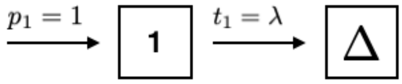

Figure 4. The phase-diagram of exponential distribution with rate parameter λ

.

Consider now the classical risk process $X(t)$. If claims' sizes have phase-type distribution, it is possible to calculate exact ruin probability of observed risk process.

Theorem 3.1. *Consider risk process $X(t)$ with $c = 1$, intensity of Poisson process α and phase-type distributed claims with transition matrix $\mathbf{T}$ and initial distribution $\mathbf{p}$. Exact ruin probability of $X(t)$ can be found as follow*

$$\Psi(u) = \mathbf{p}_+ e^{\mathbf{T}+\mathbf{t}\mathbf{p}_+}\mathbf{e},$$

where $\mathbf{p}_+ = -\alpha\mathbf{p}\mathbf{T}^{-1}$ and $\mathbf{e}$ is column vector with all elements equal to one.

Let's consider one numerical example to illustrate the process of calculation of the ruin probability of a company with phase-type distributed claims.

Example 3.1. *Suppose that company's main cash-flows can be described by risk process $X(t)$ with $c = 1$ and intensity of Poisson distribution $\alpha = 3$ and claims are phase-type distributed with parameters in Figure 5. The first moment of this distribution is equal to 0.265. Using the outcome of Theorem 3.1 let's find the ruin probability of the company if its initial capital is u.*

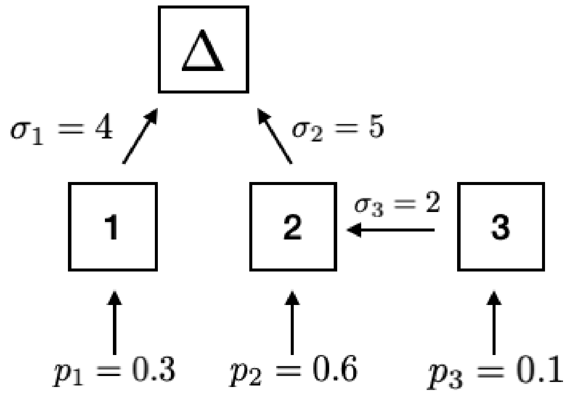

Figure 5. Markov chain generating a phase-type distribution

The distribution has three transient states. Hence, the transition matrix $\mathbf{T}$ has the following form.

$$\mathbf{T} = \begin{pmatrix} -\sigma_1 & 0 & 0 \\ 0 & -\sigma_2 & 0 \\ 0 & \sigma_3 & -\sigma_3 \end{pmatrix} = \begin{pmatrix} -4 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 2 & -2 \end{pmatrix}.$$

Therefor the exit rate vector $t = -\mathbf{Te}$ is

$$\mathbf{t} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix}.$$

From Figure 5 we can see that it is possible to reach absorbing state Δ directly form states 1 and 2 but not from the state 3.That is why the third element of vector t is equal to zero and the first two elements are σ_1 and σ_2 respectively.

Initial distribution of probabilities form which state the Markov process starts $\mathbf{p}$ is given by the vector

$$\mathbf{p} = \begin{pmatrix} 0.3 & 0.6 & 0.1 \end{pmatrix}.$$

Application of Theorem 3.1 requires positive relative safety loading ρ. Let's check if this condition fulfilled:

$$\rho = \frac{c - \alpha\mu}{\alpha\mu} = \frac{1 - 0.265 \cdot 3}{0.265 \cdot 3} = 0.258 > 0.$$

Let's now calculate the vector $\mathbf{p}_+ = -\alpha \mathbf{p} \mathbf{T}^{-1}$. Substitution of the values into the expression of $\mathbf{p}_+$ leads to

$$\mathbf{p}_+ = -3 \begin{pmatrix} 0.3 & 0.6 & 0.1 \end{pmatrix} \begin{pmatrix} -0.25 & 0 & 0 \\ 0 & -0.2 & 0 \\ 0 & -0.2 & -0.5 \end{pmatrix} = \begin{pmatrix} 0.225 & 0.42 & 0.15 \end{pmatrix}.$$

Next we calculate the matrix $\mathbf{T} + \mathbf{t}\mathbf{p}_+ =: \mathbf{Q}$ (needed in Theorem 3.1):

$$\mathbf{Q} = \begin{pmatrix} -4 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 2 & -2 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix} \begin{pmatrix} 0.225 & 0.42 & 0.15 \end{pmatrix} = \begin{pmatrix} -3.1 & 1.68 & 0.6 \\ 1.125 & -2.9 & 0.75 \\ 0 & 2 & -2 \end{pmatrix}$$

By Theorem 3.1 we need to find exponential of matrix $\mathbf{Q}u$.

Property 3.1. [7] *Exponential form of $n \times n$ matrix $\mathbf{Q}u$ can be found by the next equation*

$$e^{\mathbf{Q}u} = \phi(u) \cdot \phi(0)^{-1},$$

where $\phi(u)$ is $n \times n$ matrix which can be presented in block-partitioned form as follow

$$\phi(u) = \begin{pmatrix} v_1 e^{\lambda_1 u}, & v_2 e^{\lambda_2 u} & \ldots, & v_n e^{\lambda_n u} \\ & & & \end{pmatrix}, \tag{8}$$

where $\lambda_1, \ldots, \lambda_2$ are eigenvalues of matrix $\mathbf{Q}$ and $v_1, \ldots v_n$ are respective right eigenvectors.

In our case matrix Q (3×3) has three eigenvalues: $\lambda_1 = -4.479969$, $\lambda_2 = -2.885753$, $\lambda_3 = -0.634278$.

Respective right eigenvectors are $v_1 = \begin{pmatrix} 0.5593054 \\ -0.6452714 \\ 0.5203867 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0.5236681 \\ 0.3449783 \\ -0.7789491 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0.5050592 \\ 0.4867147 \\ 0.7127581 \end{pmatrix}$.

By its definition (8) matrix $\phi(u)$ has the following form.

$$\phi(u) = \begin{pmatrix} 0.5593054 e^{-4.479969u} & 0.5236681 e^{-2.885753u} & 0.5050592 e^{-0.634278u} \\ -0.6452714 e^{-4.479969u} & 0.3449783 e^{-2.885753u} & 0.4867147 e^{-0.634278u} \\ 0.5203867 e^{-4.479969u} & -0.7789491 e^{-2.885753u} & 0.7127581 e^{-0.634278u} \end{pmatrix}.$$

Hence, $\phi^{-1}(0)$ is

$$\phi^{-1}(0) = \begin{pmatrix} 0.7052368 & -0.8650708 & 0.09099351 \\ 0.8047468 & 0.1532572 & -0.67489563 \\ 0.3645851 & 0.7990803 & 0.59899548 \end{pmatrix}.$$

According to Property 3.1, we have

$$\begin{aligned} e^{Q\cdot t} =& e^{-4.479969t} \begin{pmatrix} 0.3944427 & -0.4838388 & 0.05089316 \\ -0.4550691 & 0.5582055 & -0.05871551 \\ 0.3669959 & -0.4501714 & 0.04735181 \end{pmatrix} + \\ &+ e^{-2.885753t} \begin{pmatrix} 0.4214202 & 0.08025592 & -0.3534213 \\ 0.2776202 & 0.05287042 & -0.2328244 \\ -0.6268568 & -0.11937957 & 0.5257093 \end{pmatrix} + \\ &+ e^{-0.634278t} \begin{pmatrix} 0.1841370 & 0.4035828 & 0.3025282 \\ 0.1774489 & 0.3889241 & 0.2915399 \\ 0.2598609 & 0.5695509 & 0.4269389 \end{pmatrix}. \end{aligned}$$

By Theorem 3.1

$$\Psi(u) = \mathbf{p}_{+} e^{\mathbf{T}+\mathbf{t}\mathbf{p}_{+}} \mathbf{e},$$

which in our case gives

$$\Psi(u) = 0.004620044 \cdot e^{-4.479969u} + 0.041298121 \cdot e^{-2.885753u} + 0.749081835 \cdot e^{-0.634278u}.$$

The dependence of the ruin probability on the value of initial capital u is visualised in Figure 6.

