Lappeenranta-Lahti University of Technology LUT
School of Engineering Science
Computational Engineering and Technical Physics
Technomathematics

Mwasi Mboya

A RISK MODEL WITH RENEWAL SHOT-NOISE COX PROCESS

Master's Thesis

Examiners: D.Sc. (Tech.) Matylda Jablonska-Sabuka
Professor Evans Gouno

Supervisors: D.Sc. (Tech.) Matylda Jablonska-Sabuka
Professor Evans Gouno

ABSTRACT

Lappeenranta-Lahti University of Technology LUT
School of Engineering Science
Computational Engineering and Technical Physics
Technomathematics

Mwasi Mboya

A risk model with renewal shot-noise Cox process

Master's Thesis

2019

95 pages, 20 figures, 13 appendices.

Examiners: D.Sc. (Tech.) Matylda Jablonska-Sabuka
Professor Evans Gouno

Keywords: renewal shot-noise Cox process, maximum likelihood method, Bayesian estimation method, Markov chain Monte Carlo

Catastrophic events may lead to sudden changes in the claim arrival rate and as a result it is difficult to predict the likelihood for the occurrence of these events. These random fluctuation cannot be modelled as a homogeneous Poisson process. Hence, the introduction of doubly stochastic Poisson process or Cox process has been an important tool in modelling such catastrophic events and reduce the probability of ruin to many of the insurance companies. This research therefore proposes a risk model with stochastic intensity function as a first step towards the process of computing the probability of ruin. Several sample paths of the model with varying parameter values are shown as a first step in understanding how the model works. A major assumption of this model is that the time of jumps follows a renewal shot-noise Cox process and that all past times are considered. This idea is essential in estimating parameters using different statistical techniques such as maximum likelihood estimation method and Bayesian estimation method in the model such as the initial intensity value, the rate of claim settlement and some parameters from

other arbitrary distributions. An algorithm to simulate such time jumps is proposed in this work via the inversion method.

PREFACE

I would like to pay my deep sense of gratitude most of all to the almighty God without whom the success and final outcome of this thesis would not be possible. I thank him for his endless blessings, knowledge and strength to accomplish this thesis.

I am highly indebted to Lappeenranta-Lahti University of Technology (LUT) for providing me with a scholarship. I thank them for their generosity in funding and supporting my masters program here at the university.

I would also like to extend my sincere thanks to my supervisors D.Sc. (Tech.) Matylda Jablonska-Sabuka and Prof. Evans Gouno for the patience, guidance, encouragement and advice they have provided throughout. I have been extremely lucky to have my supervisors who have been supportive to my work and who responded to my questions and queries so promptly.

Generally their guidance and supervision was very helpful in bringing this thesis to the fruitful end.

Lappeenranta, May 18, 2019

Mwasi Mboya

CONTENTS

List of Algorithms

LIST OF ABBREVIATIONS

MLE	Maximum Likelihood Estimation
MCMC	Markov Chain Monte Carlo
RWM	Random Walk Metropolis

1 INTRODUCTION

1.1 Background and motivation

Catastrophic events have been causing great damages and sufferings to human beings since the beginning of this world. These extremely unfortunate events range from catastrophic weather events such as, tornadoes, hurricanes, earthquakes, and tsunamis to catastrophe for the economy such as, an economic depression. Consequently, scientist have been striving to develop mathematical and scientific models that can be used to predict the likelihood for the occurrence of these events.

One of the areas that has captured the attention of mathematicians is insurance business. By simple definition, insurance is a contract whereby a person called insurer agrees in consideration of money paid as premium by the person called insured to indemnify the latter against loss resulting to him on the happening of certain event(s) [1]. In order to meet these obligations the amount of premium charged to the insured must be estimated properly to reflect the risk that the insurer has to undertake. In other words, the amount of premium charged to the insured must be sufficient enough to reflect the possibility of the future claims. An insurance claim is a formal request to an insurance company asking for a payment after the occurrence of a certain contingent event (based on the terms of the insurance policy). This implies that the continual survival of an insurance company depends largely on the accuracy of mathematical models used to predict the claim severity, frequency and arrival rates.

The classical approaches to modelling insurance losses are based on a simple assumption that claims counts over a certain period of time follow a certain distribution, particularly Poisson, binomial and negative binomial distributions. However, there are number of challenges associated with these models, including uncertainties about its accuracy and the need for historical aggregated data in order to fit these distributions. These challenges triggered the need for developing a model that can capture the time evolution of claims and also use the available data in a more effective and efficient way, that is a Poisson process.

There is one main challenge associated with this model especially when it comes to the assumption that the claim arrival rate for the insurer remain constant over a certain period of time. This applies also to a homogeneous Poisson process where the arrival rate is assumed to vary deterministically over time. Thus its ability to capture random complexity of insurance business is highly questionable. This necessitated the need for developing a stochastic model which incorporate the stochastic nature of changes in insurer's claims rate (intensity), that is doubly stochastic Poisson processes, or Cox processes. For instance, during extreme rain an insurer is more likely to receive a significant number of claims from the policy holders who have insured their building against floods. Consider a case in Tanzania where heavy rain in April, 2018 which lasted for a period of only three days resulted into a death of 20 people, with around 250 homes reportedly destroyed [2].

1.2 Purpose of the study

The main purpose of this study is to model the claim arrival rate using non homogeneous Poisson process which allow its intensity function to be stochastic process in order to capture random complexity of insurance business. This necessitated the need for developing an algorithm to simulate a renewal shot-noise Cox process which incorporate the stochastic nature of changes in insurer's claims rate (intensity).

1.3 Objectives

The specific research objectives of the thesis are as follows:

- Compare different graphical representation of the sample path of the intensity process $\lambda(t)$ while changing the exponential rate of decay of the claim settlement between catastrophic events.
- Use of maximum likelihood estimation method to derive the parameter estimates of the intensity function and other arbitrary distribution.
- Use of Bayesian estimation method to derive the posterior distribution for the intensity function and other arbitrary distribution.
- Introduce the Monte Carlo Markov Chain (MCMC) method to enable sampling the parameter estimates from the posterior distribution of the intensity function and other arbitrary distribution derived using Bayesian estimation method.
- Develop an algorithm to simulate the process with renewal shot-noise Cox intensity.

1.4 Definition of terms

In this section, we provide terminologies that is used throughout the report.

Definition 1.4.1 (Stochastic process)

A stochastic process $\{X(t),\ t \geq 0\}$ *is simply defined as a collection of random variables* $X(t)$*. There exist a special type of stochastic process named as counting process that is used to model claim frequency over time as stated in [3].*

Definition 1.4.2 (Poisson process)

The counting process $\{N(t),\ t \geq 0\}$ *is said to be Poisson process if it satisfies the following conditions. According to [4], the counting*

- $N(0) = 0$*, means that the initial arrivals is assumed to be zero.*
- *Independent increments: The random variable* $N(t)$ *has independent increments that is the number of arrivals that occurs in disjoint time interval are statistically independent. For instance, given the disjoint time* $s < t < u$ *then* $N(t) - N(s)$ *is independent to* $N(u) - N(t)$*.*
- *Stationary increments: The random variable* $N(t)$ *has stationary increments that means the intervals of the same length of time have the same statistical behaviour (probability distribution) independent of the exact location of the interval on the time axis. Therefore, random variable of arrivals with the same time interval* τ*, have the same probability distribution. For instance,* $N(\tau + s) - N(s)$ *has the same distribution function as* $N(\tau + u) - N(u)$ *with the same time interval* τ*.*

There are two main types of Poisson process which are homogeneous and inhomogeneous Poisson process.

Homogeneous Poisson process is one of the most popular Poisson process denoted as $\{\tilde{N}(t),\ t \geq 0\}$ where the intensity function is assumed to be constant over time $\lambda(t) = \lambda$. In fitting homogeneous Poisson process, the only parameter which needs to be estimated is λ and maximum likelihood estimation can be used to estimate the parameter λ. The assumption that intensity of claims remain constant over time needs to be improved in modelling claims from catastrophic events to depend on time [3].

On the other hand, an alternative model to homogeneous Poisson process is inhomogeneous Poisson process in which the intensity function $\lambda(t)$ is a non-constant deterministic function that depends on time. In insurance, the inhomogeneity in claim intensity is very useful especially in modelling the occurrence of seasonal catastrophic events [3].

Definition 1.4.3 (Doubly stochastic Poisson process or Cox process)

This is the stochastic process which is the generalization of Poisson process where the time-dependent intensity $\lambda(t)$ is itself a stochastic process [5]. In insurance modelling, Poisson process is used to model the claim arrivals. However, there have been significant number of questions on appropriateness of Poisson process in insurance modelling most especially in rainfall modelling. For catastrophic events, the assumption of claim to follow Poisson process is inadequate as it has deterministic intensity. Therefore, we will need to employ Cox process/ doubly stochastic Poisson process in which the claim intensity $\lambda(t)$ is itself a stochastic process. However, doubly stochastic Poisson process can be viewed as two step randomisation procedure [6].

Definition 1.4.4 (Shot-noise process)

Renewal process is simply an integer-valued or counting process $\{N(t),\ t \geq 0\}$ for which the inter-arrival times are independently and identically distributed random variables with arbitrary distribution [7]. This decrease continue until another catastrophic event occurs and this result into a positive jump in the shot-noise process. Hence, shot-noise process is used as a parameter of doubly stochastic Poisson process to measure the claim frequency due to catastrophic events [5].

Definition 1.4.5 (Renewal process)

Renewal process is simply an integer-valued or counting process $\{N(t),\ t \geq 0\}$ for which the inter-arrival times are independently and identically distributed random variables with arbitrary distribution [7].

Begin with the terminology point process as it is essential for the better understanding of renewal process. The process $\{N(t),\ t \geq 0\}$ denoted by $N(t)$ is said to be a point process on R_+, where the number of times T_n in the interval $(0,\ t]$ is defined as in [8] by:

$$N(t) \ = \ \sum_{n=1}^{\infty} 1(T_n \leq t),\ t \geq 0.$$

Assume this counting process $N(t)$ is finite valued for every t, which is similar to $T_n \to \infty$ $a.s$ as $n \to \infty$, where T_n are the arrival times.

However, the point process $N(t)$ is said to be simple if its arrival times are distinct: $0 < T_1 < T_2 < \ldots a.s.$ Therefore, a simple point process $N(t)$ is a renewal process if its inter-arrival times are independent and identically distributed with the same arbitrary distribution function. Examples of renewal process include random times at which: customers arrive at the queue for bank services, and the way insurance claims are settled.

However, Poisson process is simply counting process for which the inter-arrival times are independent and identically distributed with exponential distribution. Therefore, the renewal process is more general counting process than Poisson process.

1.5 Structure of the thesis

The essay is structured as follows. Chapter 2 introduces the Cox process with renewal shot-noise intensity. Further, this chapter provides illustration for different samples paths of the renewal shot-noise intensity process $\lambda(t)$ plotted using different parameters values with different distribution. This process is then used in estimation of parameters using maximum likelihood estimation method in Chapter 3. Chapter 4 introduces Bayesian estimation method and provides derivation of posterior distribution of the intensity function and other arbitrary distributions. Chapter 5 introduces MCMC method to Bayesian estimation. Furthermore, this chapter explains the most popular techniques of MCMC method and their application to obtain Bayesian estimates. The algorithm for simulating a process with a renewal shot-noise Cox intensity is given in Chapter 6. Chapter 7 presents the results and discussions obtained from the *MATLAB* codes. Finally, Chapter 8 presents a summary of conclusion and recommendation of this study.

2 RENEWAL SHOT-NOISE COX PROCESS

This chapter is devoted to renewal shot-noise Cox process. We shall present some background and basic definitions that will be essential in the course of this project.

2.1 Background of renewal shot-noise Cox process

Cox process was first introduced by [9] as a natural generalization of Poisson process by considering intensity process as the realization of random measure [10]. Cox process provides flexibility of intensity not only depending on time but allows it to be a stochastic process. It can be viewed as a two step randomisation procedure, which is able to model the stochastic nature of catastrophic loss occurrence in the real world.

A shot-noise Cox process is simply a doubly stochastic process which, conditionally on the realization of random measure is a Poisson process with intensity function $\lambda(t)$ in which initially the claim intensity was assumed to be deterministic. Therefore, shot-noise process can be used as the intensity of a Cox process to measure the number of catastrophic losses [11].

Cox models are used widely in many aspects: insurance, finance, queuing theory, statistic. For instance, the study by [6] provides the good insight about the application of Cox process with Poisson shot-noise intensity to pricing stop-loss catastrophe reinsurance contract and catastrophe insurance derivatives.

2.2 Motivation of renewal shot-noise Cox process

An insurance company insures the policyholder in case of claim occurring due to catastrophic event in exchange of some regular premium. Numerous works have clearly provided some information useful in determination of optimal premium, enabling the insurance company to charge premium that should be high enough such that there is sufficiently small ruin probability i.e. probability of an insurance company to have insufficient initial surplus to pay-off the claims arose. The claim sizes are assumed to be independent and identically distributed with the arrival times said to be jump times from Poisson process. Recently, the extension of this model considered renewal times where the inter-arrival times were no longer exponential distributed [3].

Traditionally, the claim arrival rate was modelled using homogeneous Poisson process to derive ruin probabilities. Recently to derive the estimates of ruin probabilities, the risk model was assumed to follow doubly stochastic Poisson process. This assumption made some improvement in a lot of insurance companies because it has reduced ruin probability as it provides flexibility of intensity not only depending on time but allows it to be a stochastic process.

In this chapter, there is extension of the study of arrival times with random arrival (intensity) rates. In particular, the arrival rates are said to have shot-noise features. However, this could be used to model the claim arrival in dynamic way: many claims will be reported right after the catastrophic events at which the arrival rate will be high at the beginning. Hence, the shot-noise arrival rate models such effect. For this case consider the claim caused by the catastrophic event where at the beginning of the event there will be majority of claims but as time passes the number of claims decreases till the occurrence of the next catastrophic event where the number of claims will increase again. In life insurance context, occurrence of any natural catastrophe will result to a similar pattern [3].

Claims arising from the catastrophic events differs from different time interval duration and they also depend on the time elapsed since the previous claim. Therefore, the improved model beyond Poisson process is used to model such claims arising from the catastrophic events [12].

Now, a renewal shot-noise Cox process is introduced below.

Definition 2.2.1 (Renewal shot-noise Cox process) *Renewal shot-noise Cox process is a point process $N_t \equiv \{T_j\}_{j=1,\,2,\ldots}$ on R_+ with the renewal shot-noise intensity λ_t, i.e. a non-negative shot-noise process driven by an ordinary renewal process specified as in [12] by*

$$\lambda_t \;=\; \lambda_0 e^{-\delta t} + \sum_{i=1}^{M_t} Y_i e^{-\delta(t-T_i^*)},\ t \geq 0, \tag{2.2.1}$$

where

- λ_0 *is the initial intensity;*
- $\delta > 0$ *is the constant rate of exponential decay;*
- $\{M_t\}_{t\geq 0}$ *is a renewal process with arrival times* $\{T_i^*\}_{i=1,\,2,\ldots}$ *i.e.* $M_t \equiv \{T_i^*\}_{i=1,\,2,\ldots}$;

- *t is the time period such that all arrival times are included in the process.*
- *$\{Y_i\}_{i=1,\ 2,\ldots}$ is a sequence of independent and identically distributed random variables (sizes of renewal jumps or shots) with distribution function $H(y),\ y > 0$, which is assumed to be absolutely continuous with density function $h(y)$ and independent of M_t.*

Assume M_t is Poisson process instead then λ_t is said to be classical shot-noise process [13]. In this project, it is assumed M_t to be renewal process and the process is then a special case of generalised shot-noise Cox processes [14].

Definition 2.2.2 (A risk process driven by a renewal shot-noise Cox process) *Consider an insurance company with surplus process X_t in continuous time on a probability space $(\Omega,\ F,\ P)$. We assume as in [5] that*

$$X_t \quad = \quad X_0 + ct - \sum_{j=1}^{N_t} Z_j,\ t \geq 0,$$

where

- *$X_0 \geq 0$ is the initial reserve at time $t = 0$;*
- *$c > 0$ is the constant rate of premium income;*
- *N_t is a renewal shot-noise Cox process with associated claim-arrival times $\{T_j\}_{j=1,\ 2,\ldots}$;*
- *$\{Z_j\}_{j=1,\ 2,\ldots}$ are claim sizes which are assumed to be $i.i.d.$ with distribution function $Z(z),\ z > 0$. $N(t)$ is also assumed to be independent.*

The surplus process works as the result of some realisation of ruin probability of the insurer $i.e.$ the probability that at some point in time, the aggregate claims exceed the received premium income and initial reserve. Therefore, at this particular point in time the insurance company is said to be ruin.

2.3 Representation of the intensity of renewal shot-noise Cox process

This section presents the simulation procedure of shot-noise intensity function from Equation (2.2.1) to obtain different graphical representation of the function. Below are the steps in the algorithm:

1. Generate random numbers of inter-arrival times $R_1,\ R_2, \dots,\ R_{M_t}$ from the exponential, gamma or Weibull distribution with their rate parameter, shape parameter and scale parameter depending on the distribution.

2. Calculate the arrival times from these generated random inter-arrival times.

$$T_i^* \;=\; T_{i-1}^* + R_i$$

3. Generate random sizes of claims, $Y_1,\ Y_2, \dots,\ Y_{M_t}$ from the log-normal distribution with its mean and variance, where $Y_i \sim \log \mathcal{N}(\mu,\ \sigma^2)$.

4. Thereafter, sample paths of different catastrophic time interval are described as follows; First, compute the sample path of renewal shot-noise process starting with the initial intensity λ_0 that decreases exponentially for all $t < T_1^*$ until the first arrival time of catastrophic event occurs T_1^*. Generally, after the occurrence of the first catastrophic event, compute the sample path of renewal shot-noise process between each two consecutive arrival times of a catastrophic event for the entire time period t and the computation involves cumulative process as shown in the representation (2.3.1)

   Mathematical definition of the renewal shot-noise intensity process in the simulation process is represented as:

$$\lambda_t = \begin{cases} \lambda_0 e^{-\delta t} & t \;<\; T_1^* \\ \lambda_0 e^{-\delta t} + Y_1 e^{-\delta(t-T_1^*)} & T_1^* \;\leq\; t \;<\; T_2^* \\ \lambda_0 e^{-\delta t} + Y_1 e^{-\delta(t-T_1^*)} + Y_2 e^{-\delta(t-T_2^*)} & T_2^* \;\leq\; t \;<\; T_3^* \\ \vdots & \vdots \quad \vdots \quad \vdots \\ \lambda_0 e^{-\delta t} + \sum\limits_{\{i: T_i^* < t < T_{i+1}^*\}} Y_i e^{-\delta(t-T_i^*)} & T_i^* \;\leq\; t \;<\; T_{i+1}^* \end{cases} \tag{2.3.1}$$

5. Therefore, using the assumptions above it is possible to plot different sample paths of renewal shot-noise intensity process with different values of delta parameter.

2.4 Some graphical representations through simulated parameter values

In this section, there will be provision of different figures when inter-arrival times are assumed to follow some arbitrary distribution and we shall give explanations on the observation contributed by different parameter values for each distribution.

Figure 1 shows four different representations of the sample paths of renewal shot-noise intensity process, λ_t when the inter-arrival times follow exponential distribution and jump sizes follow log-normal distribution. The assumptions used for these sample paths are as follows:

i) Assuming that claim sizes follow log-normal distribution with mean parameter, $\mu = \ln(30)$ and variance parameter, $\sigma^2 = \ln(3.5)$.

ii) Assuming that inter-arrival times follow exponential distribution with rate parameter equal to 1, $i.e.$ rate is simply a reciprocal of scale parameter. Then, determine the arrival times from these random selected inter-arrival times.

iii) Assuming that delta parameter, δ $i.e.$ exponential rate of decay varies within the vector $0.1,\ 0.5,\ 2$ and 10.

iv) Setting up the initial intensity, $\lambda_0 = 4$ and time period for the whole process to occur is given by the sequence from 0 to 20 separated by 0.25 time interval. Typically t will be quite large such that all claims are included in the process.

For the comparative study, the values for the delta parameter were $\delta = 0.1,\ \delta = 0.5,\ \delta = 2$ and $\delta = 10$. Using the above assumptions, sample paths for the renewal shot-noise intensity process are generated, $\lambda(t)$ for the times, $t = 1,\ 2, \ldots,\ 20$. The simulated sample paths of the shot noise intensity, $\lambda(t)$ for the chosen values of δ are shown in Figure 1. It can be seen that for the low δ case ($\delta = 0.1$), initially the shot-noise intensity will decay for longer time before another high shot in the intensity, thereafter the shot-noise intensity will decay for shorter time before another shot. Generally, there is an observation that as δ increases the shot noise intensity takes longer time to decay. Hence the peaks in the shot-noise intensity, $\lambda(t)$ are much more apparent (noticeable) in the case when δ is large than when δ is small.

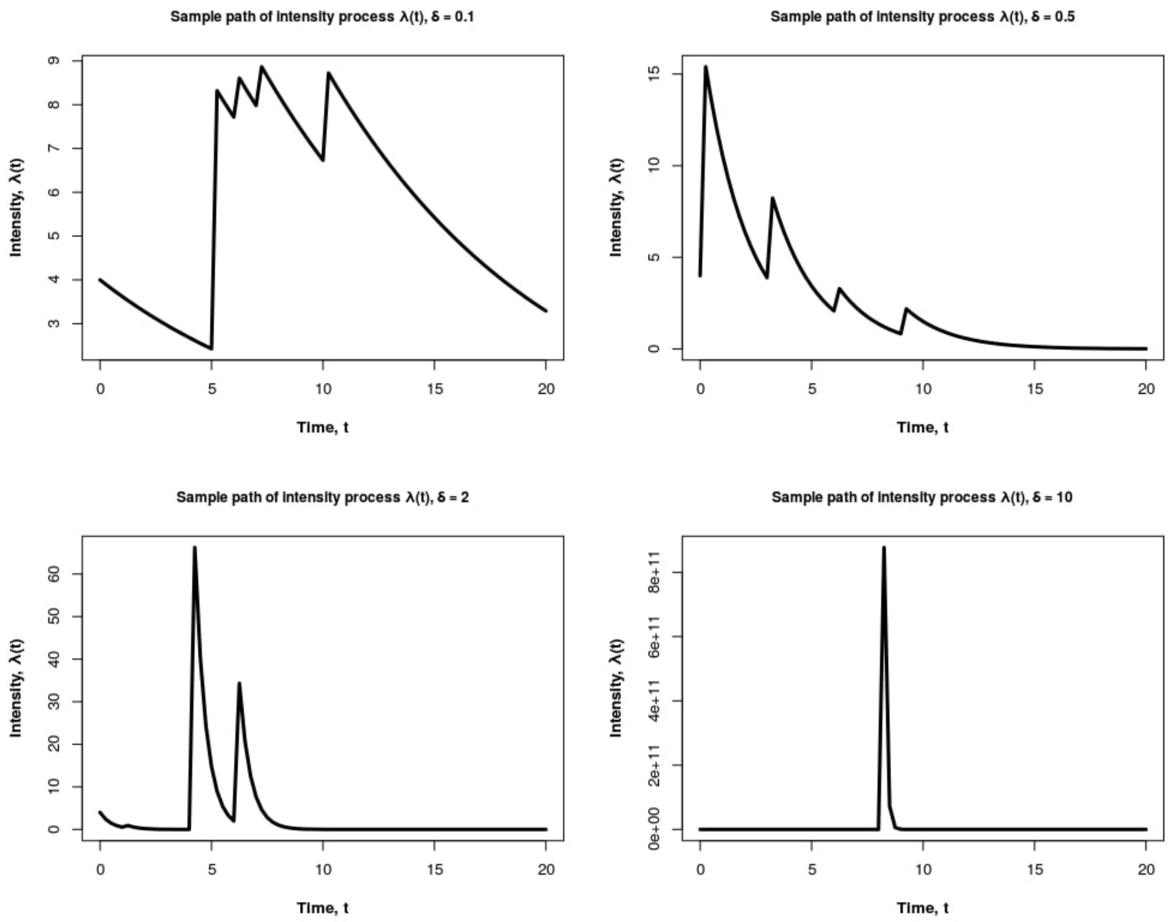

Figure 1. Sample paths of renewal shot-noise intensity process, λ_t when inter-arrival times follow exponential distribution and jump sizes follow log-normal distribution varying the delta parameter.

