FACULTY OF MATHEMATICS AND PHYSICS
Charles University

MASTER THESIS

Bc. Anna Vedyushenko

Non-homogeneous Poisson process - estimation and simulation

Department of Probability and Mathematical Statistics

Supervisor of the master thesis: RNDr. Michal Pešta, Ph.D.
Study programme: Mathematics
Study branch: Financial and Insurance Mathematics

Prague 2018

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

I would like to thank my supervisor, RNDr. Michal Pešta, Ph.D., for professional guidance, precious advice and helpfulness during the work on this thesis. In addition, I am very grateful to the Czech Insurers' Bureau for providing me with the data essential for the practical part of this thesis.

Title: Non-homogeneous Poisson process - estimation and simulation

Author: Bc. Anna Vedyushenko

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Michal Pešta, Ph.D., Department of Probability and Mathematical Statistics

Abstract: This thesis covers non-homogeneous Poisson processes along with estimation of the intensity (rate) function and some selected simulation methods. In Chapter 1 the main properties of a non-homogeneous Poisson process are summarized. The main focus of Chapter 2 is the general maximum likelihood estimation procedure adjusted to a non-homogeneous Poisson process, together with some recommendations about calculation of the initial estimates of the intensity function parameters. In Chapter 3 some general simulation methods as well as the methods designed specially for log linear and log quadratic rate functions are discussed. Chapter 4 contains the application of the described estimation and simulation methods on real data from non-life insurance. Furthermore, the considered simulation methods are compared with respect to their time efficiency and accuracy of the simulations.

Keywords: Poisson process, non-homogeneous, inhomogeneous, estimation, intensity function, simulation

Contents

Introduction

Poisson processes are widely used in practice for predicting events' arrival. The simplest and, probably, the most known model is the one with a constant rate — a homogeneous Poisson process. However, such model does not allow for any changes in the intensity of events' arrival within time. A more complex model — a non-homogeneous Poisson process (NHPP) — should be used in the cases where events arrive differently within time. In this model the rate is not constant anymore. It depends both on the starting-point and endpoint of the analysed time interval.

The purpose of this thesis is primarily to summarize the most important properties of non-homogeneous Poisson processes along with the certain estimation and simulation methods. The properties described in the thesis are very important for further understanding of the estimation and simulation procedures. Unfortunately, none of the current most known books about the topic contain such compilation.

The main properties of NHPPs are summarized in Chapter 1. Along with the basic properties, we present a very useful method to reduce complex computations about a NHPP as well as derive the conditional distribution of arrival times.

The main focus of Chapter 2 is to describe a general procedure of estimating the parameters of the rate function using maximum likelihood method. We present the likelihood equations and the Fisher Information Matrix adjusted for the case of a NHPP. Furthermore, we describe the procedure for the log linear family of rate functions in more detail.

Chapter 3 contains the theory in respect of the most widely used simulation methods for a NHPP. We consider two general algorithms which can be applied to rate functions of any form along with the special methods designed for log linear and log quadratic rate functions. In addition, we discuss the theoretical time efficiency of the described algorithms.

Chapter 4 is dedicated to practical application of the previously described methods, which is another aim of this thesis. We use the data on MTPL claims caused by uninsured drivers from the Czech Republic, provided by Czech Insurers' Bureau. We present general description of the data together with its analysis. Next, we perform certain grouping of the data and justify the choice of the grouping. Then we apply the selected estimation and simulation methods on the grouped data. Afterwards, the real time efficiency of the simulation algorithms is discussed, and the recommendations about the most suitable algorithm are provided. Furthermore, we simulate the next year payments of the selected claims. In addition, the source code of the simulation algorithms is provided in the Attachment of this thesis.

1. Non-homogeneous Poisson process and its properties

In this chapter we define a non-homogeneous Poisson process (NHPP) and state its basic properties which are important for the next chapters, where we focus on the estimation and simulation methods. Non-homogeneity of the Poisson process basically means that the distribution of the number of events between two particular points on the timeline is no longer a function depending on the difference between these points, as it is in case of a homogeneous Poisson Process (HPP). In our case it is a function of both starting-point and endpoint of the time interval.

Now let us look at the definition of a NHPP in more detail. There are more equivalent definitions of the latter; nevertheless, we state here the one described in Ross [2010].

Definition 1. *(Non-homogeneous Poisson process)*
The counting process $\{N(t),\ t \geq 0\}$ *is said to be a non-homogeneous Poisson process with intensity function* $\lambda(t),\ t \geq 0$*, if it satisfies*

- $N(0) = 0$ *almost surely;*
- $\{N(t)\}$ *has independent increments;*
- $\mathsf{P}\left[N(t+h) - N(t) \geq 2\right] = o(h)$;
- $\mathsf{P}\left[N(t+h) - N(t) = 1\right] = \lambda(t)h + o(h)$.

Function $\lambda(t)$ is sometimes also referred to as the *instantaneous arrival rate.*

Ross [2010] is exploring the relationship between the average number of events which occurred until time t and the intensity function $\lambda(t)$ of the corresponding NHPP:

$$\mathsf{E}\ N(t) = \int_0^t \lambda(u)\mathrm{d}u = \Lambda(t) - \Lambda(0) \stackrel{\text{def}}{=} \mu(t). \tag{1.1}$$

We shall further refer to $\mu(t)$ as the *expectation function* of the NHPP $\{N(t)\}$. In addition, if we consider the expected number of events between times t and $t+s$ and use a simple integration property, we can conclude from expression (1.1):

$$\mathsf{E}\left[N(t+s) - N(t)\right] = \int_t^{t+s} \lambda(u)\mathrm{d}u = \Lambda(t+s) - \Lambda(t).$$

Çinlar [2013] shows that $\mu(t)$ is, in fact, a non-decreasing right-continuous function. However, from now on we assume that $\mu(t)$ is a continuous function. Furthermore, we assume that

$$0 \leq \int_{\mathcal{R}} \lambda(u)\mathrm{d}u < \infty,$$

for all bounded subsets $\mathcal{R}$ of the state space S of the process.

In Figure 1.1 we can see expected numbers of events until time t for two NHPPs with $\lambda(t) = \exp\{\theta_0 + \theta_1 t\}$, which represents *log linear model* for the intensity function. In the first case $\theta_0 = 0.8$, $\theta_1 = 0.1$; for the second intensity function $\theta_0 = 0.8$, $\theta_1 = -0.1$.

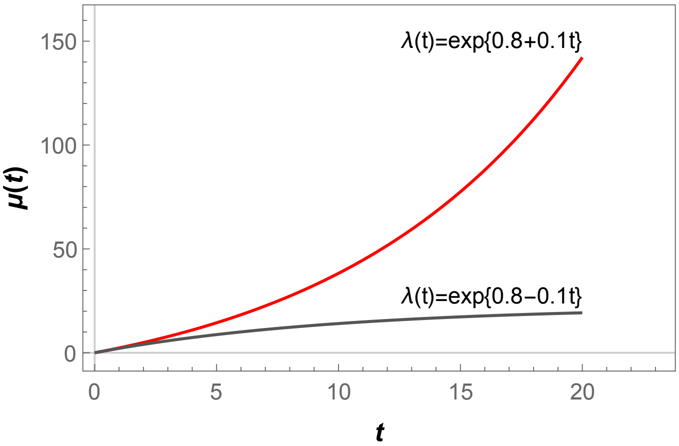


Figure 1.1: *Expected numbers of events until time t for log linear model for* $\lambda(t)$

Dealing with a NHPP can be very often computationally challenging. Nevertheless, we can easily reduce computations to the ones about a HPP, as can be seen from the next theorem stated in Çinlar [2013].

Theorem 1. *(Relation between a NHPP and a HPP)*
Let $\mu(t)$ be a continuous non-decreasing function. Then $T_1, T_2, \ldots$ are the arrival times of a NHPP $\{N(t), t \geq 0\}$ with expectation function $\mu(t)$ if and only if $\mu(T_1), \mu(T_2), \ldots$ are the arrival times of a HPP $\{M(t), t \geq 0\}$ with arrival rate $\lambda = 1$.

The previous theorem can be restated in terms of the Poisson processes $\{N(t)\}$ and $\{M(t)\}$. If we define a *time inverse* of $\mu(t)$ by

$$\tau(t) = \inf\{s : \mu(s) \geq t\}, \quad t \geq 0 \tag{1.2}$$

the number of events until time $\tau(t)$ in the NHPP $\{N(t)\}$ is then equivalent to the number of events until time t in the HPP $\{M(t)\}$, i.e.:

$$N(\tau(t)) = M(t).$$

Let us now examine the distribution of the number of events between times t and $t + s$, i.e. the distribution of increments of the NHPP $\{N(t)\}$. According to Çinlar [2013], we can derive one by making use of Theorem 1.

Theorem 2. *(Distribution of increments of a NHPP)*
Let $\{N(t), t \geq 0\}$ be a NHPP with continuous expectation function $\mu(t)$. Then for any $t, s \geq 0$ it holds:

$$\mathsf{P}\left(N(t+s) - N(t) = k\right) = \frac{[\Lambda(t+s) - \Lambda(t)]^k}{k!}\, \mathrm{e}^{-\{\Lambda(t+s) - \Lambda(t)\}}.$$

That is, the distribution of $N(t+s) - N(t)$ is, in fact, Poisson with parameter $\Lambda(t+s) - \Lambda(t)$. It is also useful to note the property of the superposition of two NHPPs', which will be used later for one of the simulation algorithms. According to Ross [2010], it holds:

Theorem 3. *(Superposition of two independent NHPPs)*
Let $\{N(t),\, t \geq 0\}$ and $\{M(t),\, t \geq 0\}$ be two independent NHPPs, with respective intensity functions $\lambda_1(t)$ and $\lambda_2(t)$. Furthermore, let $N^*(t) = N(t) + M(t)$. Then, the following are true.

1. *$\{N^*(t)\}$ is a NHPP with intensity function $\lambda_1(t) + \lambda_2(t)$;*
2. *Given that an event of the $\{N^*(t)\}$ process occurs at time t then, independently of what occurred prior to t, the event at t was from the $\{N(t)\}$ process with probability $\lambda_1(t)/(\lambda_1(t) + \lambda_2(t))$.*

Now let us examine the distribution of arrival times (or, equivalently, times to events $1, \ldots, n$) $T_1, T_2, \ldots, T_n$ of a NHPP $\{N(t)\}$ under the condition that exactly n events occurred in the time interval $(0, T]$. Firstly, let us look at the distribution function of the time to the next event in a NHPP. As Cox and Lewis [1966] state, it has the following form.

Theorem 4. *(Probability distribution of the time to the next event in a NHPP)*
Let $\{N(t),\, t \geq 0\}$ be a NHPP with continuous expectation function $\mu(t)$. Then for any $t,\, s \geq 0$

$$\begin{aligned} \mathsf{P}\left(1 \text{ or more events occurred in } (t, t+s]\right) &= 1 - \exp\left\{-\int_t^{t+s} \lambda(u)\mathrm{d}u\right\} \\ &= 1 - \mathrm{e}^{-\{\Lambda(t+s) - \Lambda(t)\}}. \end{aligned} \quad (1.3)$$

From the above statement we can derive the joint distribution of arrival times $T_1, T_2, \ldots, T_n$ and occurrence of exactly n events. To simplify the computations, we firstly find the probability density function of the time to the next event by deriving expression (1.3) with respect to s, i.e. we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathsf{P}\left(1 \text{ or more events occurred in } (t, t+s]\right) = \lambda(t+s)\,\mathrm{e}^{-\{\Lambda(t+s) - \Lambda(t)\}}.$$

Furthermore, according to Cox and Lewis [1966], if we observe events in the interval $(0, T]$ and n events occurred at times $t_1, t_2, \ldots, t_n$, the desired joint probability density function takes form

$$\begin{aligned} \lambda(t_1)\,\mathrm{e}^{-\{\Lambda(t_1) - \Lambda(0)\}} \cdot \lambda(t_2)\,\mathrm{e}^{-\{\Lambda(t_2) - \Lambda(t_1)\}} \cdot \ldots \cdot \lambda(t_n)\,\mathrm{e}^{-\{\Lambda(t_n) - \Lambda(t_{n-1})\}} \cdot \mathrm{e}^{-\{\Lambda(T) - \Lambda(t_n)\}} \\ = \mathrm{e}^{-(\Lambda(T) - \Lambda(0))} \prod_{i=1}^{n} \lambda(t_i), \end{aligned} \quad (1.4)$$

where the multiplier $\mathrm{e}^{-\{\Lambda(T) - \Lambda(t_n)\}}$ in the first part of the equation stands for the probability of no events occurring in the interval $(t_n, T]$. To find the conditional probability density function of the arrival times $T_1, T_2, \ldots, T_n$ under the condition that exactly n events occurred, we should divide expression (1.4) by the probability of exactly n events occurring, that is, by

$$\frac{[\Lambda(T)-\Lambda(0)]^n}{n!}\, \mathrm{e}^{-\{\Lambda(T)-\Lambda(0)\}}.$$

Then, in total, the conditional probability density function of $T_1, T_2, \ldots, T_n$ looks like

$$\frac{n!}{[\Lambda(T)-\Lambda(0)]^n} \prod_{i=1}^{n} \lambda(t_i), \tag{1.5}$$

where the range of t_i satisfies $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n \leq T$. Expression (1.5) is, in fact, as mentioned in Cox and Lewis [1966], the probability density function of *ordered* sample of n variables from a truncated exponential distribution. That is, by using a relation between such probability density function and marginal ordered variables, we can derive the probability function of an ordered variable $T_{(i)}$, which takes form

$$f_{T_{(i)}}(t_i \mid N(T)=n) = \frac{\lambda(t)}{\Lambda(T)-\Lambda(0)}, \quad i = 1, \ldots, n, \tag{1.6}$$

if $t_1 < t_2 < \ldots < t_n$ is satisfied. The respective distribution function can be derived from expression (1.6) by integration. That implies we have proved the following theorem.

