Stockholms universitet

Individual loss reserving with piecewise constant hazard rates

Shuhua Zhang

Masteruppsats i försäkringsmatematik
Master Thesis in Actuarial Mathematics

Masteruppsats 2019:2
Försäkringsmatematik
Februari 2019

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Master Thesis **2019:2**
http://www.math.su.se

Individual loss reserving with piecewise constant hazard rates

Shuhua Zhang*

February 2019

Abstract

The purpose of this thesis is to evaluate the application of the piecewise constant hazard rates in predicting reserves in general insurance. Two types of datasets which include only reported claims are simulated and studied separately. Three types of events are distinguished during the development of a claim. The piecewise constant hazard rates are estimated by maximum likelihood theory. As a comparison, a method for smoothing piecewise constant rates is analyzed. The result shows that piecewise constant hazard rates obtained directly from maximum likelihood estimation perform excellently. They give a remarkable fit for predicting future payments of the Reported But Not Settled (RBNS) claims. The estimators obtained from by applying the smoothning method are not suitable to predict reserves.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: navalniteflora@gmail.com. Supervisor: Mathias Lindholm.

Acknowledgements
I would like to thank my supervisor Mathias Lindholm for many hours discussions and valuable feedback.

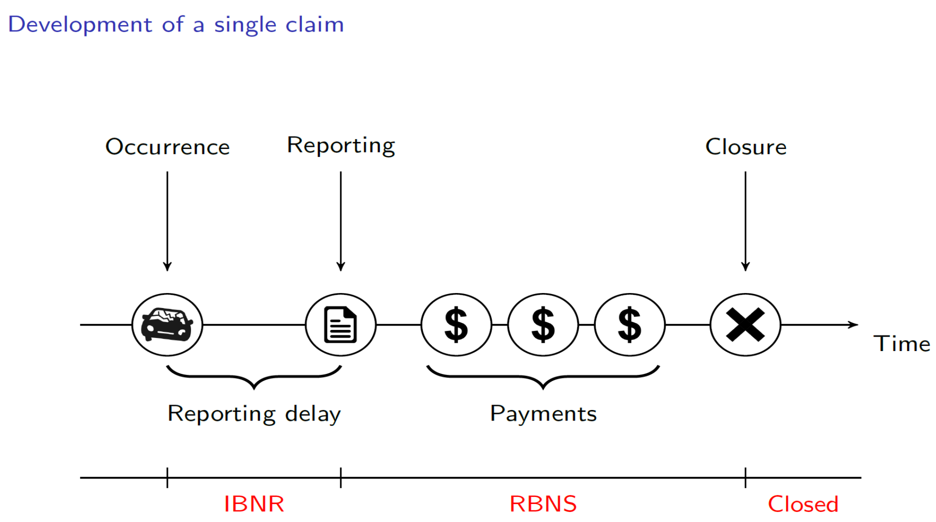
Contents

1 Introduction

Accurate loss reserves are essential for insurers to maintain adequate capital and to price their insurance products efficiently. Due to the introduction of new supervisory guidelines (Solvency 2) and reporting standards (IFRS 4 and IFRS 17), the prediction of future cash flows and their uncertainty becomes more important. Under-reserving may result in failure to meet liabilities and even insolvency of the insurers. Conversely, excessive reserves may limit the insurer's growth opportunities and weaken its competitive position in the market. Hence, actuaries should develop reserving models which can predict reserves with better quality.

Figure 1 illustrate the development process of a general insurance claim. A claim occurs at a certain time point, it is reported to the insurer (possibly

Figure 1: Development of a general insurance claim.



after a period of delay) and a series of payments follow until settlement (or closing) of the claim. The time gap between the occurrence of the claim and reporting is called "reporting delay" and the time gap between reporting and settlement is called "settlement delay". At the present date, the insurer has to determine reserves for Reported But Not Settled (RBNS) claims and Incurred But Not Reported (IBNR) claims.

1.1 Macro and Micro Methods

[3] mentions that for more than a century, actuaries have been using run-off triangles to project future payments in non-life insurance. A run-off triangle summarizes available information per arrival and development year. The earliest traditional method which formalized this technique in the '30s in

[3] and still widely used is the *chain ladder* method (CLM). Another commonly used macro-level reserving method is the expected claims method. Other macro-level models, such as the Bornhuetter-Ferguson (B-F) and the Cape-Cod method, are constructed as a blend of the chain-ladder and the expected claims technique [26].

All macro-level models are based on aggregate data in a run-off triangle. The advantages of macro-level models are that they are easy to understand, and can be mentioned in financial communication, without disclosing too much information [1]. But recently, see e.g. [4], [14], question the use of aggregate loss data. Due to aggregation, not all information in data is taken into account. Problems with triangular data are also mentioned, for instance, the problem of zero or negative cells in the triangle, see e.g. [10]. [25] discusses the robustness properties and the influence of outliers on triangular methods. The separate assessment of true IBNR and RBNS claims in run-off triangle is not straightforward, see e.g. [6], [19]. These substantial pieces of literature demonstrate that macro-level models cannot always adequately capture the complexities of stochastic reserving for general insurance.

More recently, greater interest has been expressed in estimating the loss reserving at individual claim level (i.e. micro-level). For that, it is necessary to be able to estimate a full distribution of possible outcomes, from which percentiles (or other measures) of the distribution can easily be obtained. A micro-level model allows much detailed modeling of the development processes. The quality of reserves and their uncertainty can be improved by using detailed claims data too. Furthermore, it becomes possible to predict reserve of IBNR and IBNS claims separately. In the '70s, [7] suggested using a marked point process to project future payments and quantify the reserves. During the last thirty years, more literature about micro-level models has been presented, e.g. [17], [18] and [20] where the authors use Position Dependent Marked Poisson Processes to formulate the development of individual claims. There are several studies which extend the work of [17] and [18], e.g. [9], [21]. [2] revisited the Marked Poisson Process theory with a small case-study.

1.2 Overview of the Present Research

A micro-level model often contains several blocks, each handling a part of the claim development process. For example, [9] introduce a micro-level model based on the work of [17] and [18], in which the model consists of four building blocks: A block to model the reporting delay process, a block to model the number of IBNR claims, a third block is used to model the payment process and the last one is about the payment amounts. Distributions for each block can be fitted based on available individual data.

Of particular interest is that the author uses the statistic framework of recurrent events to model the payment process. The events are specified to three types and they are modeled using a piecewise-constant specification of the hazard rate of an event. How to apply the recurrent events theory and a piecewise-constant specification in micro models has eventually been studied, see e.g. [21].

1.3 Objective

In this master thesis, we will focus on using micro models to predict reserves of RBNS claims. The recurrent events theory and survival hazard rate function combined with non-homogeneous Poisson process theory will be applied in this thesis. Most parts of the methods and applications are based on the work of [9] and [21].

Poisson processes with time-dependent rate functions will be used to simulate datasets where the rate function is the hazard rate function from a survival model. More especially, from a Weibull distribution. Two types of datasets will be simulated by adjusting parameter settings in the rate function, one type of dataset with a long settlement time and one type of dataset with a short settlement time. The reason for working with two types of datasets is that it is of interest to study the performance of piecewise-constant hazard rate functions on data with different development patterns.

The model which will be used to generate the development process of a claim is based on Poisson processes with three distinguished types: It may be a process without any payment then being settled or with a payment to the settlement, or maybe several intermediate payments conduct before settlement. These three types are mentioned most possibly occur during the development of a claim in [9]. All these possibilities will be analyzed.

We will concentrate on the situation where the rate function is estimated with the specification that it is piecewise constant in prespecified subintervals (time intervals). In this thesis, we only consider that the time intervals are of equal length. Estimators are obtained by using the maximum-likelihood-method. We will also apply a smoothing method to one of the ML estimators in each dataset. These three estimated estimator vectors will be evaluated by comparing their predictive abilities.

Cash flows and total reserves for each dataset will be predicted using the estimated parameter vectors (piecewise-constant hazard rates). Performance of parameter settings will be compared both by cash flows in each calendar year and by event types. Analysis of robustness will also be carried out.

The remainder of the paper is organized as follow. In Section 2, relevant theory, models together with simulation procedures are described. In

Section 3, the characteristics of the two types of datasets are presented. Section 4 discusses the results and in Section 5 we will summarize findings and conclusions.

2 Methodology

Section 2.1 describes the theory for Poisson processes with multi-type recurrent events. In section 2.2 the model outline of this thesis is discussed. The likelihood function in the Poisson process with totally developed data and censored data are derived in section 2.3. Section 2.4 presents two parameter estimation methods and the last two sections describe how the future events respective the datasets can be simulated.

2.1 The Statistical Models

Theories which are described follow the works and results from [16], [22], [24]. Especially, the definition of the counting process, Poisson process and their properties come from [22]. Theory about the relationship between homogeneous and non-homogeneous Poisson process is from [24]. Definitions about recurrent events, including applications of the Poisson process in this type of events are based on [16].

2.1.1 The Poisson process

The Poisson process is the most common counting process and it is widely analyzed and used in science and technology. A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ represents the total number of events that occur by time t.

From the definition of the counting process, we see that a counting process $N(t)$ must satisfy:

(i) $N(t) \geq 0$.

(ii) $N(t)$ is integer valued.

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

(iv) For $s < t$, $N(t) - N(s)$ equals the number of events that occur in the interval $(s, t]$.

For a Poisson process with constant rate $\rho > 0$, the following properties hold:

(i) $N(0) = 0$

(ii) $\{N(t), t \geq 0\}$ has independent increments

(iii) $P(N(t+h) - N(t) = 1) = \rho h + o(h)$.

(iv) $P(N(t+h) - N(t) \geq 2) = o(h)$

(v) The number of events $N(t)$ in any interval of length t is Poisson distributed with mean $\mu = \mathrm{E}(N(t)) = \rho t$.

where $o(h)$ is a function with property:

$$\lim_{h \to 0} \frac{o(h)}{h} = 0$$

A process with such constant linear mean function is said to be *homogeneous*, *non-homogeneous* otherwise, i.e. the mean value function in the Poisson process is non-linear.

The non-homogeneous Poisson process has become an important alternative to the homogeneous process. This is because its intensity function can itself be a time-dependent variable. The intensity in the process can be a function of time or varies by other time-varying variables.

The one-dimensional non-homogeneous Poisson process has many important characteristic properties such as

(i) the number of points in any interval follows a Poisson distribution

(ii) the number of points in any finite set of non-overlapping intervals are mutually independent random variables.

The intensity function $\rho(t)$ is a nonnegative integrable function in a non-homogeneous Poisson process. The mean function which is also called the cumulative intensity for a non-homogeneous Poisson process

$$\mu(t) = \int_0^t \rho(u) du \quad t > 0$$

is continuous and finite for all $t > 0$. For a Poisson process, $\rho(t)$ determines whether there is a trend in the rate of events. If $\rho(t)$ is monotone increasing or decreasing then a monotone trend is said to exist.

Poisson process models may be parametric or nonparametric. Common parametric models for $\rho(t)$ are e.g. $\rho(t; \alpha, \beta) = \exp(\alpha + \beta t)$ and $\rho(t; \alpha, \beta) = \alpha \beta t^{\beta - 1}$, see e.g. [15]. These forms are similar with hazard rate functions in survival models. For example the hazard rate function defining the Weibull distribution has the form $h(t) = \gamma p t^{p-1}$ and the cumulative rate function $H(t) = \gamma t^p$ (Definition and properties of hazard rate function will be introduced in the following section). Poisson processes with hazard rate function $h(t)$ as intensity function, namely $\rho(t) := h(t)$, $\mu(t) := H(t)$, are widely used to simulate the time to the occurrence of the next event. For instance, it can be used to generate the payment process and the settlement delay process in micro-level stochastic loss reserving, see e.g. [9], [21].

2.1.2 Multi-type recurrent events

In general insurance and clinical trial studies it may often lie interest in studying processes which generate events repeatedly over time, e.g. a policyholder may receive multiple compensations for en accident, a patient may receive multiple treatments during the treatment process, etc. Such processes are refered to as *recurrent event processes* and the data they provide are called *recurrent event data*, for more details see e.g. [16].

Assume the process starts at $t = 0$ with $N(0) = 0$, where $N(t)$ denotes the number of occurrences of some type of event over the time interval $(0, t]$ for some individual. In the continuous time setting, models for recurrent events can be specified generally by considering the probability distribution for the number of events in short intervals $[t, t + \Delta t)$, given the history of event occurrence before time t. Let $\Delta N(t) = N(t+\Delta t) - N(t)$ denote the number of events in the interval $[t, t + \Delta t)$, $B(t) = \{N(s) : 0 \leq s < t\}$, $t > 0$ denote the history of the process up until time t. We assume that two events cannot occur at exactly the same time point. Then, the intensity function is defined as:

$$\lambda(t|B(t)) = \lim_{\Delta t \to 0} \frac{P(\Delta N(t) = 1|B(t))}{\Delta t}. \tag{1}$$

It is assumed that an intensity is bounded and continuous except possibly at a finite number of points over any finite time interval. The intensity function defines an event process, and all process characteristics can be determined from it.