Figure 2 shows four different representations of the sample paths of renewal shot-noise intensity process, λ_t when the inter-arrival times follow gamma distribution and jump sizes follow log-normal distribution. The assumptions from these sample paths are as follows:

i) Assuming that claim sizes follow log-normal distribution with mean parameter, $\mu = \ln(30)$ and variance parameter, $\sigma^2 = \ln(3.5)$.

ii) Assuming inter-arrival times follow gamma distribution with fixed shape parameter equal to 2 and rate parameter equal to 1 $i.e.$ rate is simply a reciprocal of scale parameter. Then, determine the arrival times from these random selected inter-arrival times.

iii) Assuming that delta parameter, δ $i.e.$ exponential rate of decay varies within the vector $0.1,\ 0.5,\ 2$ and 10.

iv) Setting up the initial intensity, $\lambda_0 = 4$ and time period for the whole process to occur is given by the sequence from 0 to 20 separated by 0.25 time interval. Typically t will be quite large such that all claims are included in the process.

For the comparative study, the values for the delta parameter were $\delta = 0.1,\ \delta = 0.5,\ \delta = 2$ and $\delta = 10$. Using the above assumptions, sample paths for the renewal shot-noise intensity process are generated, $\lambda(t)$ for the times, $t = 1,\ 2, \ldots,\ 20$. The simulated sample paths of the renewal shot-noise intensity process, $\lambda(t)$ for the chosen values of δ are shown in Figure 2. It can be seen that for the low δ case ($\delta = 0.1$), the shot-noise intensity will decay for shorter time before another shot in the intensity while as δ increases the shot-noise intensity takes longer time to decay. Hence the peaks in the shot-noise intensity, $\lambda(t)$ are much more apparent (noticeable) in the case when δ is large than when δ is small.

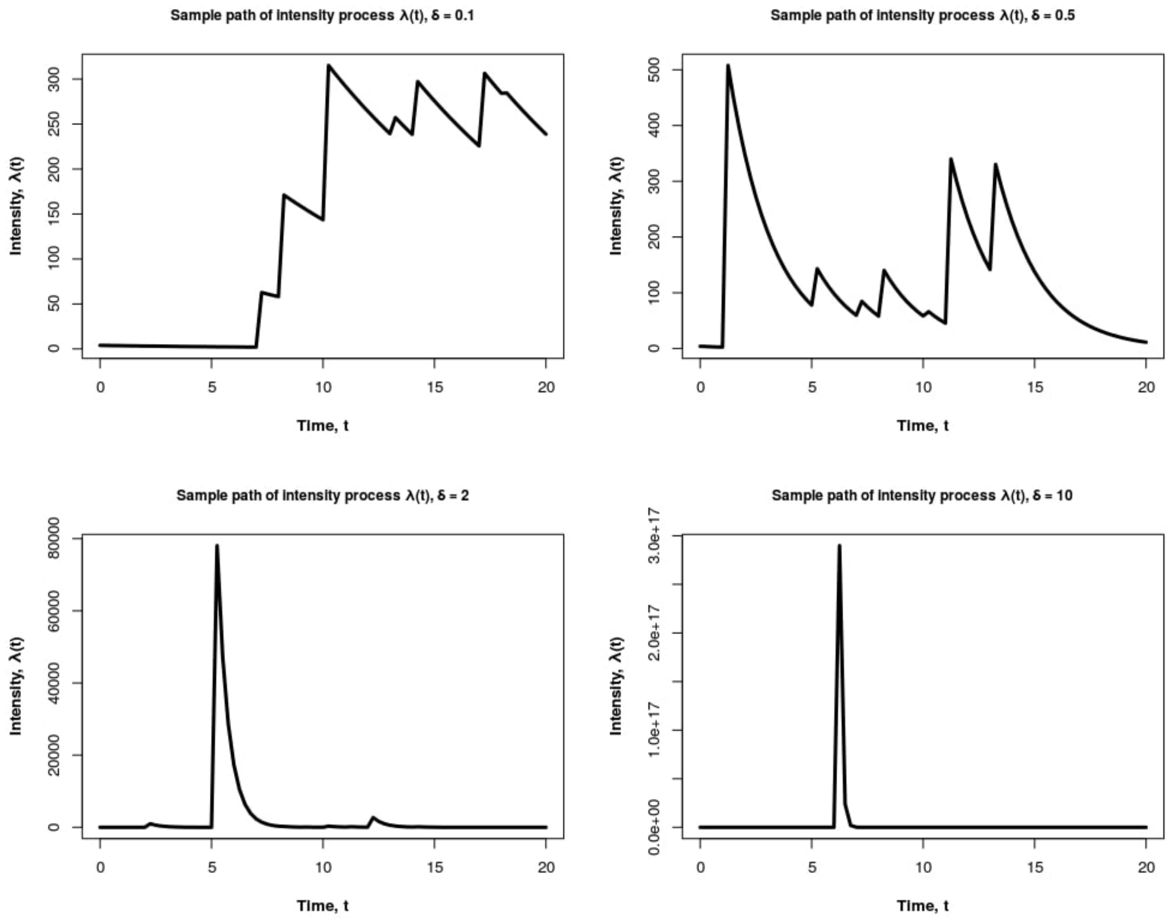

Figure 2. Sample paths of renewal shot-noise intensity process, λ_t when inter-arrival times follow gamma distribution and jump sizes follow log-normal distribution varying the delta parameter.

Figure 3 shows four different representations of the sample paths of renewal shot-noise intensity process, λ_t when the inter-arrival times follow Weibull distribution and jump sizes follow log-normal distribution. The assumptions from these sample paths are as follows:

i) Assuming that claim sizes follow log-normal distribution with mean parameter $\mu = \ln(20)$ and variance parameter $\sigma^2 = \ln(3.5)$.

ii) Assuming that inter-arrival times follow Weibull distribution with fixed shape parameter equal to 1.5 and scale parameter equal to 1 $i.e.$ rate is simply a reciprocal of scale parameter. Then, determine the arrival times from these random selected inter-arrival times.

iii) Assuming that delta parameter, δ $i.e.$ exponential rate of decay varies within the vector $0.1,\ 0.5,\ 2$ and 10.

iv) Setting up the initial intensity, $\lambda_0 = 4$ and time period for the whole process to occur is given by the sequence from 0 to 20 separated by 0.25 time interval. Typically t will be quite large such that all claims are included in the process.

For the comparative study, the values for the delta parameter were $\delta = 0.1,\ \delta = 0.5,\ \delta = 2$ and $\delta = 10$. Using the above assumptions, sample paths for the renewal shot-noise intensity process are generated, $\lambda(t)$ for the times, $t = 1,\ 2, \ldots,\ 20$. The simulated sample paths of the renewal shot-noise intensity process, $\lambda(t)$ for the chosen values of δ are shown in Figure 3. It can be seen that for the low δ case ($\delta = 0.1$), initially the shot-noise intensity remain constant for a while thereafter the shot-noise intensity will start to decay for shorter time before another shot in the intensity while as δ increases the shot-noise intensity takes longer time to decay. Hence the peaks in the shot-noise intensity, $\lambda(t)$ are much more apparent (noticeable) in the case when δ is large than when δ is small.

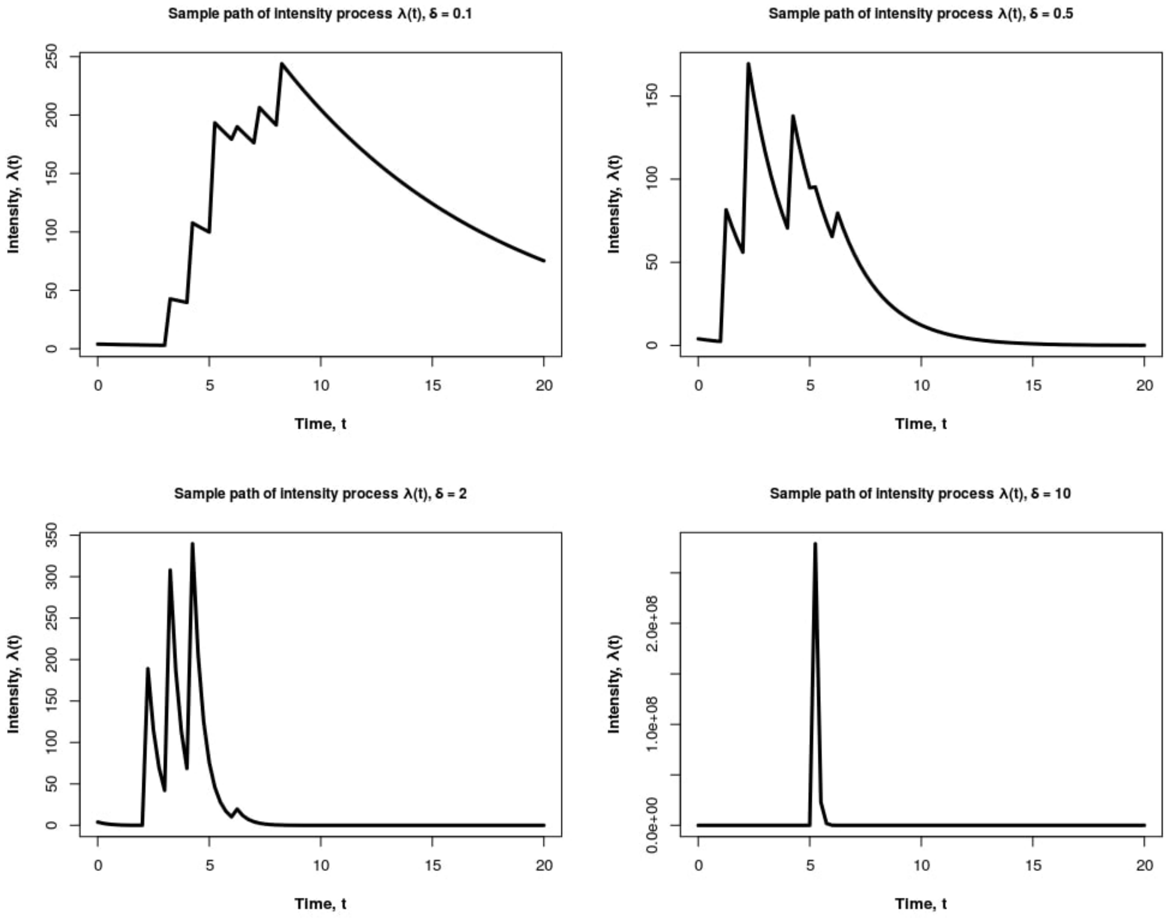

Figure 3. Sample paths of renewal shot-noise intensity process, λ_t when inter-arrival times follow Weibull distribution and jump sizes follow log-normal distribution varying the delta parameter.

Generally, Figures 1, 2 and 3 shows drastic fall in shot-noise process as the exponential rate of decay increases from $\delta = 0.1$ to $\delta = 10$. Observation shows that as time passes, shot-noise process decreases as more and more claims are settled. This decrease continues until another catastrophe occurs which results into a positive jump in the shot-noise process.

This chapter has provided the basis for building the renewal shot-noise Cox process. Specifically, it has covered the intensity function studied by simulating the renewal process, that is the i^{th} arrival times T_i^* with their corresponding i^{th} claim sizes Y_i.

3 MAXIMUM LIKELIHOOD ESTIMATION (MLE)

This chapter presents mathematical derivation of estimators for different types of distributions using the maximum likelihood estimation method. For the purpose of this study, estimators for four distributions would be presented, all of which belong to a family of positive skewed distributions, that is log-normal, exponential, gamma and Weibull that are simulated using renewal process $M(t)$. These estimators will be used to simulate the renewal shot-noise Cox process.

3.1 Introduction to MLE

Maximum likelihood estimation is the most popular estimation technique for many distributions, it mostly uses differentiation to find the parameter values that would maximize the probability of getting a particular sample. Some engaging features of maximum likelihood estimators include that they are asymptotically unbiased, and the bias approaches zero as the sample size n increases as stated in [15].

It is one of the best and efficient methods of estimating parameters given a certain distribution using some observed data. In particular, maximum likelihood estimators have excellent and usually easily determined asymptotic properties which is a benefit in dealing with large-sample data from catastrophic events.

There are several stages in the estimation of parameters;

- One of the most important stages is writing down the likelihood function,

$$L(\Theta) \;=\; \prod_{i=1}^{n} f(x_i;\, \Theta),$$

  for the random sample of $x_1,\ x_2, \ldots,\ x_n$ from the policyholder with density or probability function $f(x;\ \Theta)$. Note that Θ is the parameter whose value is to be estimated. It is said that likelihood function is the function of unknown parameter Θ. For each value of Θ, we get the corresponding values of the likelihood function. However, it is necessary to find the value of Θ that maximizes the likelihood function.

- The next stage involves taking logs of likelihood functions to simplify the procedures in determination of maximum likelihood estimator, $\widehat{\Theta}$.

- Lastly, it is in this stage where differentiation is involved. This means finding derivatives of log likelihood or likelihood with respect to the parameter Θ and derivative is set to zero as a result it gives the maximum likelihood estimator for the true parameter values $\widehat{\Theta}$.

3.2 Derivation of the estimators in the intensity function

The Hawkes process is a particular case of the doubly stochastic Poisson process: It is a linear self-exciting process. Its conditional intensity process has a particular form which depends upon the previous values of the point process $N(t)$ itself.

Definition 3.2.1 (Hawkes' process) *A point process $N(.)$ is called self-exciting if the intensity $\lambda(.)$ depends not only on the time t_i but also on the entire past times of the points process $t_1, \ldots,\ t_{i-1}$ as stated in [16].*

The renewal shot-noise Cox process is an extension of the Hawkes' process. Using $\{t_1,\ t_2, \ldots,\ t_{N(T)}\}$ to denote the observed sequence of past arrival times of the point process $N(t)$ in the interval $(0,\ T]$ with the positive influence on the current value of the intensity process. The intensity function is not depending on the point process itself but on another renewal process leading to an expression that is given by:

$$\lambda(s) \;=\; \lambda_0\, e^{-\delta s} + \sum_{\{j:T_j^* < s\}} Y_j\, e^{-\delta(s-T_j^*)}.$$

Proposition 3.2.2 *Let $N(.)$ be a regular point process on $[0, T]$ for some finite positive T and let $t_1,\ t_2, \ldots,\ t_{N(T)}$ denote a realisation of $N(.)$ over $[0,\ T]$. Then, the general expression of the likelihood L of $N(.)$ for any process from [17] is expressible in the form*

$$L \;=\; \left(\prod_{i=1}^{N(T)} \lambda(t_i)\right) \exp\left(-\int_0^T \lambda(s)ds\right).$$

From the given Theorem 4.2.2, the log-likelihood function is given by:

$$\begin{aligned}
\log(L(\lambda_0,\delta)) &= \log\left[\left(\prod_{i=1}^{N(T)}\lambda(t_i)\right)\exp\left(-\int_0^T\lambda(s)ds\right)\right] \\
&= -\int_0^T\lambda(s)ds+\int_0^T\log\lambda(t_i)dN(t_i) \\
&= -\sum_{i=1}^{N(T)}\int_{t_{i-1}}^{t_i}\lambda(s)ds+\sum_{i=1}^{N(T)}\log\lambda(t_i) \qquad (3.2.1)
\end{aligned}$$

Substituting the intensity function of renewal shot-noise Cox process into Equation (4.2.1) and obtain the expansion

$$\begin{aligned}
&= -\sum_{i=1}^{N(T)}\int_{t_{i-1}}^{t_i}\Big(\lambda_0\, e^{-\delta s}+\sum_{\{j:T_j^*<t_i\}}Y_j\, e^{-\delta(s-T_j^*)}\Big)ds \\
&+ \sum_{i=1}^{N(T)}\log\Big(\lambda_0\, e^{-\delta t_i}+\sum_{\{j:T_j^*<t_i\}}Y_j\, e^{-\delta(t_i-T_j^*)}\Big) \\
&= -\sum_{i=1}^{N(T)}\int_{t_{i-1}}^{t_i}\lambda_0 e^{-\delta s}ds-\sum_{i=1}^{N(T)}\int_{t_{i-1}}^{t_i}\sum_{\{j:T_j^*<t_i\}}Y_j\, e^{-\delta(s-T_j^*)}ds \\
&+ \sum_{i=1}^{N(T)}\log\Big(\lambda_0\, e^{-\delta t_i}+\sum_{\{j:T_j^*<t_i\}}Y_j\, e^{-\delta(t_i-T_j^*)}\Big). \qquad (3.2.2)
\end{aligned}$$

Evaluate the first integral:

$$\begin{aligned}
-\sum_{i=1}^{N(T)}\int_{t_{i-1}}^{t_i}\lambda_0 e^{-\delta s}ds &= -\lambda_0\sum_{i=1}^{N(T)}\int_{t_{i-1}}^{t_i}e^{-\delta s}ds \\
&= -\lambda_0\sum_{i=1}^{N(T)}-\frac{1}{\delta}e^{-\delta s}\Big|_{t_{i-1}}^{t_i} \\
&= \frac{\lambda_0}{\delta}\sum_{i=1}^{N(T)}(e^{-\delta t_i}-e^{-\delta t_{i-1}}).
\end{aligned}$$

Since, $\exp(-dt)$ is a decreasing function, and so the sum is a telescoping sum, giving the result.

$$-\sum_{i=1}^{N(T)} \int_{t_{i-1}}^{t_i} \lambda_0 e^{-\delta s} ds = \frac{\lambda_0}{\delta}(e^{-\delta t_{N(T)}} - e^{-\delta t_0}).$$

Assume that the initial point in time $t_0 = 0$ then obtain the first integral:

$$-\sum_{i=1}^{N(T)} \int_{t_{i-1}}^{t_i} \lambda_0 e^{-\delta s} ds = \frac{\lambda_0}{\delta}(e^{-\delta t_{N(T)}} - 1). \tag{3.2.3}$$

Evaluate the the second integral:

$$\begin{aligned} -\sum_{i=1}^{N(T)} \sum_{\{j:T_j^*<t_i\}} \int_{t_{i-1}}^{t_i} Y_j \, e^{-\delta(s-T_j^*)} ds &= -\sum_{i=1}^{N(T)} \sum_{\{j:T_j^*<t_i\}} Y_j \, e^{\delta T_j^*} \int_{t_{i-1}}^{t_i} e^{-\delta s} ds \\ &= -\sum_{i=1}^{N(T)} \sum_{\{j:T_j^*<t_i\}} Y_j \, e^{\delta T_j^*} \left(-\frac{1}{\delta} \left(e^{-\delta t_i} - e^{-\delta t_{i-1}} \right) \right). \end{aligned}$$

The second integral is given as:

$$-\sum_{i=1}^{N(T)} \sum_{\{j:T_j^*<t_i\}} \int_{t_{i-1}}^{t_i} Y_j \, e^{-\delta(s-T_j^*)} ds = -\sum_{i=1}^{N(T)} \sum_{\{j:T_j^*<t_i\}} Y_j \, e^{\delta T_j^*} \left[-\frac{1}{\delta} \left(e^{-\delta t_i} - e^{-\delta t_{i-1}} \right) \right]. \tag{3.2.4}$$

Evaluate the the third integral:

$$\sum_{i=1}^{N(T)} \log \lambda(t_i) = \sum_{i=1}^{N(T)} \log \Big(\lambda_0 \, e^{-\delta t_i} + \sum_{\{j:T_j^*<t_i\}} Y_j \, e^{-\delta(t_i - T_j^*)} \Big). \tag{3.2.5}$$

Finally, to get the log-likelihood function Equation (3.2.2) reduces to

$$\begin{aligned}\log(L(\lambda_0, \delta)) &= \frac{\lambda_0}{\delta}(e^{-\delta t_{N(T)}} - 1) - \sum_{i=1}^{N(T)} \sum_{\{j:T_j^*<t_i\}} Y_j \, e^{\delta T_j^*} \left[-\frac{1}{\delta} \left(e^{-\delta t_i} - e^{-\delta t_{i-1}} \right) \right] \\ &+ \sum_{i=1}^{N(T)} \log \left(\lambda_0 \, e^{-\delta t_i} + \sum_{\{j:T_j^*<t_i\}} Y_j \, e^{-\delta(t_i - T_j^*)} \right).\end{aligned}$$

Next, proceed to derive the formulae for the estimators of the unknown parameters $(\lambda_0,\ \delta)$ as follows. Differentiating the log-likelihood function with respect to λ_0 and setting the result to zero, and get

$$\left. \frac{\partial \log(L(\lambda_0, \delta))}{\partial \lambda_0} \right|_{(\lambda_0,\ \delta)=(\hat{\lambda_0},\ \hat{\delta})} = \frac{1}{\hat{\delta}}(e^{-\hat{\delta} t_{N(T)}} - 1) + \sum_{i=1}^{N(T)} \frac{e^{-\hat{\delta} t_i}}{\hat{\lambda_0} \, e^{-\hat{\delta} t_i} + \sum\limits_{\{j:T_j^*<t_i\}} Y_j \, e^{-\hat{\delta}(t_i - T_j^*)}} = 0.$$

Solving for $\hat{\lambda_0}$, we get that

$$\sum_{i=1}^{N(T)} \frac{e^{-\hat{\delta} t_i}}{\hat{\lambda_0} \, e^{-\hat{\delta} t_i} + \sum\limits_{\{j:T_j^*<t_i\}} Y_j \, e^{-\hat{\delta}(t_i - T_j^*)}} = -\frac{1}{\hat{\delta}}(e^{-\hat{\delta} t_{N(T)}} - 1). \qquad (3.2.6)$$

Differentiating the log-likelihood function with respect to δ and setting the result to zero, and obtain that

$$\begin{aligned}\frac{\partial \log(L(\lambda_0, \delta))}{\partial \delta}\Bigg|_{(\lambda_0,\ \delta)=(\hat{\lambda_0},\ \hat{\delta})} &= -\frac{\hat{\lambda_0}}{\hat{\delta}^2}\left(e^{-\hat{\delta} t_{N(T)}} - 1\right) + \frac{\hat{\lambda_0}}{\hat{\delta}}\left(t_{N(T)} e^{-\hat{\delta} t_{N(T)}}\right) \\ &- \sum_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j\, T_j^*\, e^{\hat{\delta} T_j^*}\left[-\frac{1}{\hat{\delta}}\left(e^{-\hat{\delta} t_i} - e^{-\hat{\delta} t_{i-1}}\right)\right] \\ &- \sum_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j\, e^{\hat{\delta} T_j^*}\left[\frac{1}{\hat{\delta}^2}\left(e^{-\hat{\delta} t_i} - e^{-\hat{\delta} t_{i-1}}\right)\right] \\ &- \sum_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j\, e^{\hat{\delta} T_j^*}\left[\frac{1}{\hat{\delta}}\left(t_i\, e^{-\hat{\delta} t_i} - t_{i-1}\, e^{-\hat{\delta} t_{i-1}}\right)\right] \\ &+ \sum_{i=1}^{N(T)} \frac{-\hat{\lambda_0}\, t_i\, e^{-\hat{\delta} t_i} + \sum_{\{j: T_j^* < t_i\}} Y_j\, (t_i - T_j^*)\, e^{-\hat{\delta}(t_i - T_j^*)}}{\hat{\lambda_0}\, e^{-\hat{\delta} t_i} + \sum_{\{j: T_j^* < t_i\}} Y_j\, e^{-\hat{\delta}(t_i - T_j^*)}} = 0. \end{aligned} \tag{3.2.7}$$

Since, explicit solutions for Equation (3.2.6) and Equation (3.2.7) does not exist, they may be solved numerically using the Newton Raphson method.