Theorem 5. *(Conditional distribution of arrival times)*
Let $\{N(t),\, t \geq 0\}$ be a NHPP with continuous expectation function $\mu(t)$. If we observe events in the interval $(0, T]$, arrival times $T_1, T_2, \ldots, T_n$, under the condition that exactly n events occurred in $(0, T]$, are distributed as the order statistics from a sample with the distribution function

$$F(t) = \frac{\Lambda(t)-\Lambda(0)}{\Lambda(T)-\Lambda(0)}, \quad 0 \leq t \leq T.$$

We have stated all basic properties of a NHPP which will be needed for the purpose of simulation algorithms. Now let us continue with Chapter 2, where we discuss intensity estimation.

2. Estimation of the intensity function

In this chapter we examine estimation of the intensity function parameters in a NHPP using maximum likelihood method. We restrict ourselves to the case of a parametric intensity. Note that its form should be chosen before performing the estimation. One can assume different models for $\lambda(t)$ and apply maximum likelihood algorithm on all of them. Subsequently, the best model can be chosen using e.g. *Akaike Information Criterion (AIC).* There are several recommended models in literature which are appropriate for practical applications. The general form of estimates obtained by maximum likelihood procedure as well as the special case of one recommended model will be mentioned here.

Common problem in practice is the case when the likelihood function is not strictly concave, hence its maximum will not be unique. Its shape depends on the chosen model for $\lambda(t)$. As the maximum likelihood estimate is, by definition, the global maximum of the likelihood function, numerical algorithms for solving the likelihood equations should be restarted with different initializations in such case. The solution with the largest likelihood is then considered to be the global maximum. We shall also describe methods suitable for finding initial estimates of the parameters of a general rate function containing a global trend as well as periodicity components, which helps to increase efficiency of the chosen numerical algorithm.

2.1 General description of maximum likelihood method for a NHPP

We consider the sample data $\mathbf{t} = (t_1, \ldots, t_n)^\top$ from a realization of a NHPP with intensity function $\lambda(t)$ on $(0, T] \subset \mathcal{S} \subset \mathbb{R}$, where $\mathcal{S}$ is the state space of the process. According to Streit [2010], the sample data are conditionally independent given their total number n. As further mentioned in the book, it implies that the points t_i are in $(0, T]$ and for that reason we are estimating the intensity only on $(0, T]$, not on the whole state space $\mathcal{S}$.

If we substitute the arrival time terms in expression (1.4) with our realization of the process, that is, $(t_1, \ldots, t_n)^\top$, we can conclude that the obtained expression is actually the joint likelihood function of the sample arrival times and the occurrence of n events. Let us further consider m parameters of the intensity function, i.e. vector of parameters $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_m)^\top$ defined on some set $\Theta \subset \mathbb{R}^m$. The intensity function in this case will be denoted as $\lambda(t \,|\, \boldsymbol{\theta})$ to emphasise its dependency on the considered set of parameters. The log-likelihood function then looks like

$$\ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta}) = \sum_{i=1}^{n} \log \lambda(t_i \,|\, \boldsymbol{\theta}) - \int_0^T \lambda(u \,|\, \boldsymbol{\theta}) \mathrm{d}u = \sum_{i=1}^{n} \log \lambda(t_i \,|\, \boldsymbol{\theta}) + \Lambda(0 \,|\, \boldsymbol{\theta}) - \Lambda(T \,|\, \boldsymbol{\theta}). \tag{2.1}$$

The maximum likelihood estimate of $\boldsymbol{\theta}$ can then be found as

$$\hat{\boldsymbol{\theta}}_{ML} \equiv \arg\max_{\boldsymbol{\theta}\in\Theta} \ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta}).$$

If we assume differentiability of $\lambda(t \,|\, \boldsymbol{\theta})$ with respect to every component of the vector of parameters $\boldsymbol{\theta}$ for each $\mathbf{t} \in (0, T]^n$, we can compute $\hat{\boldsymbol{\theta}}_{ML}$ by solving the necessary conditions, i.e. by setting to zero the partial derivatives with respect to each θ_i, or, if written more formally, that $\partial \left[\ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta})\right] / \partial \boldsymbol{\theta} = \mathbf{0}_m$. Hence we shall be solving

$$\sum_{i=1}^{n} \frac{1}{\lambda(t_i \,|\, \boldsymbol{\theta})} \frac{\partial}{\partial \boldsymbol{\theta}} \lambda(t_i \,|\, \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \int_0^T \lambda(u \,|\, \boldsymbol{\theta}) \mathrm{d}u = \frac{\partial}{\partial \boldsymbol{\theta}} \left[\Lambda(T \,|\, \boldsymbol{\theta}) - \Lambda(0 \,|\, \boldsymbol{\theta})\right].$$

It is essential to verify that the likelihood function is indeed concave at the point of the solution $\hat{\boldsymbol{\theta}}_{ML}$, which will ensure that $\hat{\boldsymbol{\theta}}_{ML}$ is at least a local maximum. We can do so by demonstrating that the Fisher Information Matrix (FIM) of the log-likelihood function is positive definite at the solution $\hat{\boldsymbol{\theta}}_{ML}$. If $\lambda(t_i \,|\, \boldsymbol{\theta}) > 0$ for all $t_i \in (0, T]$, the FIM is defined by

$$I(\boldsymbol{\theta}) = \mathsf{E}\left[-\frac{\partial^2 \ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\right].$$

Furthermore, if the regularity conditions are satisfied, confidence intervals for $\boldsymbol{\theta}$ can be derived using the property of asymptotic normality of $\hat{\boldsymbol{\theta}}_{ML}$. It holds that

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}) \xrightarrow{D} \mathsf{N}_m(0, I^{-1}(\boldsymbol{\theta}))$$

as the number of events n occurred in the interval $(0, T]$ approaches infinity. Based on the asymptotic distribution one can then construct the asymptotic confidence interval for $\boldsymbol{\theta}$ e.g. using Slutsky's theorem and the properties of a continuous transformation of a random vector.

Moreover, as stated in Streit [2010], the inverse of the FIM is, in fact, the lower bound on the covariance matrix of any unbiased estimator of $\boldsymbol{\theta}$, provided that the FIM is a regular matrix. Hence, it holds for any unbiased estimator $\hat{\boldsymbol{\theta}}$, if we fix the value of $\boldsymbol{\theta}$, that

$$\mathrm{var}\,[\hat{\boldsymbol{\theta}}] \geq I^{-1}(\boldsymbol{\theta}).$$

Log linear rate function

Now let us move to a special case of the model for the intensity function, which was considered by Lewis and Shedler [1976] as useful for practical applications of a NHPP. We assume that $\lambda(t)$ follows so-called *log linear model*, $\lambda(t) = \exp\{\theta_0 + \theta_1 t\}$. The reasons to assume an exponential model, not just a linear one, that is, $\lambda(t) = \theta_0 + \theta_1 t$, are the following. As Lewis and Shedler [1976] state, λ stays positive for all values of θ_0 and θ_1 in case of log linear model, whereas for a linear model it can be acquired only by applying non-linear restrictions on the parameters. In addition, statistical procedures performed on log linear model are simple.

Such model represents the case when the intensity function is monotonically increasing or decreasing, depending on whether θ_1 is greater or less than zero. The case $\theta_1 = 0$ gives a HPP with rate $\lambda = \mathrm{e}^{\theta_0}$. The model described here is the simplest of the general family of log linear rate functions. We shall also discuss a more complicated case with an additional term $\theta_2 t^2$ in the next chapter, where the methods for simulation of such process will be presented.

From expression (1.4) we know that in case of the considered model, the joint likelihood function takes form

$$f(\mathbf{t},\, n \,|\, \boldsymbol{\theta}) = \exp\left\{ n\theta_0 + \theta_1 \sum_{i=1}^{n} t_i - \frac{\mathrm{e}^{\theta_0}\left[\mathrm{e}^{\theta_1 T} - 1\right]}{\theta_1} \right\}. \tag{2.2}$$

However, Cox and Lewis [1966] show that, since the observations in expression (2.2) are contained only in terms $(n, \sum t_i)$, the latter are sufficient statistics for drawing any conclusions about the values of θ_0 and θ_1. Furthermore, they state that for fixed θ_1, a sufficient statistic for θ_0 is the number of events n. It follows then that for estimating θ_1 we can consider the conditional distribution of the arrival times under the condition that n events occurred. Hence using (1.5) we get

$$f(\mathbf{t} \,|\, n,\, \boldsymbol{\theta}) = \frac{n!\,\mathrm{e}^{n\theta_0}\;\mathrm{e}^{\theta_1 \sum t_i}}{\left[\mathrm{e}^{\theta_0}\;(\mathrm{e}^{\theta_1 T} - 1)\,/\,\theta_1\right]^n} = \frac{n!\,[\theta_1]^n}{(\,\mathrm{e}^{\theta_1 T} - 1)^n}\;\; \mathrm{e}^{\theta_1 \sum t_i}.$$

The log-likelihood function then takes form

$$\ell(\mathbf{t} \,|\, n,\, \theta_1) = \log n! + n\log\theta_1 + \theta_1 \sum_{i=1}^{n} t_i - n\log\left(\mathrm{e}^{\theta_1 T} - 1\right)$$

and its derivative with respect to θ_1, according to Cox and Lewis [1966], looks like

$$\frac{\partial \ell(\mathbf{t} \,|\, n,\, \theta_1)}{\partial \theta_1} \stackrel{\text{def}}{=} \ell' = \begin{cases} n/\theta_1 + \sum t_i - nT/(1 -\, \mathrm{e}^{-\theta_1 T}), & \theta_1 \neq 0, \\ -nT/2 + \sum t_i, & \theta_1 = 0. \end{cases}$$

The maximum likelihood estimate of θ_1 can be found by solving $\partial\ell(\mathbf{t} \,|\, n,\, \theta_1)/\partial\theta_1 = 0$. Numerical methods should be used to find the solution in this case, an initial estimate should be provided as well. Otherwise, the numerical solving algorithm does not necessarily have to converge. In order to verify that the likelihood function is concave at the point of the solution, we shall check for positive definiteness of the FIM. In our case the FIM takes the following form, as Cox and Lewis [1966] state:

$$I(\theta_1) = \begin{cases} n\left[1/(\theta_1)^2 - T^2\,\mathrm{e}^{-\theta_1 T}/(1 - \mathrm{e}^{-\theta_1 T})^2\right], & \theta_1 \neq 0, \\ -nT^2/12, & \theta_1 = 0. \end{cases} \tag{2.3}$$

One might consider testing for the value of the parameter θ_1, that is, we want to test for the null hypothesis

$$H_0:\;\; \theta_1 = \theta_1^*$$

against the alternative

$$H_0:\;\; \theta_1 \neq \theta_1^*.$$

Cox and Lewis [1966] show that in order to perform the test, one can consider the following test statistic

$$Z(\theta_1^*) = \frac{\ell'(\theta_1^*)}{\sqrt{I(\theta_1^*)}},$$

which has exactly zero mean and unit variance under the null hypothesis. In addition, its distribution is asymptotically normal under the null hypothesis. Therefore, we reject H_0 at the desired significance level α in case

$$|Z(\theta_1^*)| \geq c_{1-\frac{\alpha}{2}},$$

where $c_{1-\frac{\alpha}{2}}$ is $(1-\frac{\alpha}{2})$-quantile of the standard normal distribution.

One special case is when one wants to test for $\theta_1 = 0$, in other words, for no trend in the intensity function. Then, under the null hypothesis, the marginal distribution of ordered arrival times from expression (1.6) is, according to Cox and Lewis [1966], uniform on the interval $(0, T)$. Consequently, the sum of the arrival times $S = \sum T_i$ is distributed as the sum of n independent random variables having uniform distribution on $(0, T)$. Hence the distribution of the variable

$$Z_u = \frac{S - \frac{nT}{2}}{T\sqrt{\frac{n}{12}}}$$

converges fast to the standard normal distribution. Due to that fact we shall be checking for the value of the test statistic

$$z_u = \frac{\frac{\sum t_i}{n} - \frac{T}{2}}{T\sqrt{\frac{1}{12n}}}. \tag{2.4}$$

The null hypothesis should be then rejected at the level of significance α when

$$|z_u| \geq c_{1-\frac{\alpha}{2}}.$$

As soon as the value of θ_1 has been defined, we can move to the estimation of θ_0. Unlike the case of θ_1, we consider the joint likelihood function defined by (2.2). Its logarithm then takes form

$$\ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta}) = n\theta_0 + \theta_1 \sum_{i=1}^{n} t_i - \frac{e^{\theta_0}\,[e^{\theta_1 T} - 1]}{\theta_1}.$$

In order to find the maximum likelihood estimate of θ_0, we solve the equation

$$\frac{\partial \ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta})}{\partial \theta_0} = n - \frac{e^{\theta_0}\,[e^{\hat{\theta}_1 T} - 1]}{\hat{\theta}_1} = 0, \tag{2.5}$$

where $\hat{\theta}_1$ is the already obtained maximum likelihood estimate of θ_1, as described above. As a result, we arrive at

$$\hat{\theta}_0 = \log n + \log \hat{\theta}_1 - \log[e^{\hat{\theta}_1 T} - 1].$$

Now let us move to the problem of finding initial estimates of the parameters of a more general rate function.

2.2 Computing initial estimates of the parameters

In this section we consider a more general family of such rate functions, which contain a polynomial component describing a possible global trend as well as trigonometric components for explaining periodicity in the event occurrence. The methods described here are fully based on Kuhl et al. [1997].