We assume that events are observed over the time interval $[\tau_0, \tau]$ for an individual. The time τ_0 corresponds to the start time and τ is referred to as the *termination time* or *end-of-follow-up time*. Sometimes it is also called a *censoring time* in survival analysis. Conditional on $B(\tau_0)$, the probability density of the outcome "n events, at times $t_1 < t_2 < \cdots < t_n$," where $n \geq 0$, for a process with intensity as (1), over the specified interval $[\tau_0, \tau]$, is

$$\prod_{j=1}^{n} \lambda(t_j|B(t_j)) \cdot \exp\Big(-\int_{\tau_0}^{\tau} \lambda(u|B(u))du\Big), \tag{2}$$

Given history $B(t_{j-1})$ and the $(j-1)$st event occurrence time t_{j-1}, the time to jth event occur at time t has distribution function:

$$P(T_j \leq t|T_{j-1} = t_{j-1}, B(t_{j-1})) = 1 - \exp\Big(-\int_{t_{j-1}}^{t} \lambda(u|B(u))du\Big). \tag{3}$$

This function shows that we can generate the next event time by generating the cumulative function $\Lambda_j = \int_{t_{j-1}}^{t} \lambda(u|B(u))du$ which has a standard exponential distribution with survivor function $\exp(-u)$, $u > 0$; and then solving (3) for t. By repeating for $j = 1, 2, ...$ we can generate successive event times t_j. For more details, see [16].

The Poisson process is one of the canonical processes for the analysis of recurrent events which is used to describe situations where events occur randomly and the numbers of events in nonoverlapping time intervals are statistically independent. The intensity in a Poisson process is of the form

$$\lambda(t|B(t)) = \rho(t), \quad t > 0. \tag{4}$$

It is seen from (4) that the probability of en event in $(t, t + \Delta t]$ may depend on t but is independent of his histroy $B(t)$. Then the probability density in (2) can be written as:

$$\prod_{j=1}^{n} \rho(t_j) \cdot \exp\Big(- \int_{\tau_0}^{\tau} \rho(u) du \Big). \tag{5}$$

Multiple types of recurrent events arise frequently. In the multi-type recurrent events two or more different types of events may occur repeatedly over the period of observation. It is mentioned in [16] that it may be sufficient to study them separately, especially if the occurrences of each type are more or less independent, but if the occurrence of one type affects the risk of other events, models for multiple counting processes are needed. In this case it may be helpful to formulate multivariate models through assumptions of conditional independence between events given univariate or multivariate random effects.

The following analysis and derivations of multi-type recurrent events are taken from [16].

Considering a population with n individuals in which each is at risk of J different types of recurrent events. Let i indicate individual $i = 1, ..., n$, j be the event type, $j = 1, ..., J$. Let $N_{ij}(t)$ denote the number of type j events occurring over the time interval $[0, t)$ for individual i, and let $\Delta N_{ij}(t) = N_{ij}(t + \Delta t) - N_{ij}(t)$. We assume as before that at most one event can occur at any given time. The event history for individual i is defined by $B_i(t) = \{N_i(s) : 0 \leq s < t\}$ and the intensity function for type j is defined as

$$\lambda_{ij}(t|B_i(t)) = \lim_{\Delta t \to 0} \frac{P(\Delta N_{ij}(t) = 1|B_i(t))}{\Delta t}. \tag{6}$$

Let t_{ijk}, $k = 1, ..., N_{ij}(t)$, denote the times of type j events over $[0, t]$, $j = 1, ..., J$ and $t_{i1}, ..., t_{iN_{i.}(t)}$ denote the times of all types of events for individual i over $[0, t]$, with $N_{i.}(t) = \sum_{j=1}^{J} N_{ij}(t)$, $\Delta N_{i.}(t) = N_{i.}(t+\Delta t) - N_{i.}(t)$. Based on the assumption that only at most one event can occur at any given time, then

$$\begin{aligned}
&P(\Delta N_{ij}(t) = 1|B_i(t)) = \lambda_{ij}(t|B_i(t))\Delta t + o(\Delta t)\\
&P(\Delta N_{i.}(t) = 0|B_i(t)) = 1 - \sum_{j=1}^{J} \lambda_{ij}(t|B_i(t))\Delta t + o(\Delta t)\\
&P(\Delta N_{i.}(t) \geq 2|B_i(t)) = o(\Delta t).
\end{aligned}$$

If we consider a specific time interval $[0, \tau]$ for individual i, with $0 = v_0 < v_1 < ... < v_R = \tau$ and $\Delta v_r = v_{r+1} - v_r$. Let $(\Delta n_{ir})_j = \delta_{ijr}$, $j = 1, \ldots, J$, where $\delta_{ijr} \in \{0, 1\}$, such that $\sum_{j=1}^{J} \delta_{ijr} \in \{0, 1\}$ which means that in each little time interval at most one event can occur. Base on these, for each small Δv_r, the probability distribution of $N_i(v_0), ..., N_i(v_R)$ can be approximated by

$$\prod_{r=0}^{R} P(\Delta N_i(v_r) = \Delta n_{ir} \mid B_i(v_r)) =$$
$$\prod_{r=0}^{R} \Big\{ \prod_{j=1}^{J} \Big(\lambda_{ij}(v_r|B_i(v_r))\Delta v_r\Big)^{\delta_{ijr}} \Big(1 - \sum_{j=1}^{J} \lambda_{ij}(v_r|B_i(v_r))\Delta v_r\Big)^{(1-\delta_{ijr})} \Big\}, \quad (7)$$

plus terms of higher order in the Δv_r. The likelihood is obtained by dividing by $\prod_j \prod_k (\Delta t_{ijk})$ and taking the limit as $R \to \infty$ to give

$$\begin{aligned}
L_i &= \Big\{ \prod_{j=1}^{J} \prod_{k=1}^{n_{ij}} \lambda_{ij}(t_{ijk}|B_i(t_{ijk})) \Big\} \exp\Big(- \sum_{j=1}^{J} \int_0^{\tau_i} \lambda_{ij}(u|B_i(u))du \Big)\\
&= \prod_{j=1}^{J} \Big\{ \prod_{k}^{n_{ij}} \lambda_{ij}(t_{ijk}|B_i(t_{ijk})) \exp\Big(- \int_0^{\tau_i} \lambda_{ij}(u|B_i(u))du \Big) \Big\}. \quad (8)
\end{aligned}$$

The factorization in (8) reveals that the intensity functions for multi-type recurrent event processes are functionally independent and the estimation of each intensity can be solved separately by maximum likelihood.

2.1.3 Weibull Hazard Rates

In the simulation study in section 2.6, events will be generated according to Poisson processes with Weibull hazard rates. We will now briefly introduce Weibull hazard rates using definitions and form from [11], [13].

Let T be a non-negative random variable representing the time until some event occurs, $F(t)$ be the distribution function of T and $f(t)$ be the density function of T. The survival function is defined as:

$$S(t) = 1 - F(t) = P(T > t)$$

and the hazard function is defined as:

$$h(t) = \lim_{h \to 0} \frac{P(t \le T < t + h | T \ge t)}{h} = \frac{f(t)}{S(t)}$$

In many calculations, the cumulative hazard function is more easily handled and is defined as:

$$H(t) = \int_0^t h(u)du, \quad t > 0$$

For a Weibull distributed variable T, its hazard function and cumulative hazard function are

$$h(t) = \gamma p t^{p-1}. \quad \text{and} \quad H(t) = \gamma t^p, \tag{9}$$

respectively, with parameters γ and p. When $p > 1$, the hazard function is increasing; when $p < 1$ it is decreasing. When $p = 1$ the Weibull distribution reduces to the exponential distribution and the hazard rate is a constant γ over time.

The inverse function of $H(t)$ is

$$g(t) = H^{-1}(t) = \frac{t^{1/p}}{\gamma} \tag{10}$$

2.1.4 Piecewise constant rate functions in Poisson processes

To estimate a continuous rate function in a Poisson process may be computationally heavy, therefore in many studies, e.g. [8], [9], choose to use piecewise constant rate functions. Piecewise constant rate functions are easy to estimate and provide easy random-process generation.

As the name implies, under such models the rate function is assumed to be a constant over prespecified time intervals. It has the form as

$$h(t) = h_k, \quad t_{k-1} < t \le t_k$$

for $k = 1, 2, ..., K$ with K cutpoints. These models have rate functions with discontinuities at the cutpoints but can provide good approximations to various shapes of functions. In [16], it is suggested that models involving three to ten pieces with cutpoints evenly distributed over the event times are flexible enough for most practical situations. We illustrate these models and their likelihood functions in Section 2.3.

In rate estimation and process generation there are two advantages with piecewise-constant form which are stated in [5]. First, piecewise-constant rates are easily estimated: Considering a single type process, in interval

$(t_{k-1}, t_k]$, with occurrence numbers of events n_j of each with total exposed time T_j, the rate can be calculated by:

$$h_k = \frac{n_k}{T_k}.$$

Second, arrival times can be generated by the inverse transformation. The next arrival time is at time t can be simulated as (3) given the previous occurrence time t_{j-1}. The cumulative rate function Λ_j for piecewise-constant rate function is piecewise linear, so solving the next arrival time T_j in terms of the previous arrival time t_{j-1} and a uniform variable $u = P(T_j \leq t|t_{j-1})$ is straightforward.

It is mentioned in several works that how to choose the number of cut-points and length of time intervals is most relevant of estimation qualities. In this paper, we conduct only equidistant time intervals but we will apply a smoothing method to find better estimations. Theoretical description will be presented in the following sections.

2.2 Model outline

The data we use are simulated from a non-homogeneous Poisson process with continuous rate function $h(t)$, where $h(t)$ is the hazard rate function from a Weibull distribution. The Poisson process generates the development of payments and settlements of the reported claims. Three hazard rate functions $h_{se}(t), h_{sep}(t)$ and $h_p(t)$ are used to determine the type of events during the payment process. The rate functions $h_{se}(t)$ and $h_{sep}(t)$ define settlement of claims, where $h_{se}(t)$ generates events of type "se" which implies the settlement of the claim without a payment while $h_{sep}(t)$ generates events of type "sep" which refers to settlement with payment. The rate function $h_p(t)$ is used to define payments without settlement (events of type "p"), i.e. intermediate payments. The above setting has been used in [9] and [21].

One thing we may mention is that the hazard rate functions in survival models are used to determine the risk of experiencing a single event at time t, and if the event occurs the process is stopped. But in our situation, we allow the processes to "continue" when we generate intermediate payments. That means at each event time point, if the event has type "p", the simulation process will continue. It will be stopped when the event type is "se" or "sep". The process differs from a renewal process too because the time that next event occurs depends on the previous occurrence time. Details about how the next time of an event can be simulated will be introduced in Section 2.5.

With the simulated data, the focus of this paper is on estimation of piecewise-constant rate functions and to analyse their predictive abilities.

That means rate functions in the Poisson process will be estimated as piecewise constant in prespecified time intervals. We first use the maximum-likelihood method to obtain estimators. When we have obtained the ML estimators, we also test a smoothing method to see if the smoothed estimators improve the reserve predictions. In the following sections, we will describe in detail how to apply the maximum-likelihood and the smoothing method.

2.3 The likelihood function

The focus of this section is to apply the maximum likelihood method for our development processes. The analysis is based on the work from [9] and [21].

As mentioned, we distinguish between three types of events during the development of a claim: Type "se" events imply settlement of the claim without a payment. A type "sep" event is a payment with settlement at the same time. Type "p" indicates payment without a settlement. $h_{se}(t), h_{sep}(t)$ and $h_p(t)$ denote the rate functions, respectively.

Let V_{ij} denote the time point the jth event occurs of claim i. If the event follows a payment, we assume the payment amount can be specified from a distribution with density function f_P.

Suppose n reported claims are each under observation from time $t = 0$ up to time τ. Later we will use τ as the evaluation time. Then we notice that parts of observations are not yet fully developed. It corresponds to right censoring. For more on censoring, see e.g. [13].

In a continuous time setting, the likelihood of n observed/reported claims development processes is

$$\begin{aligned} L(h., \theta) =& \Big\{ \prod_{i=1}^{n} \Big(\prod_{j} h_{sep}^{\delta_{ij1}}(v_{ij}) h_{se}^{\delta_{ij2}}(v_{ij}) h_{p}^{\delta_{ij3}}(v_{ij}) \Big) \\ & \times \exp \Big(- \int_0^{\tau_i} (h_{sep}(s) + h_{se}(s) + h_p(s)) ds \Big) \Big\} \\ & \times \prod_{i=1}^{n} \prod_{j'} f_{P|V_{ij'}=v_{ij'}}(x_{ij'}; \theta) \end{aligned} \tag{11}$$

where δ_{ijk} denote the indicator function belonging to claim i defined by

$$\delta_{ijk} := \begin{cases} 1, & \text{if } j = k \\ 0, & \phantom{\text{if }} j \neq k \end{cases}$$

where $k = 1, 2, 3$ corresponds to claim type se, sep and p and $\tau_i = \min(\tau, v_i)$. j runs over all registered events in the observation period for claims while $j^{'}$ runs over all paid payments before valuation time τ_i.

If the rate function is assumed to be piecewise constant, i.e. constant over prespecified intervals we follow [21] to modify the likelihood function.