3.3 Derivation of the estimators in the log-normal distribution

A positive random variable $Y_j > 0$ is log-normally distributed, if the natural logarithm of Y_j is normally distributed such that $\ln(Y_j) \sim \mathcal{N}(\mu,\ \sigma^2)$. Probability density function of log-normal distribution is given by:

$$f(Y_j \mid \mu,\ \sigma^2) \;=\; \frac{1}{Y_j\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\ln Y_j - \mu)^2}{2\sigma^2}\right\},$$

where Y_j is the j^{th} claim size of log-normal distribution, $-\infty < \mu < \infty$ is the mean parameter of log-normal distribution and $\sigma > 0$ is the shape parameter of log-normal distribution.

The likelihood function of the log-normal distribution for a series of jump sizes $Y_j' s$ $(j = 1,\ 2, \ldots,\ M(t))$ is denoted by taking product of probability densities of the individual

$Y_j's$:

$$
\begin{aligned}
L(\mu,\ \sigma^2 \mid Y) &= \prod_{j=1}^{M(t)} f(Y_j \mid \mu,\ \sigma^2) \\
&= \prod_{j=1}^{M(t)} (2\pi\sigma^2)^{-\frac{1}{2}} Y_j^{-1} \exp\left\{-\frac{(\ln Y_j - \mu)^2}{2\sigma^2}\right\} \\
&= (2\pi\sigma^2)^{-\frac{M(t)}{2}} \exp\left\{-\frac{\sum\limits_{j=1}^{M(t)} (\ln Y_j - \mu)^2}{2\sigma^2}\right\} \prod_{j=1}^{M(t)} Y_j^{-1}.
\end{aligned}
$$

The log-likelihood function of the log-normal for the series of $Y_j's$ $(j = 1,\ 2, \ldots,\ M(t))$ is then derived by taking the natural log of the likelihood function.

$$
\begin{aligned}
\log L(\mu, \sigma^2 \mid Y) &= \ln\left[(2\pi\sigma^2)^{-\frac{M(t)}{2}} \exp\left\{-\frac{\sum\limits_{j=1}^{M(t)} (\ln Y_j - \mu)^2}{2\sigma^2}\right\} \prod_{j=1}^{M(t)} Y_j^{-1}\right] \\
&= -\frac{M(t)}{2} \ln\big(2\pi\sigma^2\big) - \sum_{j=1}^{M(t)} \ln Y_j - \frac{\sum\limits_{j=1}^{M(t)} (\ln Y_j - \mu)^2}{2\sigma^2}.
\end{aligned}
$$

Now, we need to find $\hat{\mu}$ and $\hat{\sigma^2}$ which maximize $\log L(\mu,\ \sigma^2 \mid Y)$. To do this, the gradient of $\log L(\mu,\ \sigma^2 \mid Y)$ is required and setting the result to zero.

Differentiating the log-likelihood function with respect to μ and setting the result to zero, and obtain that

$$
\left.\frac{\partial \log L(\mu,\ \sigma^2 \mid Y)}{\partial \mu}\right|_{(\mu,\ \sigma^2)=(\hat{\mu},\ \hat{\sigma}^2)} = \frac{\sum\limits_{j=1}^{M(t)} \ln Y_j}{\hat{\sigma}^2} - \frac{M(t)\hat{\mu}}{\hat{\sigma}^2} = 0.
$$

Solving for $\hat{\mu}$, we obtain the estimator for the mean parameter μ as

$$
\hat{\mu} = \frac{\sum\limits_{j=1}^{M(t)} \ln Y_j}{M(t)}. \tag{3.3.1}
$$

Differentiating the log-likelihood function with respect to σ^2 and setting the result to zero, and obtain that

$$\left.\frac{\partial \log L(\mu,\ \sigma^2 \mid Y)}{\partial \sigma^2}\right|_{(\mu,\ \sigma^2)=(\hat{\mu},\ \hat{\sigma}^2)} = -M(t) + \frac{\sum_{j=1}^{M(t)} (\ln Y_j - \hat{\mu})^2}{\hat{\sigma}^2} = 0.$$

Solving for $\hat{\sigma}^2$, we obtain the estimator for the variance parameter σ^2 as

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^{M(t)} (\ln Y_j - \hat{\mu})^2}{M(t)},$$

which simplifies to

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^{M(t)} (\ln Y_j)^2 - \frac{\left(\sum_{j=1}^{M(t)} \ln Y_j\right)^2}{M(t)}}{M(t)}. \tag{3.3.2}$$

3.4 Derivation of the estimator in the exponential distribution

A positive random variable $R_j > 0$ is exponentially distributed such that $R_j \sim \exp(\alpha)$. Probability density function of exponential distribution is given by:

$$f(R_j \mid \alpha) = \alpha \exp\left(-\alpha R_j\right),$$

where $R_j > 0$ is the j^{th} inter-arrival time of exponential distribution and α is the rate parameter of exponential distribution.

The likelihood function of the exponential distribution for a series of inter-arrival times $R_j's$ $(j = 1,\ 2, \ldots,\ M(t))$ is denoted by taking product of probability densities of the individual $R_j's$:

$$\begin{aligned}L(\alpha \mid R) &= \prod_{j=1}^{M(t)} f(R_j \mid \alpha) \\ &= \prod_{j=1}^{M(t)} \alpha \exp\left(-\alpha R_j\right) \\ &= \alpha^{M(t)} \exp\left(-\alpha \sum_{j=1}^{M(t)} R_j\right).\end{aligned}$$

The log-likelihood function is then derived by taking the natural log of the likelihood function,

$$\begin{aligned}\log L(\alpha \mid R) &= \ln\left[\alpha^{M(t)} \exp\left(-\alpha \sum_{j=1}^{M(t)} R_j\right)\right] \\ &= M(t)\ln\alpha - \alpha \sum_{j=1}^{M(t)} R_j.\end{aligned}$$

Differentiating the log-likelihood function with respect to α and setting the result to zero, and obtain that

$$\left.\frac{\partial L(\alpha \mid R)}{\partial \alpha}\right|_{\alpha=\hat{\alpha}} = \frac{M(t)}{\hat{\alpha}} - \sum_{j=1}^{M(t)} R_j = 0.$$

Solving for $\hat{\alpha}$, we obtain the estimator,

$$\hat{\alpha} = \frac{M(t)}{\sum_{j=1}^{M(t)} R_j}. \tag{3.4.1}$$

3.5 Derivation of the estimators in the gamma distribution

A positive random variable $R_j > 0$ is gamma distributed such that $R_j \sim \Gamma(\beta,\ k,\ \alpha)$. Probability density function of gamma distribution is given by:

$$f(R_j \mid \beta,\ k,\ \alpha) = \frac{\alpha R_j^{\alpha k-1}}{\beta^{\alpha k}\Gamma(k)} \exp\left\{-\left(\frac{R_j}{\beta}\right)^{\alpha}\right\},$$

where $R_j > 0$ is the j^{th} inter-arrival time of gamma distribution, $k > 0$ is the shape parameter of gamma distribution and $\beta > 0$ is the scale parameter of gamma distribution.

The likelihood function of the gamma distribution for a series of inter-arrival times $R_j's$ $(j = 1,\ 2, \ldots,\ M(t))$ is denoted by taking product of probability densities of the individual $R_j's$:

$$\begin{aligned} L(\beta,\ k,\ \alpha \mid R) &= \prod_{j=1}^{M(t)} f(R_j \mid \beta,\ k,\ \alpha) \\ &= \prod_{j=1}^{M(t)} \left\{ \frac{\alpha R_j^{\alpha k-1}}{\beta^{\alpha k}\Gamma(k)} \exp\left[-\left(\frac{R_j}{\beta}\right)^{\alpha} \right] \right\} \\ &= \alpha^{M(t)} \left(\beta^{\alpha k}\Gamma(k)\right)^{-M(t)} \exp\left(-\frac{\sum\limits_{j=1}^{M(t)} R_j^{\alpha}}{\beta^{\alpha}} \right) \prod_{j=1}^{M(t)} R_j^{\alpha k-1}. \end{aligned}$$

The log-likelihood function is then derived by taking the natural log of the likelihood function,

$$\begin{aligned} \log L(\beta,\ k,\ \alpha \mid R) &= \ln\left\{ \alpha^{M(t)} \left(\beta^{\alpha k}\Gamma(k)\right)^{-M(t)} \exp\left(-\sum_{j=1}^{M(t)} \left(\frac{R_j}{\beta}\right)^{\alpha} \right) \prod_{j=1}^{M(t)} R_j^{\alpha k-1} \right\} \\ &= M(t)\ln\alpha - M(t)\alpha k \ln\beta - M(t)\ln\Gamma(k) - \sum_{j=1}^{M(t)} \left(\frac{R_j}{\beta}\right)^{\alpha} \\ &+ (\alpha k - 1)\sum_{j=1}^{M(t)} \ln R_j. \end{aligned}$$

Differentiating the log-likelihood function with respect to β and setting the result to zero, and obtain that

$$\left.\frac{\partial \log L(\beta,\ k,\ \alpha \mid R)}{\partial \beta}\right|_{(\beta,\ k)=(\hat{\beta},\ \hat{k})} = -\frac{M(t)\alpha\hat{k}}{\hat{\beta}} + \frac{\alpha}{\hat{\beta}} \sum_{j=1}^{M(t)} \left(\frac{R_j}{\hat{\beta}}\right)^{\alpha} = 0.$$

Solving for $\hat{\beta}$, we obtain the estimator,

$$\hat{\beta} = \left(\frac{\sum_{j=1}^{M(t)} R_j^{\alpha}}{M(t)\hat{k}} \right)^{\frac{1}{\alpha}}. \tag{3.5.1}$$

Differentiating the log-likelihood function with respect to k and setting the result to zero, and obtain that

$$\left.\frac{\partial \log L(\beta,\ k,\ \alpha \mid R)}{\partial k}\right|_{(\beta,\ k)=(\hat{\beta},\ \hat{k})} = -M(t)\alpha \ln \hat{\beta} - M(t)\frac{\partial \ln \Gamma\left(\hat{k}\right)}{\partial \hat{k}} + \alpha \sum_{j=1}^{M(t)} \ln R_j = 0$$

Next, solve for $\hat{k}$ and replace $\frac{\partial \ln \Gamma\left(\hat{k}\right)}{\partial \hat{k}}$ with $\Psi(\hat{k})$ to obtain

$$\begin{aligned} M(t)\Psi(\hat{k}) &= \alpha \sum_{j=1}^{M(t)} \ln R_j - M(t)\alpha \ln \hat{\beta} \\ \Psi(\hat{k}) &= \frac{\alpha \sum_{j=1}^{M(t)} \ln R_j}{M(t)} - \alpha \ln \hat{\beta}. \end{aligned}$$

Differentiating the log-likelihood function with respect to α and setting the result to zero, and obtain that

$$\begin{aligned} \left.\frac{\partial \log L(\beta,\ k,\ \alpha \mid R)}{\partial \alpha}\right|_{(\beta,k)=(\hat{\beta},\hat{k})} &= \frac{M(t)}{\alpha} - M(t)\hat{k} \ln \hat{\beta} - \sum_{j=1}^{M(t)} \left(\frac{R_j}{\hat{\beta}}\right)^{\alpha} \ln\left(\frac{R_j}{\hat{\beta}}\right) \\ &+ \hat{k} \sum_{j=1}^{M(t)} \ln R_j = 0. \end{aligned}$$

$$\Longrightarrow \sum_{j=1}^{M(t)} \left(\frac{R_j}{\hat{\beta}}\right)^{\alpha} \ln\left(\frac{R_j}{\hat{\beta}}\right) - \hat{k}\sum_{j=1}^{M(t)} \ln R_j = \frac{M(t)}{\alpha} - M(t)\hat{k}\ln\hat{\beta}$$

$$\frac{M(t)}{\alpha} + \hat{k}\sum_{j=1}^{M(t)} \ln R_j = \frac{\sum_{j=1}^{M(t)} R_j^{\alpha}}{\hat{\beta}^{\alpha}} \left(\sum_{j=1}^{M(t)} \ln R_j - \ln\hat{\beta}\right) + M(t)\hat{k}\ln\hat{\beta}.$$

Substituting Equation (3.5.1) into the above equation, and obtain shape parameter $\hat{k}$, as follows

$$\begin{aligned} \frac{M(t)}{\alpha} + \hat{k}\sum_{j=1}^{M(t)} \ln R_j &= M(t)\hat{k}\left(\sum_{j=1}^{M(t)} \ln R_j - \frac{1}{\alpha}\ln\sum_{j=1}^{M(t)} R_j^{\alpha} + \frac{1}{\alpha}\ln\Big(M(t)\hat{k}\Big)\right) \\ &+ \frac{M(t)\hat{k}}{\alpha}\ln\sum_{j=1}^{M(t)} R_j^{\alpha} - \frac{M(t)\hat{k}}{\alpha}\ln\Big(M(t)\hat{k}\Big) \\ M(t)\hat{k}\sum_{j=1}^{M(t)} \ln R_j - \hat{k}\sum_{j=1}^{M(t)} \ln R_j &= \frac{M(t)}{\alpha} \\ \hat{k} &= \frac{M(t)}{\alpha\left((M(t)-1)\sum_{j=1}^{M(t)} \ln R_j\right)}. \end{aligned} \quad (3.5.2)$$

Now return to gamma distribution from the generalized gamma distribution by replacing the $\alpha = 1$ into Equation (3.5.2) then get the shape parameter $\hat{k}$,

$$\hat{k} = \frac{M(t)}{\left((M(t)-1)\sum_{j=1}^{M(t)} \ln R_j\right)}. \quad (3.5.3)$$

Substitute Equation (3.5.3) into Equation (3.5.1) when $\alpha = 1$ to get the scale parameter $\hat{\beta}$,

$$\hat{\beta} = \frac{\sum_{j=1}^{M(t)} R_j \left((M(t)-1)\sum_{j=1}^{M(t)} \ln R_j\right)}{M(t)^2}. \quad (3.5.4)$$

3.6 Derivation of the estimators in the Weibull distribution

A positive random variable $R_j > 0$ is Weibull distributed such that $R_j \sim \text{Weibull}(\gamma,\ \eta)$. Probability density function of Weibull distribution is given by:

$$f(R_j \mid \gamma,\ \eta) \ = \ \frac{\gamma}{\eta^\gamma} R_j^{\gamma-1} \exp\left\{ -\left(\frac{R_j}{\eta}\right)^\gamma \right\},$$

where $R_j > 0$ is the j^{th} inter-arrival time of Weibull distribution, $\gamma > 0$ is the shape parameter of Weibull distribution and $\eta > 0$ is the scale parameter of Weibull distribution.

The likelihood function of the Weibull distribution for a series of inter-arrival times $R_j's$ $(j = 1,\ 2, \ldots,\ M(t))$ is denoted by taking product of probability densities of the individual $R_j's$:

$$\begin{aligned} L(\gamma,\ \eta \mid R) &= \prod_{j=1}^{M(t)} f(R_j \mid \gamma,\ \eta) \\ &= \left(\frac{\gamma}{\eta^\gamma}\right)^{M(t)} \exp\left\{ -\frac{\sum_{j=1}^{M(t)} R_j^\gamma}{\eta^\gamma} \right\} \prod_{j=1}^{M(t)} R_j^{\gamma-1}. \end{aligned}$$

The log-likelihood function is then derived by taking the natural log of the likelihood function.

$$\begin{aligned} \log L(\gamma,\ \eta \mid R) &= M(t) \ln\left(\frac{\gamma}{\eta^\gamma}\right) - \frac{\sum_{j=1}^{M(t)} R_j^\gamma}{\eta^\gamma} + \sum_{j=1}^{M(t)} \ln R_j^{\gamma-1} \\ &= M(t)\ln\gamma - M(t)\gamma\ln\eta - \frac{\sum_{j=1}^{M(t)} R_j^\gamma}{\eta^\gamma} + (\gamma-1)\sum_{j=1}^{M(t)} \ln R_j. \end{aligned}$$

Differentiating the log-likelihood function with respect to $\hat{\eta}$ and setting the result to zero, to get

$$\left.\frac{\partial \log L(\gamma, \eta \mid R)}{\partial \eta}\right|_{(\gamma,\ \eta)=(\hat{\gamma},\ \hat{\eta})} = -\frac{M(t)\hat{\gamma}}{\hat{\eta}} + \hat{\gamma}\frac{\sum_{j=1}^{M(t)} R_j^{\hat{\gamma}}}{\hat{\eta}^{\hat{\gamma}+1}} = 0.$$

Solving for the scale parameter $\hat{\eta}$ from the above equation, we obtain that

$$\hat{\eta} \;=\; \frac{\sum_{j=1}^{M(t)} R_j}{M(t)^{\frac{1}{\hat{\gamma}}}}. \tag{3.6.1}$$

Differentiating the log-likelihood function with respect to γ and setting the result to zero, we obtain that

$$\left.\frac{\partial \log L(\gamma,\ \eta \mid R)}{\partial \gamma}\right|_{(\gamma,\eta)=(\hat{\gamma},\hat{\eta})} \;=\; \frac{M(t)}{\hat{\gamma}} - M(t)\ln \hat{\eta} - \sum_{j=1}^{M(t)} \left(\frac{R_j}{\hat{\eta}}\right)^{\hat{\gamma}} \ln\left(\frac{R_j}{\hat{\eta}}\right) + \sum_{j=1}^{M(t)} \ln R_j$$

$$\Longrightarrow \sum_{j=1}^{M(t)} \left(\frac{R_j}{\hat{\eta}}\right)^{\hat{\gamma}} \ln\left(\frac{R_j}{\hat{\eta}}\right) - \frac{M(t)}{\hat{\gamma}} \;=\; \sum_{j=1}^{M(t)} \ln R_j - M(t)\ln \hat{\eta}.$$

Substitute Equation (3.6.1) into the above equation to get shape parameter $\hat{\gamma}$,

$$\begin{aligned}\left(\sum_{j=1}^{M(t)} \ln R_j - \ln \sum_{j=1}^{M(t)} R_j + \frac{1}{\hat{\gamma}} \ln M(t)\right) M(t) - \frac{M(t)}{\hat{\gamma}} \;&=\; \sum_{j=1}^{M(t)} \ln R_j - M(t) \ln \sum_{j=1}^{M(t)} R_j \\ &+\; \frac{M(t)}{\hat{\gamma}} \ln M(t)\end{aligned}$$

$$\Longrightarrow \hat{\gamma} \;=\; \frac{M(t)}{(M(t)-1)\sum_{j=1}^{M(t)} \ln R_j}. \tag{3.6.2}$$

Substitute Equation (3.6.2) into Equation (3.6.1) to get scale parameter $\hat{\eta}$,

$$\hat{\eta} \;=\; \frac{\sum_{j=1}^{M(t)} R_j}{M(t)^{\left[\frac{(M(t)-1)\sum_{j=1}^{M(t)} \ln R_j}{M(t)}\right]}}. \tag{3.6.3}$$

4 BAYESIAN ESTIMATION METHOD

This chapter presents mathematical derivation of estimators for different types of distributions using the Bayesian method. For the purpose of this study, estimators for four distributions would be presented, all of which belong to a family of positive skewed distributions, that is log-normal, exponential, gamma and Weibull that are simulated using renewal process M(t). These estimators will be used to simulate the renewal shot-noise Cox process.

4.1 Introduction to Bayesian estimation method

In maximum likelihood estimation method, we have seen that the parameters to be estimated are said to be fixed quantities and only data is allowed to be random. But in Bayesian estimation method, the parameter estimates are allowed to be random that results in less uncertainty as the parameters are more informed due to the addition of the prior information about the model parameter.

In this case, we face the necessity to use Bayes' theorem that would be required to demonstrate the conditional posterior distribution of the model parameter given the random generated data, expressed as follows

$$\pi(\Theta \mid x_i) = \frac{L(x_i \mid \Theta)\pi(\Theta)}{\int L(x_i \mid \Theta)\pi(\Theta)d\Theta},$$

where x_i are the random generated data from a given distribution and Θ is the parameter value to be estimated. The integral part, $\int L(x_i \mid \Theta)\pi(\Theta)d\Theta$ which is known as the normalization constant is mostly inconvenient to compute it analytically for some other distribution and for the numerical integration may sometime be impossible.

With the help of Markov chain Monte Carlo (MCMC) methods, Bayesian inference of the model parameter may be found without the need to compute this difficult integral. Now, for simplicity the conditional posterior distribution of the model parameter given the random generated data is reduced to

$$\pi(\Theta \mid x_i) \propto L(x_i \mid \Theta)\pi(\Theta). \tag{4.1.1}$$

Before we go into details of MCMC method, we first present the derivation of conditional posterior distribution for the intensity function. Thereafter, we present other derivation of conditional posterior distributions from the family of positive skewed distributions, that is log-normal, exponential, gamma and Weibull.