We shall label the considered rate function EPTMP, meaning exponential-polynomial-trigonometric rate function having multiple periodicities. It is of the following form:

$$\lambda(t) = \exp\left\{\sum_{i=0}^{m} \theta_i t^i + \sum_{k=1}^{p} \gamma_k \sin(\omega_k t + \phi_k)\right\} \stackrel{\text{def}}{=} \exp\{h(t;\, m,\, p,\, \boldsymbol{\theta})\}, \tag{2.6}$$

where

$$\boldsymbol{\theta} = [\theta_0, \theta_1, \ldots, \theta_m, \gamma_1, \ldots, \gamma_p, \phi_1, \ldots, \phi_p, \omega_1, \ldots, \omega_p]$$

is the vector of parameters we want to estimate. Its first $m+1$ terms, that is, $\theta_0, \theta_1, \ldots, \theta_m$, represent the global trend in occurrence of events over time. Components of the periodic part can be understood as follows. Parameters $\gamma_1, \ldots, \gamma_p$ stand for amplitudes of oscillation, $\phi_1, \ldots, \phi_p$ describe phases (possible time shifts in the periodic components) and, finally, $\omega_1, \ldots, \omega_p$ represent frequencies. In general, the frequency of a periodic component represents the number of cycles during some chosen unit of time. For example, if we choose year as a time unit, frequency can be monthly, quarterly, semi-annually or annually (where the latter stands basically for no periodic cycles).

Frequencies are either known from the origin of the analysed data, or could be estimated along with other parameters using maximum likelihood approach. However, initial estimates of frequencies should be provided. This can be done by analysing the periodogram of historical payments. According to Lee et al. [1991], if a process indicates cycling behavior, its periodogram should contain peaks in the neighborhood of the corresponding frequency points, regardless of the possible long-term trend.

Suppose that we have a realization of a NHPP consisting of n events which occurred in the fixed time interval $(0, T]$ with observed arrival times $t_1 < t_2 < \ldots < t_n$. The log-likelihood function of the parameter vector $\boldsymbol{\theta}$ then takes form

$$\ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta}) = \sum_{i=0}^{m} \theta_i \sum_{j=1}^{n} t_j^i + \sum_{k=1}^{p} \sum_{j=1}^{n} \gamma_k \sin(\omega_k t_j + \phi_k) - \int_0^T \exp\{h(u;\, m,\, p,\, \boldsymbol{\theta})\} \mathrm{d}u.$$

Elements of $\boldsymbol{\theta}$ can be estimated by conditioning the log-likelihood function on a fixed value of m, that is, on the degree of the polynomial component. Likelihood equations should be solved for several values of m, so that we get the set of remaining parameters for each value of m. The most suitable set along with the value of m can be chosen using likelihood ratio test which will be presented later.

Initial estimates of the periodic components

Suppose that we have at our disposal either initial estimates or known values of the frequencies. Furthermore, we temporarily assume that there is no long-term evolutionary trend in the interval $(0, T]$, so the log-likelihood function reduces to

$$\ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta}) = n\theta + \sum_{k=1}^{p}\sum_{j=1}^{n} \gamma_k \sin(\omega_k t_j + \phi_k) \\ - \int_0^T \left[\prod_{k=1}^{p} \exp\left\{\theta/p + \gamma_k \sin(\omega_k u + \phi_k)\right\}\right] \mathrm{d}u, \quad (2.7)$$

where we define $\theta = \theta_0$ for simplicity.

Kuhl et al. [1997] consider that we can obtain good initial estimates if we estimate the parameters of each periodic component independently. Authors also present a useful approximation of the right-hand side of expression (2.7), which plays a key role in the presented method. This approximation can be described as

$$\int_0^T \left[\prod_{k=1}^{p} \exp\left\{\theta/p + \gamma_k \sin(\omega_k u + \phi_k)\right\}\right] \mathrm{d}u \\ \approx \mathrm{e}^{\theta} T^{-(p-1)} \prod_{k=1}^{p} \left[\int_0^T \exp\left\{\gamma_k \sin(\omega_k u + \phi_k)\right\} \mathrm{d}u\right]. \quad (2.8)$$

Reasoning behind this approximation can be found in Kuhl et al. [1997] (Appendix A). The considered approximation is suitable in cases of rather small number of periodic components p. For larger values of p the error of the approximation is compounded. Furthermore, authors state that the approximation can be poor in case when inequality $\theta/p \gg \gamma_k$ does not hold for some k, that is, if there exists some k for which θ/p is not much larger than γ_k.

If we denote the length of the cycle by $\mathcal{L}$ and consider only observations forming an interval, say $(0, T^c]$, consisting of complete cycles, we can notice that $T^c = \nu\mathcal{L}$, where ν stands for the number of complete cycles. One can also note that frequencies ω_k can be rewritten as $\omega_k = 2\pi/\mathcal{L}$. Moreover, since the integrand in each integral on the right-hand side of (2.8) has the came cyclic behavior over each subinterval $[(j-1)\mathcal{L},\, j\mathcal{L}]$ for $j = 1,\, 2,\, \ldots,\, \nu$, we can derive the following:

$$\int_0^{T^c} \exp\left\{\gamma_k \sin(\omega_k u + \phi_k)\right\} \mathrm{d}u = \nu \int_0^{\mathcal{L}} \exp\left\{\gamma_k \sin\left(\frac{2\pi}{\mathcal{L}} u + \phi_k\right)\right\} \mathrm{d}u \\ = \nu \int_0^{2\pi} \exp\left\{\gamma_k \sin(\zeta + \phi_k) \frac{\mathcal{L}}{2\pi}\right\} \mathrm{d}\zeta = \frac{T^c}{2\pi} \int_0^{2\pi} \exp\left\{\gamma_k \cos(\zeta)\right\} \mathrm{d}\zeta \\ = T^c \cdot I_0(\gamma_k),$$

where we performed the substitution $\zeta = \omega u = 2\pi u/\mathcal{L}$. The last two equalities are based on so-called modified Bessel function of the first kind, which is defined as

$$I_n(\gamma_k) = \frac{1}{\pi} \int_0^{\pi} \exp[\gamma_k \cos\zeta] \cos(n\zeta) \mathrm{d}\zeta.$$

More information about this function and its properties can be found e.g. in Abramowitz and Stegun [1965].

Hence the approximation shown in (2.8) can be rewritten as

$$\int_0^{T^c} \left\{ \prod_{k=1}^{p} \exp[\theta/p + \gamma_k \sin(\omega_k u + \phi_k)] \right\} \mathrm{d}u \approx \mathrm{e}^{\theta} T^c \prod_{k=1}^{p} I_0(\gamma_k).$$

The log-likelihood function (2.7) would then take the following form

$$\ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta}) \approx n\theta + \sum_{k=1}^{p} \gamma_k \sin(\phi_k) A(\omega_k) + \sum_{k=1}^{p} \gamma_k \cos(\phi_k) B(\omega_k) - \mathrm{e}^{\theta} T^c \prod_{k=1}^{p} I_0(\gamma_k), \tag{2.9}$$

where

$$A(\omega_k) = \sum_{j=1}^{n} \cos(\omega_k t_j), \quad B(\omega_k) = \sum_{j=1}^{n} \sin(\omega_k t_j), \quad k = 1,\, 2,\, \ldots,\, p. \tag{2.10}$$

To obtain initial estimates of the parameters, we have to compute partial derivatives of the approximate log-likelihood function (2.9) with respect to each parameter and solve the following equations:

$$\frac{\partial \ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta})}{\partial \theta} = 0;$$

$$\frac{\partial \ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta})}{\partial \gamma_k} = 0;$$

$$\frac{\partial \ell(\mathbf{t},\, n \,|\, \boldsymbol{\theta})}{\partial \phi_k} = 0$$

for $k = 1,\, 2,\, \ldots,\, p$. Then, according to Kuhl et al. [1997], we obtain the following initial estimates for ϕ_k and γ_k:

$$\hat{\phi}_k = \arctan\left[\frac{A(\omega_k)}{B(\omega_k)}\right], \quad k = 1,\, 2,\, \ldots,\, p,$$

and $\hat{\gamma}_k$ is a solution of

$$\frac{I_1(\hat{\gamma}_k)}{I_0(\hat{\gamma}_k)} = \frac{\sqrt{A^2(\omega_k) + B^2(\omega_k)}}{n_k}, \quad k = 1,\, 2,\, \ldots,\, p,$$

where n_k is the number of events in the interval $(0, \lfloor \omega_k T^c/\{2\pi\} \rfloor \cdot 2\pi/\omega_k]$, and $\lfloor x \rfloor$ stands for the floor function. In addition, the upper limit n of the summations in (2.10) is interchanged for n_k. One can note that for amplitude and phase parameters their initial estimates depend only on the frequencies, respectively, on their known values or initial estimates.

Initial estimates of the trend components

Kuhl et al. [1997] consider a generalized version of the moment matching procedure in order to compute initial estimates of the trend components. They state that the first $m+1$ moments of rate function (2.6) over the interval $(0,T]$ have the form

$$\int_0^T u^i \lambda(u) \mathrm{d}u = \int_0^T u^i \exp\{h(u;\, m,\, p,\, \boldsymbol{\theta})\} \mathrm{d}u$$

for $i = 0,\, 1,\, \ldots,\, m$. Taking the partial derivative of the log-likelihood function with respect to θ_i, setting it to zero and solving for $\sum_{j=1}^{n} t_j^i$ yields

$$\sum_{j=1}^{n} t_j^i = \int_0^T u^i \exp\{h(u;\, m,\, p,\, \boldsymbol{\theta})\} \mathrm{d}u \tag{2.11}$$

for $i = 0,\, 1,\, \ldots,\, m$.

The moment-matching procedure finds the coefficients $\{c_j : j = 0,\, 1,\, \ldots,\, m\}$ of some polynomial $\sum_{j=0}^{m} c_j u^j$ of degree m whose first $m+1$ moments are equal to those of $\exp\{h(u;\, m,\, p,\, \boldsymbol{\theta})\}$ on the interval $(0,T]$. Equation (2.11) implies that we should consider the following system of equations in order to find initial estimates of the coefficients $\{c_j \,:\, j = 0,\, 1,\, \ldots,\, m\}$:

$$\sum_{j=1}^{n} t_j^i = \int_0^T u^i \left(\sum_{j=0}^{m} c_j u^j \right) \mathrm{d}u = \sum_{j=0}^{m} \frac{c_j T^{i+j+1}}{i+j+1}$$

for $i = 0,\, 1,\, \ldots,\, m$.

Kuhl et al. [1997] introduce the following procedure for finding initial estimates of θ_i for $i = 0,\, 1,\, \ldots,\, m$. Authors consider matching the moments of the function $\log(\sum_{j=1}^{m} c_j u^j)$ to the ones of $h(u;\, m,\, p,\, \boldsymbol{\theta})$ over the interval $(0,T]$. This consideration leads us to the system

$$\int_0^T u^i \log\left(\sum_{j=0}^{m} c_j u^j \right) \mathrm{d}u = \int_0^T u^i \left[\sum_{j=0}^{m} \theta_j u^j + \sum_{k=1}^{p} \gamma_k \sin(\omega_k u + \phi_k) \right] \mathrm{d}u \tag{2.12}$$

for $i = 0,\, 1,\, \ldots,\, m$. Note that the initial estimates of the parameters of the periodic components $\{\gamma_k,\, \omega_k,\, \phi_k \,:\, k = 1,\, \ldots,\, p\}$ should be used to find the solution of (2.12).

All of the initial estimates considered in this section provide a fairly good starting point for numerical algorithms solving the likelihood equations. Nevertheless, it cannot assure that the algorithm will not diverge.

It is essential to note that this procedure can be used to find initial estimates of the parameters of degree-two exponential polynomial rate function as well. This type of a rate function is described in Section 3.4 of Chapter 3.

Choice of degree m of the polynomial representing the global trend

In this part we dedicate ourselves to finding the final estimates of all parameters along with the choice of degree m of the polynomial component representing the global trend. As mentioned above, one should firstly compute sets of parameter estimates conditionally on several different values of m. Let us denote the set of estimates corresponding to fixed degree m as $\hat{\boldsymbol{\theta}}_m$. We consider an extension of the likelihood ratio test in order to find the most suitable m and $\hat{\boldsymbol{\theta}}_m$. We test for the null hypothesis

$$H_0 : m \text{ is the true degree of (2.6)}$$

against the alternative

$$H_0 : m+1 \text{ is the true degree of (2.6).}$$

We shall be using the following test statistic

$$Z_m = 2\left[\ell(\mathbf{t},\, n \,|\, \hat{\boldsymbol{\theta}}_{m+1}) - \ell(\mathbf{t},\, n \,|\, \hat{\boldsymbol{\theta}}_m)\right] \overset{H_0}{\sim} \chi^2_1 \text{ as } T \to \infty.$$

The null hypothesis is rejected when $Z_m \geq \chi^2_1(1-\alpha)$, where $\chi^2_1(1-\alpha)$ is $(1-\alpha)$-quantile of χ^2 distribution with one degree of freedom, and α is the desired significance level. The test should be repeated until one finds the smallest m for which the null hypothesis is not rejected. The vector of final estimates of the parameters $\hat{\boldsymbol{\theta}}_m$ is then determined by the final value of m.

We have discussed intensity estimation in a NHPP using maximum likelihood method as well as methods suitable for computing initial estimates of the parameters. Now we move to the presentation of the simulation methods for a NHPP as well as comparison of their efficiency.

3. Simulation methods for a Non-Homogeneous Poisson process

In this chapter we describe several commonly used methods for simulation of a NHPP as well as perform a comparison of these algorithms with respect to their efficiency. Some of these methods can be used for a general NHPP, others are designed for specific models of the intensity function. In the practical part of this thesis we shall use the discussed methods for modelling payments on bodily injury claims caused by uninsured drivers.