Let $d_0 < d_1 < \cdots < d_K$ denote K cutpoints such that $d_0 = 0$ and $d_K = \tau$. then the rate function is given as

$$h_e(v) \equiv h_{e,l} \quad d_{l-1} < v \leq d_l, \quad e \in \{sep, se, p\}. \tag{12}$$

Let $n^{oc}_{e,l}$ denote the number of observed events of type $e \in \{se, sep, p\}$ in interval $(d_{l-1}, d_l]$. By using the above notation, the likelihood from (11) can be written as

$$\begin{aligned} L(h_., \theta) = &\prod_{l=1}^{K} \Big(h_{se,l}^{n^{oc}_{se,l}}\, h_{sep,l}^{n^{oc}_{sep,l}}\, h_{p,l}^{n^{oc}_{p,l}} \Big) \\ &\times \exp\Big(- \sum_{i=1}^{n} \sum_{l=1}^{K} \big(h_{se,l} + h_{sep,l} + h_{p,l} \big) \int_{d_{l-1}}^{d_l} \mathbf{1}(s \leq \tau_i) ds \Big) \\ &\times \prod_{i=1}^{n} \prod_{j'} f_{P|V_{ij'}=v_{ij'}}(x_{ij'}; \theta) \end{aligned} \tag{13}$$

In this thesis, we simplify the payment amount as a constant; then the last double product is equal to 1. Parameters which need to be estimated are only the piecewise constant hazard rates $h_{e,l}, e \in \{se, sep, p\}$, $l = \{1, ..., K\}$, K is the number of intervals.

2.4 Estimation of parameters

2.4.1 The Maximum likelihood method

Parameters can be estimated by using the standard maximum likehood method. With a piecewise constant specification of the hazard rate function with K cutpoints, $h_{e,l}$ with $e \in \{se, sep, p\}$, $l = \{1, \cdots, K\}$ need to be estimated.

Due to the independence between the processes, we can estimate the constant rates $h_{e,l}$ separately.

Let $\omega_l(v) = \mathbf{1}(d_{l-1} < v \leq d_l)$, $l = 1, ..., K$ indicate whether $v \in (d_{l-1}, d_l]$, $n^{oc}_{e,l}$ indicates the total number of events observed by all claims with type $e \in \{se, sep, p\}$ over time interval $(d_{l-1}, d_l]$, then

$$n^{oc}_{e,l} = \sum_{i=1}^{n_e} \sum_{j=1}^{n_i} \omega_l(v_{il}), \tag{14}$$

where n_e is the number of claims with type e, n_i is the maximum number exposured in each time interval.

Let S_l denote the total explosure time in $(d_{l-1}, d_l]$ across all individuals, then

$$S_l = \sum_{i=1}^{n} \int_{d_{l-1}}^{d_l} \mathbf{1}(s \leq \tau_i) ds \tag{15}$$

The estimate of $h_{e,l}$ can be obtained by maximizing the likelihood (13), then we get

$$\hat{h}_{e,l} = \frac{n_{e,l}^{oc}}{S_l}, \quad e \in \{se, sep, p\}. \tag{16}$$

All $\hat{h}_{e,l}$ are independent. Thus, based on (15) and (16), variances of hazard rates are estimated by $\hat{h}_{e,l}/S_{e,l} = n_{e,l}^{oc}/S_{e,l}^2$. The estimated cumulative hazard function is

$$\hat{H}_e(t) = \sum_{l=1}^{K} \hat{h}_{e,l}\, \Delta d(t_l) = \sum_{l=1}^{K} \hat{h}_{e,l} \int_{d_{l-1}}^{d_l} \mathbf{1}(t \geq s) ds \tag{17}$$

where $\Delta d_l(t)$ is the length of interval.

2.4.2 The iterative smoothing method

This smoothing method is developed by [5] and we describe it in this subsection.

It is assumed that we already have an estimated rate function vector $\hat{\boldsymbol{h}} = \{\hat{h}_1, ..., \hat{h}_K\}'$ with K cutpoints. We will now determine a new Poisson rate function $\boldsymbol{\kappa}$ which will hopefully be "better" than $\hat{\boldsymbol{h}}$. This method bases on the estimated result $\hat{\boldsymbol{h}}$ which indicates that any statistical errors in the given constant rates $\hat{\boldsymbol{h}}$ will be remained.

The method requires that for each time interval the mean number of arrivals remains unchanged, i.e.

$$\int_{t_{k-1}}^{t_k} \kappa(t) dt = \int_{t_{k-1}}^{t_k} h(t) dt.$$

It is constrained also that the solution should be nonnegative and symmetric in time. That is, reversing the subscripts on $h_1, ..., h_K$ should result in reversing the time index of κ.

The "better" or "smoother" result is evaluated by the smoothing function

$$z = \sum_{k=1}^{K-2} (h_{k+2} - 2h_{k+1} + h_k)^2 \tag{18}$$

for piecewise-constant rate functions defined by $\boldsymbol{h} = \{h_1, ..., h_K\}$. In this definition, the smaller value of z corresponds to a smoother function. Perfect smoothness is a straight line and $z = 0$.

We use only one of the smoothing methods in [5]: the number of intervals are doubled. This method is formulated as follows: Given that we have a Poisson rate function $h(t)$ in the form of K nonnegative constants $h_1, ..., h_K$ corresponding to the time intervals $(t_{k-1}, t_k]$ for $k = 1, ..., K$, then we will find a solution of returning a piecewise-constant rate function with $2K$ pieces, each of length half that of the original intervals. The new rate function is

$$\kappa(t) = \begin{cases} h_k - \gamma_k & \text{for} \quad t_{k-1} < t \leq (t_{k-1} + t_k)/2 \\ h_k + \gamma_k & \text{for} \quad (t_{k-1} + t_k)/2 < t \leq t_k \end{cases}$$

With this function, we decrease h_k in the left half part of interval by γ_k and increase the right half part by γ_k in interval $(t_{k-1}, t_k]$. Any values of the decision variables $\gamma_1, ..., \gamma_K$ satisfy the constraints will be the answer. It means any non-negative and symmetric values $\gamma_1, ..., \gamma_K$ then the expected number of arrivals is unchanged for interval.

To obtain the nonnegative solution, we must constrain $|\gamma_k| \leq h_k$ for $k = 1, ..., K$.

The idea of this method is to choose the value of γ_k in each interval k to match the slope of h near interval k. In this sense, the slope of the original rate function h in segment k is

$$\frac{h_{k+1} - h_{k-1}}{2},$$

the difference in the h values adjacent to segment k divided by the time difference between their midpoints.

Similarly, the slope of the new rate function τ is

$$\frac{(h_k + \gamma_k) - (h_k + \gamma_k)}{1/2},$$

Setting the two slopes equal yields the first-order slope-matching solution is

$$\gamma_k = \frac{h_{k+1} - h_{k-1}}{8}. \tag{19}$$

The end intervals, 1 and K are defined by $\gamma_1 = 2h_1 - h_2$ and $\gamma_K = 2h_K - h_{K-1}$, respectively.

As soon as the γ_k is obtained, we get the new hazard rates $\tau_{k,left}$ and $\tau_{k,right}$.

This method will be applied to smooth the ML-estimators. Evaluation will be done both through comparing the z values in (18) and the predicted reserves.

2.5 Simulating future events

For a claim, as we mentioned in the first section, it follows a development process from occurrence to settlement. The timeline can be divided into two parts: The first part is the time between the occurrence and reporting, the second is between the reporting and settlement. In this thesis, we focus on the second part. That is, the claims which have been acknowledged or reported to the insurance company (RBNS claims). Once a claim is reported and acknowledged, a series of payments occur until the claim is settled. At present time (say τ), parts of claims are not yet fully developed, i.e. they are not yet settled (or closed). That means the insurance company needs simulate, or predict, future payments and set a reserve for them. We will use a non-homogeneous Poisson process with piecewise-constant rate functions to do that.

To describe this in more detail, we define the following notations for claim i.

- s_i = the time since reporting at valuation time,
- τ = valuation time (time of censoring),
- V_i = time to next event,
- E_i = type of next event,
- e = {se, sep, p} group of event types.

Simulating time to next event.

The time-dependent rate functions $h_e(t), e \in \{se, sep, p\}$ are specified as hazard rate functions from the Weibull distribution. From the relation between hazard rate and distribution function the first event occurs at time v can be obtained as:

$$P(V \leq v) = 1 - \exp\big(-\int_0^v \sum_e h_e(t)dt\big) \tag{20}$$

For the reported claims, the time since reporting s is known. For claim i the next event at time v_{next} can take place at any time with condition $v_{next} > s_i$, where s_i is the time since reporting for claim i. The conditional distribution then can be written as:

$$
\begin{aligned}
P(V_i \leq v_{next}|V_i > s_i) =& \frac{F(v_{next}) - F(s_i)}{1 - F(s_i)} \\
=& \frac{\exp(-\int_0^{s_i} \sum_e h_e(t)dt) - \exp(-\int_0^{v_{next}} \sum_e h_e(t)dt)}{\exp(-\int_0^{s_i} \sum_e h_e(t)dt)} \\
=& 1 - \exp\big(-\int_{s_i}^{v_{next}} \sum_e h_e(t)dt\big).
\end{aligned} \tag{21}
$$

We know that $\int_{s_i}^{v_{next}} \sum_e h_e(t)dt$ is the cumulative rate function of all three types between time s_i and v_{next} and we denote it as $H_i(s_i, v_{next})$. To simulate v_{next} we use the inverse transformation method in [22] and draw $u \sim U(0,1)$ to get

$$u = P(V_i \leq v_{next}|V_i > s_i).$$

In our study we specify the rate functions as piecewise constant which is defined as (12). The cumulative rate function over time $(s_i, v_{next}]$ then is:

$$
\begin{aligned}
H_i(s_i, v_{next}) =& \sum_{l=1}^{L} \Big(\mathbf{1}\{d_{l-1} < s_i \leq d_l\}(d_l - s_i)\sum_e h_{e,l} \\
&+ \mathbf{1}\{s_i < d_l < v_{next}\}(d_l - d_{l-1})\sum_e h_{e,l} \\
&+ \mathbf{1}\{d_{l-1} < v_{next} \leq d_l\}(v_{next} - d_{l-1})\sum_e h_{e,l}\Big).
\end{aligned} \tag{22}
$$

In (22) the first term on the right calculates the cumulative rates between s_i and the upper limit where s_i is located, the second term adds the cumulative rates of intervals between s_i and v_{next} (not including the interval where s_i and v_{next} is located) and the third term calculates the cumulative rate in the interval where v_{next} is located.

Then (21) can be rewritten to:

$$H_i(s_i, v_{next}) = -\ln(1-u), \tag{23}$$

then the solution is

$$V_i = \{v : H_i(s_i, v) = -\ln(1-u)\} \tag{24}$$

When v_{next} is obtained, set $s_i = v_{next}$, as the time of the previous event and repeat the process. The process will be stopped when the generated event is of type "se" or "sep". In the next section we describe how to generate event types.

Simulating event types. We define three types of events: A type "se" event implies settlement without payment, a type "sep" event indicates settlement with a payment at the same time and type "p" events generate payment without settlement, i.e. intermediate payments. Remember that type "sep" and type "p" events contain a positive payment but not for type "se" events. A type "se" or a type "sep" event means settlement of claims while a type "p" event indicates the process will continue to develop. Given the next event time v_{next} , $d_{l-1} \leq v_{next} < d_l$, the probability that the next event is in type "se", "sep" or "p" can be simulated by

$$\begin{aligned} P(E_i = "se" | \text{one event at } v_{next}) &= \frac{h_{se}(v_{next})}{\sum_e h_e(v_{next})} \\ P(E_i = "sep" | \text{one event at } v_{next}) &= \frac{h_{sep}(v_{next})}{\sum_e h_e(v_{next})} \\ P(E_i = "p" | \text{one event at } v_{next}) &= \frac{h_{p}(v_{next})}{\sum_e h_e(v_{next})} \end{aligned}$$

plugging in the estimated results, we get:

$$\begin{aligned} P(E_i = "se" | v_{next} \in \{d_{l-1}, d_l\}) &= \frac{\hat{h}_{se,l}}{(\hat{h}_{se,l} + \hat{h}_{sep,l} + \hat{h}_{p,l})} \\ P(E_i = "sep" | v_{next} \in \{d_{l-1}, d_l\}) &= \frac{\hat{h}_{sep,l}}{(\hat{h}_{se,l} + \hat{h}_{sep,l} + \hat{h}_{p,l})} \\ P(E_i = "p" | v_{next} \in \{d_{l-1}, d_l\}) &= \frac{\hat{h}_{p,l}}{(\hat{h}_{se,l} + \hat{h}_{sep,l} + \hat{h}_{p,l})} \end{aligned} \tag{25}$$

Payment amount. Given that the process follow the "*sep*" and "*p*" type we should simulate a payment amount. In this thesis we simplify the process and assume the payment is 1 unit.

Summary. The development process for observed claim i which has developed s_i time units since reporting before being censored is simulated according to the following:

(i) Simulate time to next event V_{next} from (24).