4.2 Derivation of posterior distribution for the intensity function

The Hawkes process is a particular case of the doubly stochastic Poisson process: It is a linear self-exciting process. Its conditional intensity process has a particular form which depends upon the previous values of the point process $N(t)$ itself.

Definition 4.2.1 (Hawkes' process) *A point process $N(.)$ is called self-exciting if the intensity $\lambda(.)$ depends not only on the time t_i but also on the entire past times of the points process $t_1, \ldots,\ t_{i-1}$ as stated in [16].*

The renewal shot-noise Cox process is an extension of the Hawkes' process. Using $\{t_1,\ t_2, \ldots,\ t_{N(T)}\}$ to denote the observed sequence of past arrival times of the point process $N(t)$ in the interval $(0,\ T]$ with the positive influence on the current value of the intensity process. The intensity function is not depending on the point process itself but on another renewal process leading to an expression that is given by

$$\lambda(s) \;=\; \lambda_0\, e^{-\delta s} + \sum_{\{j: T_j^* < s\}} Y_j\, e^{-\delta(s - T_j^*)}.$$

Proposition 4.2.2 *Let $N(.)$ be a regular point process on $[0, T]$ for some finite positive T and let $t_1,\ t_2, \ldots,\ t_{N(T)}$ denote a realisation of $N(.)$ over $[0,\ T]$. Then, the general expression of the likelihood L of $N(.)$ for any process from [17] is expressible in the form*

$$L \;=\; \left(\prod_{i=1}^{N(T)} \lambda(t_i)\right) \exp\left(-\int_0^T \lambda(s) ds\right). \tag{4.2.1}$$

From the given Proposition 4.2.2, we substitute the intensity function of renewal shot-noise Cox process into Equation (4.2.1) and obtain the expansion

$$
\begin{aligned}
L &= \prod_{i=1}^{N(T)} \left(\lambda_0 \, e^{-\delta t_i} + \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{-\delta(t_i - T_j^*)} \right) \\
& \exp \left\{ - \int_0^T \left(\lambda_0 \, e^{-\delta s} + \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{-\delta(s - T_j^*)} \right) ds \right\}. \qquad (4.2.2)
\end{aligned}
$$

Evaluate the first section of the expansion

$$
\prod_{i=1}^{N(T)} \left(\lambda_0 \, e^{-\delta t_i} + \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{-\delta(t_i - T_j^*)} \right) = \lambda_0 \, e^{-\delta \sum_{i=1}^{N(T)} t_i} + \prod_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{-\delta(t_i - T_j^*)}.
$$

Evaluate the second section of the expansion

$$
\exp \left(- \int_0^T \lambda(s) ds \right) = \exp \left\{ - \sum_{i=1}^{N(T)} \int_{t_{i-1}}^{t_i} \left(\lambda_0 \, e^{-\delta s} + \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{-\delta(s - T_j^*)} \right) ds \right\} \qquad (4.2.3)
$$

Substitute Equation (3.2.3) and Equation (3.2.4) into Equation (4.2.3), and obtain the expansion

$$
= \exp \left\{ \frac{\lambda_0}{\delta} (e^{-\delta t_{N(T)}} - 1) - \sum_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{\delta T_j^*} \left[-\frac{1}{\delta} \left(e^{-\delta t_i} - e^{-\delta t_{i-1}} \right) \right] \right\}.
$$

Finally, to get the likelihood function, Equation (4.2.2) reduces to

$$
\begin{aligned}
L(\lambda_0, \delta \mid t^{N(T)}, Y^{N(T)}) &= \left\{ \lambda_0 \, e^{-\delta \sum_{i=1}^{N(T)} t_i} + \prod_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{-\delta(t_i - T_j^*)} \right\} \\
& \exp \left\{ \frac{\lambda_0}{\delta} (e^{-\delta t_{N(T)}} - 1) - \sum_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{\delta T_j^*} \left[-\frac{1}{\delta} \left(e^{-\delta t_i} - e^{-\delta t_{i-1}} \right) \right] \right\}. \qquad (4.2.4)
\end{aligned}
$$

The likelihood function proportional to the parameter λ_0, fixing all other quantities, is expressed as

$$L(\lambda_0 \mid t^{N(T)}, Y^{N(T)}) \;\;\propto\;\; \lambda_0 \exp\left\{ \frac{\lambda_0}{\delta}(e^{-\delta t_{N(T)}} - 1) \right\}. \qquad (4.2.5)$$

Assuming the prior distribution of the parameter λ_0 is gamma distribution with hyperparameters of shape, θ and scale, ϕ as suggested in [18], we obtain that

$$\pi(\lambda_0) \;\;\propto\;\; \lambda_0^{\theta-1} \exp\biggl(-\frac{\lambda_0}{\phi}\biggr); \quad \lambda_0 > 0. \qquad (4.2.6)$$

Based on Equation (4.1.1), the conditional posterior distribution of λ_0 is given by

$$\pi(\lambda_0 \mid t^{N(T)}, Y^{N(T)}) \;\;\propto\;\; L(\lambda_0 \mid t^{N(T)}, Y^{N(T)})\pi(\lambda_0). \qquad (4.2.7)$$

Substituting Equation (4.2.5) and Equation (4.2.6) into Equation (4.2.7) and obtain the expression

$$\pi(\lambda_0 \mid t^{N(T)}, Y^{N(T)}) \;\;\propto\;\; \lambda_0^{1+\theta-1} \exp\left\{ \lambda_0 \left(\frac{e^{-\delta}}{\delta} - \frac{1}{\delta} - \frac{1}{\phi} \right) \right\}. \qquad (4.2.8)$$

Therefore, the conditional posterior distribution of λ_0 is said to be gamma distribution with the hyperparameters of shape, $\theta' = 1 + \theta$ and scale, $\Phi' = \left(\frac{e^{-\delta}}{\delta} - \frac{1}{\delta} - \frac{1}{\phi} \right)^{-1}$.

Also, the likelihood function proportional to the parameter δ fixing all other quantities is expressed as

$$\begin{aligned} L(\delta \mid t^{N(T)}, Y^{N(T)}) \;\;\propto\;\; & \left\{ e^{-\delta \sum_{i=1}^{N(T)} t_i} + \prod_{i=1}^{N(T)} \sum_{\{j:T_j^* < t_i\}} Y_j \, e^{-\delta(t_i - T_j^*)} \right\} \\ & \exp \left\{ \frac{\lambda_0}{\delta}(e^{-\delta t_{N(T)}} - 1) - \sum_{i=1}^{N(T)} \sum_{\{j:T_j^* < t_i\}} Y_j \, e^{\delta T_j^*} \left[-\frac{1}{\delta} \left(e^{-\delta t_i} - e^{-\delta t_{i-1}} \right) \right] \right\}. \end{aligned} \qquad (4.2.9)$$

Assuming the prior distribution of the parameter δ is gamma distribution with hyperparameters of scale, n and shape, p as suggested in [18], we obtain that

$$\pi(\delta) \;\;\propto\;\; \delta^{\,n-1} \exp\biggl(-\frac{\delta}{p}\biggr); \quad \delta > 0. \qquad (4.2.10)$$

Based on Equation (4.1.1), the conditional posterior distribution of δ is given by:

$$\pi(\delta \mid t^{N(T)}, Y^{N(T)}, \lambda_0) \;\; \propto \;\; L(\delta \mid t^{N(T)}, Y^{N(T)}, \lambda_0)\pi(\delta). \qquad (4.2.11)$$

Substituting Equation (4.2.9) and Equation (4.2.10) into Equation (4.2.11) and obtain the expression

$$\begin{aligned}\pi(\delta \mid t^{N(T)}, Y^{N(T)}) \;\; \propto \;\; & \left\{ \delta^{\,n-1} e^{-\delta \sum\limits_{i=1}^{N(T)} t_i} + \prod_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j \, \delta^{\,n-1} e^{-\delta (t_i - T_j^*)} \right\} \\ & \exp \left\{ \frac{\hat{\lambda}_0}{\delta}(e^{-\delta t_{N(T)}} - 1) - \frac{\delta}{p} - \sum_{i=1}^{N(T)} \sum_{\{j: T_j^* < t_i\}} Y_j \, e^{\delta T_j^*} \left[-\frac{1}{\delta} \left(e^{-\delta t_i} - e^{-\delta t_{i-1}} \right) \right] \right\} \end{aligned} \qquad (4.2.12$$

Therefore, the conditional posterior distribution of δ is not of the closed form solution.

4.3 Derivation of posterior distribution for the log-normal distribution

Probability density function of log-normal distribution is given by:

$$f(Y_j \mid \mu, \; \Phi) \;\; = \;\; \frac{1}{Y_j} \sqrt{\frac{\Phi}{2\pi}} \exp \left\{ -\frac{\Phi}{2} (\ln Y_j - \mu)^2 \right\},$$

where Y_j is the j^{th} claim size of log-normal distribution, $-\infty < \mu < \infty$ is the mean parameter and Φ is the precision parameter but not of the log-normal distribution.

The likelihood function of the log-normal distribution for a series of jump sizes $Y_j's$ $(j = 1, \; 2, \ldots, \; M(t))$ is denoted by taking product of probability densities of the individual $Y_j's$:

$$\begin{aligned} L(\mu, \; \Phi \mid Y) \;\; &= \;\; \prod_{j=1}^{M(t)} f(Y_j \mid \mu, \; \Phi) \\ &= \;\; \prod_{j=1}^{M(t)} \left(\frac{\Phi}{2\pi} \right)^{\frac{1}{2}} Y_j^{-1} \exp \left\{ -\frac{\Phi}{2} (\ln Y_j - \mu)^2 \right\} \\ &= \;\; \left(\frac{\Phi}{2\pi} \right)^{\frac{M(t)}{2}} \exp \left\{ -\frac{\Phi}{2} \sum_{j=1}^{M(t)} (\ln Y_j - \mu)^2 \right\} \prod_{j=1}^{M(t)} Y_j^{-1}. \end{aligned}$$

4.3.1 Unknown parameter, μ

Assume that the parameter μ, is unknown and parameter Φ, is known. The likelihood function proportional to parameter μ is given by:

$$L(\mu \mid Y) \;\; \propto \;\; \exp\left\{ -\frac{M(t)\Phi}{2} \left(\frac{\sum_{j=1}^{M(t)} \ln Y_j}{M(t)} - \mu \right)^2 \right\}. \qquad (4.3.1)$$

where $\overline{Y} = \dfrac{\sum_{j=1}^{M(t)} \ln Y_j}{M(t)}$ is the mean of log data.

Assuming the conjugate prior distribution of the parameter μ is normal distribution with hyperparameters of mean, m and precision, q as suggested in [19], we obtain that

$$\pi(\mu) \;\; \propto \;\; \exp\left\{ -\frac{q}{2} (\mu - m)^2 \right\}. \qquad (4.3.2)$$

Based on Equation (4.1.1), the conditional posterior distribution of μ is given by:

$$\pi(\mu \mid Y) \;\; \propto \;\; L(\mu \mid Y)\pi(\mu). \qquad (4.3.3)$$

Substituting Equation (4.3.1) and Equation (4.3.2) into Equation (4.3.3) and obtain the expression

$$\begin{aligned}
\pi(\mu \mid Y) \;\; &\propto \;\; \exp\left\{ -\frac{M(t)\Phi}{2} \left(\overline{Y} - \mu\right)^2 - \frac{q}{2}(\mu - m)^2 \right\} \\
&\propto \;\; \exp\left\{ -\frac{1}{2} \left(M(t)\Phi\overline{Y}^2 - 2M(t)\Phi\overline{Y}\mu + M(t)\Phi\mu^2 + q\mu^2 - 2q\mu m + qm^2 \right) \right\} \\
&\propto \;\; \exp\left[-\frac{1}{2} \left\{ \mu^2(M(t)\Phi + q) - 2\mu(M(t)\Phi\overline{Y} + qm) + M(t)\Phi\overline{Y}^2 + qm^2 \right\} \right] \\
&\propto \;\; \exp\left\{ -\frac{1}{2} (M(t)\Phi + q) \left(\mu^2 - 2\mu\frac{M(t)\Phi\overline{Y} + qm}{M(t)\Phi + q} + \frac{M(t)\Phi\overline{Y}^2 + qm^2}{M(t)\Phi \; + q} \right) \right\} \\
&\propto \;\; \exp\left\{ -\frac{1}{2} (M(t)\Phi + q) \left(\mu - \frac{M(t)\Phi\overline{Y} + qm}{M(t)\Phi + q} \right)^2 \right\} \qquad (4.3.4)
\end{aligned}$$

Therefore, the conditional posterior distribution of μ is said to be normal distribution with the hyperparameters of mean, $m' = \dfrac{M(t)\Phi\overline{Y} + qm}{M(t)\Phi + q}$ and precision, $q' = M(t)\Phi + q$.

4.3.2 Unknown parameter, Φ

Assume that the parameter Φ, is unknown and parameter μ, is known. The likelihood function proportional to parameter Φ is given by:

$$L(\Phi \mid Y) \quad \propto \quad \Phi^{\frac{M(t)}{2}} \exp\left\{-\frac{M(t)\Phi}{2}\sum_{j=1}^{M(t)}(\ln Y_j - \mu)^2\right\}. \tag{4.3.5}$$

where $SS = \sum_{j=1}^{M(t)} (\ln Y_j - \mu)^2$ is the sum of squares of the residuals.

Assuming the conjugate prior distribution of the parameter Φ is gamma distribution with hyperparameters of shape, α and scale, β as suggested in [19], we obtain that

$$\pi(\Phi) \quad \propto \quad \Phi^{\,\alpha-1} \exp\left(-\frac{\Phi}{\beta}\right); \quad \Phi > 0. \tag{4.3.6}$$

Based on Equation (4.1.1), the conditional posterior distribution of Φ is given by:

$$\pi(\Phi \mid Y) \quad \propto \quad L(\Phi \mid Y)\pi(\Phi). \tag{4.3.7}$$

Substituting Equation (4.3.5) and Equation (4.3.6) into Equation (4.3.7) and obtain the expression

$$\begin{aligned} \pi(\Phi \mid Y) \quad &\propto \quad \Phi^{\alpha+\frac{M(t)}{2}-1} \exp\left(-\frac{\Phi}{2}SS - \frac{\Phi}{\beta}\right) \\ &\propto \quad \Phi^{\alpha+\frac{M(t)}{2}-1} \exp\left\{-\Phi\left(\frac{SS}{2} + \frac{1}{\beta}\right)\right\}. \end{aligned} \tag{4.3.8}$$

Therefore, the conditional posterior distribution of Φ is said to be gamma distribution with the hyperparameters of shape, $\alpha' = \alpha + \dfrac{M(t)}{2}$ and scale, $\beta' = \left(\dfrac{SS}{2} + \dfrac{1}{\beta}\right)^{-1}$.

4.4 Derivation of posterior distribution for the exponential distribution

Probability density function of exponential distribution is given by:

$$f(R_j \mid \lambda) \quad = \quad \lambda \exp\left(-\lambda R_j\right),$$

where $R_j > 0$ is the j^{th} inter-arrival time of exponential distribution and λ is the rate parameter of exponential distribution.

The likelihood function of the exponential distribution for a series of inter-arrival times $R_j^{\prime}s$ $(j = 1,\ 2, \ldots,\ M(t))$ is denoted by taking product of probability densities of the individual $R_j^{\prime}s$:

$$\begin{aligned} L(\lambda \mid R) &= \prod_{j=1}^{M(t)} f(R_j \mid \lambda) \\ &= \prod_{j=1}^{M(t)} \lambda \exp\left(-\lambda R_j\right) \\ &= \lambda^{M(t)} \exp\left(-\lambda \sum_{j=1}^{M(t)} R_j\right). \end{aligned} \tag{4.4.1}$$

Assuming the conjugate prior distribution of the parameter λ is gamma distribution with hyperparameters of shape, a and scale, b as suggested in [20], we obtain that

$$\pi(\lambda) \quad \propto \quad \lambda^{\,a-1} \exp\left(-\frac{\lambda}{b}\right); \quad \lambda > 0. \tag{4.4.2}$$

Based on Equation (4.1.1), the conditional posterior distribution of λ is given by:

$$\pi(\lambda \mid R) \quad \propto \quad L(\lambda \mid R)\pi(\lambda). \tag{4.4.3}$$

Substituting Equation (4.4.1) and Equation (4.4.2) into Equation (4.4.3) and obtain the expression

$$\pi(\lambda \mid R) \quad \propto \quad \lambda^{\,a+M(t)-1} \exp\left\{-\lambda\left(\sum_{j=1}^{M(t)} R_j + \frac{1}{b}\right)\right\}. \tag{4.4.4}$$

Therefore, the conditional posterior distribution of λ is said to be gamma distribution with the hyperparameters of shape, $a' = a + M(t)$ and scale, $b' = \left(\sum\limits_{j=1}^{M(t)} R_j + \dfrac{1}{b}\right)^{-1}$.

4.5 Derivation of posterior distribution for the gamma distribution

Probability density function of gamma distribution is given by

$$f(R_j \mid k,\ s) \ = \ \frac{R_j^{k-1}}{s^k \Gamma\left(k\right)} \exp\left(-\frac{R_j}{s}\right),$$

where $R_j > 0$ is the j^{th} inter-arrival time of gamma distribution, $k > 0$ is the shape parameter of gamma distribution and $s > 0$ is the scale parameter of gamma distribution.

The likelihood function of the gamma distribution for a series of inter-arrival times $R_j's$ $(j = 1,\ 2, \ldots,\ M(t))$ is denoted by taking product of probability densities of the individual $R_j's$

$$\begin{aligned} L(k,\ s \mid R) &= \prod_{j=1}^{M(t)} f(R_j \mid s,\ k) \\ &= \prod_{j=1}^{M(t)} \left\{ \frac{R_j^{k-1}}{s^k \Gamma\left(k\right)} \exp\left(-\frac{R_j}{s}\right) \right\} \\ &= \left(s^k \Gamma\left(k\right)\right)^{-M(t)} \exp\left(-\frac{\sum\limits_{j=1}^{M(t)} R_j}{s}\right) \prod_{j=1}^{M(t)} R_j^{k-1}. \end{aligned} \tag{4.5.1}$$

With the indefinite assumption on the prior of the two-parameter gamma distribution, let us assume that the non informative prior as stated in [21] is expressed as

$$\pi(k, s) \ \propto \ \frac{1}{s}. \tag{4.5.2}$$

Based on Equation (4.1.1), the joint posterior distribution of k and s is given by

$$\pi(k, s \mid R) \ \propto \ L(k, s \mid R)\pi(k, s). \tag{4.5.3}$$

Substituting Equation (4.5.1) and Equation (4.5.2) into Equation (4.5.3) and obtain the expression

$$\pi(k, s \mid R) \ \propto \ \frac{\left(s^k \Gamma\left(k\right)\right)^{-M(t)}}{s} \exp\left(-\frac{\sum\limits_{j=1}^{M(t)} R_j}{s}\right) \prod_{j=1}^{M(t)} R_j^{k-1}. \tag{4.5.4}$$

The conditional posterior distribution of parameter k given fixing all other quantities is given as

$$\pi(k \mid R) \quad \propto \quad \left(s^k\Gamma\left(k\right)\right)^{-M(t)} \prod_{j=1}^{M(t)} R_j^{k-1}. \tag{4.5.5}$$

Therefore, the conditional posterior distribution of k is not of the closed form solution.

Also, the conditional posterior distribution of parameter s given fixing all other quantities is given as

$$\pi(s \mid R) \quad \propto \quad s^{-(kM(t)+1)} \exp\left(-\frac{\sum_{j=1}^{M(t)} R_j}{s}\right). \tag{4.5.6}$$

Therefore, the conditional posterior distribution of s is said to be inverse-gamma distribution with the hyperparameters of shape, $k' = kM(t)$ and scale, $s' = \sum_{j=1}^{M(t)} R_j$.

4.6 Derivation of posterior distribution for the Weibull distribution

Probability density function of Weibull distribution is given by

$$f(R_j \mid \gamma,\, \eta) \quad = \quad \frac{\gamma}{\eta^\gamma} R_j^{\gamma-1} \exp\left\{-\left(\frac{R_j}{\eta}\right)^\gamma\right\},$$

where $R_j > 0$ is the j^{th} inter-arrival time of Weibull distribution, $\gamma > 0$ is the shape parameter of Weibull distribution and $\eta > 0$ is the scale parameter of Weibull distribution.

The likelihood function of the Weibull distribution for a series of inter-arrival times $R_j's$ $(j = 1,\ 2, \ldots,\ M(t))$ is denoted by taking product of probability densities of the individual $R_j's$

$$\begin{aligned} L(\gamma,\, \eta \mid R) \quad &= \quad \prod_{j=1}^{M(t)} f(R_j \mid \gamma,\, \eta) \\ &= \quad \left(\frac{\gamma}{\eta^\gamma}\right)^{M(t)} \exp\left\{-\frac{\sum_{j=1}^{M(t)} R_j^\gamma}{\eta^\gamma}\right\} \prod_{j=1}^{M(t)} R_j^{\gamma-1}. \end{aligned} \tag{4.6.1}$$

With the indefinite assumption on the prior of the two-parameter Weibull distribution, let us assume that the non informative prior as stated in [22] is expressed as

$$\pi(\gamma,\ \eta) \quad \propto \quad \frac{1}{\gamma\eta}. \tag{4.6.2}$$

Based on Equation (4.1.1), the joint posterior distribution of γ and η is given by

$$\pi(\gamma,\eta \mid R) \quad \propto \quad L(\gamma,\eta \mid R)\pi(\gamma,\eta). \tag{4.6.3}$$

Substituting Equation (4.6.1) and Equation (4.6.2) into Equation (4.6.3) and obtain the expression

$$\pi(\gamma,\eta \mid R) \quad \propto \quad \frac{1}{\gamma\eta}\left(\frac{\gamma}{\eta^{\gamma}}\right)^{M(t)} \exp\left\{-\frac{\sum_{j=1}^{M(t)} R_i^{\gamma}}{\eta^{\gamma}}\right\}\prod_{j=1}^{M(t)} R_j^{\gamma-1}. \tag{4.6.4}$$

The conditional posterior distribution of parameter γ given fixing all other quantities is given as

$$\pi(\gamma \mid R) \quad \propto \quad \frac{1}{\gamma}\left(\frac{\gamma}{\eta^{\gamma}}\right)^{M(t)} \exp\left\{-\frac{\sum_{j=1}^{M(t)} R_j^{\gamma}}{\eta^{\gamma}}\right\}\prod_{j=1}^{M(t)} R_j^{\gamma-1}. \tag{4.6.5}$$

Therefore, the conditional posterior distribution of γ is not of the closed form solution.