3.1 Time-scale transformation of a NHPP

The algorithm considered in this section is analogical to inverse transform method for simulation of continuous non-uniform random variables. It is, as Klein and Roberts [1984] state, an exact method of simulation a NHPP. According to Çinlar [2013], Theorem 1 is used to transform a NHPP into a HPP with rate 1. However, the algorithm requires simulating of a random variable from the uniform distribution on $(0, 1)$ in each step.

Firstly, we construct time inverse $\tau(t)$ of expectation function $\mu(t)$ as defined by expression (1.2). If $\mu(t)$ is strictly monotone, it implies then that it is also invertible. Function $\tau(t)$ would be simply an inverse of $\mu(t)$ in such case. If one cannot assume strict monotonicity of $\mu(t)$, calculation of $\tau(t)$ could be challenging and numerical methods have to be involved in some cases.

In the next step we are going to simulate times between events in a HPP with rate 1: if we already know arrival time t^*_{i-1} of the previous event, then adding the simulated time to the next event to t^*_{i-1} would give us exactly the arrival time of the i-th event, that is, t^*_i. Thereafter, we transform the time to the event i of a HPP into the one of a NHPP via the inverse of $\mu(t)$. From the general properties of a HPP we know that times between events are independent random variables having exponential distribution with parameter λ, in our case with parameter 1. Let us denote the time between $(i-1)$-th and i-th event in a HPP with rate 1 by S_i. Then, using the general inverse transform method for generating a continuous non-uniform random variable, we can state that

$$F_{S_i}(s_i) = 1 - \mathrm{e}^{-s_i} = u,$$

where u is a random variable from the uniform distribution on $(0, 1)$. After several algebraic operations, we arrive at

$$s_i = -\log(1-u)$$

Then, given that the time to the event $i-1$ is t^*_{i-1}, we can find the time to the event i by $t^*_i = t^*_{i-1} + s_i$, which can be further transformed to the time to the i-th event t_i of a NHPP via expression (1.2).

Now let us assume that we want to simulate a NHPP with expectation function $\mu(t)$ in the fixed interval $(0, T]$. To summarize the steps stated above, we describe the simulation algorithm using time-scale transformation in the following way:

1. Set $s_0 = 0$, $t_0^* = 0$, $t_0 = 0$, $i = 0$.

2. Simulate a variable u_{i+1} from the uniform distribution on $(0, 1)$.

3. Set $s_{i+1} = -\log(1 - u_{i+1})$.

4. If $t_i^* + s_{i+1} > \mu(T)$, go to step 6.

5. Otherwise, set $t_{i+1}^* = t_i^* + s_{i+1}$, $t_{i+1} = \tau(t_{i+1}^*)$ and $i = i + 1$. Go to step 2.

6. Return t_1, $t_2, \ldots, t_n$, where $n = i$, and also n.

The number of events simulated in $(0, T]$ would be then the current value of i before the exit. The algorithm is relatively simple to apply; however, inverse $\tau(t)$ could be computationally challenging, depending on the form of $\mu(t)$. Furthermore, the calculation time of $\mu(t)$ depends on the form of $\lambda(t)$ as well, as we have to integrate the rate function in this case.

3.2 Thinning of a NHPP

The algorithm described in this section is rather simple and can be applied to any form of rate function $\lambda(t)$. It is fully based on Lewis and Shedler [1979]. Firstly, we should find some Poisson process whose rate function dominates the given rate function on the whole time interval. Afterwards, we delete certain points of such process according to the particular criterion. The method is based on the following theorem.

Theorem 6. *(Thinning of a NHPP)*
Consider a NHPP $\{N^*(t), t \geq 0\}$ *with rate function* $\lambda^*(t)$ *and expectation function* $\mu^*(t) = \Lambda^*(t) - \Lambda^*(0)$. *Let* T_1^*, $T_2^*, \ldots, T_{N*(T)}^*$ *be the arrival times of* $\{N^*(t)\}$ *in the fixed interval* $(0, T]$. *Suppose that for every* t *satisfying* $0 \leq t \leq T$ *it holds that* $\lambda(t) \leq \lambda^*(t)$. *For* $i = 1$, $2, \ldots, N^*(T)$, *delete the point* T_i^* *with probability* $1 - \lambda(T_i^*)/\lambda^*(T_i^*)$. *Then the remaining points form a NHPP* $\{N(t), t \geq 0\}$ *with rate function* $\lambda(t)$ *in the time interval* $(0, T]$.

The proof of this theorem can be found again in Lewis and Shedler [1979].

Let us now describe the general algorithm of simulation of a NHPP using thinning in the fixed interval $(0, T]$. Its steps are the following:

1. Generate the arrival times of a NHPP $\{N(t)^*\}$ with rate function $\lambda^*(t)$ in the interval $(0, T]$. If the number of generated points n^* equals 0, exit the algorithm $\rightarrow$ there are no events of the process $\{N(t)\}$ in $(0, T]$.

2. Denote the generated *ordered* arrival times by t_1^*, $t_2^*, \ldots, t_{n^*}^*$. Set $i = 1$ and $k = 0$.

3. Generate a variable u_i from the uniform distribution on $(0, 1)$.
If $u_i \leq \lambda(t_i^*)/\lambda^*(t_i^*)$, set $k = k + 1$ and $t_k = t_i^*$.

4. Set $i = i + 1$. If $i \leq n^*$, go to step 3.

5. Return t_1, $t_2, \ldots, t_n$, where $n = k$, and also n.

The main source of time inefficiency of the algorithm is the computation of $\lambda(t)$. In case when $\{N^*(t)\}$ is a HPP with rate function $\lambda^*(t) = \lambda^*$ and the minimum of $\lambda(t)$, denoted by $\tilde{\lambda}$, is known, one can note that t_i^* is always accepted if $u \leq \tilde{\lambda}/\lambda^*$. This could speed up the computations, as in some cases one would not need to calculate $\lambda(t)$.

The simplest form of the method of thinning, i.e. if we choose $\lambda^*(t)$ according to $\lambda^*(t) = \lambda^*$, where $\lambda^* \geq \max_{0 \leq t \leq T} \lambda(t)$, can be used to simulate a NHPP on an interval-by-interval basis. In this case the process $\{N^*(t)\}$ is actually a HPP. The method is based on the fact that times between events in a HPP are independent random variables which are exponentially distributed with parameter λ^*. If we consider generation of a NHPP with rate function $\lambda(t)$ in the fixed interval $(0, T]$, the steps of such simplified algorithm would be the following:

1. Set $t_0 = 0$, $t_0^* = 0$, $E_0^* = 0$, $i = 0$, $k = 0$.

2. Generate a variable E_{i+1}^* from the exponential distribution with parameter λ^* and a variable u_{i+1} from the uniform distribution on $(0, 1)$.

3. If $t_i^* + E_{i+1}^* > T$, go to step 5. Otherwise, set $t_{i+1}^* = t_i^* + E_{i+1}^*$.

4. If $u_{i+1} \leq \lambda(t_{i+1}^*)/\lambda^*$, set $k = k + 1$, $t_k = t_{i+1}^*$ and $i = i + 1$. Go to step 2.

5. Return t_1, $t_2, \ldots, t_n$, where $n = k$, and also n.

The method of thinning described in this section can be applied to any rate function without the necessity of numerical integration or simulation of Poisson variables, which, however, diminishes efficiency of the algorithm. It can be used for more complex rate functions, in case of which numerical integration or inverse of $\lambda(t)$ could be very challenging. In order to increase efficiency, one should choose $\lambda^*(t)$ as close as possible to $\lambda(t)$, taking into account also the difficulty of generating the process $\{N^*(t)\}$.

3.3 NHPP with log linear rate function

In this section we consider rate functions having the particular form, which was already mentioned in the previous chapter. In this case the rate function is of the form $\lambda(t) = \exp\{\theta_0 + \theta_1 t\} \stackrel{\text{def}}{=} \tilde{\theta}_0 \, \mathrm{e}^{\theta_1 t}$. The natural logarithm of the considered rate function is linear in the parameters, hence the name "log linear rate function". General properties of such rate function were described in Chapter 2. This section is fully based on Lewis and Shedler [1976].

Considering the process in a fixed interval $(0, T]$, the following relationship for its expectation function holds

$$\mu(t) = \Lambda(t) - \Lambda(0) = \int\limits_0^t \lambda(u) \mathrm{d}u = \begin{cases} \tilde{\theta}_0(\mathrm{e}^{\theta_1 t} - 1)/\theta_1, & \theta_1 \neq 0, \\ \tilde{\theta}_0 t, & \theta_1 = 0. \end{cases}$$

In this thesis we restrict ourselves to the case $\theta_1 \neq 0$ only. Hence the variable $N(T)$ indicating the number of events which occurred in the interval $(0, T]$ is Poisson distributed with parameter $\mu(T) = \tilde{\theta_0}(e^{\theta_1 T} - 1)/\theta_1$. According to Theorem 5, the arrival times of such process, under the condition that exactly n events occurred in $(0, T]$, are distributed as order statistics from the distribution

$$F(t) = \frac{\Lambda(t) - \Lambda(0)}{\Lambda(T) - \Lambda(0)} = \frac{e^{\theta_1 t} - 1}{e^{\theta_1 T} - 1}, \quad 0 \leq t \leq T, \quad \theta_1 \neq 0.$$

For $\theta_1 \neq 0$ we can also invert $F(t)$ and get

$$t = F^{-1}(p) = \log\left[1 + p(e^{\theta_1 T} - 1)\right]/\theta_1, \quad 0 \leq p \leq 1.$$

Because the inverse of $F(t)$ is known explicitly, one way of simulating such NHPP is to use the inverse probability transform method directly. The algorithm can be described in the following way.

Algorithm 1 ($\theta_1 \neq 0$)

1. Generate a variable n from the Poisson distribution with parameter $\mu(T) = \tilde{\theta_0}(e^{\theta_1 T} - 1)/\theta_1$. If n equals 0, exit $\rightarrow$ there are no events in $(0, T]$.
2. Otherwise, generate n variables from the uniform distribution on $(0, 1)$ and order them to get $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$.
3. Calculate $\log\left[1 + U_{(i)}(e^{\theta_1 T} - 1)\right]/\theta_1$ for all $i = 1, \ldots, n$ and set

$$T_1 = \log\left[1 + U_{(1)}(e^{\theta_1 T} - 1)\right]/\theta_1, \;\; T_2 = \log\left[1 + U_{(2)}(e^{\theta_1 T} - 1)\right]/\theta_1, \ldots$$

4. Return $T_1, T_2, \ldots, T_n$ and n.

The described algorithm requires generation of one Poisson variable, n uniform variables and their subsequent ordering as well as calculation of n logarithms.

Compared to the algorithm which uses time-scale transformation of a NHPP, the one described above is considered by Lewis and Shedler [1976] to be more efficient. However, the method using gap statistics, which is discussed further, is considered to be even more efficient than the latter two.

The next method we are going to describe is designed specially for the log linear family of rate functions and does not require ordering of the generated variables. Let us consider the case $\theta_1 < 0$ in the next part, the case of $\theta_1 > 0$ will be described later. The simulation algorithm for the fixed time interval $(0, T]$ is based on the following two theorems.

Theorem 7. *(Gap statistics for exponential distribution)*
Let $Y_1, Y_2, \ldots, Y_m$ be independent random variables having exponential distribution with parameter β with order statistics $Y_{(1)}, Y_{(2)}, \ldots, Y_{(m)}$. Let us define the gap statistics as $D_1 = Y_{(1)}$, $D_2 = Y_{(2)} - Y_{(1)}, \ldots, D_m = Y_{(m)} - Y_{(m-1)}$. Then the gap statistics are independent random variables having exponential distribution with means $\mathsf{E}\,(D_i) = \beta/(m + 1 - i)$, $i = 1, \ldots, m$.

The proof can be found e.g. in Cox and Lewis [1966].

Theorem 8. *(The gap process)*
Let m be a realization of a random variable M having Poisson distribution with parameter $\mu^* = -\tilde{\theta}_0/\theta_1$, and let us set $\beta = -\theta_1 > 0$. Then the gap process is a NHPP with rate function $\lambda(t) = \tilde{\theta}_0\, \mathrm{e}^{\theta_1 t}$ on $(0, \infty)$.

A computation showing the relations between the parameters is shown in Lewis and Shedler [1976]. The main result can be summarized in the following algorithm.

Algorithm 2a) $(\theta_1 < 0)$

1. Set $t_0 = 0$, $E_0 = 0$, $i = 0$.
2. Generate a random variable m from the Poisson distribution with parameter $\mu = -\tilde{\theta}_0/\theta_1$. If m equals 0, exit $\rightarrow$ there are no events in $(0, T]$.
3. Generate a random variable E_{i+1} from exponential distribution with rate parameter 1.
4. If $E_{i+1}/[\beta(m-i)] + t_i > T$, where $\beta = -\theta_1$, go to step 5. Otherwise, set $t_{i+1} = E_{i+1}/[\beta(m-i)]$, $i = i+1$ and go to step 3.
5. Return t_1, $t_2, \ldots,$ t_n, where $n = i$, and n.

To summarize, $n \leq m$ events are generated and n is Poisson distributed with parameter $\mu = \tilde{\theta}_0(\mathrm{e}^{\theta_1 T} - 1)/\theta_1$, as stated in Lewis and Shedler [1976]. Authors also note that *Algorithm 2*, unlike *Algorithm 1*, can use fast exponential generators and requires neither ordering of uniform variables nor logarithms calculation.