(ii) Given the next event time v_i, calculate the probability of type "se", "sep" or "p" from (25);

(iii) Set $p_1 = P(\text{type "se"})$, $p_2 = P(\text{type "se"}) + P(\text{type "sep"})$. Draw a random number u from U(0,1), if

$0 \leq u < p_1 \implies$ type "se" event, stop the process.

$p_1 < u \leq p_2 \implies$ type "sep" event, set 1 unit and stop the process.

$u > p_2 \implies$ type "p" event, set 1 unit, let$s_i = v_i$, go to step (i).

2.6 Simulating datasets

Different from [9] and [21], we use simulated data in this analysis. A non-homogeneous Poisson process with a mean function $\mu(t)$ can be generated by using the following theory which is described in [24].

The homogeneous and the non-homogeneous Poisson process are very closely related [24]. The relationship is defined as follows:

Let N be a Poisson process on $[0, \infty)$ with mean function $\mu(t)$. Let $\widetilde{N}$ be a standard homogeneous Poisson process, i.e. a homogeneous Poisson process with intensity $\rho = 1$. Then:

- The process $(\widetilde{N}(\mu(t))_{t \geq 0}$ is Poisson process with mean value function μ.
- If μ is continuous, increasing and the invers μ^{-1} exists then $N(\mu^{-1}(t))_{t \geq 0}$ is a standard homogeneous Poisson process.

With these relationships a non-homogeneous Poisson process $N(t)$ with mean value function $\mu(t)$ can be interpreted as a time change of a standard homogeneous Poisson process $\widetilde{N}$:

$$(N(t))_{t \geq 0} \stackrel{d}{=} (\widetilde{N}(\mu(t)))_{t \geq 0}.$$

Therefore the occurrence time of events from a non-homogeneous Poisson process with mean value function μ have representation

$$T_n = \mu^{-1}(\widetilde{T}_n), \quad \widetilde{T}_n = \widetilde{W}_1 + \cdots + \widetilde{W}_n, \quad n \geq 1, \quad \widetilde{W}_i \sim \text{Exp}(1). \tag{26}$$

When assuming a continuous Weibull hazard rate, i.e. $\rho(t) = \gamma p t^{p-1}$, see Section 2.1.3. the expected cumulative number of events at time t is given by

$$\mu(t) = H(t) = \gamma t^p. \tag{27}$$

By using this, it is not difficult to calculate the inverse of $\mu(t)$:

$$g(T) = \mu^{-1}(T) = \big(\frac{T}{\gamma}\big)^{1/p}. \tag{28}$$

Given the values of parameters γ and p, using (26) and (28) we can now illustrate the occurrence time of events. We vary values of γ and p to generate three parallel Poisson processes. They compete with each other to decide how long the process will be developed and which settlement type will be generated.

Each sample contains 10 000 reported claims during 5 accident years with on average 2 000 claims each accident year. We assume all claims are reported to the insurance company immediately after the accident, i.e. no reporting delay. Let $i = 1, ..., 10\,000$ index each claim. For claim i, we will generate the full development process with $\{V_{ij}, E_{ij} : j = 1, ..., J_i\}$ where V_{ij}, E_{ij} are the occurrence time and event type of the jth event, respectively.

For claim i, the sampling procedure can be summarized as follows.

(i) Simulate a time point T_{sep}. Draw $W_1 \sim \text{Exp}(1)$ and calculate $T_{sep} = g(W_1)$ with parameters γ_{sep}, p_{sep} according (26) & (28).

(ii) Simulate a time point T_{se} by using the same mechanism as in (i) but with parameters γ_{se} and p_{se}.

(iii) Simulate the first time point $T_{p,1}$ by using the same method as in (i), but with prameters γ_p and p_p.

(iv) Stop or continue the process. Set $T_{\text{stop}} = \min(T_{sep}, T_{se})$ and compare with $T_{p,1}$, if $T_{p,1} > T_{\text{stop}}$, the procedure is stopped, otherwise make a payment 1, and continue the process.

(v) Simulate time points in case of process will continue. Draw $W_2 \sim \text{Exp}(1)$ and obtain $T_{p,2}$ by calculating $T_{p,2} = g(W_1 + W_2)$ from (26) & (28). Thereafter Compare $T_{p,2}$ with T_{stop}. If $T_{p,2} < T_{\text{stop}}$ repeat this simulating procedure, so we get $T_{p,3}, \cdots T_{p,i}$ until $T_{p,i} \geq T_{\text{stop}}$ and the procedure is stopped.

(vi) Settlement time in case of process is stopped. Set T_{stop} as the settlement time. If $T_{\text{stop}} = T_{se}$ it indicates that claim i is settled without payment and the settlement time is T_{se} otherwise claim i is settled at time T_{sep} with a last payment.

(vii) Repeat the sampling procedure for all claims.

(viii) Draw a random number from a uniform distribution in time interval $(0, \tau)$, where τ denotes the available time for data (60 months in our case), to place claims into different accident years.

3 Data

Two types of datasets with sample size $n = 10\,000$ in each are generated by selecting values of parameters γ and p from continuous hazard rate functions $h(t) = \gamma p t^{p-1}$. Set $\gamma = a^{-b}, p = b$, where a and b are the scale and shape parameters in the Weibull model, respectively, see [11]. The mean and variance in the Weibull distribution are $a\Gamma(1+1/b)$ and $a^2\big[\Gamma(1+2/b) - \big(\Gamma(1+1/b)\big)^2\big]$. This thus corresponds to the expected value and variance of the time to the first event of a type of event. Here $\Gamma(.)$ means the Gamma function. For more information about the Gamma function, see e.g. [23]. Two types of datasets are assumed to be available for five years back in time (uniformly distributed under five years). These two simulated types of datasets reflect two types of insurance products: one with long settlement time and another with short settlement time. It is important to analyze two distinct lines of business since the settlement time has an important effect on the work of reserving. Therefore it is also of interest to investigate if the method works differently on the long-settlement-time data and short-settlement-time data. Details about the parameters used to generate each dataset are documented in Table 1. For simplicity, we call the long-settlement-time data "Data1" and the short-settlement-time data "Data2".

Table 1: Parameter settings: γ_{se}, p_{se} are used to generate Type "se" events; γ_{sep}, p_{sep} are used to generate Type "sep" events; γ_p, p_p are used to generate events with Type "p".

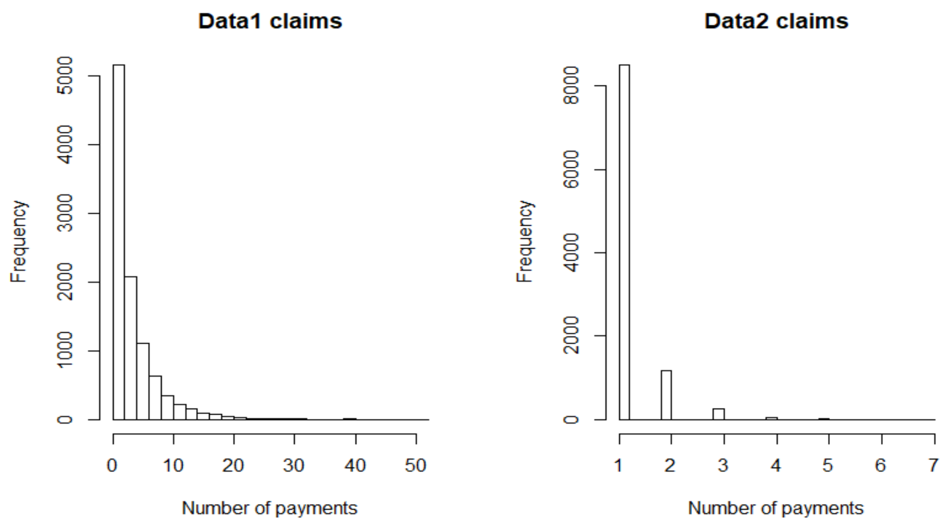
	γ_{se}	γ_{sep}	γ_p	p_{se}	p_{sep}	p_p
Data1	0.65	0.5	0.8	0.15	0.06	0.32
Data2	0.55	0.2	0.4	0.63	1.4	0.35

We present descriptive statistics for two simulated types of datasets with complete development process in Table 2. Number of payments and settlement delay are shown in Figure 2 and 3, respectively. Claims from Data1 have a payment pattern with more intermediate payments and longer settlement time. The average settlement time is 18 months and 93% of events are settled after five years. Maximum number of payments can be more than 40 times before settlement. With maximal observation time $s = 60$ months, it remains about 25% active claims from all five accident years. Typical products with similar development pattern as claims from Data1 are e.g. bodily injury liability insurance, errors & omissions liability insurance. These insurance products are complicated to predict future payments due to big variation and heavier right tail. It is also common with extreme payment amounts.

Table 2: Initial categories: Average number of payment, average settlement time, number of settlement during five development years.

	Data1	Data2
Average no. of payments	3.98	1.19
Average settlement time (month)	18.12	1.34
No. of settl. Y1 (%)	61.9	99.2
No. of settl. Y2 (%)	15.6	0.6
No. of settl. Y3 (%)	7.7	0.1
No. of settl. Y4 (%)	4.5	0.07
No. of settl. Y5 (%)	3	0.04
No. of settl. $>$ Y6 (%)	7.3	0

Figure 2: A histogram of number of payments based on 10 000 simulated data: Data1 (left) and Data2 (right).



In contrast, claims from Data2 have a relatively simple process and develop quickly to settlement. Most of claims are settled within 12 months (99%) and future payments are small. With a maximal observation time of 60 months, there are no active claims anymore. Insurance products which have similar development pattern as claims from Data2 are plenty, for instance, material damage insurance [9] and animal insurance.

We follow the studies of [9] and [21] and distinguish three types of events during the development of a claim for both datasets. With parameter settings in Table 1, for claims from Data1, process with type "se", i.e. settlement without payment has an expected settlement time of 25 months while the process with type "sep" (settlement with payment) has 500 months. The first three expected payment times which simulated from

Figure 3: A histogram of settlement delay : Data1 (left) and Data2 (right).

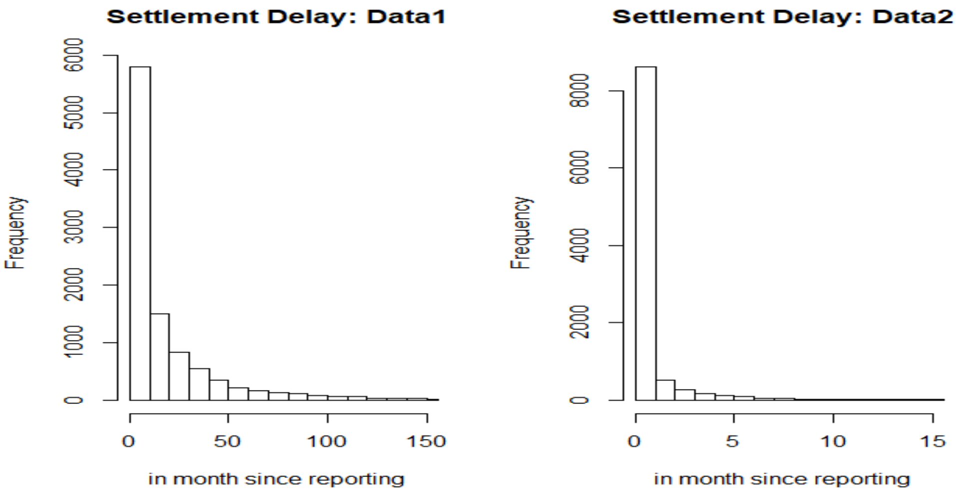
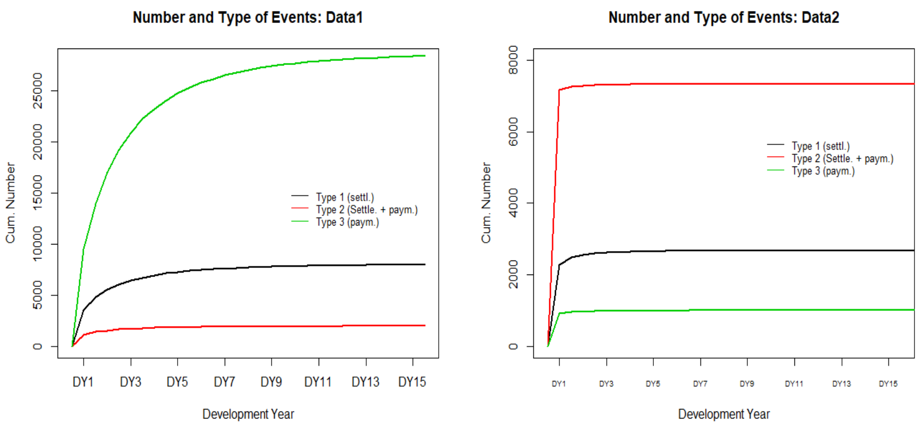


Figure 4: Cumulative number of events as a function of development years: Data1 (left) and Data2 (right).



the process with parameters γ_p and p_p are 2, 17 and 62 months, respectively. The left panel of Figure 4 gives the cumulative number of events with the three types over development years for Data1 claims.