Also, the conditional posterior distribution of parameter η given fixing all other quantities is given as

$$\pi(\eta \mid R) \quad \propto \quad \frac{1}{\eta^{\gamma M(t)+1}} \exp\left\{-\frac{\sum_{j=1}^{M(t)} R_j^{\gamma}}{\eta^{\gamma}}\right\}. \tag{4.6.6}$$

Therefore, the conditional posterior distribution of η is not of the closed form solution.

5 MCMC METHODS TO BAYESIAN ESTIMATION

This chapter presents the most popular techniques of MCMC methods. For the purpose of this study, the techniques and their respective algorithms will be presented. These algorithms will provide assistance to the Bayesian estimation with the help of MCMC methods.

5.1 Introduction to MCMC methods

MCMC methods have been proposed to generate a sequence of samples, $\theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ whose distribution approaches the posterior distribution as the sample size, n increases. The Monte Carlo term is used to describe methods that involve random generation of numbers. Based on the sequence of samples generated, the new sample point, θ_{i+1} depend on the previous sample point, θ_i and hence the samples form a Markov Chain [23].

With MCMC method, the model parameters can be estimated without the need to estimate the difficult integral. Metropolis-Hastings algorithm together with Gibbs sampling turns out to be the most popular techniques of MCMC methods.

5.2 Metropolis-Hastings algorithm

This is one of the most widely used MCMC algorithms as it simply works by generating the candidate sample points from the proposal distribution followed by either accepting or rejecting the proposal candidate points generated using a simple rule. This simple rule can be implemented by generating the uniform random number, u and comparing it to the acceptance probability. The algorithm has proved to be more efficient since it deals with non-symmetric proposal distributions. This algorithm still works well even with cases where the conditional posterior distribution does not follow any closed distribution [24].

The Metropolis-Hastings algorithm can be written as described in [25] as follows:

Algorithm 1 : Metropolis-Hastings algorithm

Input: Initial parameter θ_1, proposal distribution $q(. \mid \theta)$, chain length N

Output: Posterior sample parameters $\{\theta_n\}_{n=1}^{N}$

while *the number of iteration is less than the chain length* **do**

- Generate a new candidate point: $\hat{\theta} \sim q(\hat{\theta} \mid \theta_n)$

- Compute $\alpha(\theta_n, \hat{\theta})$ = min$(1, \dfrac{\pi(\hat{\theta})q(\theta_n \mid \hat{\theta})}{\pi(\theta_n)q(\hat{\theta} \mid \theta_n)})$

- Generate a uniform random variable: $u \sim U(0, 1)$

if $u < \alpha(\theta_n, \hat{\theta})$ **then**

Accept: Set $\theta_{n+1} = \hat{\theta}$

else

Reject: Set $\theta_{n+1} = \theta_n$

end

end

To implement the Metropolis-Hasting algorithm, we need to specify the proposal distribution, $q(\hat{\theta} \mid \theta_n)$ in the algorithm that would be easy to sample from it.

5.2.1 Random walk Metropolis (RWM)

The RWM is said to be one of the best choice in generating the candidate samples since it explores the neighbourhood of the current candidate value, θ_n to propose the new candidate value, $\hat{\theta}$ as mentioned in [24].

This works by assuming that the proposal distribution, $q(\hat{\theta} \mid \theta_n)$ is symmetric and the probability of generating the movement from θ_n to $\hat{\theta}$ only depend on difference of its movement, $q(\hat{\theta} \mid \theta_n) = g(\left|\theta_n - \hat{\theta}\right|)$ where g is the symmetric distribution [24].

To implement this case, we need to generate $\hat{\theta}$ by setting $\hat{\theta} = \theta_n + \epsilon$ where ϵ are said to be independent and identically distributed (i.i.d) with symmetric distribution such as normal distribution with the mean of 0 and variance of $\sigma^2_{\theta_n}$. Here, the choice of proposal variance, $\sigma^2_{\theta_n}$ is very important as it has a great influence on the efficiency of the algorithm.

The modification of Metropolis-Hasting algorithm to random walk Metropolis algorithm can be written as described in [25] as follows:

Algorithm 2 : Random Walk Metropolis algorithm

Input: Initial parameter θ_1, proposal variance $\sigma^2_{\theta_n}$, chain length N

Output: Posterior sample parameters $\{\theta_n\}_{n=1}^N$

while *the number of iteration is less than the chain length* **do**

- Generate a new candidate point: $\hat{\theta} \sim N(\theta_n, \sigma^2_{\theta_n})$

- Compute $\alpha(\theta_n, \hat{\theta}) = \min(1, \frac{\pi(\hat{\theta})}{\pi(\theta_n)})$

- Generate a uniform random variable: $u \sim U(0, 1)$

if $u < \alpha(\theta_n, \hat{\theta})$ **then**

Accept: Set $\theta_{n+1} = \hat{\theta}$

else

Reject: Set $\theta_{n+1} = \theta_n$

end

end

5.3 Gibbs sampling

In the above algorithm, the candidate sample points are all proposed at the same time. However, sometimes in high-dimensional problems it is actually difficult to get a good multivariate proposal distributions. In this case, Gibbs sampling is introduced to allow sampling from one-dimensional proposal distribution meaning each parameter is sampled one at a time given that the remaining parameters are kept fixed [25].

The algorithm is only efficient, if the conditional posterior distribution of each parameter is of known form or closed distribution. Gibbs sampling is the special case of Metropolis-Hastings algorithm where the proposal distributions are the conditional posterior distributions and all the proposal candidate sample points are always accepted.

The Gibbs sampler can be written as described in [25] as follows:

Algorithm 3 : Gibbs sampler

Input: Initial parameter $\theta_0 = ({\theta_1}^0, {\theta_2}^0, \dots, {\theta_p}^0) \sim q(\theta)$, parameter length p, chain length N

Output: Posterior sample parameters $\{\Theta_n\}_{n=1}^{p}$

for *iteration* $i = 1, 2, \dots, N$ **do**

$\quad {\theta_1}^i \sim \pi(\Theta_1 = \theta_1 \mid \Theta_2 = {\theta_2}^{i-1}, \Theta_3 = {\theta_3}^{i-1}, \dots, \Theta_p = {\theta_p}^{i-1})$

$\quad {\theta_2}^i \sim \pi(\Theta_2 = \theta_2 \mid \Theta_1 = {\theta_1}^{i}, \Theta_3 = {\theta_3}^{i-1}, \dots, \Theta_p = {\theta_p}^{i-1})$

$\quad \vdots$

$\quad {\theta_p}^i \sim \pi(\Theta_p = \theta_p \mid \Theta_1 = {\theta_1}^{i}, \Theta_2 = {\theta_2}^{i}, \dots, \Theta_p = \theta_{p-1}^{i})$

end

6 A SIMULATION STUDY

This chapter introduces the simulation study of the renewal shot-noise Cox process. To accomplish this study, the intensity function introduced and built in Chapter 2 will be used to simulate the point process $N(t)$ with intensity of the renewal shot-noise Cox process.

6.1 Derivation of general expressions used in simulation techniques

In this section, derivation of general expressions used in the simulation process is provided. However, first familiarize ourselves with some basic concepts and definitions that would assist in the derivation procedures.

Definition 6.1.1 (Point process) *A point process $N(t)$ is a stochastic, or random process composed of a time series of binary events that occur in continuous time. However, there are properties of binary time series obtained from the underlying point process in terms of a natural parameter function of the point process. This parameter is called the zero probability function since it is computed as the probability that no events of the point process occur in a given time set. They are used to describe data that are localized at a finite set of time points. As opposed to continuous-valued processes, which can take on any of countless values at each point in time, a point process can take on only one of two possible values, indicating whether or not an event occurs at that time [13].*

Proposition 6.1.2 *Let the point process $N(t)$ be an non-homogeneous Poisson process where the intensity function $\lambda(t)$ is a non-constant deterministic function that depends on time then the process is said to follow Poisson distribution with the parameter $\int\limits_0^t \lambda(s)\, ds$ as stated in [16]. Therefore, we have:*

$$P(N(t) = n) \quad = \quad \frac{\left(\int\limits_0^t \lambda(s)ds\right)^n \exp\left(-\int\limits_0^t \lambda(s)\, ds\right)}{n!}.$$

Proposition 6.1.3 *The number of events in any interval of length t is Poisson distributed with mean $\int\limits_0^t \lambda(s)\, ds$ as proved in [16] without taking into consideration on the starting* *point i.e, stationary increments. That is, for all* $s,\ t \geq 0$

$$P(N(t+s) - N(s) = n) \quad = \quad \frac{\left(\int\limits_0^t \lambda(s)ds\right)^n \exp\left(-\int\limits_0^t \lambda(s)\, ds\right)}{n!}.$$

Consider a Poisson process, and let T_1 denote the time of the first event. Further, for $i \geq 1$ we let T_i denote the time between the $(i-1)st$ and the ith event. The sequence $\{T_i,\ i \geq 1\}$ is called the sequence of inter-arrival times.

Now, determine the distribution of the T_i but to do so first take into consideration $\{T_1 > t\}$ that takes place if and only if, no events of the Poisson process occur in the interval $[0,\ t]$, and thus

$$P(T_1 > t) = P(N(t) = 0) = \exp\left\{-\int\limits_0^t \lambda(s)\, ds\right\}.$$

Hence, T_1 has an exponential distribution with mean value function $\Lambda(t) = \int\limits_0^t \lambda(s)\, ds < \infty$ that is the expected number of events of the non-homogeneous Poisson process on the time interval $(0,\ t)$. To obtain the distribution of T_2 conditional on T_1, we proceed as follows:

$$\begin{aligned} P(T_2 > t | T_1 = s) &= P(N(t+s) - N(s) = 0 | T_1 = s) \\ &= P(N(t+s) - N(s) = 0) \text{ (by independent increments)}, \\ &= \exp\left\{-\int\limits_0^t \lambda(s)\, ds\right\} \text{ (by stationary increments)}. \end{aligned} \quad (6.1.1)$$

Therefore, conclude that T_2 is also exponential random variable with mean value function $\Lambda(t) = \int\limits_0^t \lambda(s)\, ds < \infty$ and further more T_2 is independent of T_1. Hence, $T_i,\ i = 1,\ 2, \ldots$ are independent identically distributed exponential random variables having the mean value function $\Lambda(t) = \int\limits_0^t \lambda(s)\, ds < \infty$ as stated in [26].

Using the basic concept on Bayes' theorem i.e. $P(A \cap B) = P(A|B) \times P(B)$ and the stated expression in Equation (6.1.1), then get the general expression of jump times and their probabilities to belong in a particular interval for $i = 1, \ldots,\ N(T)$ to assist in derivation procedures.

Suppose that we are dealing with the stochastic intensity function $\lambda(t)$ that is given by:

$$\lambda(t) = \lambda_0 \, e^{-\delta t} + \sum_{\{j:T_j^* < t\}} Y_j \, e^{-\delta(t - T_j^*)}.$$

Then

$$\begin{aligned} P(T_{i-1}^* < T < T_i^*) &= P(T < T_i^* \mid T > T_{i-1}^*) P(T > T_{i-1}^*) \\ &= \big(1 - P(T > T_i^* \mid T > T_{i-1}^*)\big) \, P(T > T_{i-1}^*) \end{aligned}$$

$$\begin{aligned} \Longrightarrow P(T > T_i^* \mid T > T_{i-1}^*) &= P(N(T_{i-1}^*, T_i^*) = 0) \\ &= \exp\left\{ - \int_{T_{i-1}^*}^{T_i^*} \big(\lambda_0 \, e^{-\delta t} + \sum_{\{j:T_j^* < T_i^*\}} Y_j \, e^{-\delta(t - T_j^*)}\big) \, dt \right\} \end{aligned}$$

$$\begin{aligned} \Longrightarrow P(T > T_{i-1}^*) &= P(N(0, T_{i-1}^*) = 0) \\ &= \exp\left\{ - \int_{0}^{T_{i-1}^*} \big(\lambda_0 \, e^{-\delta t} + \sum_{\{j:T_j^* < T_{i-1}^*\}} Y_j \, e^{-\delta(t - T_j^*)}\big) \, dt \right\}, \end{aligned}$$

where the jump time t generated by the Cox process $N(t)$ with the intensity function depending on the renewal process is said to be realization of the random variable T.

In this simulation process, fix in values for time T, initial intensity λ_0 and exponential rate of decay δ.

6.1.4 Finding the probability of i^{th} jump times from the point process $N(t)$ belonging to a given interval

1. First, find the probability P_1 of i^{th} jump times from the point process $N(t)$ belonging to an interval $[0,\ T_1^*]$.

$$
\begin{aligned}
P(T < T_1^*) &= 1 - P(T > T_1^*) \\
&= 1 - \exp\left\{ - \int_0^{T_1^*} \lambda_0\, e^{-\delta t}\, dt \right\} \\
\Longrightarrow P_1 &= 1 - \exp\left\{ \frac{\lambda_0}{\delta} \left(e^{-\delta T_1^*} - 1\right) \right\}.
\end{aligned}
$$

2. Then, find the probability P_2 of i^{th} jump times from the point process $N(t)$ belonging to an interval $[T_1^*,\ T_2^*]$ that is given by:

$$
\begin{aligned}
P(T_1^* < T < T_2^*) &= P(T < T_2^* | T > T_1^*) P(T > T_1^*) \\
&= \left[1 - \exp\left\{ \frac{\lambda_0}{\delta} \left(e^{-\delta T_2^*} - e^{-\delta T_1^*}\right) + \frac{Y_1}{\delta} \left(e^{-\delta (T_2^* - T_1^*)} - 1\right) \right\} \right] \\
&\times \ \exp\left\{ \frac{\lambda_0}{\delta} \left(e^{-\delta T_1^*} - 1\right) \right\} \\
\Longrightarrow P_2 &= \exp\left\{ \frac{\lambda_0}{\delta} \left(e^{-\delta T_1^*} - 1\right) \right\} - \exp\left\{ \frac{\lambda_0}{\delta} \left(e^{-\delta T_2^*} - 1\right) + \frac{Y_1}{\delta} \left(e^{-\delta (T_2^* - T_1^*)} - 1 \right. \right.
\end{aligned}
$$

3. Formulation of the general expression of the probability P_i for $i = 1,\ 2, \ldots,\ N(C)$ is given by:

$$
\begin{aligned}
P_i &= \exp\left\{ \frac{\lambda_0}{\delta} \left(e^{-\delta T_{i-1}^*} - 1\right) + \sum_{j=1}^{i-2} \frac{Y_j}{\delta} \left(e^{-\delta (T_{i-1}^* - T_j^*)} - e^{\delta T_j^*} \right) \right\} \\
&- \exp\left\{ \frac{\lambda_0}{\delta} \left(e^{-\delta T_i^*} - 1\right) + \sum_{j=1}^{i-1} \frac{Y_j}{\delta} \left(e^{-\delta (T_i^* - T_j^*)} - e^{\delta T_j^*} \right) \right\}. \qquad (6.1.2)
\end{aligned}
$$

and the probability expression of $P_{N(C)+1}$ is given as follows:

$$\begin{aligned} P_{N(C)+1} &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T^*_{N(C)}}-1\right)+\sum_{j=1}^{N(C)-1}\frac{Y_j}{\delta}\left(e^{-\delta(T^*_{N(C)}-T^*_j)}-e^{\delta T^*_j}\right)\right\} \\ &- \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T^*_{N(C)+1}}-1\right)+\sum_{j=1}^{N(C)}\frac{Y_j}{\delta}\left(e^{-\delta(T^*_{N(C)+1}-T^*_j)}-e^{\delta T^*_j}\right)\right\}. \end{aligned} \tag{6.1.3}$$

6.1.5 Finding the value of summation of the probabilties q_i depending on the interval in which the value of t_i lies

1. Initially, find the value of q_1 assuming that the value of time t_1 lies in any of the interval between 0 and T. Hence, the value of q_1 is given by:

$$\begin{aligned} q_1 &= \sum_{i=1}^{N(C)} P_i \\ &= \sum_{i=1}^{N(C)}\left[\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T^*_{i-1}}-1\right)+\sum_{j=1}^{i-2}\frac{Y_j}{\delta}\left(e^{-\delta(T^*_{i-1}-T^*_j)}-e^{\delta T^*_j}\right)\right\}\right. \\ &- \left.\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T^*_i}-1\right)+\sum_{j=1}^{i-1}\frac{Y_j}{\delta}\left(e^{-\delta(T^*_i-T^*_j)}-e^{\delta T^*_j}\right)\right\}\right]. \end{aligned} \tag{6.1.4}$$

2. Then, find the value of q_2 given time t_1 assuming time t_2 will belong in the time interval $[t_1,\ T_1^*]$. Start with evaluation of the probability in the interval $[t_1,\ T_1^*]$ that is given by:

$$\begin{aligned} P(t_1 < T < T_1^*) &= P(T < T_1^* | T > t_1)P(T > t_1) \\ &= \left[1-\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_1^*}-e^{-\delta t_1}\right)\right\}\right]\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_1}-1\right)\right\} \\ &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_1}-1\right)\right\}-\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_1^*}-1\right)\right\}. \end{aligned}$$

Hence, the value of q_2 is given by:

$$\begin{aligned} q_2 &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_1}-1\right)\right\} - \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_1^*}-1\right)\right\} + \sum_{i=2}^{N(C)} P_i \\ &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_1}-1\right)\right\} - \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_1^*}-1\right)\right\} \\ &+ \sum_{i=2}^{N(C)}\left[\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_{i-1}^*}-1\right)+\sum_{j=1}^{i-2}\frac{Y_j}{\delta}\left(e^{-\delta(T_{i-1}^*-T_j^*)}-e^{\delta T_j^*}\right)\right\}\right. \\ &- \left.\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_i^*}-1\right)+\sum_{j=1}^{i-1}\frac{Y_j}{\delta}\left(e^{-\delta(T_i^*-T_j^*)}-e^{\delta T_j^*}\right)\right\}\right]. \end{aligned}$$

3. Continuing in this way, find the general expression of q_i for $i = 1,\ 2, \ldots,\ N(C)$ is given by:

$$\begin{aligned} q_i &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_{i-1}}-1\right)+\sum_{j=1}^{i-2}\frac{Y_j}{\delta}\left(e^{-\delta(t_{i-1}-T_j^*)}-e^{\delta T_j^*}\right)\right\} \\ &- \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_{i-1}^*}-1\right)+\sum_{j=1}^{i-2}\frac{Y_j}{\delta}\left(e^{-\delta(T_{i-1}^*-T_j^*)}-e^{\delta T_j^*}\right)\right\} \\ &+ \sum_{j=i}^{N(C)}\left[\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_{j-1}^*}-1\right)+\sum_{k=1}^{j-2}\frac{Y_k}{\delta}\left(e^{-\delta(T_{j-1}^*-T_k^*)}-e^{\delta T_k^*}\right)\right\}\right. \\ &- \left.\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_j^*}-1\right)+\sum_{k=1}^{j-1}\frac{Y_k}{\delta}\left(e^{-\delta(T_j^*-T_k^*)}-e^{\delta T_k^*}\right)\right\}\right]. \end{aligned} \tag{6.1.5}$$

6.1.6 Finding the i^{th} jump times from the point process $N(t)$ using an inversion method

1. First, find the value of t_1 belonging to an interval $[0,\ T_1^*]$ and let $P(T > t_1) = 1 - F(t_1)$ that is derived as follows:

$$\begin{aligned} P(T > t_1) &= P(N(0, t_1) = 0) \\ &= \exp\left\{-\int_0^{t_1} \lambda_0\, e^{-\delta s}\, ds\right\} \\ \Longrightarrow 1 - F(t_1) &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_1}-1\right)\right\}. \end{aligned}$$

Using, inversion method find the value of t_1 as follows

$$\begin{aligned}
\log\left(1 - F(t_1)\right) &= \frac{\lambda_0}{\delta}\left(e^{-\delta t_1} - 1\right) \\
t_1 &= -\frac{1}{\delta}\log\left[\frac{\delta}{\lambda_0}\log\left(1 - F(t_1)\right) + 1\right] \\
F^{-1}(U_1) &= -\frac{1}{\delta}\log\left[\frac{\delta}{\lambda_0}\log\left(1 - U_1\right) + 1\right].
\end{aligned}$$

2. Next, find the value of t_2 belonging to an interval $[T_1^*,\ T_2^*]$ and let $P(T > t_2) = 1 - F(t_2)$ that is derived as follows:

$$\begin{aligned}
P(T > t_2) &= P(N(0, t_2) = 0) \\
&= \exp\left\{-\int_0^{t_2}\left(\lambda_0\, e^{-\delta s} + Y_1\, e^{-\delta(s - T_1^*)}\right)\, ds\right\} \\
\Longrightarrow 1 - F(t_2) &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_2} - 1\right) + \frac{Y_1}{\delta}\left(e^{-\delta(t_2 - T_1^*)} - e^{\delta T_1^*}\right)\right\}.
\end{aligned}$$

Using, inversion method find the value of t_2.

$$\begin{aligned}
\log\left(1 - F(t_2)\right) &= \frac{1}{\delta}\left[\lambda_0\left(e^{-\delta t_2} - 1\right) + Y_1\left(e^{-\delta(t_2 - T_1^*)} - e^{\delta T_1^*}\right)\right] \\
t_2 &= -\frac{1}{\delta}\log\left[\frac{\delta\ \log\left(1 - F(t_2)\right) + \lambda_0 + Y_1\, e^{\delta T_1^*}}{\lambda_0 + Y_1\, e^{\delta T_1^*}}\right] \\
F^{-1}(U_2) &= -\frac{1}{\delta}\log\left[\frac{\delta\ \log\left(1 - U_2\right) + \lambda_0 + Y_1\, e^{\delta T_1^*}}{\lambda_0 + Y_1\, e^{\delta T_1^*}}\right].
\end{aligned}$$

3. Therefore, the general expression of the value of t_i for $i = 1,\ 2, \ldots,\ N(C)$ is given

by:

$$
\begin{aligned}
t_i &= -\frac{1}{\delta}\log\left[\frac{\delta\ \log\left(1-F(t_2)\right)+\lambda_0+\sum_{j=1}^{i-1} Y_j\ e^{\delta T_j^*}}{\lambda_0+\sum_{j=1}^{i-1} Y_j\ e^{\delta T_j^*}}\right] \\
F^{-1}(U_i) &= -\frac{1}{\delta}\log\left[\frac{\delta\ \log\left(1-U_i\right)+\lambda_0+\sum_{j=1}^{i-1} Y_j\ e^{\delta T_j^*}}{\lambda_0+\sum_{j=1}^{i-1} Y_j\ e^{\delta T_j^*}}\right]. \qquad (6.1.6)
\end{aligned}
$$

6.2 Simulating a renewal shot-noise Cox process

In this section, there will be proposition of an algorithm for simulating a renewal shot-noise Cox process. Then arrival times of a non-homogeneous Poisson process can be simulated from the renewal shot-noise Cox process in order to generate some arrival times for the renewal shot-noise process. Then the arrival times are aggregated to represent the number of events per time interval $[0,\ T]$.