Now let us adjust *Algorithm 2* for the case $\theta_1 > 0$ using the so-called time-reversal technique. Let us denote by $\xi(t)$ the time measured backwards from T, i.e. $\xi(t) = T - t$ for all $0 \leq t \leq T$. Then $N(\xi(t))$, that is, the number of events in $(T - \xi(t), T]$, or, equivalently, in $(t, T]$, is Poisson distributed with the mean $\Lambda(T) - \Lambda(T - \xi(t))$.

The rate function can be then derived as follows:

$$\lambda^*(\xi(t)) = \frac{\mathrm{d}}{\mathrm{d}\xi(t)}[\Lambda(T) - \Lambda(T - \xi(t))] = \lambda(T - \xi(t)) = \exp\{\theta_0 + \theta_1 T + (-\theta_1)\xi(t)\}$$
$$\stackrel{\text{def}}{=} \tilde{\theta}_0^{\,*}\, \mathrm{e}^{\theta_1^* \xi(t)}.$$

Note that the coefficient of $\xi(t)$ is negative. Hence the adjusted algorithm can be described as follows.

Algorithm 2b) ($\theta_1 > 0$)

1. Set $t_0 = 0$, $t_0^* = 0$, $E_0^* = 0$, $i = 0$.
2. Generate a random variable m from Poisson distribution with parameter $\mu^* = -\tilde{\theta_0}^*/\theta_1^*$, where $\tilde{\theta_0}^* = \exp\{\theta_0 + \theta_1 T\}$ and $\theta_1^* = -\theta_1$. If m equals 0, exit $\rightarrow$ there are no events in $(0, T]$.
3. Generate a random variable E_{i+1}^* from exponential distribution with rate parameter 1.
4. If $E_{i+1}^*/[\beta(m-i)] + t_i^* > T$, where $\beta = -\theta_1^* = \theta_1$, go to step 5. Otherwise, set $t_{i+1}^* = E_{i+1}/[\beta(m-i)]$, $i = i+1$ and go to step 3.
5. Calculate $t_j = T - t_{i-j+1}^*$ for all $j = 1, \dots, i$.
6. Return $t_1, t_2, \dots, t_n$, where $n = i$, and n.

Algorithm 2a) and *Algorithm 2b)* need, as stated in Lewis and Shedler [1976], only one Poisson variable and, on average, $\Lambda(T) - \Lambda(0) + 1$ exponential variables. On the contrary, *Algorithm 1* needs one Poisson variable, an ordering and, on average, $\Lambda(T) - \Lambda(0)$ logarithms and uniform variables. Thus authors consider *Algorithms 2a)* and *2b)* to be almost about twice as fast as *Algorithm 1* or time-scale transformation via the inverse of $\mu(t)$.

3.4 NHPP with degree-two exponential polynomial rate function

In this section we consider an extension of log linear rate function which is basically log quadratic, meaning we are going to analyse the rate function of the type

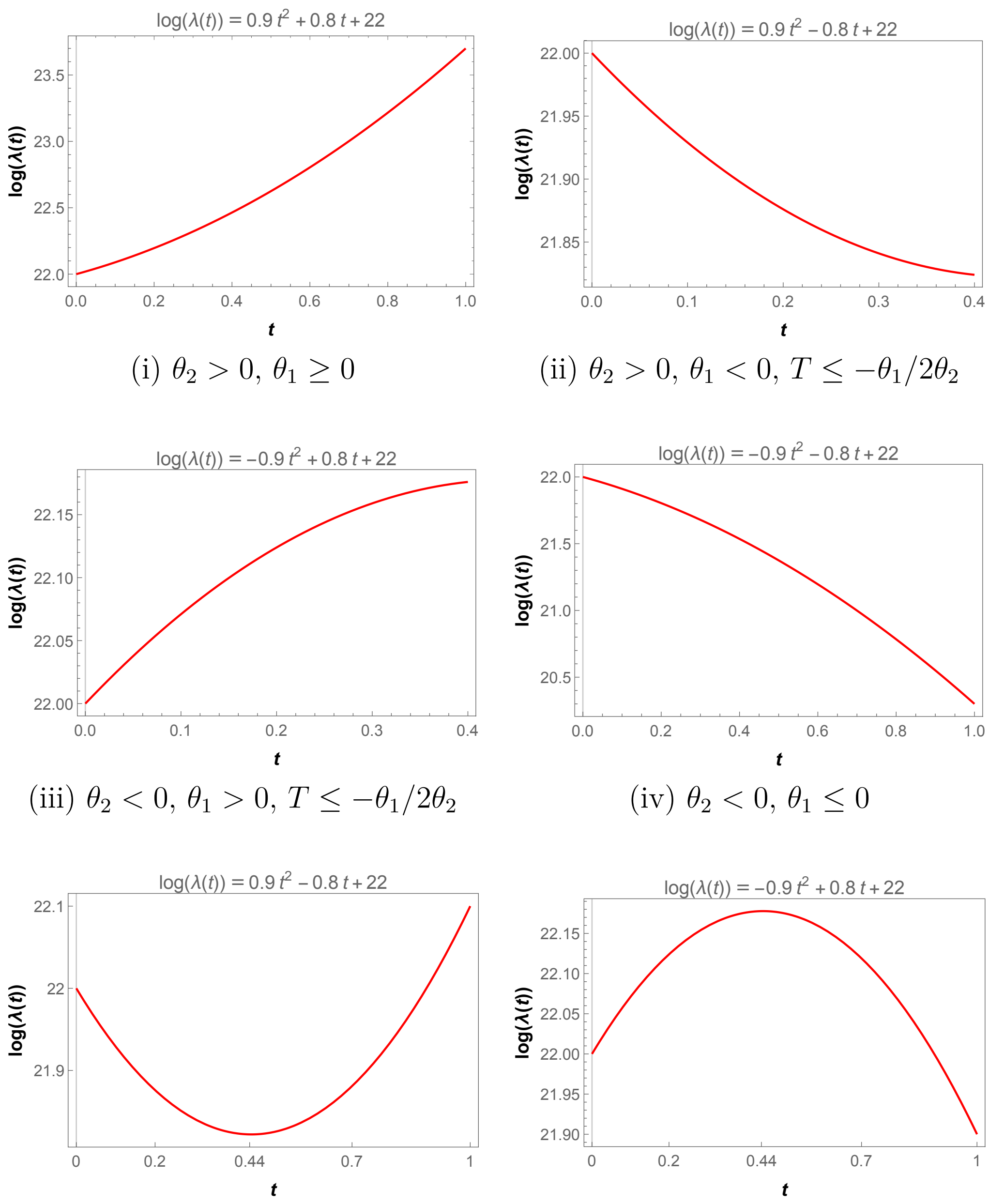
$$\lambda(t) = \exp\{\theta_0 + \theta_1 t + \theta_2 t^2\}. \tag{3.1}$$

Such rate functions allow for events whose arrival rate changes direction within time, i.e. it can either increase until it reaches the peak and after that monotonically decrease, or vice versa. Again, exponentiation ensures that the rate function is always positive.

The method we are going to describe is examined in Lewis and Shedler [1979] and is based on decomposition of a NHPP with rate function (3.1) into two NHPPs, i.e. we decompose rate (3.1) in the following way

$$\lambda(t) = \tilde{\lambda}(t) + (\lambda(t) - \tilde{\lambda}(t)) = \tilde{\lambda}(t) + \lambda^\star(t), \quad a < t \leq b. \tag{3.2}$$

Hence the analyzed process $\{N(t)\}$ with rate function $\lambda(t)$ is decomposed as $N(t) = \tilde{N}(t) + N^\star(t)$, where rate functions $\tilde{\lambda}(t)$, resp. $\lambda^\star(t)$ correspond to the processes $\tilde{N}(t)$, resp. $N^\star(t)$. The reasoning behind the consideration of an interval $(a, b]$ instead of $(0, T]$ is that in case of the changing direction rate function interval $(0, T]$ has to be decomposed into two disjoint intervals. In Figure 3.1 all possible cases of the direction of the logarithm of rate function (3.1) are displayed,



(i) $\theta_2 > 0,\ \theta_1 \geq 0$

(ii) $\theta_2 > 0,\ \theta_1 < 0,\ T \leq -\theta_1/2\theta_2$

(iii) $\theta_2 < 0,\ \theta_1 > 0,\ T \leq -\theta_1/2\theta_2$

(iv) $\theta_2 < 0,\ \theta_1 \leq 0$

(v) $\theta_2 > 0,\ \theta_1 < 0,\ T > -\theta_1/2\theta_2$

(vi) $\theta_2 < 0,\ \theta_1 > 0,\ T > -\theta_1/2\theta_2$

Figure 3.1: *Function* $\log(\lambda(t))$ *displayed for all possible cases of direction of* $\lambda(t)$.

depending on the signs of θ_1 and θ_2 as well as on the value of T. In case of the changing direction rate function the inflexion point equals $-\theta_1/2\theta_2$.

One can note that by design of (3.2), $\tilde{\lambda}(t) \leq \lambda(t)$ on the interval $(a, b]$, and $\tilde{\lambda}(t)$ has the form

$$\tilde{\lambda}(t) = \exp\{\gamma_0 + \gamma_1 t\}, \quad \gamma_1 \neq 0,\ a < t \leq b,$$

where γ_0 and γ_1 are chosen in the way that the mean number of events in $(a, b]$, $\tilde{\mu} = \tilde{\Lambda}(a) - \tilde{\Lambda}(b)$, is largest possible. According to Lewis and Shedler [1979], $\tilde{\mu}$ is maximized by maximizing the area under $\tilde{\lambda}(t)$ for $a < t \leq b$. Geometric considerations related to this problem are discussed in detail in Lewis and Shedler [1979]. In this thesis we reckon only the final suggestions of the authors which will be summarized in separate tables.

The process $\{\tilde{N}(t)\}$ with rate $\tilde{\lambda}(t)$ can be generated efficiently by Algorithm 1 or 2 described in Section 3.3. Regarding the process $\{N^\star(t)\}$, Lewis and Shedler [1979] consider the method of thinning for the simulation of this process. This method has already been discussed in Section 3.2. Authors suggest to choose the majorizing rate $\lambda^*(t)$ according to $\lambda^*(t) = \lambda^*$, where $\lambda^* \geq \max_{a<t\leq b} \lambda^\star(t)$, to minimize the number of deleted points.

Let us now consider two separate simulation algorithms for the cases of the rate function with change of direction and without. These algorithms can be then summarized in the following steps:

Algorithm 1 (no change of the rate direction, i.e. cases (i)-(iv) in Figure 3.1)

1. Set γ_0 and γ_1 according to the corresponding case from Table 3.1. Set $a = 0$ and $b = T$.
2. Using one of the algorithms described in Section 3.3, simulate n arrival times of the process $\{\tilde{N}(t)\}$ with rate function $\tilde{\lambda}(t) = \exp\{\gamma_0 + \gamma_1 t\}$ on $(a, b]$ to obtain times
$$t'_1,\, t'_2,\, \ldots,\, t'_n.$$
3. Set λ^* according to
$$\lambda^* \geq \max_{a<t\leq b} \lambda^\star(t) = \max_{a<t\leq b} \Big[\exp\{\theta_0 + \theta_1 t + \theta_2 t^2\} - \exp\{\gamma_0 + \gamma_1 t\}\Big].$$
4. Using the algorithm described in Section 3.2, simulate m events of the process $\{N^\star(t)\}$ with majorizing rate λ^* on $(a, b]$ to obtain times
$$t''_1,\, t''_2,\, \ldots,\, t''_m.$$
5. If $n + m = 0$, exit $\rightarrow$ there are no events in $(0, T]$.
6. If $m = 0$, return $t'_1,\, t'_2,\, \ldots,\, t'_n$ as the required arrival times and exit.
7. If $n = 0$, return $t''_1,\, t''_2,\, \ldots,\, t''_m$ as the required arrival times and exit.
8. Otherwise, merge $t'_1 \leq t'_2 \leq \ldots \leq t'_n$ and $t''_1 \leq t''_2 \leq \ldots \leq t''_m$, order and return as the required arrival times, then exit.

Table 3.1: *Values of* γ_0 and γ_1 *for the rate without direction change*

	θ_1	θ_2	T	γ_0	γ_1
(i)	≥ 0	> 0	–	$\theta_0 - \theta_2 T^2$	$\theta_1 + 2\theta_2 T$
(ii)	< 0	> 0	$\leq -\theta_1/2\theta_2$	θ_0	θ_1
(iii)	> 0	< 0	$\leq -\theta_1/2\theta_2$	θ_0	$\theta_1 + \theta_2 T$
(iv)	≤ 0	< 0	–	θ_0	$\theta_1 + \theta_2 T$

Algorithm 2 (change of the rate direction, i.e. cases (v)-(vi) in Figure 3.1)

First part, t from the interval $(0, -\theta_1/2\theta_2]$

1. Set γ_0 and γ_1 according to the corresponding case (v.1) or (vi.1) from Table 3.2. Set $a = 0$ and $b = -\theta_1/2\theta_2$.

2. Using one of the algorithms described in Section 3.2, simulate n_1 arrival times of the process $\{\tilde{N}(t)\}$ with rate function $\tilde{\lambda}(t) = \exp\{\gamma_0 + \gamma_1 t\}$ on $(a, b]$ to obtain times
$$t'_1, t'_2, \ldots, t'_{n_1}.$$

3. Set λ^* according to
$$\lambda^* \geq \max_{a<t\leq b} \lambda^\star(t) = \max_{a<t\leq b} \left[\exp\{\theta_0 + \theta_1 t + \theta_2 t^2\} - \exp\{\gamma_0 + \gamma_1 t\}\right].$$

4. Using the algorithm described in Section 3.2, simulate m_1 events of the process $\{N^\star(t)\}$ with majorizing rate λ^* on $(a, b]$ to obtain times
$$t''_1, t''_2, \ldots, t''_{m_1}.$$

Second part, t from the interval $(-\theta_1/2\theta_2, T]$

5. Set γ_0 and γ_1 according to the corresponding case (v.2) or (vi.2) from Table 3.2. Set $a = -\theta_1/2\theta_2$ and $b = T$.