The expected settlement times for claims from Data2 with types "se" and "sep" are 13 and 22 months respectively, while the first three average payment times are 14, 99 and 316 months. Comparing the average settlement and each payment time, the differences between the development patterns of the two datasets are obvious. The cumulative number of events for Data2 claims are shown in the right panel of Figure 4.

Both type "sep" and "p" events generate a payment. In this thesis we are only interested in quantifying the effect on the best estimate reserve when using piecewise constant hazard rates in a non-homogeneous Poisson process. Due to this we will not simulate the payment amount but let it queal to 1 for both types of events and datasets.

4 Analysis and result

In this section, we will go through a more detailed description of parameter estimation, reserve prediction as well as uncertainty test.

4.1 Data1

4.1.1 Estimating hazard rates

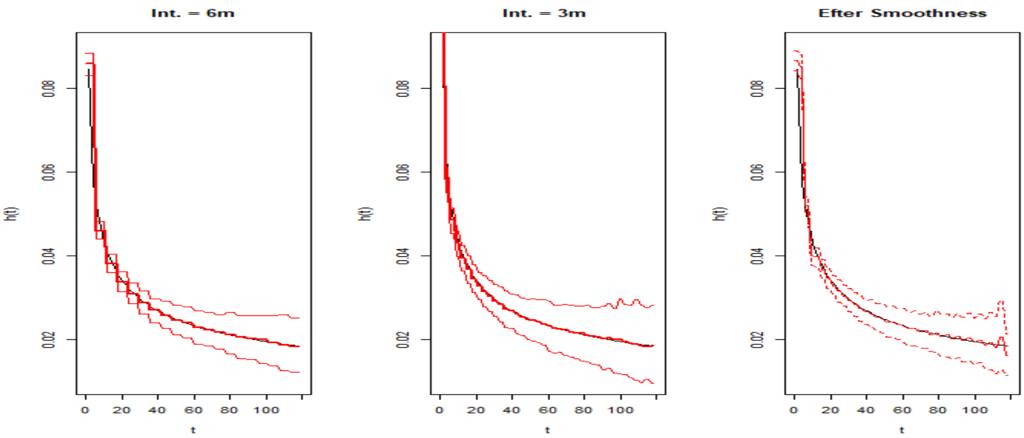
A key point of estimating piecewise-constant rate function is the choice of time intervals. In practice, as noticed by [15] for Poisson regression, *"the choice of time intervals should generally be guided by subject matter aspects, but the numbers of events and numbers at risk within intervals may also be considered when specifying the number and lengths of the intervals.* In our situation, we have no prior knowledge for the choice of the time intervals so a reasonable choice is to use equidistant time intervals. Two lengths of time intervals for Data1 claims will be analysed: six months and three months. Times of events, therefore, are divided into multiple non-overlapping subgroups. In case of hazard rates that are constant on six month intervals, the times of events with type e, $e \in \{sep, se, p\}$ are subsetted into 21 intervals, namely $[0, 6)$ months, $[6, 12)$ months,..., $[114, 120)$ months and ≥ 120 months, generating a parameter vector $\boldsymbol{h_e^{6m}}$. For hazard rates that are specified as constant on three month intervals, the time intervals are set by $[0, 3)$ months, ..., $[117, 120)$ months and ≥ 120 months, producing a total of 41 intervals. we store these in the vector $\boldsymbol{h_e^{3m}}$.

ML estimates of piecewise-constant hazard rates is straightforward and can be obtained as (16).

Furthermore, since we are estimating a continuous function of time, it is of interest to analyse the effect of smoothing the ML-estimates. In particular, we want to see how the smoothed estimators perform w.r.t. the reserve prediction. We will apply the iterative smoothing method from Section 2.4.2 to the six-month-constant estimators. The estimators are in vector $\boldsymbol{h_e^{smooth}}$. Notice that this smoothing method can be applied iteratively, see [5], but we only do one step in this thesis.

Figure 5 presents the mean of the fitted hazard rates of type "se" events as well as the 5% and 95% empirical percentiles (simulation size $m = 500$). Results of hazard rates with type "sep" and type "p" can be found in Appendix 1.

Figure 5: Fitted hazard rates of process with type "se". *Bold red lines* describe the mean of estimates; *dashed lines* are the 5% and 95% empirical percentiles based on 500 simulations. *Black lines* show the true continuous hazard rate functions. Left panel: hazard rates are constant on six month intervals; Middle panel: hazard rates are constant on three month intervals; Right panel: smoothed results based on the six-month-constant hazard rates.



The left panel of Figure 5 shows the estimated results with the specification that hazard rates are constant on six-month intervals. It seems that the estimated hazard rate in the first time interval is overestimated but works quite well in other ranges. The 90% confidence interval widens during the development year which indicates a increased uncertainty for estimates during the development time.

The middle panel of Figure 5 shows the estimated hazard rates with a three-month constant specification. The mean values perform better than six-month-constant hazard rates, but the 90% confidence intervals widen larger. The differences appear more clearly as the process develops. It is not a surprise. More cut points yields an increased number of subgroups together with a decreased number of observations in each subgroup. Fewer observations of course leads to a more substantial uncertainty.

The right panel of Figure 5 shows the results when we apply the smoothing method to the estimated six-month-constant hazard rates. The figure clearly shows that the ML-estimators have become smoother, but at the same time, producing wider confidence intervals. One thing we should note is that the smoothed hazard rate in the first time interval is still overestimated.

We see from Appendix 1 the estimated hazard rates of type "sep" and "p" events perform similarly to those of type "se" events.

As we noticed, the variation increases over time in all three panels due to the decreased number of observations for larger observation times. So it may be useful to decrease variation by aggregating those time intervals with too few observations. For Data1 we choose to aggregate the intervals after 60 months. Reasons for this are first after 60 months there are less than 3% of total events in each interval which makes it hard to estimate the hazard rates. Then the total numbers of events after 60 months distribute 7% of the total events which we believe that it does not loss much information after the aggregation. Estimates for type "sep" and type "p" events perform similarly to the type "se" events therefore the same aggregating procedure will be applied to these two types too.

After aggregation, the performance of the estimated results of the three types are presented in Figure 6. The fitting of mean value in the aggregated interval is not much worse while the variation has decreased dramatically.

The values of z in (18) for corresponded $\hat{\boldsymbol{h}}_{\boldsymbol{e}}^{\boldsymbol{6m}}$ and $\hat{\boldsymbol{h}}_{\boldsymbol{e}}^{\boldsymbol{smooth}}$ based on the aggregated data are presented in Table 3. Remember that the smaller value of z corresponds to a smoother result. The result in Table 3 shows that the smoothing method has not significantly smoothed the ML-estimators after we aggregate the time intervals.

Table 3: Smoothing function value z for hazard rates estimators.

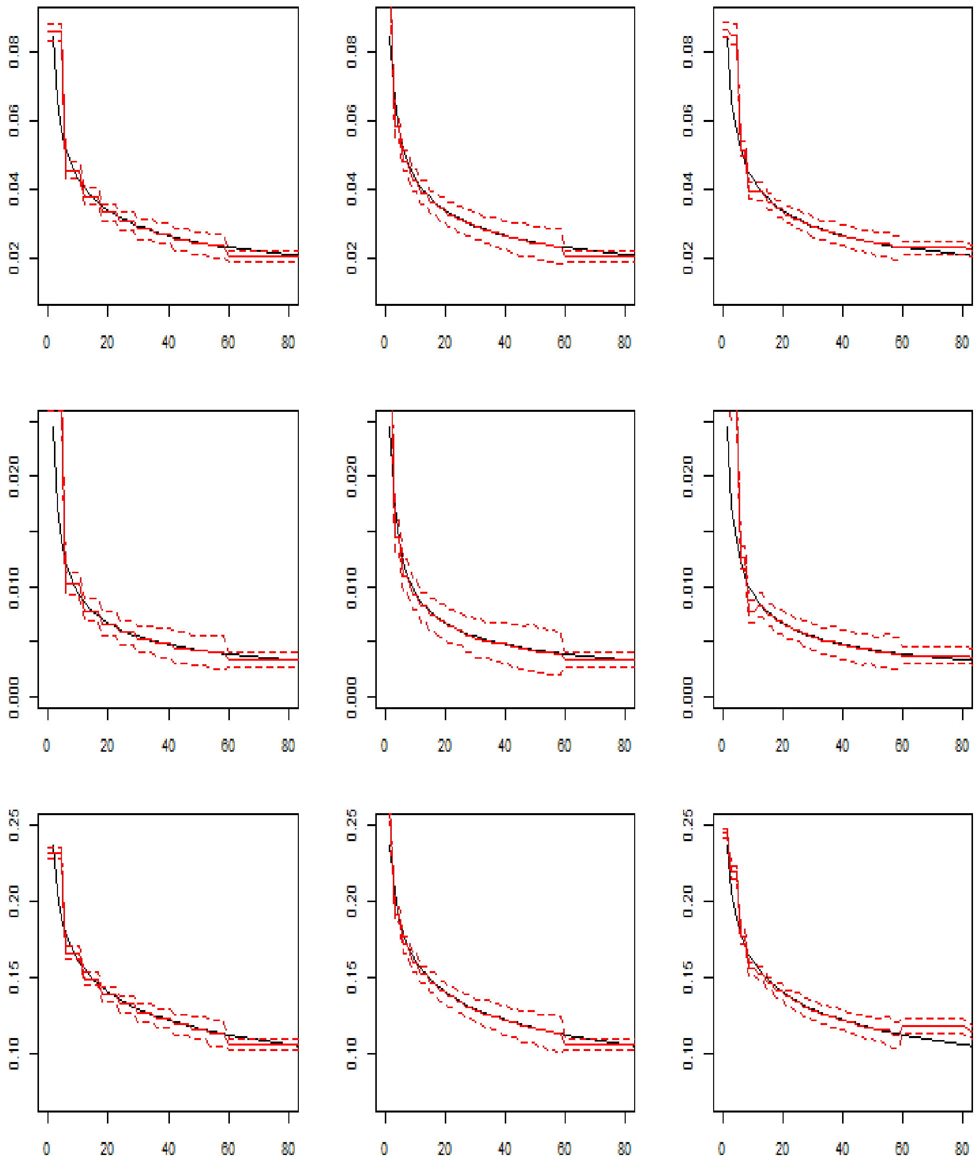
	$\hat{\boldsymbol{h}}^{\boldsymbol{6m}}$	$\hat{\boldsymbol{h}}^{\boldsymbol{smooth}}$
Type "se"	0.0015	0.0021
Type "sep"	0.00019	0.00024
Type "p"	0.0029	0.0024

4.1.2 Effect of sample size and simulating size on estimation errors

The reported results above are based on 10 000 claims. To explore the impact of the number of claims on the estimated results, we will also compare the estimations by using 5 000, 50 000 and 100 0000 claims. For reducing the computational load, only the events with type "se" are experimented and the estimated hazard rates in the first six months are presented in Figure 7. (Estimations in the interval [54, 60) are shown in Appendix 2). For each sample size, there are not any significant changes in the expected values of the estimations, the standard deviations decrease proportionally with the square root of the number of claims.

Another possible factor which may affect uncertainty of estimation is the simulation size. Considering both the computational load and the results seen in Figure 7, we decide to fix the sample size to 10 000 and experiment 500, 5 000 and 10 000 estimations to explore this effect.

Figure 6: Estimated hazard rates together with 5% and 95% empirical percentiles based on 500 simulations. *Columns*: hazard rates are constant on six month intervals; hazard rates are constant on three month intervals; Smoothed results based on the six-month-constant hazard rates. Rows: type "se", type "sep" and type "p" events. All results are assumed with no further changes efter 60 months in the hazard rates. The solid black lines show the true continuous hazard rates



As shown in Figure 8, it has no significant changing in the mean values and standard deviations when we increase the simulation size. This phenomenon indicates that if the sample size is sufficiently large simulation

Figure 7: Histogram of piecewise constant hazard rates estimated from events with type "se" on the first six months. The sample size is 5 000, 10 000, 50 000 and 100 000. The solid black lines present the mean value of simulations.

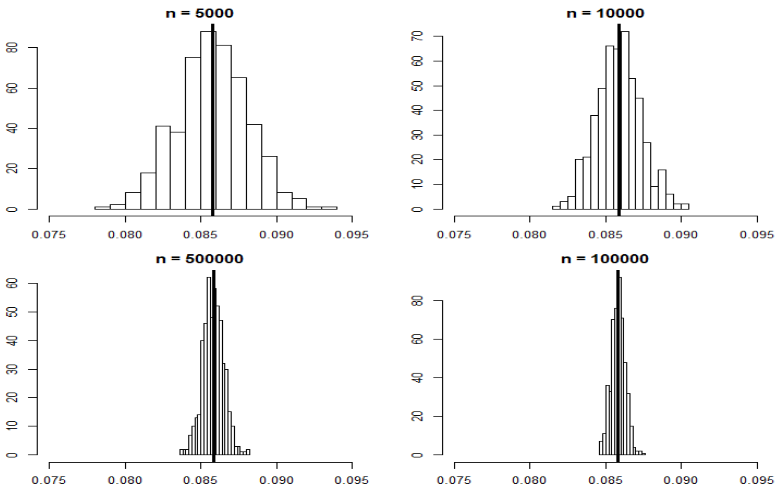
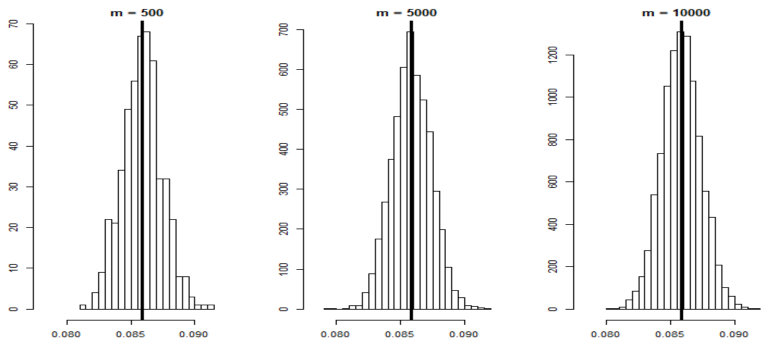


Figure 8: Histogram of piecewise constant hazard rates estimated from events with type "se" on the first six months. Sample size is 100 000, simulation size m is 500, 5 000 and 50 000. The solid black lines present the mean value of estimations.



size will not affect the estimation uncertainty to any greater extent.