The algorithm for the generation of simulated data from renewal shot-noise Cox process is thus described as follows:

1. Draw the random number U_i from the uniform distribution on $[0,\ q_i]$ given t_{i-1}. Initially, assume that $t_0 = 0$ to compute q_1 and later continue updating the expression of q_i depending on the interval in which the value of t_i lies, where the general expression derived in Equation (6.1.5) is given by:

$$
\begin{aligned}
q_i &= \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta t_{i-1}}-1\right)+\sum_{j=1}^{i-2}\frac{Y_j}{\delta}\left(e^{-\delta(t_{i-1}-T_j^*)}-e^{\delta T_j^*}\right)\right\} \\
&- \exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_{i-1}^*}-1\right)+\sum_{j=1}^{i-2}\frac{Y_j}{\delta}\left(e^{-\delta(T_{i-1}^*-T_j^*)}-e^{\delta T_j^*}\right)\right\} \\
&+ \sum_{j=i}^{N(C)}\left[\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_{j-1}^*}-1\right)+\sum_{k=1}^{j-2}\frac{Y_k}{\delta}\left(e^{-\delta(T_{j-1}^*-T_k^*)}-e^{\delta T_k^*}\right)\right\}\right. \\
&- \left.\exp\left\{\frac{\lambda_0}{\delta}\left(e^{-\delta T_j^*}-1\right)+\sum_{k=1}^{j-1}\frac{Y_k}{\delta}\left(e^{-\delta(T_j^*-T_k^*)}-e^{\delta T_k^*}\right)\right\}\right].
\end{aligned}
$$

2. Thereafter, obtain a uniform random number U_i and find out the interval in which the random number U_i lies between any of the interval $[Q_{i-1},\ Q_i]$ as a result obtain the interval at which t_i lies with

$$Q_i = \sum_{j=1}^{i} P_j$$

and

$$Q_0 = 0,$$

where

$$\begin{aligned} P_j &= \exp\left\{ \frac{\lambda_0}{\delta}\left(e^{-\delta T^*_{j-1}} - 1\right) + \sum_{k=1}^{j-2} \frac{Y_k}{\delta}\left(e^{-\delta(T^*_{j-1}-T^*_k)} - e^{\delta T^*_k}\right) \right\} \\ &- \exp\left\{ \frac{\lambda_0}{\delta}\left(e^{-\delta T^*_{j}} - 1\right) + \sum_{k=1}^{j-1} \frac{Y_k}{\delta}\left(e^{-\delta(T^*_{j}-T^*_k)} - e^{\delta T^*_k}\right) \right\}. \end{aligned}$$

3. Then, using the general expression derived in the Equation (6.1.6) that is used to simulate T between $[T^*_{i-1},\ T^*_i]$ and the inversion method to get the value of t_i depending on the interval it lies in, where the general expression is given as follows:

$$t_i = F^{-1}(U_i) = -\frac{1}{\delta}\log\left[\frac{\delta\ \log(1-U_i) + \lambda_0 + \sum_{j=1}^{i-1} Y_j\ e^{\delta T^*_j}}{\lambda_0 + \sum_{j=1}^{i-1} Y_j\ e^{\delta T^*_j}}\right].$$

The algorithm for the inversion method is given as follows:

- Draw the uniform random number U_i on $[0, 1]$.
- Compute the value of t_i $i.e.$ $F^{-1}(U_i)$.

and go back to step 1 to repeat the entire algorithm several times till time T.

The above algorithm generates the jump times t_i for $i = 1,\ 2, \ldots,\ N(C)$ of renewal shot-noise Cox process.

In this chapter, the proposed algorithm for simulating the renewal shot-noise Cox process will be presented. Further, the proposed algorithm is structured in such a way that it allows for the time-dependent intensity function to be a stochastic process. As described in the previous chapters this increases reliability of the estimates in insurance business.

7 RESULTS AND DISCUSSION

This chapter summarizes the results together with the deep discussion of the results discovered using both statistical techniques as illustrated in the Chapter (3) and Chapter (4) with the help of *MATLAB* software. Later, we will then accomplish the study by comparing the results from both statistical techniques using the algorithm built in Chapter (6) to simulate the point process $N(t)$ with intensity of the renewal shot-noise Cox process and generate the jump times at which future catastrophic events will happen.

This involves generation of different parameters using the sample real data from an insurance company in Poland as the result of occurrence of catastrophic events over the time period from 7.1.1990 to 10.12.1992 and later on use the generated parameter values to generate their respective inter-arrival times and claim sizes that will be useful in the algorithm to simulate the point process $N(t)$ with intensity of the renewal shot-noise Cox process to get reasonable results of jump times.

7.1 The maximum likelihood estimators

This is one of the statistical techniques that is used to generate fixed parameter values from different distributions that are important in the simulation of the point process $N(t)$ with intensity of the renewal shot-noise Cox process.

We generate the claim sizes using the maximum likelihood estimators of the log-normal distribution $i.e.$ the mean parameter, $\hat{\mu}$ and the variance parameter, $\hat{\sigma}^2$ as derived in Equation (3.3.1) and Equation (3.3.2) respectively, we get the mean parameter, $\hat{\mu} = 18.0203$ and the variance parameter, $\hat{\sigma}^2 = 1.5963$.

However, in the generation of the inter-arrival times we proposed three arbitrary distributions $i.e.$ exponential, gamma and Weibull. First, using the maximum likelihood estimator of the exponential distribution $i.e$ the rate parameter, $\hat{\alpha}$ as derived in Equation (3.4.1) to obtain the rate parameter, $\hat{\alpha} = 0.0974$. Then, using the maximum likelihood estimators of the gamma distribution $i.e$ the shape parameter, $\hat{k}$ and the scale parameter, $\hat{\beta}$ as derived in Equation (3.5.2) and Equation (3.5.1) respectively, we get the shape parameter, $\hat{k} = 0.0054$ and use the result of the estimated shape parameter to obtain the scale parameter, $\hat{\beta} = 1.8861e + 03$. Lastly, using the maximum likelihood estimators of the Weibull distribution $i.e.$ the shape parameter, $\hat{\gamma}$ and the scale parameter, $\hat{\eta}$ as derived in Equation

(3.6.2) and Equation (3.6.3) respectively, we get the shape parameter, $\hat{\gamma} = 0.0054$ and use the result of the estimated shape parameter to obtain the scale parameter, $\hat{\eta} = 0$.

Using the generated claim sizes from the maximum likelihood estimators of the log-normal distribution and generated inter-arrival times from the maximum likelihood estimators of three different arbitrary distributions $i.e.$ exponential, gamma and Weibull to obtain the initial intensity, $\hat{\lambda_0}$ and exponential rate of decay, $\hat{\delta}$ as derived from their derivatives of the log-likelihood function in Equation (3.2.6) and Equation (3.2.7) respectively with the help of Newton Raphson method. First, using the inter-arrival times that are assumed to be exponentially distributed, we get the numerical values of initial intensity, $\hat{\lambda_0} = 3.8548$ and use the result of the estimated initial intensity to obtain the exponential rate of decay, $\hat{\delta} = 9.1621e - 06$. Then, using the inter-arrival times that are assumed to be gamma distributed, we get the numerical values of initial intensity, $\hat{\lambda_0} = 113.2740$ and use the result of the estimated initial intensity to obtain the exponential rate of decay, $\hat{\delta} = 8.9878e - 06$. Lastly, using the inter-arrival times that are assumed to be Weibull distributed, we get the numerical values of initial intensity, $\hat{\lambda_0} = 115.7068$ and use the result of the estimated initial intensity to obtain the exponential rate of decay, $\hat{\delta} = 8.9904e - 06$.

7.2 The Bayesian estimators

This is another statistical techniques that is used to generate random parameter values using MCMC method from different posterior distributions that are important in the simulation of the point process $N(t)$ with intensity of the renewal shot-noise Cox process.

Figure 4 shows the sample paths of the MCMC runs of 10000 realizations for both mean parameter and precision parameter respectively, where jump sizes follow log-normal distribution. The sample paths show that the mixing of the chain is optimal after 10000 realizations.

Here, we generate the claim sizes using the Bayesian estimators of the log-normal distribution $i.e.$ the mean parameter, $\hat{\mu}$ and the precision parameter, $\hat{\Phi}$ as derived from their posterior distributions in Equation (4.3.4) and Equation (4.3.8) respectively with the help of MCMC method using Gibbs sampling algorithm, we get the mean parameter, $\hat{\mu} = 1.4478$ and the precision parameter, $\hat{\Phi} = 0.0043$ from their respective mean values of the MCMC chain parameter values after 10000 realizations.

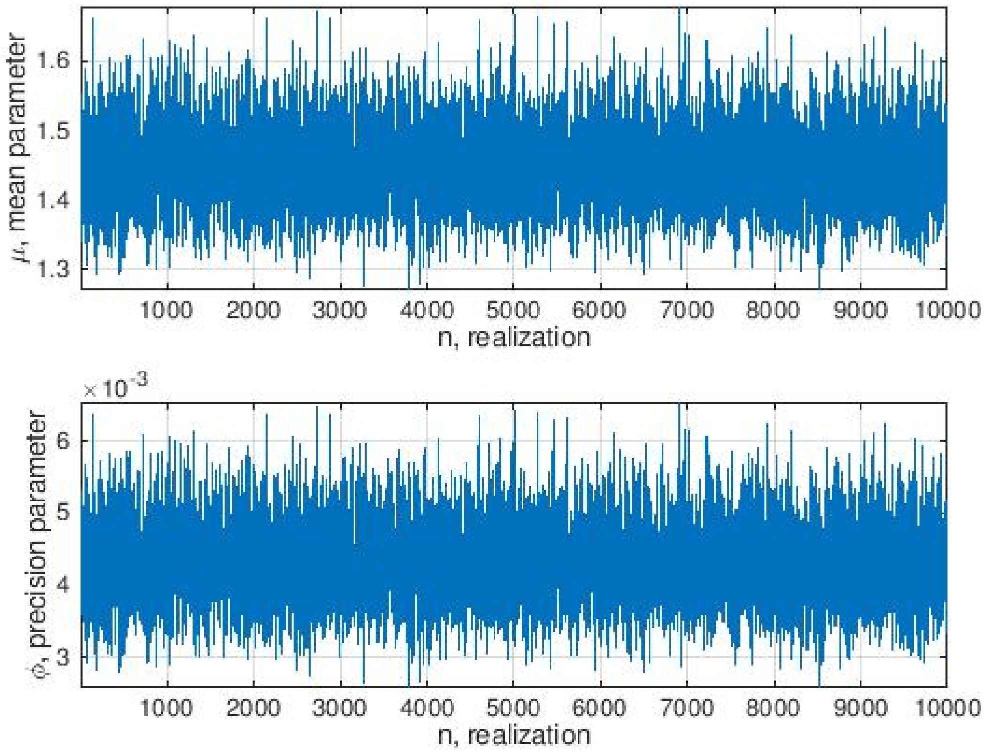

Figure 4. Sample paths of the MCMC runs for both mean parameter and precision parameter, where jump sizes follow log-normal distribution.

Figure 5 shows the posterior distribution of the parameters after 10000 realizations for both mean parameter and precision parameter, where jump sizes follow log-normal distribution. This shows that there is zero correlation between these parameters after 10000 realizations $i.e.$ non-existence of relationship between parameters.

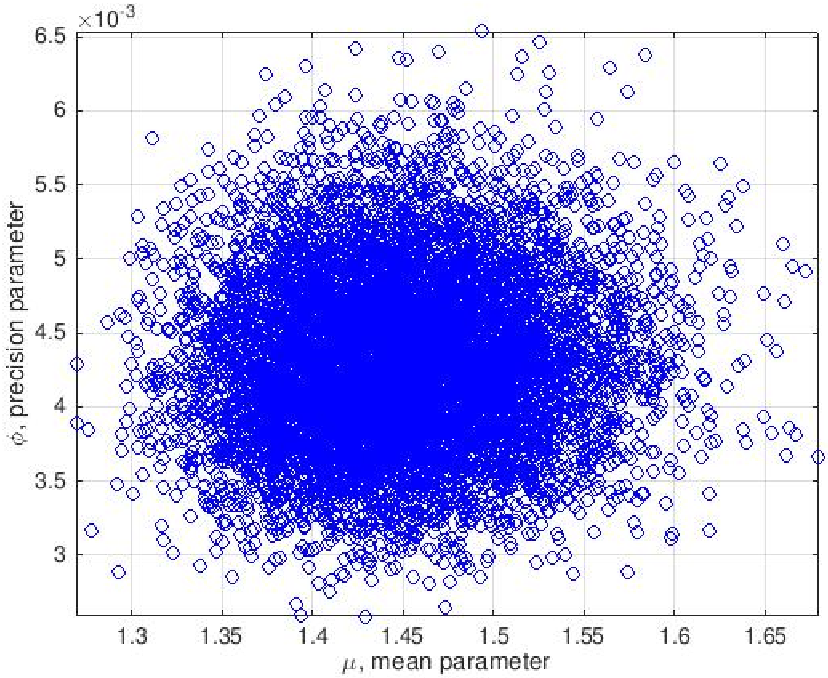

Figure 5. Posterior distribution of the parameters, where jump sizes follow log-normal distribution.

Figure 6 shows the sample paths of the MCMC runs of 10000 realizations for rate parameter, where inter-arrival times follow exponential distribution. The sample paths show that the mixing of the chain is somewhat not optimal after 10000 realizations.

Here, we generate the inter-arrival times using the Bayesian estimators of the exponential distribution $i.e.$ the rate parameter, $\hat{\lambda}$ as derived from its posterior distribution in Equation (4.4.4) with the help of MCMC method using random walk Metropolis algorithm, we get the rate parameter, $\hat{\lambda} = 0.1068$ from its mean value of the MCMC chain parameter values after 10000 realizations.

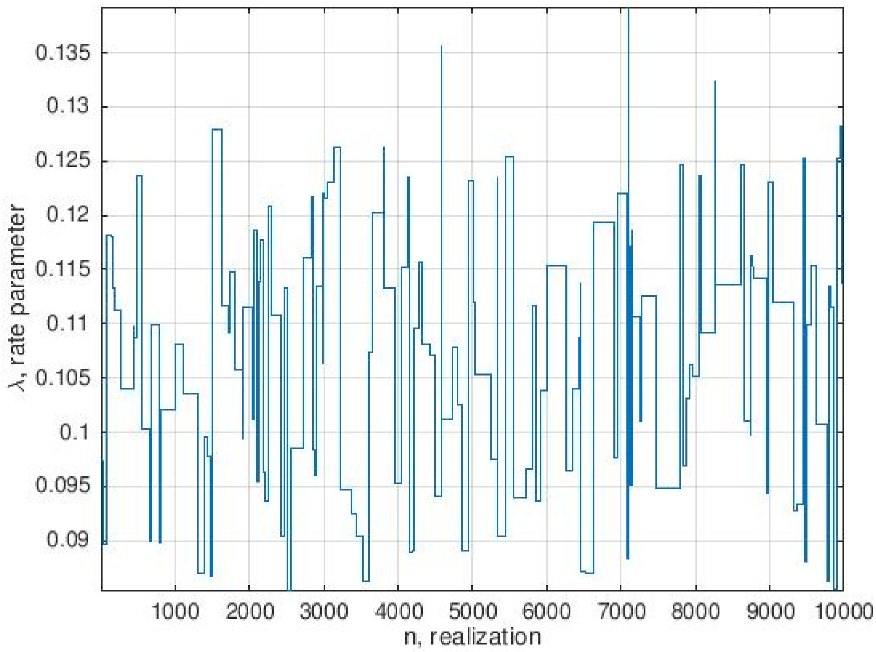

Figure 6. Sample paths of the MCMC runs for rate parameter, where inter-arrival times follow exponential distribution.

Figure 7 shows the sample paths of the MCMC runs of 10000 realizations for both scale parameter and shape parameter respectively, where inter-arrival times follow gamma distribution. The sample paths show that the mixing of the chain is optimal after 10000 realizations.

Here, we generate the inter-arrival times using the Bayesian estimators of the gamma distribution $i.e.$ the scale parameter, $\hat{s}$ and the shape parameter, $\hat{k}$ as derived from their posterior distribution in Equation (4.5.6) and Equation (4.5.5) respectively with the help of MCMC method using both Gibbs sampling and random walk Metropolis algorithm, we get the scale parameter, $\hat{s} = 5.1600$ and the shape parameter, $\hat{k} = 1.6201$ from their respective mean values of the MCMC chain parameter values after 10000 realizations.

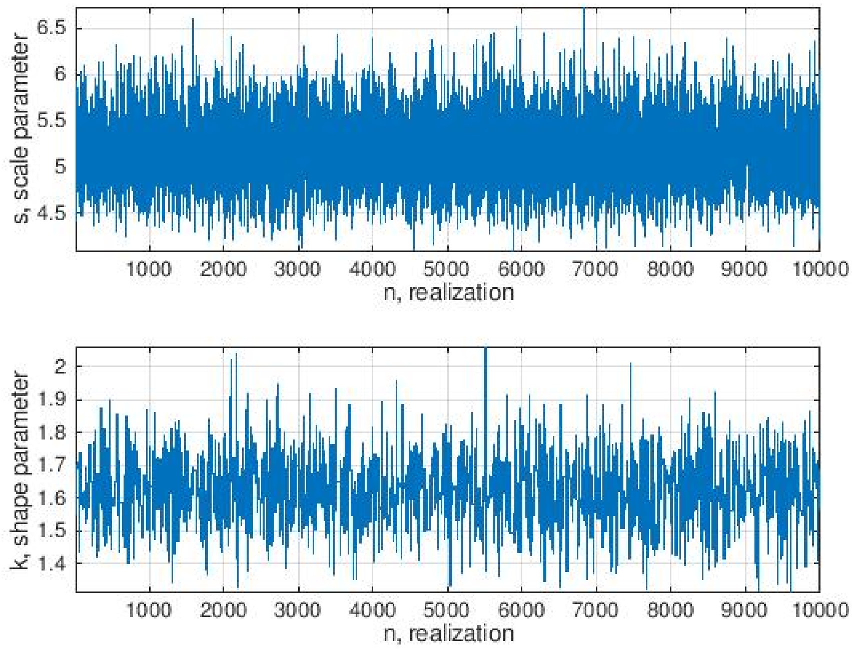

Figure 7. Sample paths of the MCMC runs for both scale parameter and shape parameter, where inter-arrival times follow gamma distribution.

Figure 8 shows the posterior distribution of the parameters after 10000 realizations for both scale parameter and shape parameter, inter-arrival times follow gamma distribution. This shows that there is zero correlation between these parameters after 10000 realizations $i.e.$ non-existence of relationship between parameters.

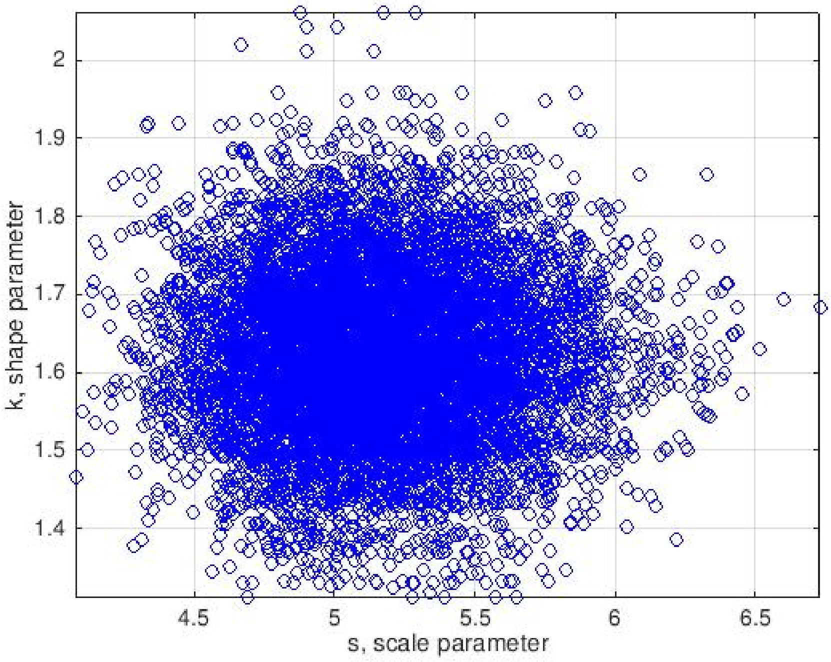

Figure 8. Posterior distribution of the parameters, where inter-arrival times follow gamma distribution.

Figure 9 shows the sample paths of the MCMC runs of 10000 realizations for both scale parameter and shape parameter respectively, where inter-arrival times follow Weibull distribution. The sample paths show that the mixing of the chain is not optimal after 10000 realizations.

Here, we generate the inter-arrival times using the Bayesian estimators of the Weibull distribution $i.e.$ the scale parameter, $\hat{\eta}$ and the shape parameter, $\hat{\gamma}$ as derived from their posterior distribution in Equation (4.6.6) and Equation (4.6.5) respectively with the help of MCMC method using random walk Metropolis algorithm, we get the scale parameter, $\hat{\eta} = 4.8550$ and the shape parameter, $\hat{\gamma} = 1.7932$ from their respective mean values of the MCMC chain parameter values after 10000 realizations.

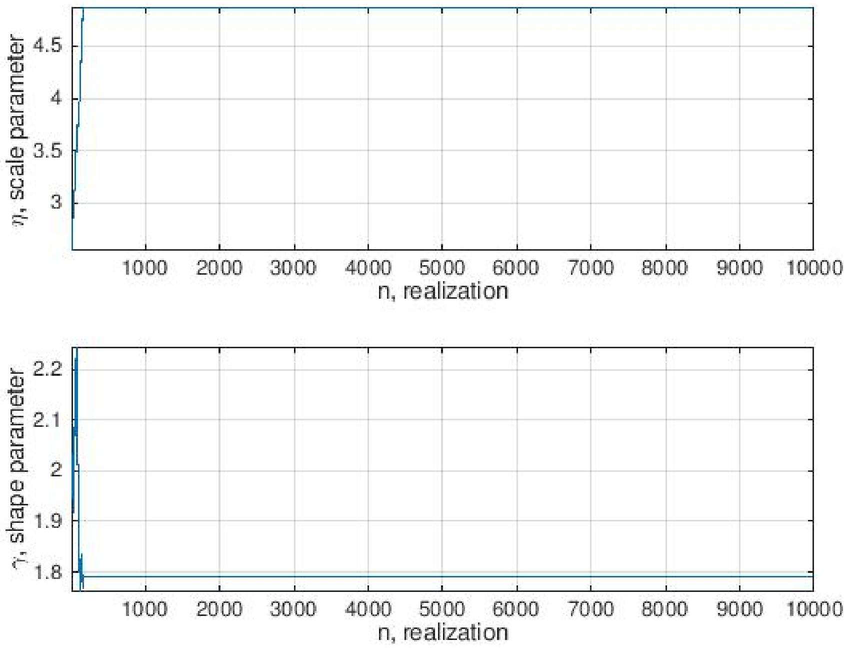

Figure 9. Sample paths of the MCMC runs for both scale parameter and shape parameter, where inter-arrival times follow Weibull distribution.