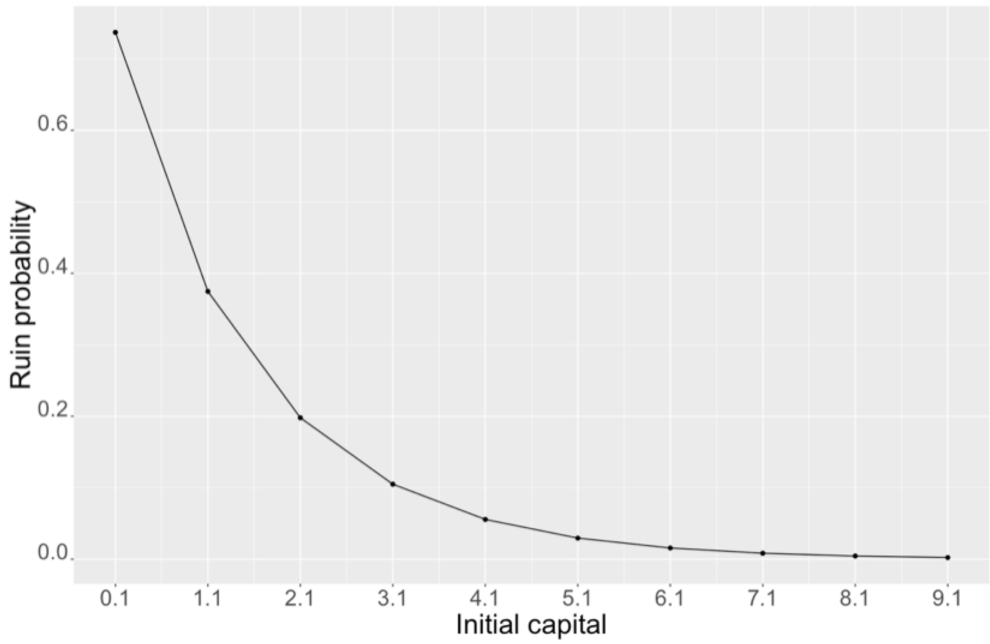

Figure 6. Ruin probability of phase-type distribution for different values of initial capital

As it was shown in Example 3.1, evaluating process of exact formula of the ruin probability for a risk process with phase-type distributed claims is not very difficult, but calculations are more complicated in comparison with exponentially distributed claims. Hence, in what follows, for the calculation of the ruin probability self-written R-script is used (Appendix 1).

An important property of phase-type distributions is the formula for the calculation of moments.

Property 3.2. *Consider the phase-type distribution B with the states space E, the transition matrix $\mathbf{T}$ and the initial distribution $\mathbf{p}$. Then n-th moment of B is $(-1)^n n\,!\mathbf{p}\mathbf{T}^{-n}\mathbf{e}$, where $\mathbf{e}$ is $n \times 1$ column vector with all elements equal to one.*

In the next sections we modify De Vylder's approximation by replacing exponential distribution with some special cases of phase-type distributions.

3.2 Erlang distribution

Erlang distribution E_p is a special case of Gamma distribution with parameter p meaning the number of phases. This corresponds to the convolution of p exponential densities with the same rate σ_1. The phase-diagram of E_p is presented in Figure 7.

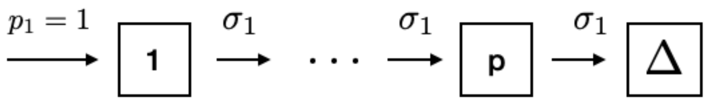

Figure 7. Phase-diagram of Erlang distrbution with p phases

Let's consider a risk process $X(t)$ with claim distribution for which it is not possible to calculate exact ruin probability. Using the idea of De Vylder's method we replace initial risk process $X(t)$ with a new risk process $\hat{X}(t)$ which has Erlang-distributed claims. Risk process with claims having E_p distribution (with fixed p) can be described by three parameters: gross premium rate c, intensity of Poisson process α and rate of Erlang distribution σ_1. It is sufficient to match the first three moments of $X(t)$ and $\hat{X}(t)$ (like in case of exponential distribution). In other words we need to solve the following system of three equations.

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot E[Z] \\ \alpha \cdot \zeta_2 = \hat{\alpha} \cdot E[Z^2] \\ \alpha \cdot \zeta_3 = \hat{\alpha} \cdot E[Z^3] \end{cases} \tag{9}$$

where Z is the size of claims having Erlang distribution with p phases and rate σ_1.

3.2.1 Erlang distribution with two phases

Firstly, assume that claims of the risk process $\hat{X}(t)$ have Erlang distribution with two phases (E_2). The transition matrix $\mathbf{T}$ and initial distribution vector $\mathbf{p}$ in case of E_2 have the following forms:

$$\mathbf{T} = \begin{pmatrix} -\sigma_1 & \sigma_1 \\ 0 & -\sigma_1 \end{pmatrix},$$

$$\mathbf{p} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

The first three moments of Erlang distribution with two phases and rate σ_1 can be found by Property 3.2.

$$\begin{aligned} E[Z] &= -1 \cdot \mathbf{pT^{-1}e}, \\ E[Z^2] &= 2 \cdot \mathbf{pT^{-2}e}, \\ E[Z^3] &= -6 \cdot \mathbf{pT^{-3}e}, \end{aligned}$$

where $\mathbf{e}$ is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Substitution of $\mathbf{p}$ and $\mathbf{T}$ of E_2 leads to

$$\begin{aligned} E[Z] &= \frac{2}{\sigma_1}, \\ E[Z^2] &= \frac{6}{\sigma_1^2}, \\ E[Z^3] &= \frac{24}{\sigma_1^3}. \end{aligned}$$

Hence, the system of equations (9) in case of E_2 is the following:

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot \frac{2}{\sigma_1} \\ \alpha \cdot \zeta_2 = \hat{\alpha} \cdot \frac{6}{\sigma_1^2} \\ \alpha \cdot \zeta_3 = \hat{\alpha} \cdot \frac{24}{\sigma_1^3} \end{cases}.$$

Solving this system of equations analogically to the system (7) in Section 2.1 we get

$$\begin{aligned} \hat{\sigma_1} &= \frac{4\zeta_2}{\zeta_3}, \\ \hat{\alpha} &= \frac{8\zeta_2^3}{3\zeta_3^2} \cdot \alpha, \\ \hat{c} &= \frac{4\zeta_2^2}{3\zeta_3} \cdot \alpha - \alpha\zeta_1 + 1. \end{aligned}$$

We will apply these formulas in Chapter 4.

3.2.2 Erlang distribution with three phases

In the same way, we can estimate parameters $\hat{c}$, $\hat{\alpha}$, $\hat{\sigma_1}$ of the risk process $\hat{X(t)}$ in case when the claims have E_3 distribution.

The transition matrix $\mathbf{T}$ and initial distribution $\mathbf{p}$ for the Erlang distribution with three phases are the following:

$$\mathbf{T} = \begin{pmatrix} -\sigma_1 & \sigma_1 & 0 \\ 0 & -\sigma_1 & \sigma_1 \\ 0 & 0 & -\sigma_1 \end{pmatrix},$$

$$\mathbf{p} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}.$$

Substituting this matrices into the outcome of Property 3.2 the first three moments of E_3 are

$$\begin{aligned} E[Z] &= \frac{3}{\sigma_1}, \\ E[Z^2] &= \frac{12}{\sigma_1^2}, \\ E[Z^3] &= \frac{60}{\sigma_1^3}. \end{aligned}$$

Hence, the estimation of parameters of risk process $\hat{X(t)}$ with E_3-distributed claims is possible to be found from the following system of three equations.

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot \frac{3}{\sigma_1} \\ \alpha \cdot \zeta_2 = \hat{\alpha} \cdot \frac{12}{\sigma_1^2} \\ \alpha \cdot \zeta_3 = \hat{\alpha} \cdot \frac{60}{\sigma_1^3} \end{cases}.$$

As a result, we get

$$\begin{aligned} \hat{\sigma_1} &= \frac{5\zeta_2}{\zeta_3}, \\ \hat{\alpha} &= \frac{25\zeta_2^3}{12\zeta_3^2} \cdot \alpha, \\ \hat{c} &= \frac{5\zeta_2^2}{4\zeta_3} \cdot \alpha - \alpha\zeta_1 + 1. \end{aligned}$$

We will apply these formulas in Chapter 4.

3.3 Hypoexponential distribution with two phases

A generalization of the Erlang distribution is so called hypoexponential distribution. Hypoexponential distribution with p phases is a convolution of p exponential densities with the intensity rates $\sigma_1,\ldots,\sigma_p$. This is graphically presented in a phase-diagram (Figure 8)

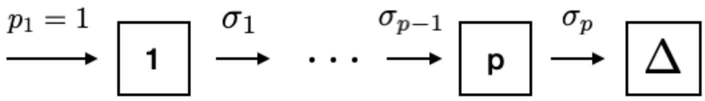

Figure 8. Phase-diagram of hypoexponential distribution with p phases

The number of parameters describing risk process $\hat{X}(t)$ with hypoexponentially distributed claims varies depending on the number of phases of the distribution. For example, if there is only one phase, $\hat{X}(t)$ can be described by three parameters, hence in this case hypoexponential distribution turns into simple exponential distribution. But if the number of states is equal to two, $\hat{X}(t)$ is described by four parameters: gross premium rate $\hat{c}$, intensity of Poisson process $\hat{\alpha}$, transition intensities from state "1" to state "2" and from state "2" to absorbing state, $\hat{\sigma}_1$ and $\hat{\sigma}_2$, respectively. Further we consider a risk process $\hat{X}(t)$ with two-phases hypoexponentially distributed claims.