4.1.3 Reserves

In this section, we will study the performance of the three estimated hazard rates by comparing predicted reserves. Five-year-ahead cash flows, total reserves as well as predictions of each type are obtained according to the *Simulating procedure* described in Section 2.5. Three estimated parameter vectors $\hat{\boldsymbol{h}}_{\boldsymbol{e}}^{\boldsymbol{6m}}$, $\hat{\boldsymbol{h}}_{\boldsymbol{e}}^{\boldsymbol{3m}}$ and $\hat{\boldsymbol{h}}_{\boldsymbol{e}}^{\boldsymbol{smooth}}$ will be used to simulate the future payments.

To evaluate the performance of a reserving model, we use the percentage reserve error *(RE)* which is defined by

$$RE = \frac{\hat{R} - R^{sim.}}{R^{sim.}}. \tag{29}$$

Where $\hat{R}$ means the predicted results obtained from constant hazard rates, $R^{sim.}$ denotes the actual future payments from the simulation. The percentage reserve error is easily calculated and its standard deviation can also be estimated empirically. The disadvantage of RE is that we lose information about prediction size and distribution of cash flows in total reserve which we may be interested in. Thus we choose to analyse the predicted cash flows/reserves $\hat{R}$ and the actual future payments $R^{sim.}$ in each year. Due to the actual future payments are not unchanged when we each time simulate claims, the mean value of simulations $\bar{R}^{sim.}$ will be calculated.

As we mentioned before, we have simplified this micro model by assuming the payment amount is 1 unit. Then the cash flows in calendar year k can be obtained by:

$$R_k = N_k = N_{sep,k} + N_{p,k} = \sum_{j\geq 0} \mathbf{1}\Big(k \leq T_j^i \leq k+1, E_j^i \in \{sep, p\}\Big) \tag{30}$$

where T_j^i is the time of jth event of claim i since reporting. E_j^i denotes type of jth events for claim i. Only events with type "sep" and type "p" contribute to the reserve predictions. The total reserve for five years is

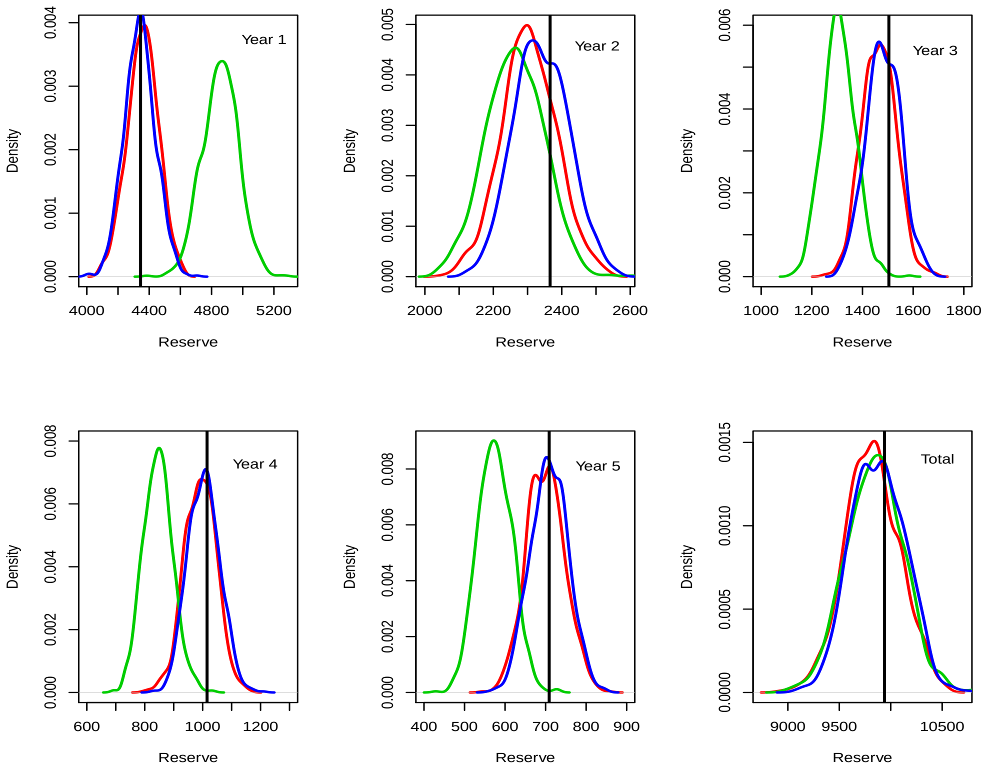
$$R_{tot.} = \sum_{k=1}^{5} R_k \tag{31}$$

We use RE to perform the comparison of three types while we use $\hat{R}$ and $\bar{R}^{sim.}$ to illustrate the development of cash flows and total reserves.

Cash flows and total reserves in five years are shown in Figure 9. The red lines present kernel densities of cash flows/reserves with six-month-constant hazard rates as intensities in the Poisson process. The blue lines show the kernel densities of cash flows/reserves with three-month-constant rate function. The green lines show results obtained from the smoothed hazard rates. The solid black lines correspond the mean value of future payments $\bar{R}^{sim.}$ from the simulations.

It seems that the predictive kernel densities obtained from the six-month-constant and three-month-constant hazard rates are more realistic than those from the smoothed estimators. The best estimate with six-month-constant and three-month-constant hazard rates in year 1, 4 and 5 are very close to the "true" reserve, but in year 2 and 3 are underestimated.

Figure 9: Predicted cash flows and total reserves for five years. First five results (from left to right, from top to bottom) are cash flows in five years. The last panel (bottom row on the right) shows the total reserves. It is based on 1 000 simulations and 10 000 claims. Red line: six-month-constant hazard rates; Blue line: three-month-constant hazard rates; Green line: smoothed hazard rates. Solid black line: the mean of the actual future payments obtained from the simulation.



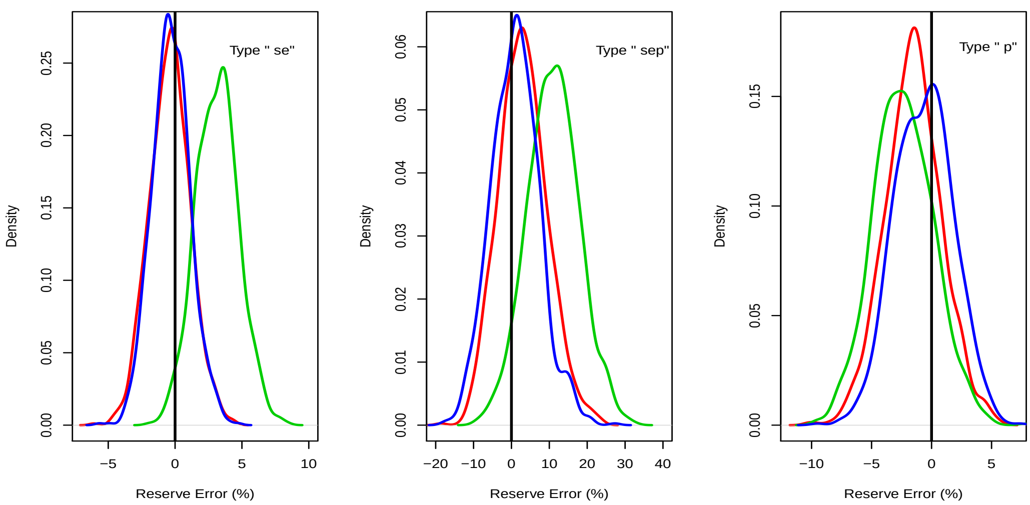
According to the results from the smoothed hazard rates, it seems that this method has not improved the performance of ML-estimators. It results in an obvious bias for each calendar year. It overstates the reserve in year 1 and understates in the other years.

For the total reserves, from the right panel on the bottom row in Figure 9, it is seen that all three parameter settings $\hat{\boldsymbol{h}}_e^{\boldsymbol{6m}}$, $\hat{\boldsymbol{h}}_e^{\boldsymbol{3m}}$ and $\hat{\boldsymbol{h}}_e^{\boldsymbol{smooth}}$ predict well. Even though the smoothed estimates do not give a good fit to the cash flows in each year, it results in a quite good performance in total reserves.

If we focus on the kernel densities of cash flows in total reserve, it shows that the cash flow in year 1 distributes around 45% of the total reserve which almost corresponds to the total sum of the remaining years. That is maybe an explanation of why the smoothed estimates give unrealistic predictions of cash flows but this is not shown in the total reserve.

To qualify the predictive ability of constant hazard rates in each type, we illustrate the kernel densities of the percentage reserve errors in Figure 10. Note that for type "sep" and type "p" events, the results correspond the reserve errors, but for type "se" event, it does not generate reserve so the results are just the percentage predictive errors of the numbers of settlement where no payment occurs. We still use name "Reserve error" for type "se" in Figure 10 just for consistency considering.

Figure 10: Kernel densities of percentage Reserve Errors by event type. Red line: six-month-constant hazard rates; Blue line: three-month-constant hazard rates; Green line: smoothed hazard rates. The black lines present value zero. Note that for type "se" event, it does not generate the reserve but the predictive errors of the numbers of events.



The six-month-constant hazard rates give a mild positive bias for type "sep" events and a negative bias for type "p" events but it gives outstanding predicted results of events with type "se". Its best estimate matches almost the actual value and the variation is little.

Three-month-constant hazard rate works well too. It produces a slightly larger variation of reserve errors with type "p" events than those with the six-month-constant setting but they are not significant when we compare the errors with type "se" and type "sep" events.

The smoothed hazard rates do not perform as good as the other two. The best estimates of errors for type "se" and type "sep" are positive which mean the overstatement while the best estimate is negative for type "p" events, i.e. it results in fewer payments than the actual one. The positive results for type "sep" and the negative value for type "p" offset each other which leads to the total predicted reserve is much closer the actual value (the bottom-right panel in Figure 9). The percentage reserve errors are larger than those obtained from six-month-constant hazard rates. It seems that the smoothing method enlarges the estimation errors what the ML estimates have.

We see that predictions from type "se" and "p" have smaller variation than type "sep". It may depend on the number of observations in each type. For Data1 claims, The number of events with type "sep" is the least among all types.

From Figure 9 and 10 we can conclude that the piecewise constant hazard rates with six-month and three-month specifications provide a remarkably good fit for the simulated claims. The predicted reserves from the smoothed hazard rates are either over- or underestimated. It can be stated that the smoothing method has not improved ML-estimators on the predictive ability.

Thus the results both in Section 4.1.1 and 4.1.3 show that the smoothing method fails not only to give a smoother estimate but also to improve the predictive ability.

4.2 Data2

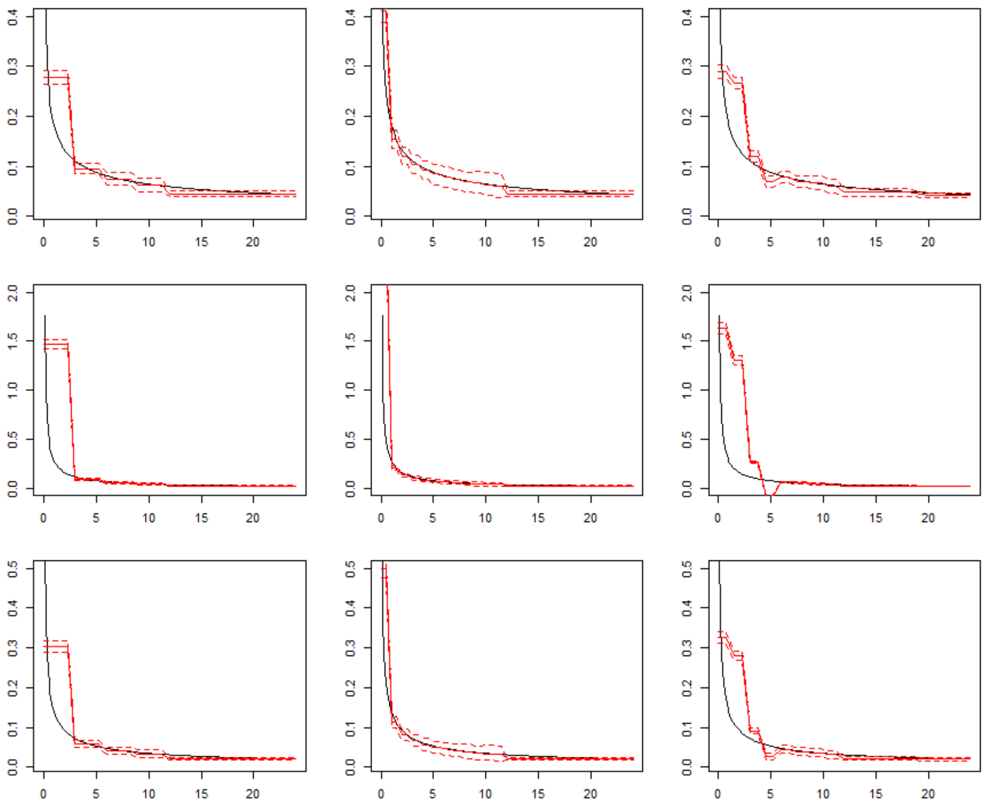
Data2 has another development pattern than Data1. Claims in this type of dataset develop quickly into the settlement and most of them are settled with a single payment. Settlement times are mostly short. 99% of claims have been settled within the reported year. So the active claims and future payments are relatively low.