6. Using one of the algorithms described in Section 3.2, simulate n_2 arrival times of the process $\{\tilde{N}(t)\}$ with rate function $\tilde{\lambda}(t) = \exp\{\gamma_0 + \gamma_1 t\}$ on $(a, b]$ to obtain times
$$t'_{n_1+1}, t'_{n_1+2}, \ldots, t'_{n_2}.$$

7. Set λ^* according to
$$\lambda^* \geq \max_{a<t\leq b} \lambda^\star(t) = \max_{a<t\leq b} \left[\exp\{\theta_0 + \theta_1 t + \theta_2 t^2\} - \exp\{\gamma_0 + \gamma_1 t\}\right].$$

8. Using the algorithm described in Section 3.2, simulate m_2 events of the process $\{N^\star(t)\}$ with majorizing rate λ^* on $(a, b]$ to obtain times
$$t''_{m_1+1}, t''_{m_1+2}, \ldots, t''_{m_2}.$$

Merging part

9. If $n = n_1 + n_2 = 0$ and $m = m_1 + m_2 = 0$, exit $\rightarrow$ there are no events in $(0, T]$.

10. If $m = 0$, return $t'_1, t'_2, \ldots, t'_n$ as the required arrival times and exit.

11. If $n = 0$, return $t''_1, t''_2, \ldots, t''_m$ as the required arrival times and exit.

12. Otherwise, merge $t'_1 \leq t'_2 \leq \ldots \leq t'_n$ and $t''_1 \leq t''_2 \leq \ldots \leq t''_m$, order and return as the required arrival times, then exit.

Table 3.2: *Values of* γ_0 and γ_1 *for the rate with direction change*

	θ_1	θ_2	γ_0	γ_1	t
(v.1)	< 0	> 0	θ_0	θ_1	$t \in (0, \leq -\theta_1/2\theta_2]$
(v.2)	< 0	> 0	$\theta_0 - \theta_2 T^2$	$\theta_1 + 2\theta_2 T$	$t \in (-\theta_1/2\theta_2, T]$
(vi.1)	> 0	< 0	θ_0	$\theta_1/2$	$t \in (0, \leq -\theta_1/2\theta_2]$
(vi.2)	> 0	< 0	$\theta_0 + (\theta_1/2)T$	$(\theta_1/2) + \theta_2 T$	$t \in (-\theta_1/2\theta_2, T]$

3.5 Simulation of a NHPP with EPTMP rate function

Now let us consider the rate function of EPTMP type. Its general description along with the estimation of the parameters is described in detail in Section 2.2. As soon as the final estimates of the rate function parameters are obtained, we can proceed with the simulation of the events from the process. General algorithms such as described in Sections 3.1 and 3.2 can be used. Kuhl et al. [1997] consider the method of piecewise inversion.

Suppose that we are analysing a NHPP with rate function $\lambda(t \,|\, \boldsymbol{\theta})$ in the fixed interval $(0, T]$ and that we have already obtained the final estimate $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$. Let us denote the rate function containing the final parameter estimates by $\lambda(t \,|\, \hat{\boldsymbol{\theta}})$. Then the distribution function of the arrival time T_{i+1} of the next event conditionally on the observed arrival time $T_i = t_i$ of the previous event takes form

$$F_{T_{i+1}\,|\,T_i}(t \,|\, t_i) \equiv \mathsf{P}\left[T_{i+1} \leq t \,|\, T_i = t_i\right] = \begin{cases} 1 - \exp\left\{ - \int\limits_{t_i}^{t} \lambda(u \,|\, \hat{\boldsymbol{\theta}}) \mathrm{d}u \right\}, & t \geq t_i, \\ 0, & \text{otherwise.} \end{cases}$$

This yields that in order to generate a value t_{i+1} of the variable T_{i+1} given $T_i = t_i$ we should generate a variable U_{i+1} from the uniform distribution on $(0, 1)$ and compute

$$t_{i+1} = F^{-1}_{T_{i+1}\,|\,T_i}(U_{i+1} \,|\, t_i),$$

or, equivalently, solve the equation

$$\int\limits_{t_i}^{T_{i+1}} \lambda(u \,|\, \hat{\boldsymbol{\theta}}) \mathrm{d}u = -\log(1 - U_{i+1})$$

for T_{i+1}. Let us summarize the described procedure in the following algorithm.

1. Set $s_0 = 0$, $t_0 = 0$, $i = 0$.
2. Simulate a variable u_{i+1} from the uniform distribution on $(0, 1)$.
3. Solve $\int\limits_{t_i}^{s_{i+1}} \lambda(u \,|\, \hat{\boldsymbol{\theta}}) \mathrm{d}u = -\log(1 - u_{i+1})$ for s_{i+1}.
4. If $s_{i+1} > T$, go to step 6.

5. Otherwise, set $t_{i+1} = s_{i+1}$ and $i = i + 1$. Go to step 2.

6. Return t_1, $t_2, \dots$, t_n, where $n = i$, and also n.

Lee et al. [1991] consider also the method of piecewise thinning for simulation of a NHPP with rate function of EPTMP type, see the article for more details.

4. Application to non-life insurance data

In this chapter we are going to use the previously described methods to estimate the number of claim payments for the data from non-life insurance. Only exponential-polynomial models up to degree 2 are considered in the practical part due to their multipurpose use and appealing properties. Higher degrees of the polynomial are not considered due to complex estimation of the parameters as well as further simulation of the estimated process. Moreover, we do not assume any models with cycling components due to the lack of observations for judging about periodicity. Even if one disregards that, the analysis of the plots of the empirical autocorrelation function for different lags did not confirm presence of any periodicity in the data as there were no peak points on the plots; the points induced monotonic behavior. One could apply more sophisticated assumptions on the rate function; however, in practice, this should be done in cooperation with the claims department since they are able to provide useful information about every questionable claim itself as well as its settlement.

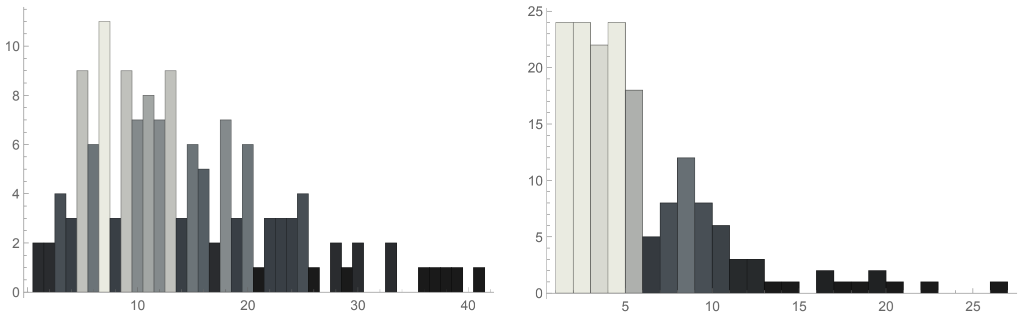
4.1 Data description

The data were provided by Czech Insurers' Bureau (CIB) and consist of incremental claim payments and RBNS reserves (reserves for reported but not settled claims) at 31.12.2015 for the claims occurred during years 2000 and 2015. The claims correspond to MTPL (Motor Third Party Liability) line of business, which is a part of non-life insurance business. These claims are handled by CIB due to the fact that they were caused by uninsured drivers. In spite of MTPL insurance is obligatory for any driver, there are still some without such policy. The purpose of CIB is to assure that the victim of the accident is paid in time, as in many cases payments for bodily injuries include expensive daily care, income loss as well as pain and suffering. Naturally, most of the drivers cannot afford paying such amounts to the victim in the short time interval.

CIB is itself financed by the insurers' contributions which are obligatory by the current legislation. Every Czech insurance company underwriting MTPL business must become a member of CIB. If contributions are not enough to meet the obligations of CIB, the organisation can request additional payments from its members. Non-proportional reinsurance programs can be also used by CIB in order to cover extreme claims. After the claim settlement or the potential court decision, CIB enforces the claim payment from the driver guilty for the accident.

The data distinguish payments for material damage, bodily injury claims, annuities (regular payments including loss of income and daily care), technical costs, loss of income for legal entities (a very rare case in the data) and others, where the claim type is specific (e.g. so called Green Card claims, representing the claims caused by uninsured drivers from Czech Republic and which occurred in the countries participating in the Greed Card System). For the purpose of our research we consider annuity and bodily injury payments since these are rather frequent and for that reason could be modeled by a NHPP. Moreover, only claims for which either cumulative payments or RBNS at 31.12.2015 exceed 1 million CZK are considered for maximum likelihood method to estimate the rate function parameters. Such limitation is due to the fact that we want to model relatively large claims, where either previous number of payments or the expected one is substantial because of a large amount to be paid out to the victim.

The considered annuity claims consist of 139 observations which occurred during years 2000 to 2015. As for the bodily injury claims, we have 168 claims meeting our requirements, also from years 2000 to 2015. Histograms of numbers of payments for both claim types can be found in Figure 4.1. One can notice that most of the annuity claims have up to 13 payments, majority of the bodily injury claims, on the contrary, up to 5 payments. This may be caused by different occurrence years of the claims – for older claims we have longer history of payments. Nevertheless, it is not always the case – some of the newer claims have already much more payments than the majority of the claims. The difference in payments between the annuity and bodily injury claims is natural since the annuity claims by their definition result in more payments than the bodily injury claims.



(i) Annuities (ii) Bodily injury

Figure 4.1: *Histogram of the numbers of payments for the annuities and bodily injury claims.*

Boxplot of both types of claims can be found in Figure 4.2. We can conclude that the bodily injury claims have in general less payments than the annuities which is natural due to the reasons described above. Median, 25%- and 75%-quantiles as well as maximum are lower for the bodily injury claims. We can arrive at the same conclusion by looking at Table 4.1 containing descriptive statistics for both claim types. Annuities have higher standard deviation meaning higher volatility in the numbers of payments. For convenience we use rounding up to one decimal point in the following table, despite that later we consider rounding up to two decimal points due to the better differentiation of the data.

Table 4.1: *Descriptive statistics of the numbers of payments for the annuities and bodily injury claims.*

	Min.	**Max.**	**Mean**	**25% q.**	**Med.**	**75% q.**	**Std. d.**
Annuities	1.0	41.0	14.1	7.0	12.0	19.0	8.6
Bodily injury	1.0	26.0	5.4	2.0	4.0	8.0	4.5

We can also examine delays in the first payment for both claim types, meaning

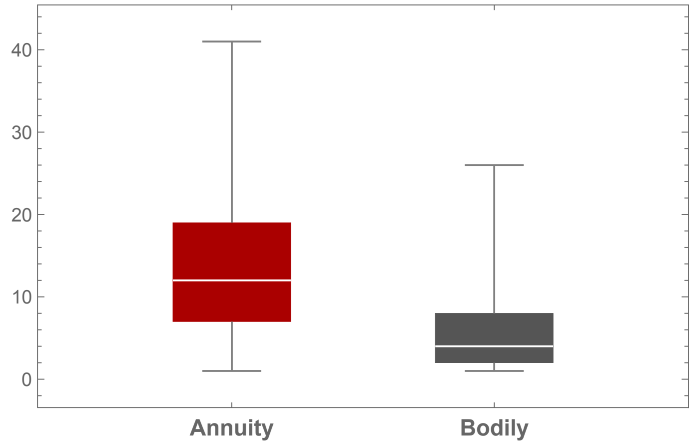


Figure 4.2: *Boxplot of the numbers of payments for the annuities and bodily injury claims.*

the number of days passed since a claim was reported until the moment of the first payment on this claim. From the boxplot of delays depicted in Figure 4.3 we can observe that the minimum and maximum delays are roughly the same for both claim types. Quantiles are lower for the bodily injuries. One can notice that the box of the annuity claims is wider which gives evidence for higher volatility in the delays of the annuity payments. This may be caused by the fact that annuity claims are often related to the victim's disability which implies more complex and time-consuming claim settlement.

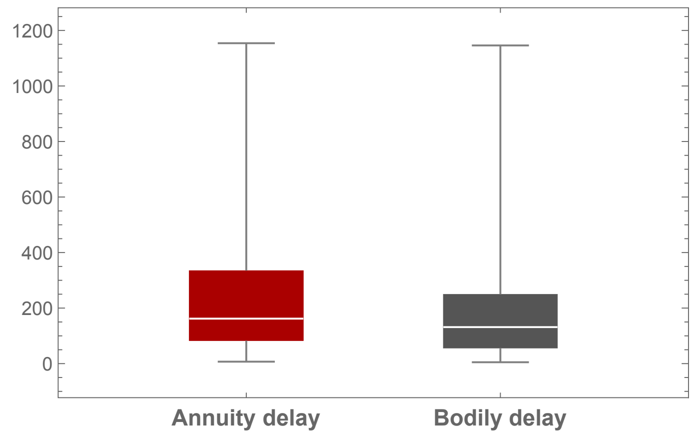


Figure 4.3: *Boxplot of the delays in the first payment for the annuities and bodily injury claims.*

4.2 Parameter estimation

Due to the claim origin we can assume independence of the moments of payments for different claims. Nevertheless, annuity and bodily injury claims should be analysed separately because of their different behavior. Our log-likelihood function is then constructed as the sum of log-likelihood functions of each claim owing to the previously mentioned assumption of independence. The function from expression (2.1) then corresponds to the log-likelihood function of each claim; the combined log-likelihood function then takes the following form:

$$\ell(\tilde{\mathbf{t}},\, N \,|\, \boldsymbol{\theta}) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} \log \lambda(t_j \,|\, \boldsymbol{\theta}) - \sum_{i=1}^{m}\left[\int_0^{T_i} \lambda(u \,|\, \boldsymbol{\theta})\mathrm{d}u\right],$$

where $\tilde{\mathbf{t}} = (t_1^1, \ldots, t_{n_1}^1, t_1^2, \ldots, t_{n_2}^2, \ldots, t_1^m, \ldots, t_{n_m}^m)$ stands for the combined times to payments of all claims, m is the overall number of the analysed claims and $N = n_1 + \ldots + n_m$ is the overall number of payments occurred. The claim report date is selected as the starting-point for evaluating times to payments since it is the moment from which we start observing the claim; in the likelihood function the starting-point then equals 0. Finally, T_i corresponds to the time from the claim report date until 31.12.2015, i.e. until the last day of observing the claims. Time differences are calculated in days. Nevertheless, to simplifying further work with the data, we select a year with 365 days as the time unit.