Suppose now that the risk process $X(t)$ has a claim distribution, for which there is no exact formula for ruin probability. Analogically to previous section we use the idea of De Vylder's method to replace initial risk process $X(t)$ with a new risk process $\hat{X}(t)$ which has two-phases hypoexponentially distributed claims. As it was mentioned above, $\hat{X}(t)$ can be described by four parameters. So in this case, it is necessary to match the first four moments of $X(t)$ and $\hat{X}(t)$ to find the estimation $\hat{X}(t)$ parameters.

The first three moments of classical risk process are evaluated in Section 2.1. The fourth moment can be found analogically by Property 2.1(adding fourth term into Taylor series of e^{-ivZ_k}).

$$E[X(t)^4] = 2\alpha\zeta_4 t + (c-\alpha\zeta_1)^4 t^4 + 6\alpha(c-\alpha\zeta_1)^2\zeta_2 t^3 - 4\alpha(c-\alpha\zeta_1)\zeta_3 t^2 + 3\alpha^2\zeta_2^2 t^2.$$

Hence, to match the first four moments of $X(t)$ and $\hat{X}(t)$, it is sufficient to solve the following system of equations

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot E(Z) \\ \alpha \cdot \zeta_2 = \hat{\alpha} \cdot E(Z^2) \\ \alpha \cdot \zeta_3 = \hat{\alpha} \cdot E(Z^3) \\ \alpha \cdot \zeta_4 = \hat{\alpha} \cdot E(Z^4) \end{cases} . \tag{10}$$

The transition matrix $\mathbf{T}$ and initial distribution $\mathbf{p}$ of hypoexponential distribution with two phases are the following

$$\mathbf{T} = \begin{pmatrix} -\sigma_1 & \sigma_1 \\ 0 & -\sigma_2 \end{pmatrix},$$

$$\mathbf{p} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

By Property 3.2 the first four moments of this distribution are

$$\begin{aligned} E(Z) &= \frac{1}{\sigma_1} + \frac{1}{\sigma_2}, \\ E(Z^2) &= 2 \cdot \left(\frac{1}{\sigma_1\sigma_2} + \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right), \\ E(Z^3) &= 6 \cdot \left(\frac{1}{\sigma_1^2\sigma_2} + \frac{1}{\sigma_1\sigma_2^2} + \frac{1}{\sigma_1^3} + \frac{1}{\sigma_2^3} \right), \\ E(Z^4) &= 24 \cdot \left(\frac{1}{\sigma_1^3\sigma_2} + \frac{1}{\sigma_1^2\sigma_2^2} + \frac{1}{\sigma_1\sigma_2^3} + \frac{1}{\sigma_1^4} + \frac{1}{\sigma_2^4} \right). \end{aligned}$$

Substitution of the moments of hypoexponential distribution into the system (10) leads to

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot \left(\frac{1}{\sigma_1} + \frac{1}{\sigma_2} \right) \\ \alpha \cdot \zeta_2 = 2 \cdot \hat{\alpha} \cdot \left(\frac{1}{\sigma_1\sigma_2} + \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \\ \alpha \cdot \zeta_3 = 6 \cdot \hat{\alpha} \cdot \left(\frac{1}{\sigma_1^2\sigma_2} + \frac{1}{\sigma_1\sigma_2^2} + \frac{1}{\sigma_1^3} + \frac{1}{\sigma_2^3} \right) \\ \alpha \cdot \zeta_4 = 24 \cdot \hat{\alpha} \cdot \left(\frac{1}{\sigma_1^3\sigma_2} + \frac{1}{\sigma_1^2\sigma_2^2} + \frac{1}{\sigma_1\sigma_2^3} + \frac{1}{\sigma_1^4} + \frac{1}{\sigma_2^4} \right) \end{cases} .$$

It is not mathematically easy to solve this system manually. So, to estimate parameters of $\hat{X}(t)$ Maxima software was used (Appendix 2). The relations obtained between the parameters of $X(t)$ and $\hat{X}(t)$ are too long and complex, thus they are not presented in the text. We will apply obtained formulas in Chapter 4.

3.4 Hyperexponetial distribution with two phases

Another popular case of phase-type distributions is hyperexponential distribution H_p with p phases. It is defined as a mixture of p exponential distributions with rates $\sigma_1, \sigma_2, \ldots, \sigma_p$. The phase-diagram of H_p is presented in Figure 9. Weights $p_1, \ldots, p_p$ of hyperexponential distribution H_p are such that $\sum_{k=1}^{p} p_k = 1$. In the framework of this thesis we consider two-phases hyperexponential distribution.

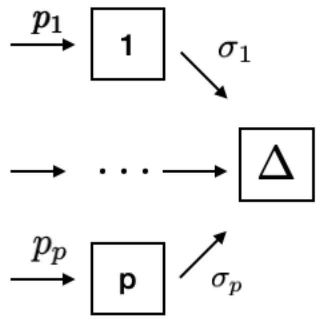

Figure 9. Phase-diagram of hyperexponential distrbution with p phases

Analogically to the previous section, let us assume that it is not possible to calculate exact ruin probability for the risk process $X(t)$. Our aim is to replace $X(t)$ with a suitable risk process $\hat{X}(t)$ which has two-phases hyperexponenntially distributed claims. Such risk process can be described by six parameters: weights $\hat{p_1}$ and $\hat{p_2}$, transition intensities $\hat{\sigma_1}$ and $\hat{\sigma_2}$, gross premium rate $\hat{c}$ and intensity of Poisson distribution $\hat{\alpha}$. Since the sum of the weights of hyperexponential distribution is always equal to one, it is sufficient to know only one of the weights. For example, if estimation of the first weight is $\hat{p_1}$, then estimation of the second one is $\hat{p_2} = 1 - \hat{p_1}$. Hence, it is sufficient to match the first five moments of $X(t)$ and $\hat{X}(t)$.

The fifth moment of classical risk process is evaluated by Property 2.1 (adding fifth term into Taylor series of e^{-ivZ_k}).

$$\begin{aligned} E[X(t)^5] &= (c - \alpha\zeta_1)^5 t^5 + 10a(c - \alpha\zeta_1)^3\zeta_2 t^4 - 10\alpha(c - \alpha\zeta_1)^2\zeta_3 t^3 \\ &\quad + 15\alpha^2(c - \alpha\zeta_1)\zeta_2^2 t^3 + 10\alpha(c - \alpha\zeta_1)\zeta_4 t^2 - 10\alpha^2\zeta_2\zeta_3 t^2 - \alpha\zeta_5 t. \end{aligned}$$

Hence, to match the first five moments of $X(t)$ and $\hat{X}(t)$ it is needed to solve the following system of equations

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot E(Z) \\ \alpha \cdot \zeta_2 = \hat{\alpha} \cdot E(Z^2) \\ \alpha \cdot \zeta_3 = \hat{\alpha} \cdot E(Z^3) \\ \alpha \cdot \zeta_4 = \hat{\alpha} \cdot E(Z^4) \\ \alpha \cdot \zeta_5 = \hat{\alpha} \cdot E(Z^5) \end{cases} \tag{11}$$

The transition matrix $\mathbf{T}$ and initial distribution $\mathbf{p}$ in case of H_2 distributions are

$$\mathbf{T} = \begin{pmatrix} -\sigma_1 & 0 \\ 0 & -\sigma_2 \end{pmatrix},$$

$$\mathbf{p} = \begin{pmatrix} p_1 & 1 - p_1 \end{pmatrix}.$$

By Property 3.2 the first five moments of hyperexponential distribution with two phases are the following:

$$E(Z) = \frac{p_1}{\sigma_1} + \frac{1 - p_1}{\sigma_2},$$

$$E(Z^2) = 2 \cdot \left(\frac{p_1}{\sigma_1^2} + \frac{1 - p_1}{\sigma_2^2} \right),$$

$$E(Z^3) = 6 \cdot \left(\frac{p_1}{\sigma_1^3} + \frac{1 - p_1}{\sigma_2^3} \right),$$

$$E(Z^4) = 24 \cdot \left(\frac{p_1}{\sigma_1^4} + \frac{1 - p_1}{\sigma_2^4} \right),$$

$$E(Z^5) = 120 \cdot \left(\frac{p_1}{\sigma_1^5} + \frac{1 - p_1}{\sigma_2^5} \right).$$

Substitution of the moments of hyperexponential distribution into the system (11) leads to

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot \left(\frac{p_1}{\sigma_1} + \frac{1-p_1}{\sigma_2} \right) \\ \alpha \cdot \zeta_2 = 2 \cdot \hat{\alpha} \cdot \left(\frac{p_1}{\sigma_1^2} + \frac{1-p_1}{\sigma_2^2} \right) \\ \alpha \cdot \zeta_3 = 6 \cdot \hat{\alpha} \cdot \left(\frac{p_1}{\sigma_1^3} + \frac{1-p_1}{\sigma_2^3} \right) \\ \alpha \cdot \zeta_4 = 24 \cdot \hat{\alpha} \cdot \left(\frac{p_1}{\sigma_1^4} + \frac{1-p_1}{\sigma_2^4} \right) \\ \alpha \cdot \zeta_5 = 120 \cdot \hat{\alpha} \cdot \left(\frac{p_1}{\sigma_1^5} + \frac{1-p_1}{\sigma_2^5} \right) \end{cases}.$$

This system was solved using Maxima software (Appendix 2). The relations obtained between the parameters of $X(t)$ and $\hat{X}(t)$ are too long and complex, thus they are not presented in the text. We will apply obtained formulas in Chapter 4.