Based on the above information, we should consider a new feasible number of cut points and length of time intervals which may differ from the settings of Data1 claims. As we mentioned, there are less than 1% of claims that are active after 12 months, so it seems reasonable that we aggregate the events which occur after 12 months and assume no further changes in the hazard rates after 12 months. To be able to capture the development pattern under the 12 months, we test both three-month-constant and one-month-constant specifications. Furthermore, we continue to study how the smoothning method performs in this type of dataset by smoothing the ML estimators with the three-month specification. Conclusions about the quality of each method will be drawn based on both performances in parameter fit and predictive abilities.

4.2.1 Estimating hazard rates

Estimated hazard rates and the 90% confidence intervals obtained from the empirical percentiles are shown in Figure 11. The first row presents the

Figure 11: Estimated hazard rates together with 5% and 95% empirical percentiles based on 500 simulations. *Columns*: hazard rates are constant in three month intervals; hazard rates are constant in one month interval; Smoothed results based on the three-months-constant hazard rates. Rows: type "se", type "sep" and type "p" events. All results are assumed with no further changes after 12 months in the hazard rates. The solid black lines show the true continuous hazard rate functions.



estimated results with type "se" events. It seems that the three-month-constant hazard rates (left panel) cannot describe the continuous change of hazard rates. The gap between the first and second interval is big which is problematic. The reason for this is that the continuous rate functions for Data2 decrease rapidly in the first three months, therefore to assume the hazard rate is constant in this interval is not realistic. Thence one-month-constant hazard rates which are shown on the middle panel in Figure 11 fit much better. The mean values of the estimated hazard rates capture better the accurate trend. But at the same time, a larger uncertainty is expected (wider 90% confidence intervals on the middle panel). From the right panel in Figure 11 we find that the gap between the first and second interval has been reduced but the fit is still not good. It, in this case, may be necessary to apply the smoothing method iteratively. But we stay in the first step and take this result to further analysis.

The estimated hazard rates in the process with type "sep" are presented on the second rows. The three-month specification and the smoothed results perform similarly on those in the process with type "se". The one-month-constant hazard rates give a better adaptation, but the fit in the first two months are still unrealistic. It is due to the hazard rate function in this type declines even faster in the first several months than those with type "se" and "p" which leads to an even poorer adaptation.

Hazard rates from type "p" events give a similar result as those from type "se".

Value of the smoothing function z for $\hat{\boldsymbol{h}}_{\boldsymbol{e}}^{\boldsymbol{3m}}$ and $\hat{\boldsymbol{h}}_{\boldsymbol{e}}^{\boldsymbol{smooth}}$ are shown in Appendix 3. It shows the smoothed estimators give almost the same results as the ML estimators.

4.2.2 Reserve

For Data2 claims, we only forecast cash flows and total reserves in three years.

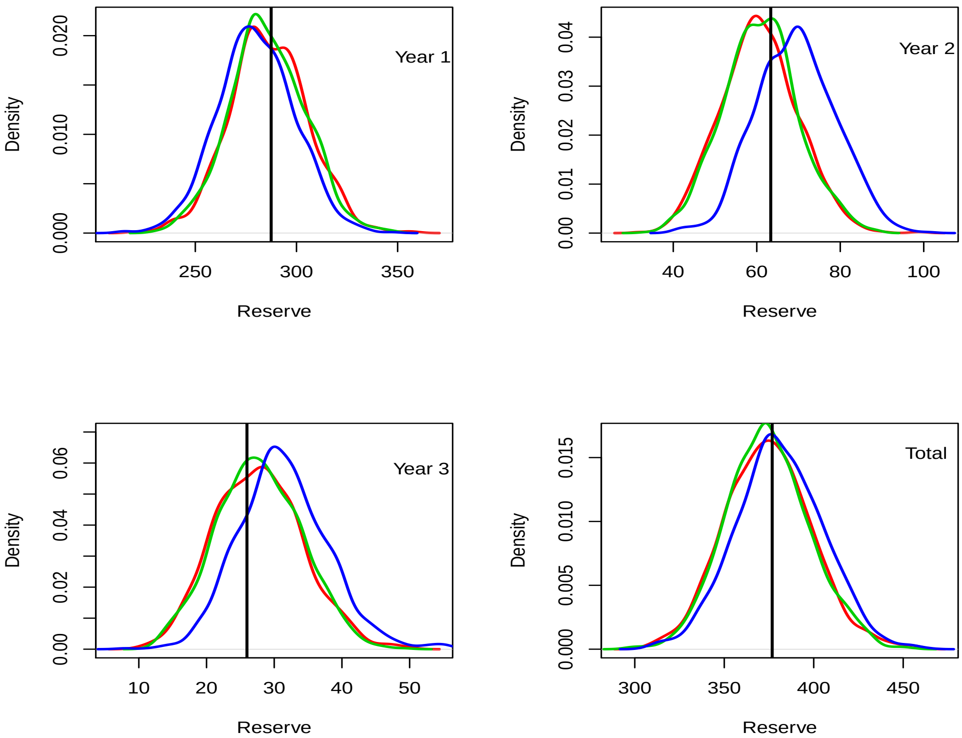
The left panel on the top row of Figure 12 shows the one-year-ahead cash flows. Kernel densities of reserves show that all three parameter settings give quite good predictions. The best estimates are slightly lower than the actual one. For cash flows in year 2 and year 3 (on the top right and bottom left in Figure 12), the one-month-constant hazard rates generate overestimation while the three-month-constant and smoothed hazard rates result in an excellent prediction. Total reserves from all three parameter settings are entirely close to the actual one.

Comparing the reserve amounts (x-axis value), we find that the best estimate of the cash flows in year 1 (around $280 \sim 290$) account for almost 80% of the total reserve ($370 \sim 380$).

Same as for Data 1, we illustrate the kernel densities of the percentage reserve errors in Figure 13. From Figure 13, we see that the kernel densities from three hazard rates vectors are different. It shows from the left panel that the smoothed hazard rates for type "se" events (the green line) generate considerable negative bias and total kernel density locates on the left of zero which means that this model dramatically understates the reserve. The three-month-constant hazard rates for type "sep" events (the red line) perform poorly as well. It has a smaller negative bias than the smoothed hazard rates but most parts of the kernel density are still on the left of zero. One-year-constant hazard rates show the best predictive ability. The best estimate is closer to zero and the line seems symmetric.

Moving to the middle panel in Figure 13, we see that the results are in contrast to the left panel. The three-month-constant specification, as

Figure 12: Predicted cash flows and total reserves in three years. First three results (from left to right, from top to bottom) are cash flows in three years. The last panel (bottom row on the right) shows the total reserves. The results are based on 1 000 simulations and 10 000 claims. Red line: three-month-constant hazard rates; Blue line: one-month-constant hazard rates; Green line: smoothed hazard rates. Solid black line: the mean of the actual future payments obtained from simulation.



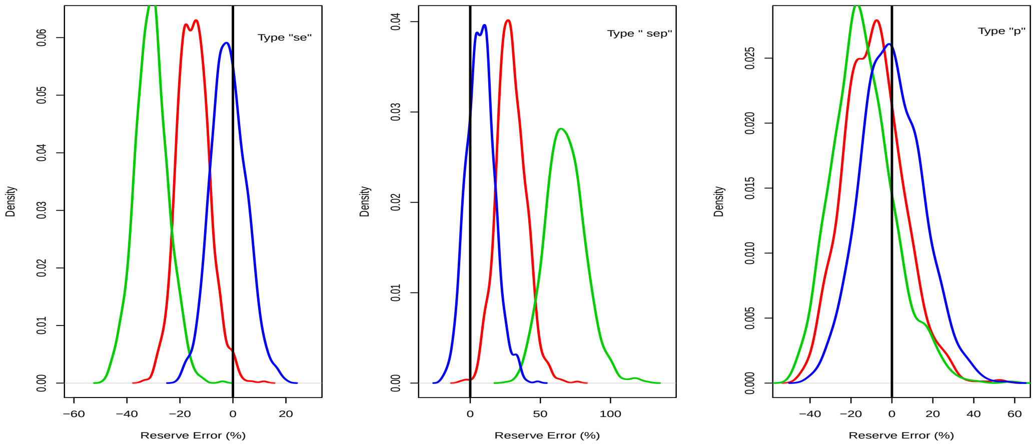
well as the smoothed hazard rates, overstate the number of events. The kernel density obtained from the one-month-constant hazard rates is more realistic than the others. Its best estimate is closer to zero.

For the prediction to the events with type "p" (see the right panel in Figure 13), the three-month-constant and the smoothed hazard rates simulate fewer numbers of events while one-month-constant generates a more realistic result.

4.2.3 Effect of sample size on reserve errors

Differences of percentage reserve errors by type of three parameter settings are significant. We may wonder if it is dependent on the number of active claims. A robustness check is necessary. Data2 in above analyses are based

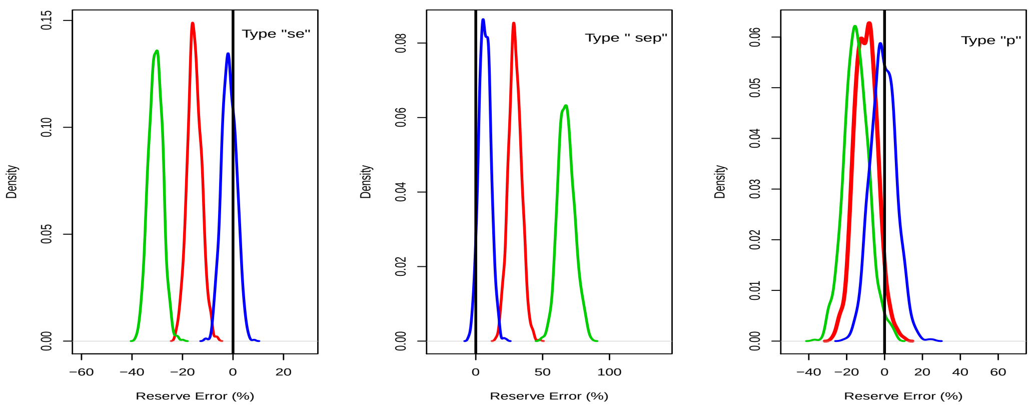
Figure 13: Kernel densities of the Percentage Reserve Errors by event type. Red line: three-month-constant hazard rates; Blue line: one-month-constant hazard rates; Green line: smoothed hazard rates. The black lines present value zero. Note that for type "se" event, it does not generate the reserve but the predictive errors of the numbers of events.



on 10 000 reported claims which are distributed into five accident years uniformly. As mentioned, there are only 1% claims are not settled after 12 months which means the prediction is based on approximately 100 active claims. To explore the impact of the number of claims on the results, we also experiment by using 50 000 claims in Data2 in which around 500 claims are needed to set reserves. We present the results by type from which we can see how the predictions are affected.

Results are shown in Figure 14. Comparing results in Figure 13 and 14, we see the best estimates are not significantly changed while the standard deviations decrease when the number of claims in the dataset increase. It tells us that the adaptation of the model does not be substantially affected by the sample size. Whether the model is suitable for data or not can be captured even for a small data size in the micro-level model study.

Figure 14: Kernel densities of the Percentage Reserve Errors by event type. Sample size is 50 000 and simulating size is 1 000. Red line: with three-month-constant hazard rates; Blue line: with one-month-constant hazard rates; Green line: with smoothed hazard rates. The black lines present value zero. Note that for type "se" event, it does not generate the reserve but the predictive errors of the numbers of events.



5 Discussion

The purpose of this thesis is to study how the piecewise constant hazard rates perform in a micro-level model in general insurance. We choose to simulate two types of datasets whose development patterns are entirely different. One contains complex processes with multiple payments and long settlement times. The development processes in the other dataset are much simpler, in which most of the claims are closed with a single payment and 99% of the claims are settled during the first accident year. By using piecewise constant hazard rates on both types of datasets, we would like to see if the model works differently on two distinct lines of business.