Figure 10 shows the posterior distribution of the parameters after 10000 realizations for both scale parameter and shape parameter, inter-arrival times follow Weibull distribution. This shows that there is zero correlation between these parameters after 10000 realizations $i.e.$ non-existence of relationship between parameters.

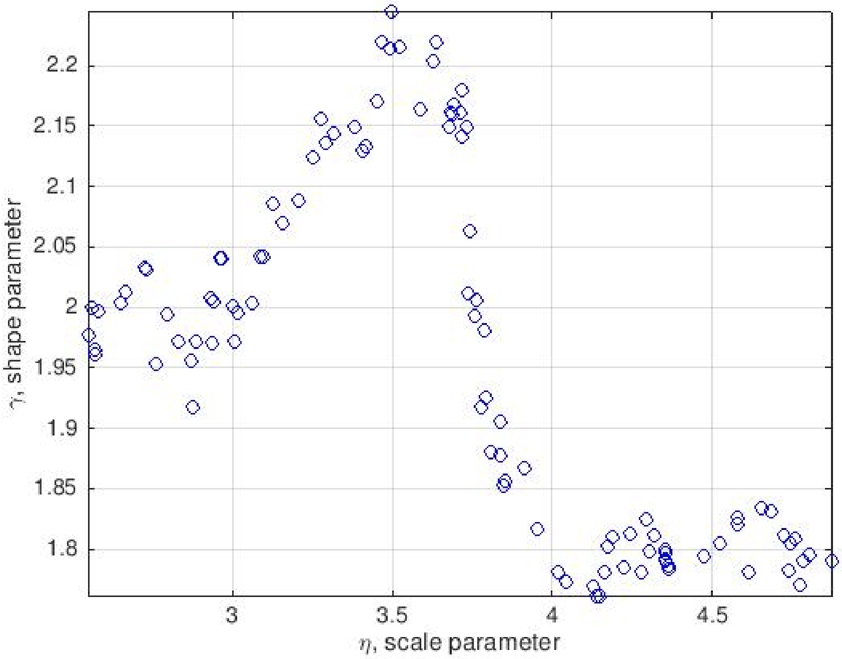

Figure 10. Posterior distribution of the parameters, where inter-arrival times follow Weibull distribution.

Figure 11 shows the sample paths of the MCMC runs of 10000 realizations for the initial intensity parameter. The sample paths show that the mixing of the chain is optimal after 10000 realizations.

Here, we generate the initial intensity parameter using the posterior distribution as derived in Equation (4.2.8) with the help of MCMC method using random walk Metropolis algorithm, we get the initial intensity parameter, $\lambda_0 = 4$ from its respective mean value of the MCMC chain parameter values after 10000 realizations.

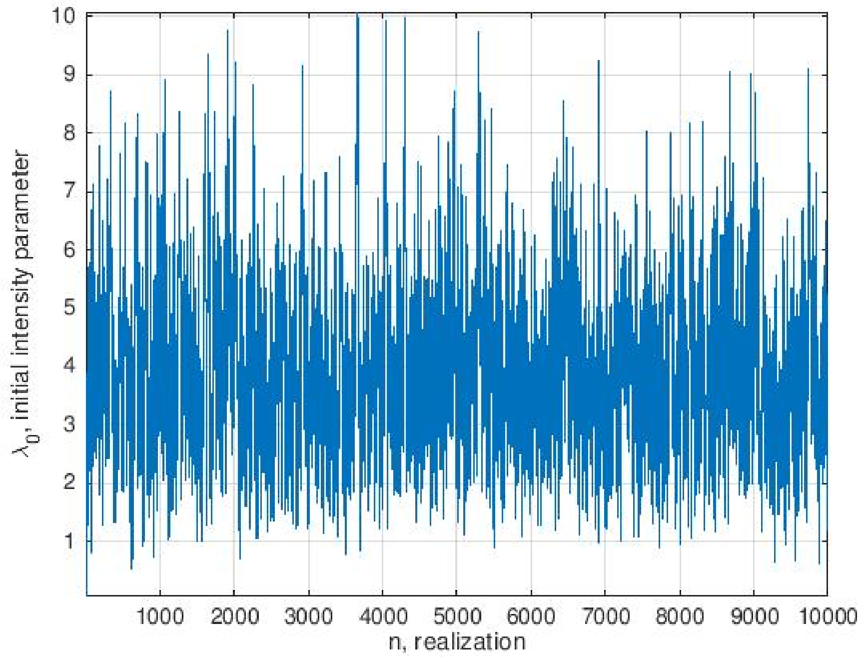

Figure 11. Sample paths of the MCMC runs for the initial intensity parameter.

Figure 12 shows the sample paths of the MCMC runs of 10000 realizations for the exponential rate of decay parameter. The sample paths show that the mixing of the chain is optimal after 10000 realizations.

Here, we generate the exponential rate of decay parameter using the posterior distribution as derived in Equation (4.2.12) with the help of MCMC method using random walk Metropolis algorithm, we get the exponential rate of decay parameter, $\delta = 0.3555$ from its respective mean value of the MCMC chain parameter values after 10000 realizations, where the inter-arrival times are assumed to be exponentially distributed.

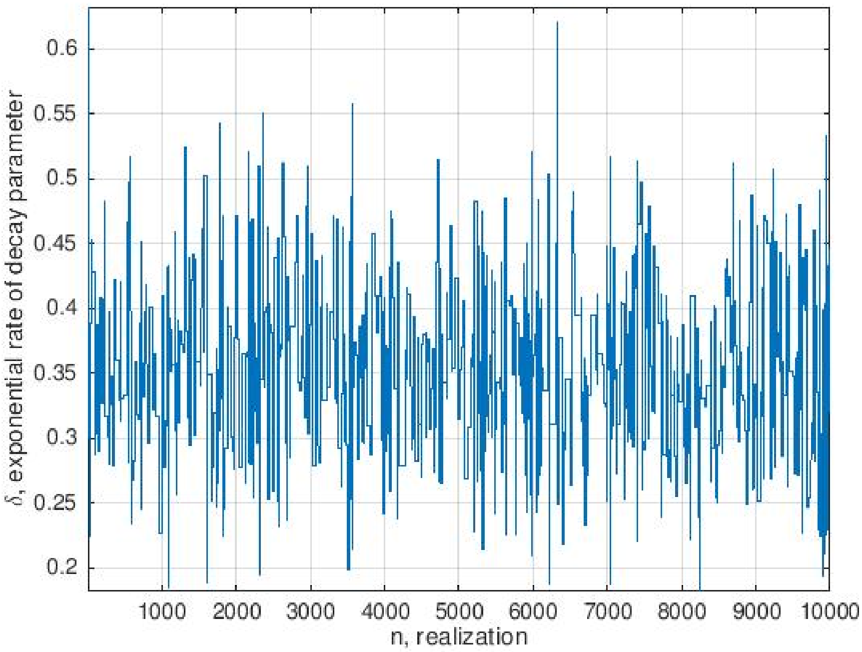

Figure 12. Sample paths of the MCMC runs for the exponential rate of decay parameter, where inter-arrival times follow exponential distribution.

Figure 13 shows the posterior distribution of the parameters after 10000 realizations for both initial intensity parameter and exponential rate of decay parameter, where inter-arrival times are assumed to be exponentially distributed. This shows that there is positive correlation between these parameters after 10000 realizations $i.e.$ existence of positive relationship between parameters. The possibility of such behavior could be due to the reason that the prior distribution of both initial intensity parameter and exponential rate of decay parameter are assumed to follow gamma distribution in Bayesian estimation method as the result they behave in a similar manner.

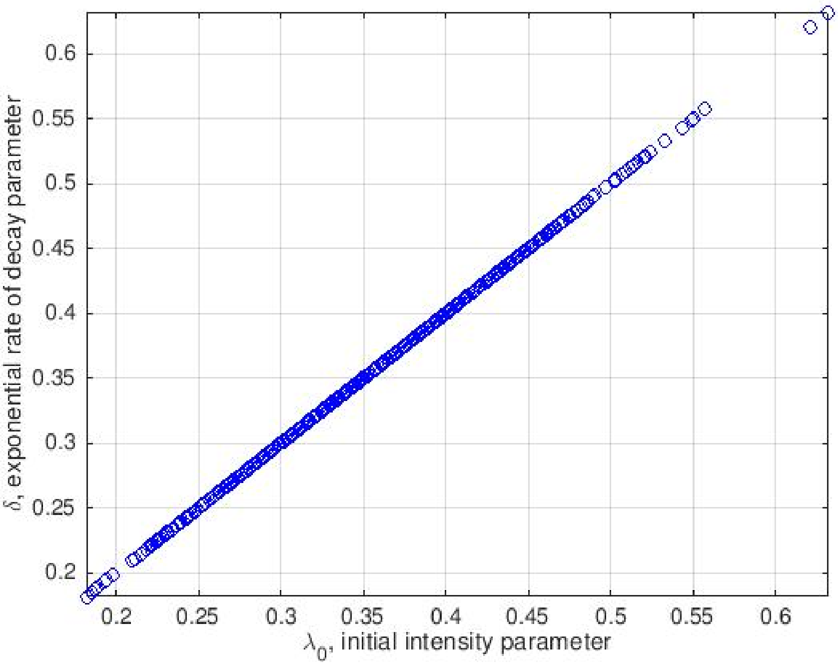

Figure 13. Posterior distribution of the parameters, where inter-arrival times follow exponential distribution.

Figure 14 shows the sample paths of the MCMC runs of 10000 realizations for the exponential rate of decay parameter. The sample paths show that the mixing of the chain is somewhat not optimal after 10000 realizations.

Here, we generate the exponential rate of decay parameter using the posterior distribution as derived in Equation (4.2.12) with the help of MCMC method using random walk Metropolis algorithm, we get the exponential rate of decay parameter, $\delta = 0.1364$ from its respective mean value of the MCMC chain parameter values after 10000 realizations, where the inter-arrival times are assumed to be gamma distributed.

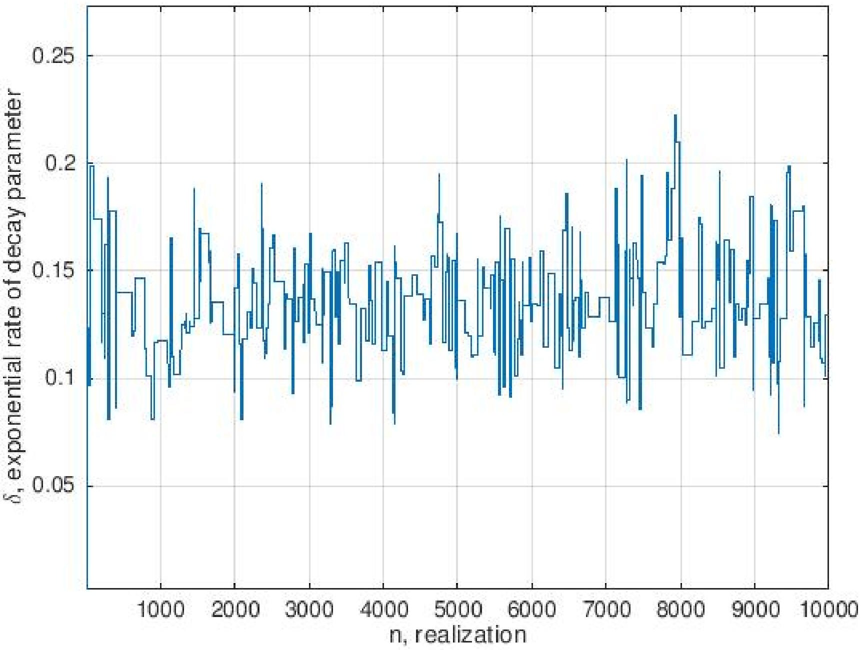

Figure 14. Sample paths of the MCMC runs for the exponential rate of decay parameter, where inter-arrival times follow gamma distribution.

Figure 15 shows the posterior distribution of the parameters after 10000 realizations for both initial intensity parameter and exponential rate of decay parameter, where inter-arrival times are assumed to be gamma distributed. This shows that there is positive correlation between these parameters after 10000 realizations $i.e.$ existence of positive relationship between parameters. The possibility of such behaviour could be due to the reason that the prior distribution of both initial intensity parameter and exponential rate of decay parameter are assumed to follow gamma distribution in Bayesian estimation method as the result they behave in a similar manner.

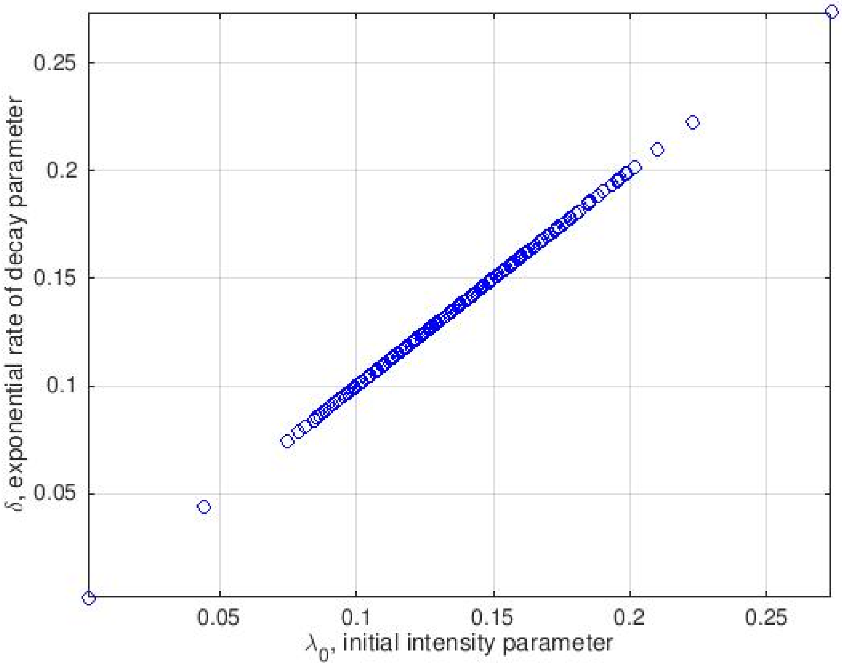

Figure 15. Posterior distribution of the parameters, where inter-arrival times follow gamma distribution.

Figure 16 shows the sample paths of the MCMC runs of 10000 realizations for the exponential rate of decay parameter. The sample paths show that the mixing of the chain is somewhat not optimal after 10000 realizations.

Here, we generate the exponential rate of decay parameter using the posterior distribution as derived in Equation (4.2.12) with the help of MCMC method using random walk Metropolis algorithm, we get the exponential rate of decay parameter, $\delta = 0.1453$ from its respective mean value of the MCMC chain parameter values after 10000 realizations, where the inter-arrival times are assumed to be Weibull distributed.

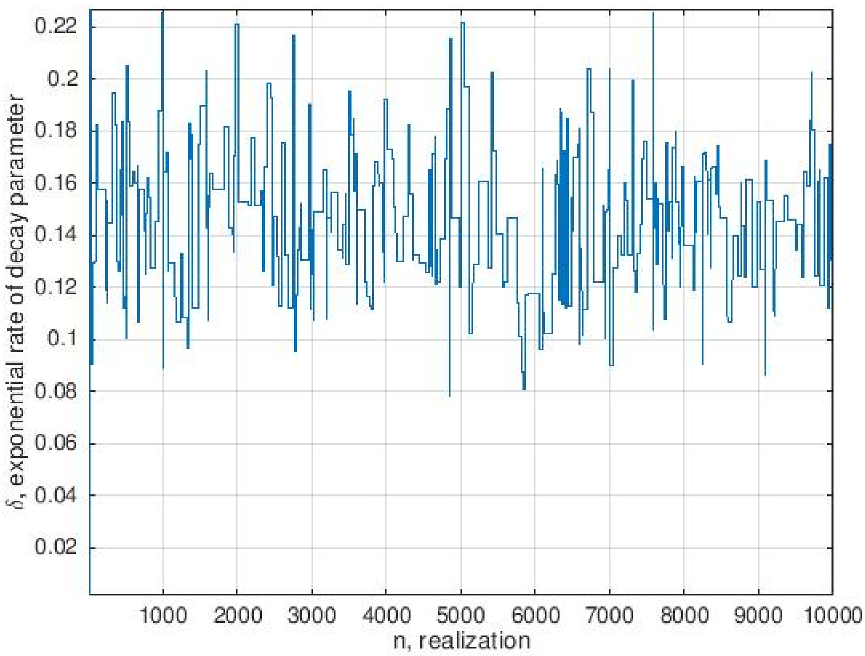

Figure 16. Sample paths of the MCMC runs for the exponential rate of decay parameter, , where inter-arrival times follow Weibull distribution.

Figure 17 shows the posterior distribution of the parameters after 10000 realizations for both initial intensity parameter and exponential rate of decay parameter, where inter-arrival times are assumed to be Weibull distributed. This shows that there is positive correlation between these parameters after 10000 realizations $i.e.$ existence of positive relationship between parameters. The possibility of such behaviour could be due to the reason that the prior distribution of both initial intensity parameter and exponential rate of decay parameter are assumed to follow gamma distribution in Bayesian estimation method as the result they behave in a similar manner.

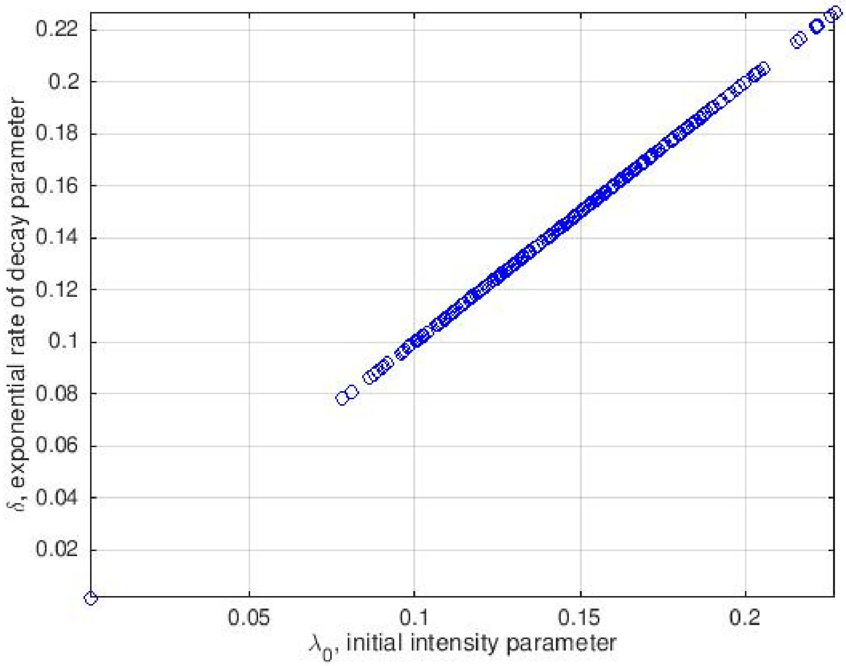

Figure 17. Posterior distribution of the parameters, where inter-arrival times follow Weibull distribution.

7.3 The jump times, t_i

The main purpose of this thesis is to generate the jump times t_i for $i = 1,\ 2, \ldots,\ N(C)$ of renewal shot-noise Cox process using the algorithm derived in Chapter (6) in the simulation of the point process $N(t)$ with intensity of the renewal shot-noise Cox process using the statistical techniques $i.e.$ maximum likelihood estimation method and Bayesian estimation method for results comparison.

Note that we will use the estimators from both statistical techniques $i.e.$ maximum likelihood estimation method and Bayesian estimation method that will assist in the simulation process using the formulated algorithm as stated in Section (6.2).

Figure 18 shows the sample paths by comparing the jump times generated using the main algorithm formulated in Section (6.2) to the arrival times generated using exponential distribution with both statistical techniques $i.e.$ maximum likelihood estimation method and Bayesian estimation method over the entire period of the occurrence of catastrophic events.

Here, we observe that the predicted jump times generated using maximum likelihood estimation method are somewhat far apart from the arrival times of the actual occurrence of catastrophic events generated using real data as compared to the ones generated using Bayesian estimation method, where inter-arrival times follow exponential distribution. This shows that there is greater accuracy of the predicted jump times generated using the Bayesian estimation method to the ones generated using maximum likelihood estimation method.

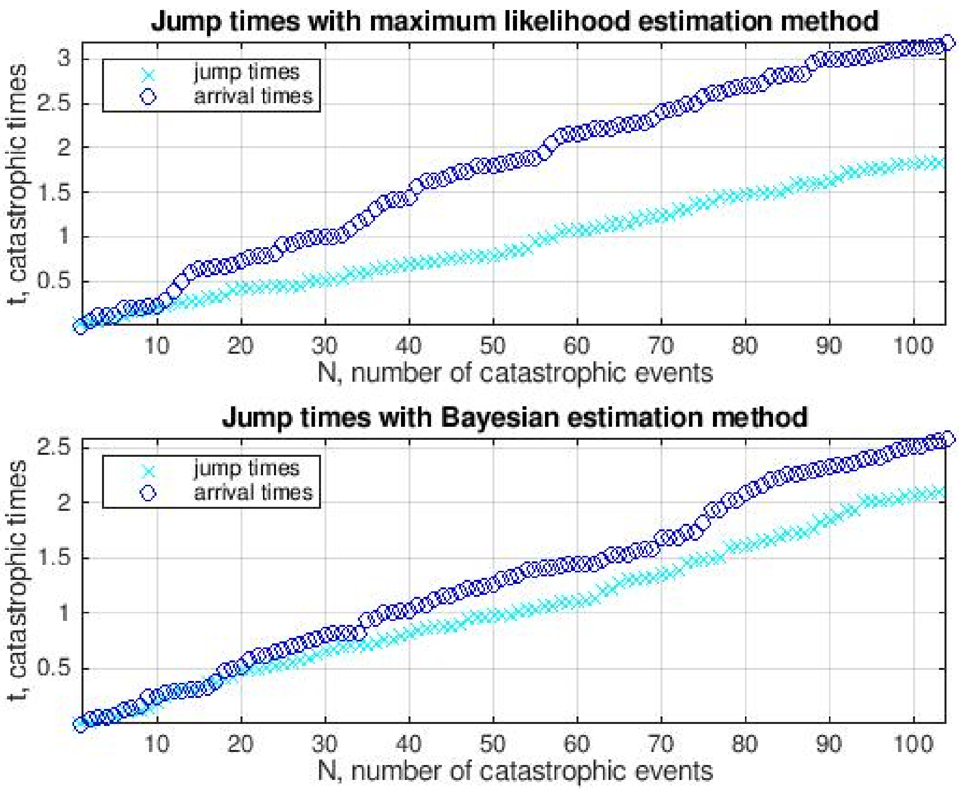

Figure 18. Sample paths of the catastrophic times for both maximum likelihood estimation method and Bayesian estimation method, where inter-arrival times follow exponential distribution.