3.5 Coxian distributions

Coxian distribution is very popular class of phase-type distributions in applied literature. This distribution has the following form of the phase-diagram (Figure 10).

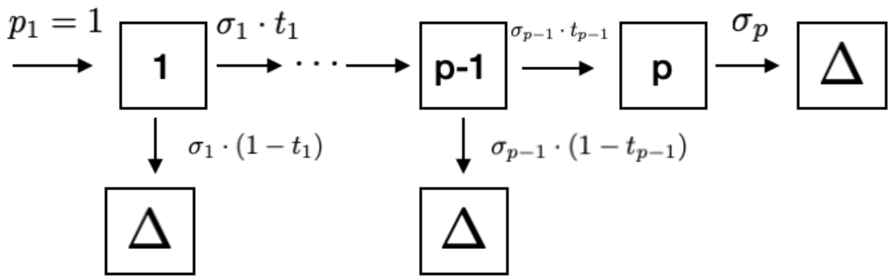

Figure 10. Phase-diagram of Coxian distrbution with p phases

Coxian distribution is a generalization of hypoexpoential distribution. Instead of only being able to reach the absorbing state from the final phase p it can be reached from any phase. Parameters $t_1, \ldots, t_{p-1} \in [0, 1]$. If all t_i are equal to one, then Coxian distribution is exactly hypoexponential distribution.

Analogically to the previous sections we calculate the ruin probability for risk process $X(t)$ by replacing initial risk process $X(t)$ with a new risk process $\hat{X}(t)$ with Coxian distribution of claims. In this section we will consider two cases of Coxian distribution.

3.5.1 Simplified two-phases Coxian distribution

First of all, let's assume that the risk process $\hat{X}(t)$ has two-phases Coxian distibution of claims, such that $\sigma_1 = \sigma_2$, i.e. its phase-diagam has the form presented in Figure 11.

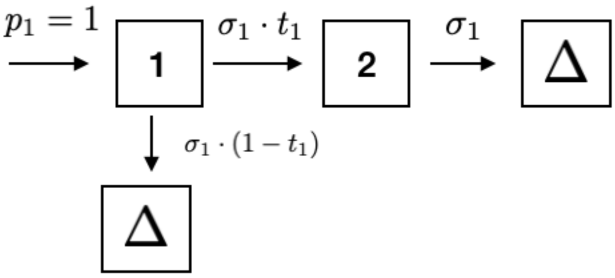

Figure 11. Phase-diagram of simplified Coxian distribution with two phases

This risk process can be described by four parameeters: transition intensity $\hat{\sigma_1}$ and t_1, gross premium rate $\hat{c}$, intensity of Poisson distribution $\hat{\alpha}$. Hence, in order to estimate all unknown parameters of this special case of Coxian distribution it is needed to match the first four moments of $X(t)$ and $\hat{X}(t)$.

The transition matrix $\mathbf{T}$ and initial distribution $\mathbf{p}$ in case of considered case of Coxian distributions are the following:

$$\mathbf{T} = \begin{pmatrix} -\sigma_1 & \sigma_1 \cdot t_1 \\ 0 & -\sigma_1 \end{pmatrix},$$

$$\mathbf{p} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

By Property 3.2 the first four moments of simplified Coxian distribution with two phases are the following:

$$E(Z) = \frac{t_1 + 1}{\sigma_1},$$
$$E(Z^2) = 2 \cdot \left(\frac{2t_1 + 1}{\sigma_1^2} \right),$$
$$E(Z^3) = 6 \cdot \left(\frac{3t_1 + 1}{\sigma_1^3} \right),$$
$$E(Z^4) = 24 \cdot \left(\frac{4t_1 + 1}{\sigma_1^4} \right).$$

Hence, in order to match the first four moments of initial risk process $X(t)$ and new risk process $\hat{X}(t)$ with Coxian distribution of claim, parameters of $\hat{X}(t)$ must satisfy the following system of equations:

$$\begin{cases} c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot \left(\frac{t_1+1}{\sigma_1}\right) \\ \alpha \cdot \zeta_2 = 2\hat{\alpha} \cdot \left(\frac{2t_1+1}{\sigma_1^2}\right) \\ \alpha \cdot \zeta_3 = 6\hat{\alpha} \cdot \left(\frac{3t_1+1}{\sigma_1^3}\right) \\ \alpha \cdot \zeta_4 = 24\hat{\alpha} \cdot \left(\frac{4t_1+1}{\sigma_1^4}\right) \end{cases} . \tag{12}$$

This system has two solutions:

$$\begin{aligned} \hat{t}_1 &= \frac{3 - \frac{9\zeta_2\zeta_4}{4\zeta_3^2} \pm \sqrt{1 - \frac{3\zeta_2\zeta_4}{4\zeta_3^2}}}{\frac{27\zeta_2\zeta_4}{4\zeta_3^2} - 8}, \\ \hat{\sigma}_1 &= \frac{4\zeta_3(1+4\hat{t}_1)}{\zeta_4(1+3\hat{t}_1)}, \\ \hat{\alpha} &= \frac{\zeta_3\hat{\sigma}_1^3}{6(1+3\hat{t}_1)} \cdot \alpha, \\ \hat{c} &= 1 - \alpha\zeta_1 + \hat{\alpha} \cdot \left(\frac{\hat{t}_1+1}{\hat{\sigma}_1}\right). \end{aligned}$$

We will apply these formulas in Chapter 4.

3.5.2 General two-phases Coxian distribution

Now let us assume that claims of the risk process $\hat{X}(t)$ have general Coxian distribution with two phases i.e. there is no assumption that $\sigma_1 = \sigma_2$. In this case $\hat{X}(t)$ has five unknown parameters: $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{t}_1$, $\hat{c}$, $\hat{\alpha}$. Hence, it is needed to match the first five moments of $X(t)$ and $\hat{X}(t)$ to find the estimations of unknown parameters.

The transition matrix $\mathbf{T}$ and initial distribution $\mathbf{p}$ for general Coxian distributions are defined as follow

$$\mathbf{T} = \begin{pmatrix} -\sigma_1 & \sigma_1 \cdot t_1 \\ 0 & -\sigma_2 \end{pmatrix},$$

$$\mathbf{p} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

By Property 3.2 the first four moments of simplified Coxian distribution with two phases are the following:

$$
\begin{aligned}
E(Z) &= \frac{t_1}{\sigma_2} + \frac{1}{\sigma_1}, \\
E(Z^2) &= 2 \cdot \left(\frac{t_1}{\sigma_1 \sigma_2} + \frac{t_1}{\sigma_2^2} + \frac{1}{\sigma_1^2} \right), \\
E(Z^3) &= 6 \cdot \left(\frac{t_1}{\sigma_1^2 \sigma_2} + \frac{t_1}{\sigma_1 \sigma_2^2} + \frac{t_1}{\sigma_2^3} + \frac{1}{\sigma_1^3} \right), \\
E(Z^4) &= 24 \cdot \left(\frac{t_1}{\sigma_1^3 \sigma_2} + \frac{t_1}{\sigma_1^2 \sigma_2^2} + \frac{t_1}{\sigma_1 \sigma_2^3} + \frac{t_1}{\sigma_2^4} + \frac{1}{\sigma_1^4} \right), \\
E(Z^4) &= 120 \cdot \left(\frac{t_1}{\sigma_1^4 \sigma_2} + \frac{t_1}{\sigma_1^3 \sigma_2^2} + \frac{t_1}{\sigma_1^2 \sigma_2^3} + \frac{t_1}{\sigma_1 \sigma_2^4} + \frac{t_1}{\sigma_2^5} + \frac{1}{\sigma_1^5} \right).
\end{aligned}
$$

The estimation of the parameters of $\hat{X}(t)$ can be found from the following system of equations.

$$
\begin{cases}
c - \alpha \cdot \zeta_1 = \hat{c} - \hat{\alpha} \cdot \left(\frac{t_1}{\sigma_2} + \frac{1}{\sigma_1} \right) \\
\alpha \cdot \zeta_2 = 2 \cdot \hat{\alpha} \cdot \left(\frac{t_1}{\sigma_1 \sigma_2} + \frac{t_1}{\sigma_2^2} + \frac{1}{\sigma_1^2} \right) \\
\alpha \cdot \zeta_3 = 6 \cdot \hat{\alpha} \cdot \left(\frac{t_1}{\sigma_1^2 \sigma_2} + \frac{t_1}{\sigma_1 \sigma_2^2} + \frac{t_1}{\sigma_2^3} + \frac{1}{\sigma_1^3} \right) \\
\alpha \cdot \zeta_4 = 24 \cdot \hat{\alpha} \cdot \left(\frac{t_1}{\sigma_1^3 \sigma_2} + \frac{t_1}{\sigma_1^2 \sigma_2^2} + \frac{t_1}{\sigma_1 \sigma_2^3} + \frac{t_1}{\sigma_2^4} + \frac{1}{\sigma_1^4} \right) \\
\alpha \cdot \zeta_5 = 120 \cdot \hat{\alpha} \cdot \left(\frac{t_1}{\sigma_1^4 \sigma_2} + \frac{t_1}{\sigma_1^3 \sigma_2^2} + \frac{t_1}{\sigma_1^2 \sigma_2^3} + \frac{t_1}{\sigma_1 \sigma_2^4} + \frac{t_1}{\sigma_2^5} + \frac{1}{\sigma_1^5} \right)
\end{cases}
$$

This system of equations was solved using Maxima software (Appendix 2). We will apply obtained formulas in Chapter 4.