Data are generated using non-homogeneous Poisson processes. The Poisson processes have hazard rate functions from survival models as intensities. We specify that the hazard rates are from the Weibull distribution. In the processes, each claim is assumed to have been reported to the insurance company and, once acknowledged, a series of payments are made until the claim is settled or the process is closed. When modelling payments, three types are used: Type "se" event means the claim is settled but no payment occur, a type "sep" event indicate the claim is settled with a last payment and type "p" events corresponds to intermediate payments. All events are simulated by their corresponding hazard rate functions h_{se}, h_{sep} and h_p.

With these two types of datasets, we examine how well the piecewise constant specification can present the continuous rate functions. We analyze the two types of data separately. For data with long settlement time (called "Data1"), we test two lengths of time intervals: six months and three months. For Data2 (the data with quick settlement time), we apply three-month and one-month intervals which we believe are the most suitable. Note that for Data1, only few events occur after approximately 60 months. Due to this, it is hard to estimate hazard rates after 60 months. We therefore choose to assume no changing in hazard rates after 60 months for Data1. Similarly for Data2, we assume that hazard rates are unchanged after 12 months for Data2.

The Estimates of parameter vectors $\boldsymbol{h_e} = \{h_{1,e}, ..., h_{K,e}\}$, where $e \in \{se, sep, p\}$ and K is the number of intervals, can be obtained by using maximum likelihood theory. Based on ML estimators we also test to smooth one of the ML-estimators ($\boldsymbol{\hat{h}_e^{6m}}$ for Data1 and $\boldsymbol{\hat{h}_e^{3m}}$ for Data2) by using an iterative smoothing method which is introduced in [2], where $\boldsymbol{\hat{h}_e^{6m}}$ and $\boldsymbol{\hat{h}_e^{3m}}$ mean the ML-estimators with six-month-constant respective three-month-constant specification. The level of smoothness is evaluated by comparing the smoothing function value z which is presented in Table 3 for Data1 and Table 4 for Data2.

The results show that the smoothing method has not efficiently smoothed the ML-estimators for neither of the datasets.

Another evaluation of piecewise-constant specification is to see their performance in prediction. This is done for Data1 in Section 4.1.3 and in Section 4.2.2 for Data2. The results show that reserves from the three-month-constant hazard rates for Data1 are much closer to the actual one. It gives a stable prediction regardless of cash flows, total reserves or types. The six-month-constant hazard rates in Data1 perform well too. The predictive ability of the six-month-constant hazard rates does not notably differ from those of the three-month-constant hazard rates. The smoothed results create a significant bias in the prediction of cash flows and each type of events, but it results in a good prediction on the total reserve for Data1 claims.

For Data2 claims, kernel densities of reserves obtained from three-month-constant specification reflect actual cash flows and total reserve. The smoothed estimators give a good fit in the prediction of cash flows and total reserve too. But these two parameter settings perform critically wrong when they are used to predict events by types. The one-month-constant hazard rates create a slight bias on the prediction of cash flows but perform more stable to predict the events in each type.

In Section 4.1.2 we study the effect of sample size and simulation size on estimation errors. It shows that the expected value of estimates is not significantly changed on the changing of the sample size while the standard deviation decreases proportionally with the number of claims. By testing different simulation size, we conclude that if the sample size is sufficiently large the simulation size will not affect the estimation uncertainty to any greater extent.

Due to the predictive results by types differ obviously, we study even whether the prediction can be improved by increasing the number of active claims in Data2. The results tell us that the best estimate of simulation is not affected much by the sample size. It only decreases the standard deviation.

Due to time limitations, we only studied one basic rate function in the non-homogeneous Poisson process. The rate function, more generally, depends not only on time t but also on other internal or external covariates. To apply the piecewise constant specification to such Poisson process may be more complex and the conclusions may differ.

We may discuss the limitation of the smoothing method in [2] too. This method is based on several constraints. One of them is that it is used to smooth an estimated rate function vector, and the smoothed estimators will "inherit" the statistical errors from the original estimated rate function. It means that this method can not "correct" the issues but it looks like somehow enlarged them in the prediction, i.e. created a larger bias than the ML estimators.

6 References

[1] A. Charpentier, M. Pigeon (2016) Macro vs. Micro Methods in Non-Life Claims Reserving (an Econometric Perspective). *Risks*, 4(2).

[2] C.R. Larsen. (2007) An individual claims reserving model. *ASTIN Bulletin*, 37(1):113–132.

[3] E. Astersan, (1938) *Les réserves techniques des sociétés d'assurances contre les accidents automobiles*;Librairie générale de droit et de jurisprudence.

[4] G. Taylor, G. McGuire, and J. Sullivan. (2008) Individual claim loss reserving conditioned by case estimates. *Annals of Actuarial Science*, 3(1-2):215-256.

[5] H. Chen, B. Schmeiser. (2011) I-smoothness: Iteratively smoothing piecewise-constant poisson-process rate functions. *Proceeding of the 2011 Winter Simulation Conference.*

[6] H. Liu and R. Verrall. (2009) Predictive distributions for reserves which separate true IBNR and IBNER claims. *ASTIN Bulletin*, 39(1):35–60.

[7] J.E. Karlsson. (1976) The expected value of IBNR claims. *Scandinavian Actuarial Journal 1976*, 1976(2), 108-110.

[8] J.F. Lawless and M. Zhan (1989) Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *The Canadian Journal of Statistics*. Vol.26, No. 4: 549-565.

[9] K. Antonio, R. Plat. (2014) Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, 7, 649-669.

[10] M. Kunkler. (2004) Modelling zeros in stochastic reserving models. *Insurance: Mathematics and Economics*, 34(1):23–35.

[11] M.V. Wüthrich (2017) *Non-life insuance: mathematics & statistics* SSRN Electronic Journal.

[12] O. Bouaziz, G. Nuel. (2016) L0 regularization for the estimation of piecewise constant hazard rates in survival analysis.

[13] O.O. Aalen, Ø. Borgan, H.K. Gjessing (2008) *Survival and Event History Analysis: A Process Point of View.* Springer.

[14] P.D. England and R.J. Verrall. (2002) Stochastic claims reserving in general insurance. *British Actuarial Journal*, 8:443–544.

[15] R. Grøn, , T. A. Gerds, P. K. Andersen. (2016) Misspecified Poisson regression models for large-scale registry data: infrerence for "large n and small p". *Stat Med 35*. 1117-29.

[16] R.J. Cook, J.F. Lawless. (2006) *The statistical analysis of recurrent Events*. Springer.

[17] R. Norberg.(1993) Prediction of outstanding liabilities in non-life insurance. *ASTIN Bulletin*, 23(1):95–115.

[18] R. Norberg.(1999) Prediction of outstanding liabilities II. Model extensions variations and extensions. *ASTIN Bulletin*, 29(1):5–25.

[19] R. Schnieper. (1991) Separating true IBNR and IBNER claims. *ASTIN Bulletin*, 21(1):111–127.

[20] S. Haastrup and E. Arjas. (1996) Claims reserving in continuous time: a nonparametric Bayesian approach. ASTIN Bulletin, 26(2):139–164.

[21] J. Sigmundsdóttir, M. Lindholm. (2017) One-Year Non-Life Solvency II Risk Calculations for Point Process Micro Models: Methods and Applications. *Mathematical Statistics Stockhold University Reserch Report* 10.

[22] S.M. Ross. (2014) *Introduction to probability models*, 11th edition. Academic press.

[23] The editors of encyclopaedia Britannica. *www.britannica.com/science.*

[24] T. Mikosch (2008) *Non-Life Insurance Mathematics: An Introduction with the Poisson Process*. Springer.

[25] T. Verdonck, M. Van Wouwe, and J. Dhaene. (2009) A robustification of the chain–ladder method. *North American Actuarial Journal*, 13(2):280–298.

[26] X. Jin, E.W. Frees (2013) Comparing micro- and macro-level loss reserving models. Working paper. Available at *https://sites.google.com/site/xiaolijin2013/research/working-paper1* (Downloaded January 8, 2019)

7 Appendix

Appendix 1: Hazard rates estimates of type ”sep” and type ”p” events for Data1 claims

Figure 15: Fitted hazard rates of process with type ”sep”. *Bold red lines* describe the mean of estimates; *dashed lines* are the 5% and 95% empirical percentiles based on 500 simulations. *Black lines* show the true continuous hazard rate functions. Left panel: hazard rates are constant on six month intervals; Middle panel: hazard rates are constant on three month intervals; Right panel: smoothed results base on six-month-constant hazard rates.

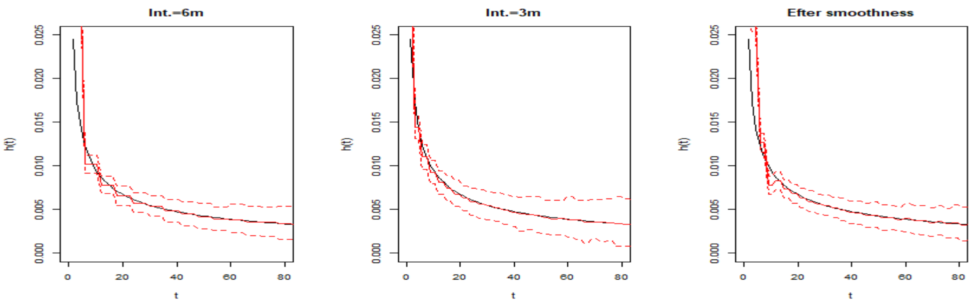
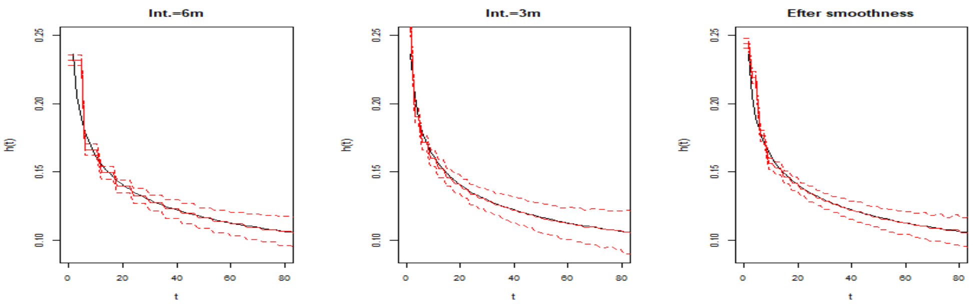
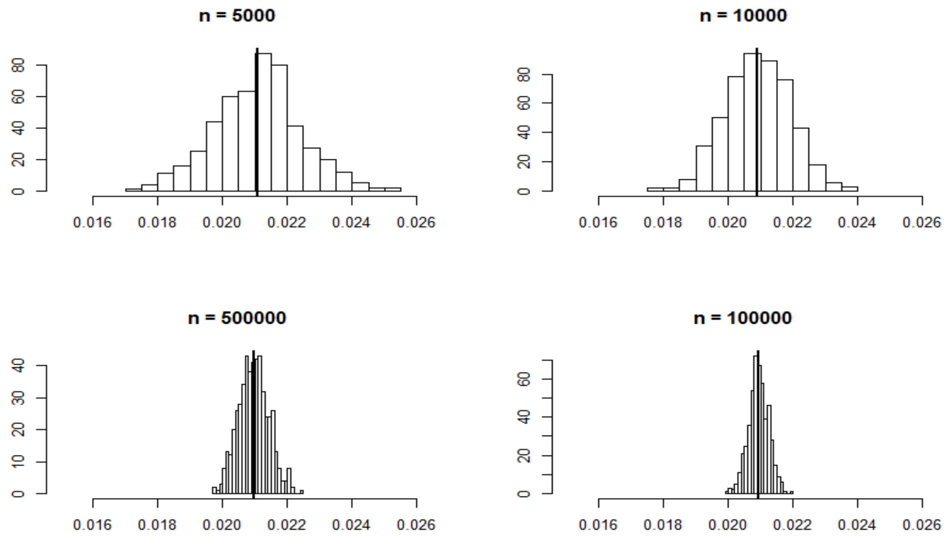


Figure 16: Fitted hazard rates of process with type ”p”. *Bold red line* describe the mean of estimates; *dashed lines* are the 5% and 95% empirical percentiles based on 500 simulations. *Black lines* show the true continuous hazard rate functions. Left panel: hazard rates are constant on six month intervals; Middle panel: hazard rates are constant on three month interval; Right panel: smoothed results base on six-month-constant hazard rates.



Appendix 2: Effect of sample size on estimation errors. Hazard rates in interval [54, 60) months

Figure 17: Histogram of piecewise constant hazard rates estimated from events with type "se" in the time interval [54,60) months. The sample size is 5 000, 10 000, 50 000 and 100 000. The solid black lines present the mean value of the simulations.



Appendix 3: Smoothing function value z for hazard rates estimators for Data2.

Table 4: Smoothing function value z for hazard rates estimators.

	$\hat{h}^{3m}$	$\hat{h}^{smooth}$
Type "se"	0.0019	0.003
Type "sep"	0.0003	0.0004
Type "p"	0.0035	0.0021