We firstly estimate the parameters of log linear rate function as well as constant rate separately for the annuity and bodily injury claims. For this purposes we consider version 11.1 of the software `Mathematica` developed by Wolfram Research, Inc. [2017]. We choose this software due to several reasons. Firstly, it is convenient for working with databases and we have a set of claims, each containing times of payments, and the number of payments differs within each claim record. Secondly, it provides a lot of efficient numerical methods for maximizing more complex functions. The function `NMaximize` is used in order to numerically maximize the log-likelihood for both cases. This function also returns the maximized value of the log-likelihood.

After obtaining the estimates we have to check whether the FIM is positive in each estimate. Considering the homogeneous case as well as the case of parameter θ_0 in the log linear model, the FIM is always positive in these points due to the way of construction of the log-likelihood function. The homogeneous case is straightforward; as for the second case, if we take the derivative of expression 2.5, we can see that it is always negative, implying that the FIM is always positive. That means we have to check for positiveness of the FIM in each estimate of the parameter θ_1 according to expression (2.3).

Finally, the likelihood ratio test (or, alternatively, the test statistic considered in expression 2.4) should be performed to determine whether parameter θ_1 representing the non-homogeneity can be set to 0. We want to test the null hypothesis

$$H_0: \;\theta_1 = 0, \text{ or, equivalently, } \lambda(t) = \theta \equiv \text{ const.}$$

against the alternative

$$H_0: \;\theta_1 \neq 0, \text{ or, equivalently, } \lambda(t) = \exp\{\theta_0 + \theta_1 t\}.$$

The test statistic is of the following form

$$S = 2 \cdot \left[\ell\left(\tilde{\mathbf{t}}, N \,|\, \hat{\theta}_0, \hat{\theta}_1\right) - \ell\left(\tilde{\mathbf{t}}, N \,|\, \hat{\theta}\right)\right], \tag{4.1}$$

where $\hat{\theta}_0$ and $\hat{\theta}_1$ are the estimates of the rate function parameters in log linear model, $\hat{\theta}$ is the estimate of the rate in the homogeneous model. S follows the

χ^2_1 distribution under the null hypothesis. The null hypothesis is then rejected if $S > \chi^2_1(1-\alpha)$, where $\chi^2_1(1-\alpha)$ is $(1-\alpha)$-quantile of χ^2_1 distribution, and α is the considered level of significance. We have the χ^2 distribution with 1 degree of freedom due to the fact that the alternative model has only one more parameter than the null model. From now on, we consider α to be 5 %.

For both annuity and bodily injury claims the FIM is positive in the points of the estimates and the likelihood ratio is significantly larger than 3.84 which is approximately the 95%-quantile of the χ^2_1 distribution. We received S equal to 310.29 for annuity claims and 647.94 for bodily injury claims which implies that we reject the hypothesis of homogeneity of the processes.

Next, we shall perform the same steps for log quadratic rate function. The likelihood-ratio test again speaks in favor of the more complex model – we arrived at S equal to 72.88 for the annuity model and 49.59 for the bodily injury model. The summarized results which include parameter estimates and their estimated standard deviation calculated as $\hat{\mathrm{sd}}(\hat{\theta}_i) = \sqrt{I^{-1}(\hat{\theta}_i)}$, can be found in Table 4.2. One can note that the estimated standard deviations are relatively low; this is caused by the fact that, firstly, the values of the parameter estimates themselves are not high and, secondly, by having the substantial number of payments accomplished.

Table 4.2: *Estimates of the parameters of log quadratic rate function as well as their estimated standard deviation.*

Annuity claims

	$\hat{\theta}_i$	$\hat{\mathrm{sd}}(\hat{\theta}_i)$
θ_0	0.83	0.02
θ_1	0.08	0.01
θ_2	-0.02	< 0.01

BI claims

	$\hat{\theta}_i$	$\hat{\mathrm{sd}}(\hat{\theta}_i)$
θ_0	1.23	0.03
θ_1	-0.64	0.01
θ_2	0.03	< 0.01

In conclusion, we consider log quadratic rate model as the final one for both annuity and bodily injury claims. It is essential to note that other approaches to grouping claims for the parameter estimation could be applied, e.g. grouping of claims by year, which means assuming a different process of payments for each year. Choice of the correct approach depends on the data origin; as for MTPL claims on uninsured drivers, they are rather standardized within the years as they do not depend e.g. on the tariff or limit policy of an insurance company. Moreover, MTPL business is mainly coordinated by the legislation of the country where the claim occurred. Only changes in the legislation may possibly cause the difference in payments between different years. For instance, there can be a switch from annuities to lump sums, when the claim is settled in a single payment containing the sum of estimated discounted future payments, as it is in the Italian market. Since we aim mainly to demonstrate the use of the methods, we assume no difference in payments within the years.

4.3 Simulation

In this part we focus on application of the simulation methods described in Chapter 3. We are going to use `RStudio` environment (developed by RStudio Team

[2017], version 1.0.143) of the software R (version 3.4.0) developed by R Core Team [2017] to demonstrate the methods since it is available free of charge for any user. Nevertheless, the considered algorithm can be basically implemented in any software which provides random variable generators.

We consider 10 annuity and 10 bodily injury claims with the highest values of the RBNS reserve at 31.12.2015. These values in CZK can be found in Table 4.3, where we do not use decimal points as there are no such in the provided data. The reason these claims are interesting for analysis is that there is substantial future uncertainty due to extremely high (considering the Czech market) expected recoveries.

Table 4.3: *10 largest annuity and bodily injury claims by RBNS at 31.12.2015.*

Annuity claims

Claim	*RBNS at 31.12.2015*
#1_A	28 954 770
#2_A	22 458 408
#3_A	21 926 000
#4_A	20 165 000
#5_A	17 182 000
#6_A	17 075 000
#7_A	16 005 000
#8_A	15 000 000
#9_A	13 400 000
#10_A	12 912 000

BI claims

Claim	*RBNS at 31.12.2015*
#1_BI	10 091 099
#2_BI	9 950 000
#3_BI	9 281 500
#4_BI	9 075 000
#5_BI	8 568 313
#6_BI	8 165 000
#7_BI	6 327 735
#8_BI	6 080 000
#9_BI	4 134 000
#10_BI	4 021 603

We are going to simulate instants of payments during the year 2016, i.e. the simulation should be in the interval $(T_i, T_i + 1]$, where T_i corresponds to the time from the claim report date until 31.12.2015 divided by 365. Therefore, the simulation methods should be slightly modified – the starting-point no longer equals 0. The source code of the algorithms can be found in the Attachment of this thesis.

Now let us compare all described methods with respect to their time efficiency and ease of their implementation. Each of the methods has been firstly tested on a set of claims, where we considered a separate process with log linear rate for each claim. 10 000 simulations of the year 2016 payments were performed for each claim. The average number of payments per year was then compared to the theoretical value $\Delta = \Lambda(T_i+1) - \Lambda(T_i)$ which was calculated based on the values of the estimated parameters.

4.3.1 Comparison of the methods

Time-scale transformation

(Section 3.1 of Chapter 3)

This method is very easy to implement; however, in case of not strictly monotone rate function where the inverse function is not defined explicitly, simulation took up to 5 times longer than for the other methods. Due to the fact that even for log linear rate the simulation time was extremely high, we are not going to use this method for the final simulations – we have much more efficient methods at our disposal. Nevertheless, the method is accurate because the deviation of the simulated average numbers of payments from the theoretical ones is rather minor.

Thinning

(Section 3.2 of Chapter 3)

The thinning method is easy to implement if one chooses $\lambda^*(t)$ according to $\lambda^*(t) = \lambda^*$, where $\lambda^* \geq \max_{0 \leq t \leq T} \lambda(t)$. We used the same logic as well since a homogeneous process is easy to simulate, and finding the maximum of the rate function is not that time-consuming. The method appeared to be time efficient as well as accurate by means of the deviation from the theoretical average.

Direct simulation of processes with log linear rate function

(Section 3.3 of Chapter 3)

The method turned out to be very efficient for simulating a process with log linear rate function. 10 000 simulations took roughly 2 times less than in case of thinning. The method is accurate as well. We consider it the best solution for a process with log linear rate function.

Gap statistics for processes with log linear rate function

(Section 3.3 of Chapter 3)

Time efficiency of this method is roughly the same as in case of the direct simulation, but it is not as easy to implement. Nevertheless, this method was not accurate when the theoretical expected number of events was less than or very close to 1. The deviations for higher expected numbers of events were small (up to 10 %); however, in the first case they could be e.g. even 300 %. For that reason we are not going to consider this method for the final simulations since in many cases the simulated averages should be less than 1. This also implies that we prefer to use the direct simulation for the algorithm designed for log quadratic rate functions.

Simulation of processes with log quadratic rate function

(Section 3.4 of Chapter 3)

This method was the most complex in terms of its implementation since one has to allow for several different cases of the direction of the rate function as well as time interval $(T_i, T_i + 1]$. The method appeared to be accurate. However, it is less time efficient comparing to the method of thinning for processes with log quadratic rate – it was roughly 15 % slower. Due to this fact the method of thinning is considered to be the most appealing and time-efficient algorithm for simulating NHPPs. Moreover, it is very easy to handle in any software.

4.3.2 Final simulations

Considering the final simulations, we performed 10 000 simulations of the year 2016 payments for the thinning method and the method for log quadratic rates for each of the claims considered above, taking into account the final parameter estimates of the log quadratic rate function along with the values of T_i. We point out that in practice some actuarial softwares use two-step simulation, when the value of the parameter itself is firstly simulated considering also its estimated standard deviation. Then we continue with the simulation algorithm using the generated values of the parameters. In this thesis we do not take into account the parameters' standard deviations.

The output contains 10 000 developments of the year 2016 payments for each method for 10 annuity and 10 bodily injury claims. The payment time instants themselves can be used in the combined model for tariff setting in order to calculate the discounted cashflows if the simulation comprises more years. In this practical part we focus only on the number of payments during the year 2016.

In Tables 4.4, 4.5, 4.6 and 4.7 we show the descriptive statistics of the simulated numbers of payments along with their deviations from the theoretical expected numbers of payments. Note that we use rounding up to 2 decimal points only for the means, deviations and standard deviations for better visual representation of the numbers, since other statistics do not contain the decimal part. We can note that the deviations are rather small in all cases. Both sum of the squared deviations and the standard deviations are roughly the same for both methods. We conclude that the method of thinning is more preferable due to its multipurpose use, simple implementation and time efficiency.

Table 4.4: *Descriptive statistics of the numbers of payments for the annuities simulated using the thinning method (Section 3.2, Chapter 3).*

	Min.	**Max.**	**25% q.**	**Med.**	**75% q.**	**Mean**	Δ	**Std. d.**
#1_A	0	5	0	0	1	0.61	-1.59%	0.79
#2_A	0	6	0	1	1	0.88	-0.32%	0.94
#3_A	0	6	0	1	1	0.76	0.44%	0.88
#4_A	0	8	0	1	2	1.25	-0.12%	1.12
#5_A	0	10	1	2	3	1.97	-0.93%	1.42
#6_A	0	8	1	2	3	1.99	1.42%	1.41
#7_A	0	7	0	1	1	0.96	1.11%	0.99
#8_A	0	6	0	1	1	0.78	-0.70%	0.88
#9_A	0	11	1	2	3	2.40	-0.51%	1.53
#10_A	0	10	1	2	3	2.21	-0.73%	1.49

Table 4.5: *Descriptive statistics of the numbers of payments for the annuities simulated using the method for degree-two exponential polynomial rate functions (Section 3.4, Chapter 3).*

	Min.	**Max.**	**25% q.**	**Med.**	**75% q.**	**Mean**	**Δ**	**Std. d.**
#1_A	0	6	0	0	1	0.63	0.84%	0.80
#2_A	0	7	0	1	1	0.89	0.27%	0.96
#3_A	0	6	0	1	1	0.77	1.68%	0.88
#4_A	0	7	0	1	2	1.25	0.44%	1.12
#5_A	0	10	1	2	3	1.99	-0.10%	1.42
#6_A	0	9	1	2	3	1.97	0.14%	1.40
#7_A	0	6	0	1	1	0.95	0.14%	0.98
#8_A	0	5	0	1	1	0.78	-0.45%	0.88
#9_A	0	10	1	2	3	2.43	0.43%	1.53
#10_A	0	10	1	2	3	2.22	-0.17%	1.50

Table 4.6: *Descriptive statistics of the numbers of payments for the bodily injury claims simulated using the thinning method (Section 3.2, Chapter 3).*

	Min.	**Max.**	**25% q.**	**Med.**	**75% q.**	**Mean**	**Δ**	**Std. d.**
#1_BI	0	4	0	0	0	0.18	2.32%	0.43
#2_BI	0	7	0	1	2	1.31	-0.95%	1.15
#3_BI	0	7	0	1	2	1.31	1.37%	1.15
#4_BI	0	5	0	0	1	0.39	0.54%	0.62
#5_BI	0	6	0	0	1	0.53	-1.86%	0.74
#6_BI	0	7	0	1	2	1.09	1.33%	1.05
#7_BI	0	7	0	1	2	1.31	2.28%	1.15
#8_BI	0	6	0	1	2	1.23	0.31%	1.10
#9_BI	0	9	1	2	3	1.93	0.45%	1.38
#10_BI	0	6	0	1	1	0.81	-1.27%	0.89

Table 4.7: *Descriptive statistics of the numbers of payments for the bodily injury claims simulated using the method for degree-two exponential polynomial rate functions (Section 3.4, Chapter 3).*

	Min.	**Max.**	**25% q.**	**Med.**	**75% q.**	**Mean**	**Δ**	**Std. d.**
#1_BI	0	3	0	0	0	0.18	3.73%	0.43
#2_BI	0	8	0	1	2	1.32	-0.66%	1.15
#3_BI	0	8	0	1	2	1.29	-0.12%	1.15
#4_BI	0	4	0	0	1	0.39	-0.06%	0.62
#5_BI	0	5	0	0	1	0.53	-2.42%	0.72
#6_BI	0	7	0	1	2	1.08	0.43%	1.04
#7_BI	0	7	0	1	2	1.27	-0.99%	1.14
#8_BI	0	7	0	1	2	1.23	0.71%	1.12
#9_BI	0	10	1	2	3	1.92	0.17%	1.37
#10_BI	0	5	0	1	1	0.81	-0.20%	0.89

Conclusion

The purpose of this thesis was to summarize the main properties of a NHPP and describe the widely used estimation and simulation methods together with demonstration of their use on real data.