4 Numerical comparison of De Vylder's approximation and phase-type approximations

The goal of this chapter is numerical comparison of the accuracy of ruin probability's approximations based on De Vylder's method and phase-type distributions considered in the previous chapter. In Section 2.2 De Vylder's approximation was applied on three examples: Gamma distribution (Example 2.1), mixed exponential distribution (Example 2.2) and lognormal distribution (Example 2.3). In this chapter we calculate ruin probabilities for the same examples using phase-type approximations and compare relative errors of all methods.

Sometimes claims of risk processes have complicated, high-order phase-type distributions with many states. It is not technically simple to calculate the ruin probability in such cases, even when exact formula is known. Hence, we got an idea to check if it is possible to estimate accurately the ruin probability in such cases using claims with a simple, low-order phase-type distributions. In order to research this question, one more numerical example is considered.

4.1 Gamma distribution

Consider the risk process $X(t)$ described in Example 2.1. It means claims are having Gamma distribution with parameters $\alpha' = \frac{1}{100}$ and $\beta' = \frac{1}{100}$, intensity of the Poisson process $\alpha = \frac{10}{11}$, relative safety loading $\rho = 5\%$ and gross premium rate c is assumed to be equal to 1.

As it was shown in Section 2.2 (Table 1), de Vylder's approximation works well in case of claims having Gamma distribution. The summary of absolute relative errors for all considered methods is presented in Figure 13. Here and in the next sections the following notations are used:

- De Vylder - De Vylder's approximation
- Erlang2 - approximation based on two-phases Erlang distribution, described in Section 3.2.1.
- Erlang3 - approximation based on three-phases Erlang distribution, described in Section 3.2.2.

- Hypo2 - approximation based on two-phases hypoexpontial distribution, described in Section 3.3.

- Hyper2 - approximation based on two-phases hyperexponential distribution, described in Section 3.4.

- Coxian1 - approximation based on simplified two-phases Coxian distribution, described in Section 3.5.1.

- Coxian2 - approximation based on general two-phases Coxian distribution, described in Section 3.5.2.

Table 4. Absolute relative errors (Gamma distributed claims)

u	De Vylder	Erlang2	Erlang3	Hyper2	Hypo2	Coxian1	Coxian2
300	0,268%	0,409%	0,491%	0,013%	0,599%	0,050%	0,013%
600	0,195%	0,288%	0,341%	0,000%	0,135%	0,014%	0,000%
900	0,097%	0,142%	0,167%	0,001%	0,018%	0,004%	0,001%
1200	0,000%	0,003%	0,007%	0,001%	0,007%	0,003%	0,001%
1500	0,089%	0,140%	0,172%	0,007%	0,013%	0,007%	0,007%
1800	0,205%	0,304%	0,363%	0,012%	0,011%	0,010%	0,012%
2100	0,300%	0,446%	0,533%	0,009%	0,016%	0,006%	0,009%
2400	0,414%	0,607%	0,723%	0,026%	0,041%	0,021%	0,026%
2700	0,522%	0,763%	0,906%	0,037%	0,060%	0,031%	0,037%
3000	0,487%	0,775%	0,946%	0,095%	0,065%	0,104%	0,095%

In case of Gamma distributed claims the most accurate approximation is given by hyperexponential (Hyper2) and general Coxian distributions (Coxian2). As can be seen from Table 4, these approximations gives exactly the same absolute relative errors.

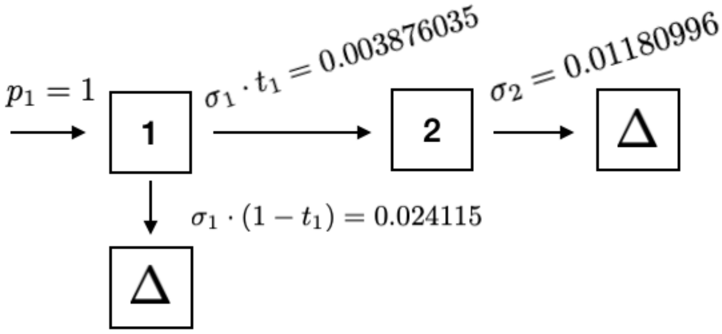

Figure 12. Estimation of parameters of general Coxian distribution

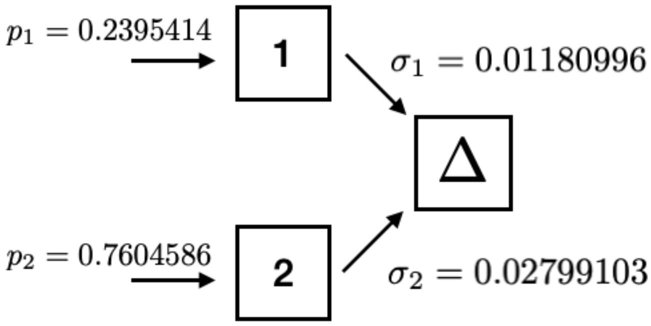

Figure 13. Estimation of parameters of hyperexponential distribution

Estimated rate parameters of these two distributions are equal ($\sigma 1_{\text{Coxian2}} = \sigma 2_{\text{Hyper2}}$ and $\sigma 2_{\text{Coxian2}} = \sigma 1_{\text{Hyper2}}$). As a result, Coxian2 and Hyper2 give exactly the same formula of the ruin probability in case of Gamma distributed claims.

$$\Psi(u)_{\text{H2}} = \Psi(u)_{\text{Coxian2}} = 0.01970989e^{-0.019107186u} + 0.87942839e^{-0.001745007u}$$

There is a clear theoretical reason for such a coincidence. Namely, it is known phase-type distribution that has no cycles in the phase-diagram can equivalently be represented as a Coxian phase-type distribution.[11] The reader can notice the same phenomena in the following examples too.

Simplified Coxian (Coxian1) and hypoexpontial (Hypo2) distributions give very accurate results too. Absolute relative errors are slight greater than in case of Hyper2 and general Coxian2 but still smaller than errors of De Vylder's approximations for most of the values of initial capital u. Erlang distributions with two (Erlang2) and three (Erlang3) phases give the worst approximations compare to other methods, but still work well and relative errors do not exceed 1%. The plot of relative errors for all seven approximations is presented in Figure 14

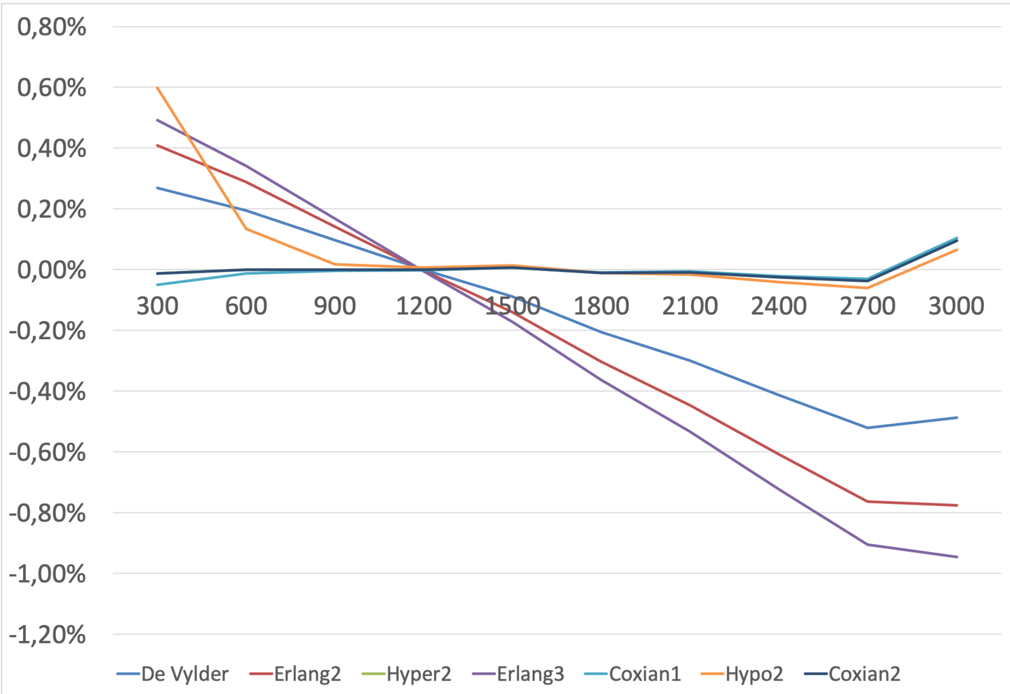

Figure 14. Relative error of ruin probabilities of approximations (Gamma distribution)

4.2 Mixed exponential distribution

Assume now that the sizes of the claims of the risk process $X(t)$ have mixed exponential (hyperexponenial) distribution with the distribution function $F(z)$.

$$F(z) = 1 - 0.0039793 \cdot e^{-0.014631z} - 0.1078392 \cdot e^{-0.190206z} - 0.8881815 \cdot e^{-5.514588z}.$$

We consider $\rho = 5\%; 10\%; 15\%; 20\%; 25\%; 30\%; 100\%$. Gross premium rate c is taken to be equal to one.

Table 5. Absolute relative errors (Mixed exponentially distributed claims)

ρ	*u*	DV	Erlang2	Erlang3	Hyper2	Coxian1	Hypo2	Coxian2
5%	10	3,209%	4,217%	4,885%	0,004%	2,252%	2,042%	0,004%
5%	100	0,373%	0,480%	0,583%	0,007%	0,595%	0,199%	0,007%
5%	1000	0,010%	0,013%	0,013%	0,011%	0,008%	0,036%	0,011%
10%	10	5,424%	7,131%	8,249%	0,019%	3,777%	3,446%	0,019%
10%	100	1,113%	1,514%	1,838%	0,006%	1,200%	0,016%	0,006%
10%	1000	0,945%	1,436%	1,729%	0,083%	0,155%	0,355%	0,083%
15%	10	6,998%	9,216%	10,659%	0,031%	4,828%	4,422%	0,031%
15%	100	1,915%	2,683%	3,268%	0,007%	1,690%	0,618%	0,007%
20%	10	8,148%	10,758%	12,446%	0,025%	5,568%	5,113%	0,025%
20%	100	2,711%	3,862%	4,719%	0,012%	2,085%	1,483%	0,012%
25%	10	8,978%	11,899%	13,779%	0,031%	6,065%	5,578%	0,031%
25%	100	3,380%	4,907%	6,030%	0,014%	2,316%	2,603%	0,014%
30%	10	9,581%	12,756%	14,792%	0,042%	6,391%	5,880%	0,042%
30%	100	3,987%	5,871%	7,247%	0,014%	2,475%	3,838%	0,014%
100%	10	10,665%	15,324%	18,258%	0,069%	5,772%	5,178%	0,069%
100%	100	7,209%	12,032%	15,553%	0,056%	1,240%	21,075%	0,056%

The results of approximations are analogical to the previous example with Gamma distribution. The smallest absolute relative errors are obtained by approximations with hyperexponential and general Coxian distributions. Note that these two approximations give almost perfect result (errors do not exceed 0.083%). In this example such result is not surprising. Initial distribution of claims is hyperexponential distribution with three phases. Moreover the weight of the first state is small enough ($w_1 = 0.0039793$). Hence distribution of claims of $X(t)$ is close to two-phases hyperexponential distribution. That is why Hyper2 and hence Coxian2 give almost perfect approximation.

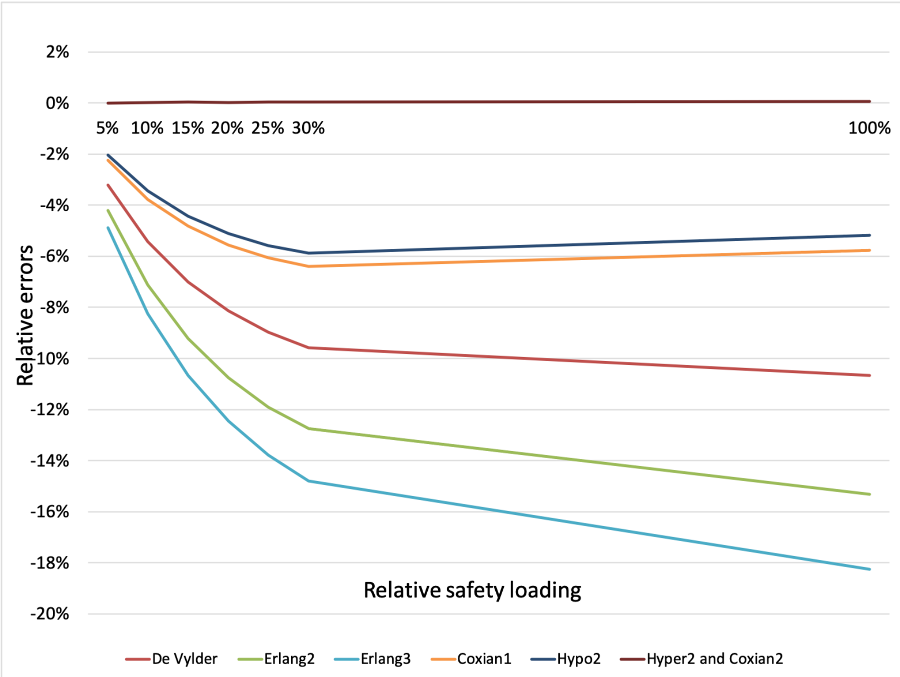

Figure 15. Relative errors for different values of ρ ($u = 10$).

Let's compare relative errors of the approximations depending on the values of the relative safety loading. From Figure 15 it is seen that Hyper2 and Coxian2 approximations are stable and give almost perfect approximation independently on ρ, but for all other methods there is the same trend in case of $\rho \in [5\%, 30\%]$. All methods underestimate the ruin probability and increasing the value of relative safety loading absolute relative errors of approximations are logarithmically increasing. Comparing relative errors in case of $\rho = 30\%$ and $\rho = 100\%$, we can see that Coxian1 and Hypo2 approximations give a bit more exact estimation of ruin probability for higher value of relative safety loading, while errors of others approximations continue increasing.

4.3 Lognormal distribution

As it was shown in Example 2.3 (Section 3.2), De Vylder's method can not precisely estimate ruin probability of a risk process if claims have lognormal distribution. Consider the same risk process $X(t)$ with lognormally distributed claims ($\sigma_L^2 = 3.24$, $\mu_L = -1.62$) and compare the accuracy of phase-type approximations.

Table 6. Absolute relative errors (Lognormal distribution)

ρ	u	De Vylder	Erlang2	Erlang3	Hyper2	Hypo2	Coxian1	Coxian2
5%	100	20,616%	21,606%	22,393%	19,326%	Approximamtion does not work		19,326%
5%	1000	55,092%	55,613%	55,700%	51,477%			51,477%
10%	100	19,482%	20,467%	21,350%	17,863%			17,863%
10%	1000	85,534%	82,542%	79,902%	75,590%			75,590%
15%	100	14,228%	15,185%	16,134%	12,373%			12,373%
15%	1000	79,663%	72,660%	67,085%	66,426%			66,426%
20%	100	8,089%	9,025%	10,035%	6,032%			6,032%
20%	1000	68,698%	58,711%	50,984%	53,958%			53,958%
25%	100	2,056%	2,976%	4,041%	0,180%			0,180%
25%	1000	59,163%	47,049%	37,795%	43,694%			43,694%
30%	100	3,529%	2,623%	1,509%	5,922%			5,922%
30%	1000	51,819%	38,133%	27,752%	35,947%			35,947%
100%	100	41,729%	40,938%	39,509%	45,122%			45,122%

From Table 6 we can see, that absolute relative errors are big in most of the cases. Risk process with any phase-type distribution of claims has exponentially decreasing ruin probability. Hence, it is not surprising that considered methods can not estimate ruin probability in case of lognormally distributed claims. Hence, relatively good estimation of the ruin probability, for example, in case of $\rho = 25\%$ and $u = 100$ can be considered as accidental.

Interesting is the fact that in case of Hypo2 and Coxian1 methods estimation of parameters do not give an adequate result. For example, one of the rate parameters of two phases hypoexponential distribution is complex number with negative real part ($\hat{\sigma}_2 = -6.419022 \cdot 10^{-4} + 8.972437 \cdot 10^{-4} i$). Complex number as an estimation of parameters of a risk process $\hat{X}(t)$ has been already met in case of hypoexponentialy distributed claims in previous examples but the real part of the number was positive and as result, it does not affect the estimation of ruin probability. In this example negativity of the real part of parameter estimation causes non-adequate ruin probability formula (probability is increasing if initial capital increases).