In Chapter 1 we presented the main properties of NHPPs. Furthermore, we discussed the method to simplify the work with a NHPP using a HPP. In addition, we derived the conditional distribution of arrival times, which was of much use while describing the simulation procedures.

Chapter 2 focused on the general description of the maximum likelihood method adapted for a NHPP. Moreover, we discussed the uniqueness of the estimate along with its asymptotic properties. Finally, we presented a special case of log linear rate function, where we described the estimation procedure in more detail.

In Chapter 3 we discussed two general methods for simulation of a NHPP — the time-scale transformation technique and the method of thinning. Afterwards, we discussed the methods designed specially for log linear and log quadratic rate functions. All of the methods were compared with respect to their theoretical time efficiency.

Chapter 4 presents another aim of this thesis — application of the discussed estimation and simulation methods on real data. As NHPPs can be widely used in actuarial practice, we used the real data from non-life insurance. The data contain payments and RBNS until the year 2015 for MTPL claims caused by uninsured drivers from the Czech Republic. The data were provided by Czech Insurers' Bureau which manages such claims.

We performed the analysis of the data along with their grouping by annuity and bodily injury claims. In the next step we estimated the parameters of constant, log linear and log quadratic rate functions and selected the most suitable model by conducting the likelihood ratio test. Then the simulation methods were tested on a set of claims, where we assumed a separate process for each claim. After that, the methods were compared with respect to their real time efficiency and accuracy of the simulations. Time-scale transformation method appeared to be time-inefficient and for that reason it was not considered in the final simulations. The method for log linear rate functions based on gap statistics was inaccurate for low expected numbers of payments. Therefore, it was not considered in the final simulations as well.

Finally, we chose 10 largest annuity and bodily injury claims by their RBNS value at the end of the observation period. These claims are interesting for analysis since they are not expected to be settled in the short time interval. We performed 10 000 simulations of the year 2016 payment instants for each claim using all considered methods. Then we compared the simulated average numbers of payments to the theoretical ones.

There are still a lot of possibilities for research in the field of NHPPs, since most of the articles used in the thesis are from 1970-1990s, when there was no powerful software to test the efficiency of the simulation methods properly. Now software is much more functional; therefore, more sophisticated simulation methods could be designed.

Bibliography

M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables.* Dover Publications, Mineola, USA, 1965.

E. Çinlar. *Introduction to stochastic processes.* Dover Publications, Mineola, USA, 2013.

D. R. Cox and P. A. W. Lewis. *The statistical analysis of series of events.* John Wiley and Sons, New York, USA, 1966.

R. W. Klein and S. D. Roberts. A time-varying poisson arrival process generator. *Simulation*, 43(4):193–195, 1984.

M. E. Kuhl, J. R. Wilson, and M. A. Johnson. Estimating and simulating poisson processes having trends or multiple periodicities. *IIE transactions*, 29(3):201–211, 1997.

S. Lee, J. R. Wilson, and M. M. Crawford. Modeling and simulation of a nonhomogeneous poisson process having cyclic behavior. *Communications in Statistics — Simulation and Computation*, 20(2-3):777–809, 1991.

P. A. W. Lewis and G. S. Shedler. Simulation of nonhomogeneous poisson processes with log linear rate function. *Biometrika*, 63(3):501–505, 1976.

P. A. W. Lewis and G. S. Shedler. Simulation of nonhomogeneous poisson processes by thinning. *Naval Research Logistics*, 26(3):403–413, 1979.

R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2017.

S. M. Ross. *Introduction to probability models.* 10th ed. Academic press, Boston, USA, 2010.

RStudio Team. *RStudio: Integrated Development Environment for R.* RStudio, Boston, USA, 2017.

R. L. Streit. *Poisson point processes: imaging, tracking, and sensing.* Springer Science+Business Media, New York, USA, 2010.

Wolfram Research, Inc. *Mathematica 11.1.* Champaign, USA, 2017.

R source code of the simulation algorithms

Time-scale transformation

Both starting point and endpoint of the desired simulation interval together with expectation function $\mu(t) \equiv \Lambda(t) - \Lambda(0)$ should be provided. The algorithm returns 0 in the case of no events, and the arrival times otherwise.

```
simulation_Time_Scale <- function(t0,t_max,Lambda){

  Lambda_inv <- function(s){
    v <- seq(t0,t_max, length.out = 1000)
    min(v[Vectorize(Lambda)(v)>=s])
  }

  t1 <- Lambda(t0)
  lt_max<-Lambda(t_max)
  t<-0
  s <- 0
  X <- numeric()

  while(t1 <= lt_max){
    u <- runif(1)
    s <- -log(1-u)
    t1 <- t1+s
    if (t1>lt_max) {
      break
      }
    t<-Lambda_inv(t1)
    X <- c( X, t)
  }
  if (length(X)>0) {
    return(X)
  } else  {
    0
  }
}
```

Thinning

Both starting point and endpoint of the desired simulation interval together with rate function $\lambda(t)$ should be provided. The algorithm returns 0 in the case of no events, and the arrival times otherwise.

```
simulation_Thinning <- function(t0,t_max,lambda){

  t1<-t0
  lambda_star <- max(sapply(seq(t0, t_max,length.out=1000), lambda))
```

```
  X <- numeric()

  while(t1 <= t_max){
    e<-rexp(1,lambda_star)
    u <- runif(1)
    t1<-t1+e
    if (t1>t_max) {
      break
    }
    if(u < lambda(t1)/lambda_star) {
      X <- c(X,t1)
    }
  }
  if (length(X)>0) {
    return(X)
  } else  {
    return(0)
  }
}
```

Direct simulation of a NHPP with log linear rate

Both starting point and endpoint of the desired simulation interval along with the parameters θ_0 and θ_1 of the rate function should be provided. The algorithm returns 0 in the case of no events, and the arrival times otherwise.

```
simulation_Log_Linear_1 <- function(par1,par2,t0,t_max){

  lpoiss<-exp(par1)*(exp(par2*t_max)-exp(par2*t0))/par2
  inv<-function(u) {
    log(exp(par2*t0)+u*(exp(par2*t_max)-exp(par2*t0)))/par2
  }
  X <- numeric()
  p<-rpois(1,lpoiss)

  if (p>0) {
    for (i in 1:p)  {
      u <- runif(1)
      t<-inv(u)
      X<-c(X,t)
    }
  if (length(X)>0) {
  return(sort(X,decreasing=FALSE))
  } else {
    return(0)
  }
  } else  {
    return(0)
  }
}
```

Gap statistics for a NHPP with log linear rate

Both starting point and endpoint of the desired simulation interval along with the parameters θ_0 and θ_1 of the rate function should be provided. The algorithm returns 0 in the case of no events, and the arrival times otherwise.

```
simulation_Log_Linear_2 <- function(par1,par2,t0,t_max){

  e<-0
  X <- numeric()

  if (par2<0) {
    t<-0
    lpoiss<- -exp(par1)/par2
    m<-rpois(1,lpoiss)

    if (m>0) {
      for (i in 0:m-1)  {
        e <- rexp(1,1)
        if (t+e/(-par2*(m-i))>1) {
          break
        }
        t<-t+e/(-par2*(m-i))+t0
        X<-c(X,t)
      }
      if (length(X)>0) {
        return(X)
      } else {
        0
      }

    } else  {
      0
    }
  } else {
    lpoiss<- exp(par1+par2*t_max)/par2
    m<-rpois(1,lpoiss)
    t1<-0

    if (m>0) {
      for (i in 0:m-1)  {
        e <- rexp(1,1)
        if (t1+e/(par2*(m-i))>1) {
          break
        }
        t1<-t1+e/(par2*(m-i))
        X<-c(X,t1)
      }
      if (length(X)>0) {
        Ts<-rep(t_max,ct)
```

```
        return(Ts-rev(X))
      } else {
        0
        }
    } else  {
      0
      }
    }
}
```

Simulation of a NHPP with log-quadratic rate function

Both starting point and endpoint of the desired simulation interval along with the parameters θ_0, θ_1 and θ_2 of the rate function should be provided. The algorithm returns 0 for each part of the algorithms in the case of no events, and the merged arrival times otherwise. Zero values can be further deleted during the simulation process.

```
simulation_Log_Quadratic <- function(par1,par2,par3,tStart,tEnd){

  gamma0<-0
  gamma1<-0
  res<-0
  t0<-tStart
  t_max <- tEnd
  crit<- -par2/(2*par3)

  'check whether it is case (i)-(iv)'

  if ((t_max<crit)|(par2*par3>0)) {
    if (par2>0) {
      if (par3>0) {

      'case (i)'

        gamma0<-par1-par3*(tStart+1)^2
        gamma1<-par2+2*(tStart+1)*par3
      } else {

        'case (iii)'

        gamma0<-par1
        gamma1<-par2+(tStart+1)*par3
      }
    } else {
      if (par3<0) {

      'case (iv)'
```

```
      gamma0<-par1
      gamma1<-par2+(tStart+1)*par3
    } else {

    'case (ii)'

      gamma0<-par1
      gamma1<-par2
    }
  }
  lambda2<-function(u) {
    exp(par1+par2*u+par3*u^2)-exp(gamma0+gamma1*u)
  }
  t0<-tStart
  t_max <- tEnd
  x<-simulation_Log_Linear_1(gamma0,gamma1,t0,t_max)
  y<-simulation_Thinning(lambda2,t0,t_max)
  return(sort(c(x,y),decreasing=FALSE))

} else {
  'check whether it is case (iv) or (v)  but with the
  starting point larger than the inflexion point -
  there is no rate direction change in such a simulated interval'

  if (t0>crit) {
    if (par2<0) {

    'case (v.2)'

      gamma0<-par1-par3*(tStart+1)^2
      gamma1<-par2+2*(tStart+1)*par3
    } else {

    'case (vi.2)'

      gamma0<-par1+(par2/2)*(tStart+1)
      gamma1<-(par2/2)+par3*(tStart+1)
    }

    lambda2<-function(u) {
      exp(par1+par2*u+par3*u^2)-exp(gamma0+gamma1*u)
    }
    t0<-tStart
    t_max <- tEnd
    x<-simulation_Log_Linear_1(gamma0,gamma1,t0,t_max)
    y<-simulation_Thinning(lambda2,t0,t_max)
    return(sort(c(x,y),decreasing=FALSE))

  } else {
```

```
      'we have to take into account the change of the direction of
       the rate function'
      'part 1 - simulation before the inflexion point of the rate
       function'
      if (par2<0) {

      'case (v.1)'
        gamma0<-par1
        gamma1<-par2
      } else {

      'case (vi.1)'
        gamma0<-par1
        gamma1<-par2/2
      }
      lambda2<-function(u) {
        exp(par1+par2*u+par3*u^2)-exp(gamma0+gamma1*u)
      }
      t0<-tStart
      t_max<--par2/(2*par3)
      x1<-simulation_Log_Linear_1(gamma0,gamma1,t0,t_max)
      y1<-simulation_Thinning(lambda2,t0,t_max)

      'part 2 - simulation after the inflexion point of the rate
       function'

      if (par2<0) {

      'case (v.2)'
        gamma0<-par1-par3*(tStart+1)^2
        gamma1<-par2+2*(tStart+1)*par3
      } else {

      'case (vi.2)'
        gamma0<-par1+(par2/2)*(tStart+1)
        gamma1<-(par2/2)+par3*(tStart+1)
      }
      lambda2<-function(u) {
        exp(par1+par2*u+par3*u^2)-exp(gamma0+gamma1*u)
      }
      t0<--par2/(2*par3)
      t_max<-tStart+1
      x2<-simulation_Log_linear_1(gamma0,gamma1,t0,t_max)
      y2<-simulation_Thinning_1(lambda2,t0,t_max)
      return(sort(c(x1,x2,y1,y2),decreasing=FALSE))
    }
  }
}
```

List of Figures

List of Tables