Two-phases hypoexponential and simplified Coxian distributions are, of course, general cases of two-phases Erlang distribution and that is why the question arises: why Hypo2 and Coxian1 can not adequately estimate the parameters if Erlang2 can handle it. The nature of the problem is based on the number of restrictions in case of each claim distribution. As it was mentioned before risk process $\hat{X}(t)$ with Erlang distribution of claims has three parameters. Hence, to estimate them it is needed to match the first three moments of initial risk process and $\hat{X}(t)$. In case of two-phases hypoexponential and simplified Coxian distributions there are four and five parameters respectively. Hence, because of greater number of restrictions it is not mathematically possible to obtain same result as in case of Erlang2.

4.4 Phase-type distribution with many states

Assume now that claims of a risk process $X(t)$ has phase-type distribution with many states. The goal of this section is to study if it possible to accurately estimate the ruin probability of $X(t)$ using approximations considered in Chapter 3.

Example 4.1.

Let distribution of claims of a risk process $X(t)$ is phase-type with ten states. Phase-diagram of this distribution is presented in Figure 16

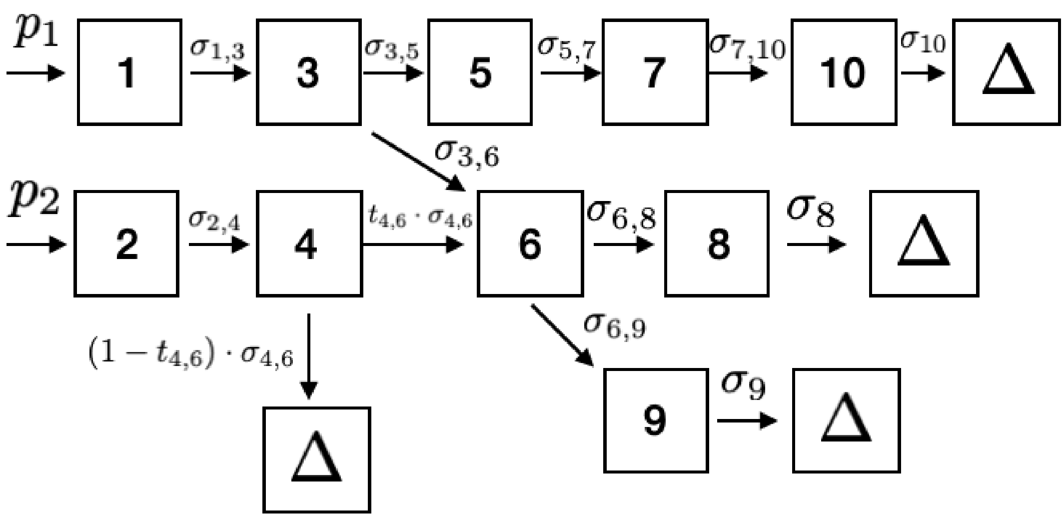

Figure 16. Phase-diagram of considered distribution

Assume that $p_1 = 0.6$; $p_2 = 0.4$; $\sigma_{1,3} = 1.3$; $\sigma_{2,4} = 2.4$; $\sigma_{3,5} = 3.5$; $\sigma_{3,6} = 3.6$; $\sigma_{4,6} = 4.6$; $t_{4,6} = 0.2$; $\sigma_{5,7} = 1.7$; $\sigma_{6,8} = 3.8$; $\sigma_{6,9} = 2.9$; $\sigma_{7,10} = 1.10$; $\sigma_8 = 0.8$; $\sigma_9 = 0.9$; $\sigma_{10} = 0.10$. Then we can estimate ruin probabilities of the risk process $X(t)$ by considered in the previous chapter methods.

Table 7. Absolute relative errors (Phase-type distribution with many states, example 4.1)

ρ	*u*	De Vylder	Erlang2	Erlang3	Hyper2	Coxian1	Hypo2	Coxian2
5%	10	0,01319%	0,04778%	0,07680%	0,00022%	0,05709%	0,03533%	0,00021%
5%	100	0,01080%	0,03945%	0,05639%	0,00005%	0,00092%	0,00005%	0,00000%
5%	200	0,00655%	0,02406%	0,03436%	0,00009%	0,00070%	0,00004%	0,00000%
10%	10	0,04905%	0,02094%	0,04657%	0,00043%	0,12138%	0,08126%	0,00042%
10%	100	0,02555%	0,09644%	0,13747%	0,00004%	0,00539%	0,00032%	0,00000%
10%	200	0,00366%	0,01069%	0,01569%	0,00008%	0,00251%	0,00017%	0,00000%
15%	10	0,09992%	0,05782%	0,05810%	0,00064%	0,18906%	0,13601%	0,00063%
15%	100	0,02727%	0,10980%	0,15551%	0,00004%	0,01313%	0,00084%	0,00000%
15%	200	0,05789%	0,20657%	0,29642%	0,00008%	0,00080%	0,00010%	0,00000%
20%	10	0,16050%	0,17253%	0,21439%	0,00085%	0,25766%	0,19795%	0,00084%
20%	100	0,00754%	0,04662%	0,06332%	0,00004%	0,02192%	0,00145%	0,00000%
20%	200	0,16792%	0,61276%	0,87743%	0,00008%	0,01151%	0,00065%	0,00000%
25%	10	0,22706%	0,31184%	0,40611%	0,00106%	0,32559%	0,26561%	0,00105%
25%	100	0,03702%	0,10849%	0,16120%	0,00003%	0,02889%	0,00195%	0,00001%
25%	200	0,33660%	1,24557%	1,78062%	0,00008%	0,04163%	0,00260%	0,00000%
30%	10	0,29700%	0,46755%	0,62164%	0,00126%	0,39183%	0,33772%	0,00125%
30%	100	0,10683%	0,36017%	0,52462%	0,00003%	0,03123%	0,00212%	0,00001%
30%	200	0,56172%	2,10129%	2,99842%	0,00008%	0,09586%	0,00624%	0,00001%
100%	10	1,16871%	2,75266%	3,83440%	0,00327%	1,04719%	1,43696%	0,00326%
100%	100	2,59861%	10,37450%	14,76820%	0,00017%	1,06465%	0,07575%	0,00014%
100%	200	6,63671%	25,80040%	35,03800%	0,00081%	4,25676%	0,30656%	0,00076%

As we can see from Table 7, the hyperexponential and general Coxian distributions with only two phases perfectly fit the high-order phase-type distribution for all values of initial capital and relative safety loading. Hypo2 is also very accurate approximation but it works worse in case of small initial capital when ruin probability is higher. De Vylder's approximation, Erlang2, Erlang3 and Coxian1 give really poor results in case of small ruin probability (high ρ and big u). To conclude, we can say that all approximations give relatively accurate estimations of the ruin probability of $X(t)$. This depends, of course, on the values of parameters of claim distribution.

Example 4.2.

Consider now same type of claim distribution, but with another values of parameters. Assume that $p_1 = 0.6$; $p_2 = 0.4$; $t_{4,6} = 0.2$ and $\sigma_{1,3} = \sigma_{2,4} = \sigma_{3,5} = \sigma_{3,6} = \sigma_{4,6} = \sigma_{5,7} = \sigma_{6,8} = \sigma_{6,9} = \sigma_{7,10} = \sigma_8 = \sigma_9 = \sigma_{10} = 1$.

In this case Hypo2 approxiamtion can not adequately estimate parameters and give incorrect formula of the ruin probability (probability is increasing if initial capital increases). Hence, the only approximations which work correctly in this case are De Vylder's approximation, Erlang2, Erlang3, Hyper2, Coxian1 and Coxian2.

Table 8. Absolute relative errors (Phase-type distribution with many states, example 4.2)

ρ	u	DV	Erlang2	Erlang3	Hyper2	Coxian1	Coxian2	Hypo2
0,05	10	0,030000%	0,011280%	0,000130%	0,000025%	0,000200%	0,000024%	Approximation does not work
0,05	100	0,002660%	0,001020%	0,000010%	0,000003%	0,000060%	0,000001%	
0,05	200	0,038950%	0,014650%	0,000120%	0,000010%	0,000090%	0,000000%	
0,1	10	0,101420%	0,037990%	0,000610%	0,000030%	0,001320%	0,000030%	
0,1	100	0,134980%	0,050730%	0,000480%	0,000006%	0,000560%	0,000001%	
0,1	200	0,398260%	0,149300%	0,001680%	0,000057%	0,002600%	0,000048%	
0,15	10	0,192790%	0,071930%	0,001510%	0,000046%	0,003770%	0,000046%	
0,15	100	0,533000%	0,199670%	0,002510%	0,000102%	0,004490%	0,000097%	
0,15	200	1,345550%	0,502150%	0,006930%	0,000416%	0,013570%	0,000406%	
0,2	10	0,289070%	0,107420%	0,002810%	0,000277%	0,007630%	0,000277%	
0,2	100	1,284360%	0,479340%	0,007260%	0,000530%	0,015320%	0,000525%	
0,2	200	3,061620%	1,135070%	0,018380%	0,001676%	0,040670%	0,001667%	
0,25	10	0,379920%	0,140520%	0,004470%	0,000733%	0,012810%	0,000733%	
0,25	100	2,445810%	0,908550%	0,015880%	0,001685%	0,036810%	0,001680%	
0,25	200	5,679400%	2,086700%	0,038390%	0,004758%	0,091760%	0,004749%	
0,3	10	0,458430%	0,168680%	0,006520%	0,001378%	0,019220%	0,001379%	
0,3	100	4,055280%	1,497710%	0,029420%	0,004098%	0,072520%	0,004093%	
0,3	200	9,310520%	3,381200%	0,069110%	0,010856%	0,174130%	0,010847%	
1	10	0,561700%	0,278580%	0,032780%	0,053888%	0,112110%	0,053888%	
1	100	87,427220%	26,459200%	1,009750%	0,571241%	3,066190%	0,571237%	
1	200	274,330380%	63,621510%	2,178150%	1,273417%	6,710260%	1,273410%	

From Table 8 we can see that working approximation methods give very accurate results in most of cases. Hyper2 and Coxian2 with only two phases perfectly fit the high-order phase-type distribution for all values of initial capital and relative safety loading. Since all transition rates of this phase-type distribution are equal, Erlang3 also returns very accurate estimation of the ruin probability. The worst approximation (in comparison with other approximations) is De Vylder's method. Note that in case of high ρ and big u (ruin probability is small in this case) De Vylder's method returns very poor result. For example, if $\rho = 100\%$ and $c = 200$ absolute relative error is $274,33\%$.

To conclude, approximation based on low-order phase-type distribution can accurately estimate the ruin probabilities of a risk process, in case if its claims have complicated, high-order phase-type distributions with many states. De Vylder's approximation returns an accurate estimation of the ruin probability based on considered examples, but absolute relative errors of this method are several times bigger than errors of Hyper2 and Coxian2.

Conclusion

The main goal of the thesis was to find a method of approximation of the ruin probability that is more accurate than famous De Vylder's method but at the same time is not technically too complicated. In the framework of this thesis six approximations were considered based on phase-type distributions for calculation of the ruin probability. Accuracy of each approximation were compared based on four examples of risk processes with different distributions of claims.

In this thesis we examined mostly approximations based on phase-type distributions with two phases. The number of parameters describing a risk process with Coxian, hypoexponential or hyperexponential distribution of claims depends on the number of phases. Even in case of two phases these approximations need to solve systems of four or five equations which are much more difficult than in case of De Vylder's approximation. Hence, the number of phases was limited by two for these distributions. Three-phases case was considered for Erlang distribution, since the number of unknown parameters describing a risk process with Erlang distribution is always three and does not depend on the numbers of phases.

Comparison of absolute relative errors of ruin probability's estimations showed that hyperexponential and general Coxian distributions give the most accurate results based on examples of risk process with Gamma and mixed exponentially distributed claims. In these cases named approximations fit the ruin probability perfectly and relative errors are negligible. The biggest errors are seen if approximation is based on two- or three-phases Erlang distribution. Their absolute relative errors are even higher than in case of De Vylder's method. Moreover, increasing of the number of phases of Erlang distributions worsens the result.

Risk processes with lognormally distributed claims are badly fitted by De Vylder's approximation. It is well understood since in lognormal case the ruin probability follows the asymptotic which significantly differes from exponential asymptotic. In this thesis it was shown that phase-type distributions can not correctly describe it too, since they return exponentially decreasing ruin probability.

To conclude, our new approximation methods based on simple phase-type distributions (such as hyperexponential and general Coxian) almost always give more accurate ruin probabilities than well-known De Vylder's method. At the same time, the price for the increased accuracy is only minimal, since ruin probabilities for phase-type distributed claims can still be calculated via explicit formulas, without any time consuming iterations.

References

[1] Pärna, K. (2018). Risk Theory. Lecture notes.

[2] Grandell, J. (1991). Aspects of Risk Theory. New-York: Springer-Verlag.

[3] Albrecher, H. and Asmussen, S. (2010). Ruin Probabilities. Published by World Scientific Publishing Co. Pte. Ltd. (second edition).

[4] Pärna, K. (2018). Stochastic Models. Lecture notes.

[5] Käärik, M. (2019). Non-life Insurance Mathematics. Lecture notes.

[6] Cramer, H. (1946). Mathematical Methods of Statistics. Bombay: Asia Publishing House

[7] Mattuck A., Miller H., Orloff J. and Lewis J. (2011) Differential Equations. Massachusetts Institute of Technology: MIT OpenCourseWare, https://ocw.mit.edu.

[8] Ushakov, N. G. (1999) Selected topics in characteristic functions. Published by De Gruyter.

[9] Panchenko, D. (2006) Statistics for Applications. Massachusetts Institute of Technology: MIT OpenCourseWare, https://ocw.mit.edu.

[10] Papadopolous, H. T., Heavey, C. and Browne, J. (1993). Queueing Theory in Manufacturing Systems Analysis and Design. London: Chapman & Hall.

[11] Dehon M. and Latouche G. (1982) A Geometric Interpretation of the Relations between the Exponential and Generalized Erlang Distributions. Advances in Applied Probability Vol. 14, No. 4.

[12] Anderson, W. J. (1991). Continuous-Time Markov Chains: An Applications-Oriented Approach. New-York: Springer-Verlag

Appendixes

Appendix 1.

R-script calculating the ruin probability for a risk process with phase-type distributed claims.

```
toenausus2<-function(T,alpha,betha){
  t <- -rowSums(T)
  T.inverse <- solve(T)
  alpha.plus <- -betha*alpha%*%T.inverse
  T.plus <- T+t%*%alpha.plus
  lambd<-eigen(T.plus)$values
  vect<-eigen(T.plus)$vectors
  f0<-vect
  f01<-solve(f0)
  kordaja<-c()
  for(i in 1:length(lambd)){
    Values<-matrix(ncol = length(lambd),nrow = length(lambd))
    for(row in 1:length(lambd)){
      for(col in 1:length(lambd)){
        vector<-f0[row,]*f01[,col]
        Values[row,col]<-vector[i]
      }
    }
    kordaja<-c(kordaja,sum(alpha.plus%*%Values))
  }
  result<-cbind(lambd, kordaja)
  colnames(result)<-c("power","multiplier")
  return(result)
}
```

Appendix 2.

Maxima script estimating the parameters of the risk process $\hat{X}(t)$ with two-phases hypoexponential distirbution of claims from Section 3.3.

```
k1:1/s1;
k2:1/s2;
Solver:solve([1-b*e1=c1-b1*(k1+k2),
    b*e2=2*b1*(k1*k2+k2^2+k1^2),
    b*e3=6*b1*(k1^2*k2+k1*k2^2+k2^3+k1^3),
    b*e4=24*b1*(k1^3*k2+k1^2*k2^2+k1*k2^3+k2^4+k1^4)
    ],[c1,b1,k1,k2]);
```

Maxima script estimating the parameters of the risk process $\hat{X}(t)$ with two-phases hyperexponential distirbution of claims from Section 3.4.

```
Solver:solve([1-b*e1=c1-b1*(p/s1+(1-p)/s2),
    b*e2=2*b1*(p/s1^2+(1-p)/s2^2),
    b*e3=6*b1*(p/s1^3+(1-p)/s2^3),
    b*e4=24*b1*(p/s1^4+(1-p)/s2^4),
    b*e5=120*b1*(p/s1^4+(1-p)/s2^4)
    ],[c1,b1,p,s1,s2]);
```

Maxima script estimating the parameters of the risk process $\hat{X}(t)$ with two-phases Coxian distirbution of claims from Section 3.5.2

```
k1:1/s1;
k2:1/s2;
Solver:solve([1-b*e1=c1-b1*(p*k2+k1),
    b*e2=2*b1*(p*k1*k2+p*k2^2+k1^2),
    b*e3=6*b1*(p*k1^2*k2+p*k1*k2^2+p*k2^3+k1^3),
    b*e4=24*b1*(p*k1^3*k2+p*k1^2*k2^2+p*k1*k2^3+p*k2^4+k1^4),
   b*e5=120*b1*(p*k1^4*k2+p*k1^3*k2^2+p*k1^2*k2^3+p*k1*k2^4+p*k2^5+k1^5)
    ],[c1,b1,p,k1,k2]);
```

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Kirill Smirnov,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose *Laostumistõenäosuse lähendamine faasi-tüüpi jaotuste abil,*

mille juhendaja on Kalev Pärna,

reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 3.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Kirill Smirnov
26.05.20