Figure 19 show the sample paths by comparing the jump times generated using the main algorithm formulated in Section (6.2) to the arrival times generated using gamma distribution with both statistical techniques *i.e.* maximum likelihood estimation method and Bayesian estimation method over the entire period of the occurrence of catastrophic events.

Here, we observe that the predicted jump times generated using maximum likelihood estimation method are completely far apart from the arrival times of the actual occurrence of catastrophic events generated using real data as compared to the ones generated using Bayesian estimation method in which the predicted jump times are somewhat close to the arrival times of the actual occurrence of catastrophic events, where inter-arrival times follow gamma distribution. This shows that there is greater accuracy of the predicted jump times generated using the Bayesian estimation method to the ones generated using maximum likelihood estimation method.

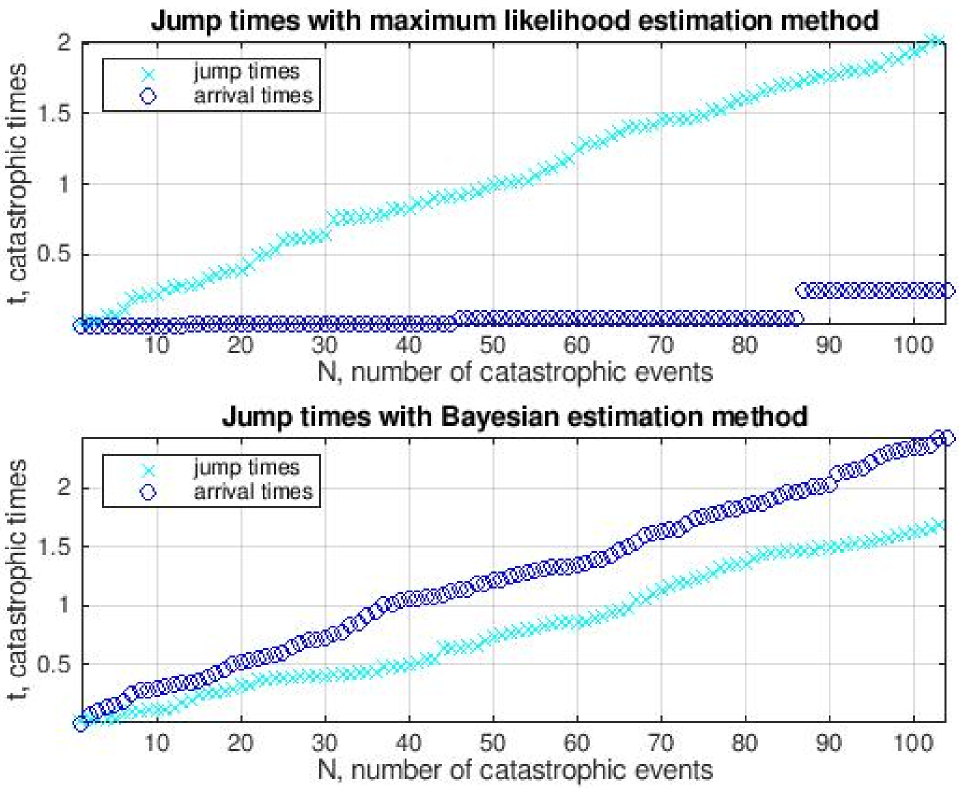

Figure 19. Sample paths of the catastrophic times for both maximum likelihood estimation method and Bayesian estimation method, where inter-arrival times follow gamma distribution.

Figure 20 shows the sample paths by comparing the jump times generated using the main algorithm formulated in Section (6.2) to the arrival times generated using Weibull distribution with both statistical techniques $i.e.$ maximum likelihood estimation method and Bayesian estimation method over the entire period of the occurrence of catastrophic events.

Here, we observe that the predicted jump times generated using maximum likelihood estimation method are completely far apart from the arrival times of the actual occurrence of catastrophic events generated using real data as compared to the ones generated using Bayesian estimation method in which the predicted jump times are very close to the arrival times of the actual occurrence of catastrophic events, where inter-arrival times follow Weibull distribution. This shows that there is greater accuracy of the predicted jump times generated using the Bayesian estimation method to the ones generated using maximum likelihood estimation method.

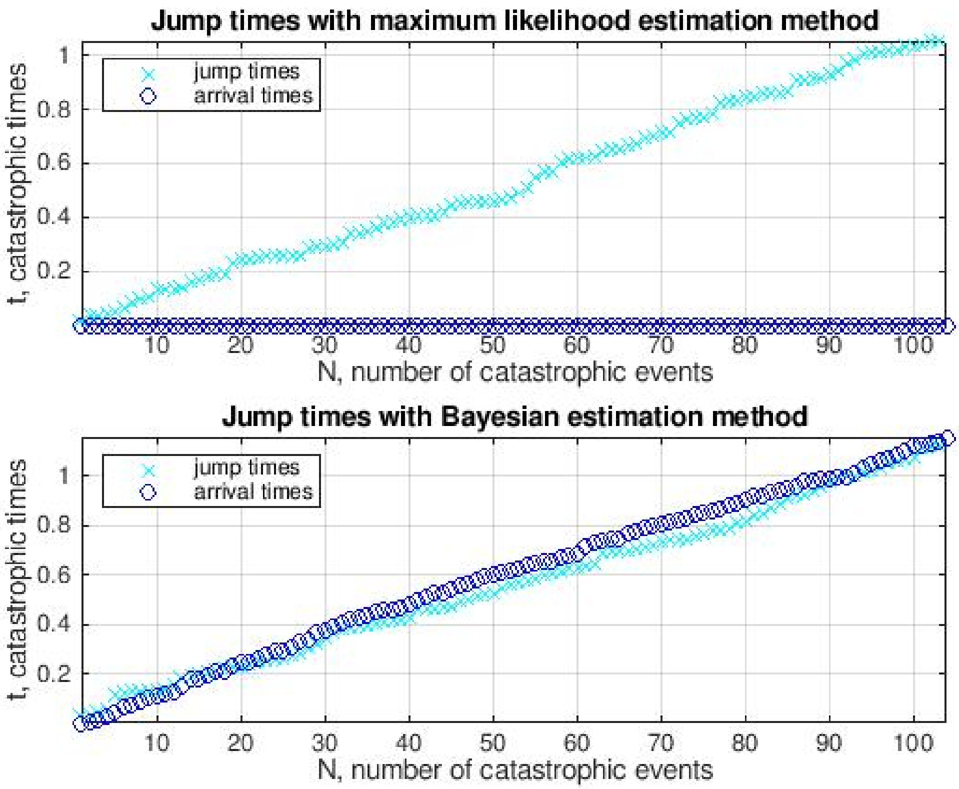

Figure 20. Sample paths of the catastrophic times for both maximum likelihood estimation method and Bayesian estimation method, where inter-arrival times follow Weibull distribution.

Generally, we have observed that the Bayesian estimation method seems to have generated the predicted jump times that appear to be more accurate as compared to the ones generated using maximum likelihood estimation method. Also, we have observed that in the maximum likelihood estimation method that there is small divergence of predicted jump times to the arrival times of the actual occurrence of catastrophic events when the inter-arrival times are assumed to follow exponential distribution as compared to the remaining arbitrary distributions.

8 CONCLUSION AND RECOMMENDATION

In this chapter, we provide the conclusion and recommendation of this study.

8.1 Conclusion

In this project, we explored the use of renewal shot-noise Cox process in the modelling of claim arrivals process as it measures frequency, magnitude and time period to determine the effects of catastrophic events through shot-noise process, such a model also allows the time dependent intensity function to be a stochastic process. Although the renewal shot-noise Cox process has been an important tool in modelling claim arrivals process, in this project we included renewal shot-noise Cox process in modelling catastrophic events taking into consideration the frequency of claim counts.

We have derived estimators for the parameters in the intensity function of renewal shot-noise Cox process and other arbitrary distributions. In this case, inter-arrival times are assumed to follow gamma, Weibull and exponential distribution while claim sizes are assumed to follow log-normal distribution using different statistical techniques $i.e.$ maximum likelihood estimation method and Bayesian estimation method for comparison purposes. Through a simulation study, we have proposed an algorithm of simulating a renewal shot-noise process $\lambda(t)$ of the point process $N(t)$ to generate the jump times at which future catastrophic events will happen.

Basing on the results of the simulated jump times generated with the help of *MATLAB* software, we have drawn a conclusion that the Bayesian estimation method proves to be a more efficient statistical technique in comparison to the maximum likelihood estimation method.

8.2 Recommendation

We recommend that the insurance companies should think about putting into practice the implementation of the algorithm built in Chapter (6) to simulate the point process $N(t)$ with intensity of the renewal shot-noise Cox process and generate the appropriate jump times at which future catastrophic events will happen using Bayesian estimation method. Hence, result to the improvement in a lot of insurance companies because it will reduce ruin probability as it provides flexibility of intensity not only depending on time but allows it to be a stochastic process. Therefore, it is very useful in determination of optimum premium, enabling them to charge the policyholder enough premium such that there is sufficiently small ruin probability.

REFERENCES

[1] George E Rejda. *Principles of risk management and insurance*. Pearson Education India, 2011.

[2] Tanzania -9 Dead as Flooding Hits Dar Es Salaam. URL, http://floodlist.com/africa/tanzania-floods-dar-es-salaam-april-2018, Accessed April 2018.

[3] Chung-Yu Liu. *Claim count modelling with shot noise Cox processes*. PhD thesis, Thesis, University of New South Wales, 2012.

[4] Dimitri P Bertsekas and John N Tsitsiklis. *Introduction to probability*, volume 1. Athena Scientific Belmont, MA, 2002.

[5] Angelos Dassios, Jiwook Jang, and Hongbiao Zhao. A risk model with renewal shot-noise Cox process. *Insurance: Mathematics and Economics*, 65:55–65, 2015.

[6] Angelos Dassios and Ji-Wook Jang. Pricing of catastrophe reinsurance and derivatives using the Cox process with shot-noise intensity. *Finance and Stochastics*, 7(1):73–95, 2003.

[7] Goyal Babita. Renewal Theory and Processes. 2012.

[8] Richard Serfozo. Renewal and Regenerative Processes. In *Basics of Applied Stochastic Processes*, pages 99–167. Springer, 2009.

[9] David R Cox. Some statistical methods connected with series of events. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 129–164, 1955.

[10] Jesper Møller. Shot noise Cox processes. *Advances in Applied Probability*, 35(3):614–640, 2003.

[11] Zailei Cheng and Youngsoo Seol. Precise deviations for Cox processes with shot noise. *arXiv preprint arXiv:1706.07864*, 2017.

[12] Jiwook Jang, Angelos Dassios, and Hongbiao Zhao. Moments of renewal shot-noise processes and their applications. *Scandinavian Actuarial Journal*, pages 1–26, 2018.

[13] David Roxbee Cox and Valerie Isham. *Point processes*, volume 12. CRC Press, 1980.

[14] Jesper Møller and Giovanni Luca Torrisi. Generalised shot noise Cox processes. *Advances in Applied Probability*, 37(1):48–74, 2005.

[15] Brenda Faith Ginos. Parameter estimation for the lognormal distribution. 2009.

[16] Yuanda Chen. Thinning Algorithms for Simulating Point Processes. 2016.

[17] Patrick J Laub, Thomas Taimre, and Philip K Pollett. Hawkes processes. *arXiv preprint arXiv:1507.02822*, 2015.

[18] Fabrizio Ruggeri, Refik Soyer, and CNR IMATI. Advances in bayesian software reliability modelling. *Advances in Mathematical Modelling for Reliability*, pages 149–157, 2008.

[19] Daniel Fink. A compendium of conjugate priors. *See http://www. people. cornell. edu/pages/df36/CONJINTRnew% 20TEX. pdf*, 46, 1997.

[20] Colton Gearhart and D Kasturiratna. Implementation of gibbs sampling within bayesian inference and its applications in actuarial science. *SIAM Journal*, pages 219–231, 2018.

[21] Young Sook Son and Mira Oh. Bayesian estimation of the two-parameter gamma distribution. *Communications in Statistics-Simulation and Computation*, 35(2):285–293, 2006.

[22] Arto Luoma. Introduction to bayesian analysis. *University of Tampere, Finland*, 2014.

[23] Andrew Gelman, Hal S Stern, John B Carlin, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 2013.

[24] Erlandson F Saraiva and Adriano K Suzuki. Bayesian computational methods for estimation of two-parameters weibull distribution in presence of right-censored data. *Chilean Journal of Statistics (ChJS)*, 8(2), 2017.

[25] S. Sharma. Markov Chain Monte Carlo methods for Bayesian data analysis in astronomy. *Annual Review of Astronomy and Astrophysics*, pages 213–259, 2017.

[26] Sheldon M Ross. Stochastic processes. 1996, 1996.

Appendix 1. Histogram plots of the sample paths of MCMC runs

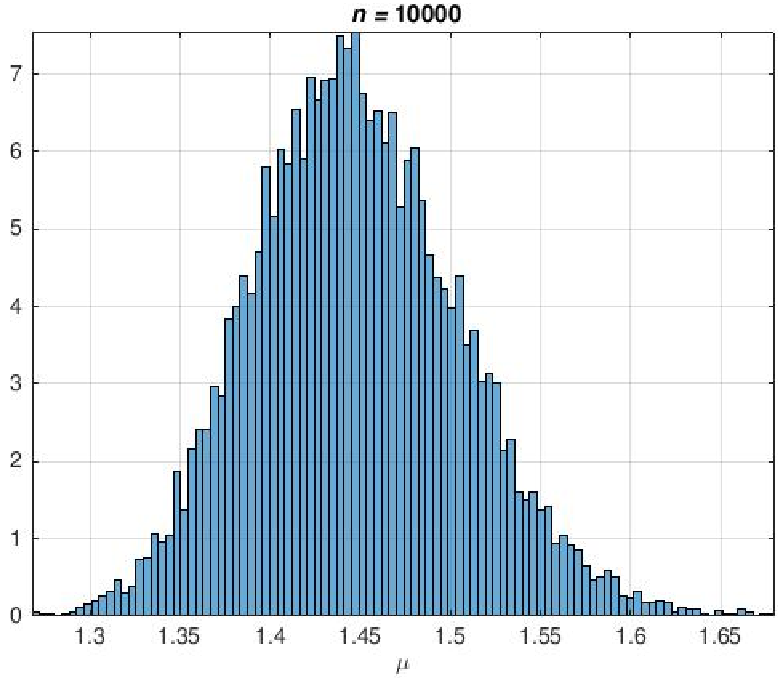

Figure A1.1. Histogram plot of the sample paths of MCMC runs for the mean parameter, where jump sizes follow log-normal distribution.

(continues)

Appendix 1. (continued)

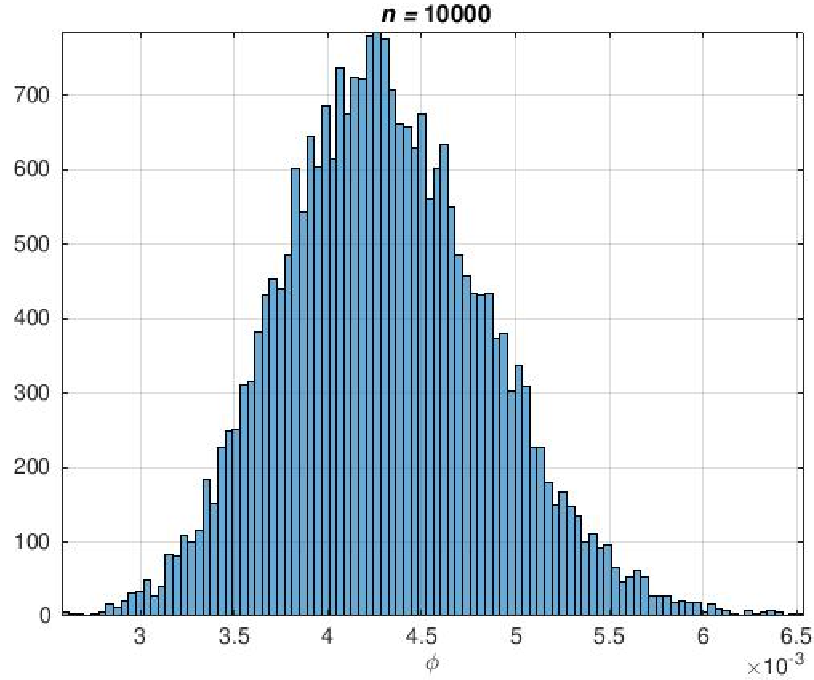

Figure A1.2. Histogram plot of the sample paths of MCMC runs for the precision parameter, where jump sizes follow log-normal distribution.

(continues)

Appendix 1. (continued)

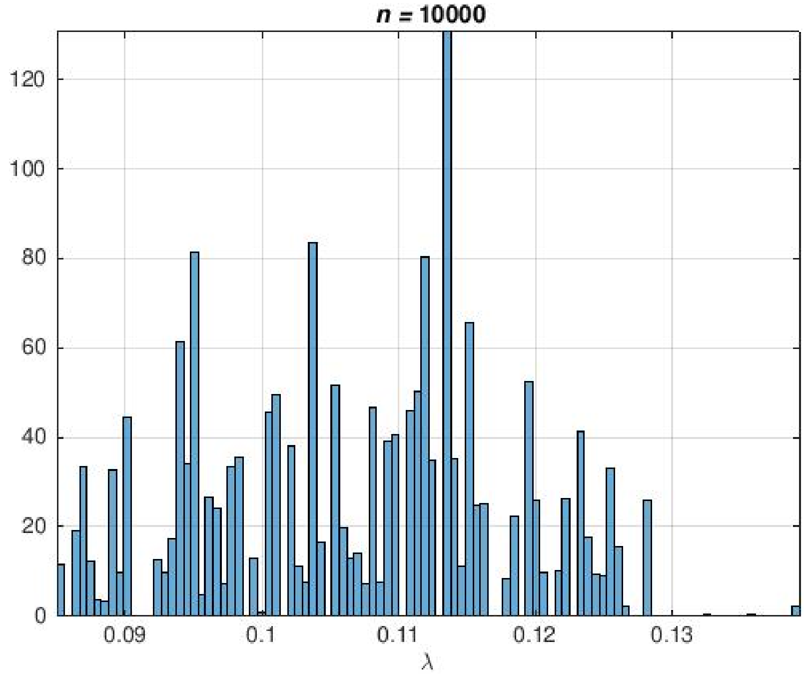

Figure A1.3. Histogram plot of the sample paths of MCMC runs for the rate parameter, where inter-arrival times follow exponential distribution.

(continues)

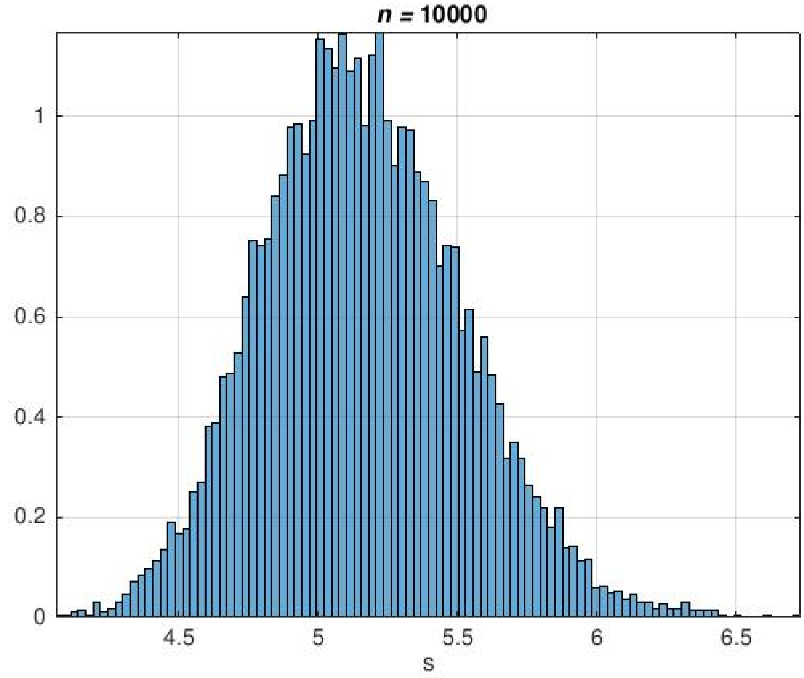

Figure A1.4. Histogram plot of the sample paths of MCMC runs for the scale parameter, where inter-arrival times follow gamma distribution.

(continues)

Appendix 1. (continued)

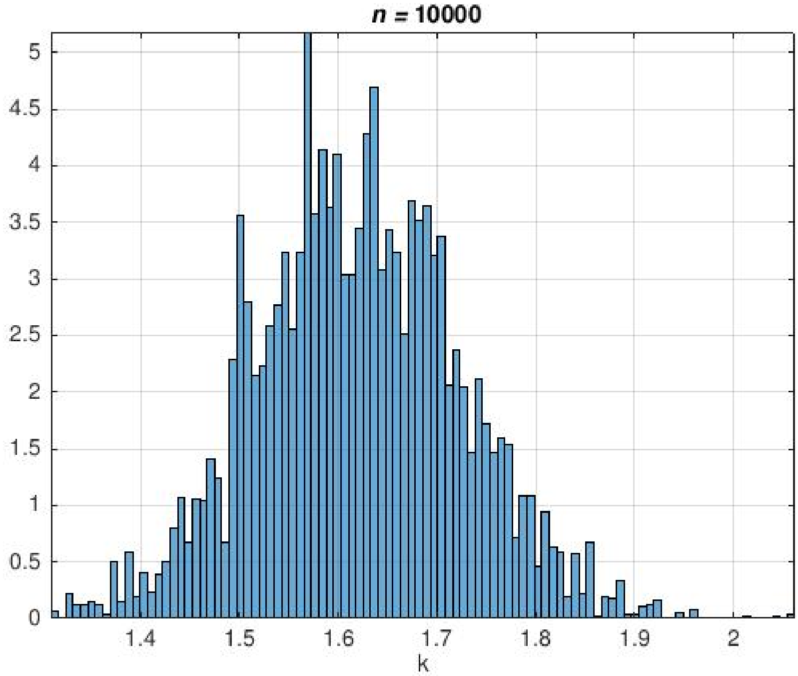

Figure A1.5. Histogram plot of the sample paths of MCMC runs for the shape parameter, where inter-arrival times follow gamma distribution.

(continues)

Appendix 1. (continued)

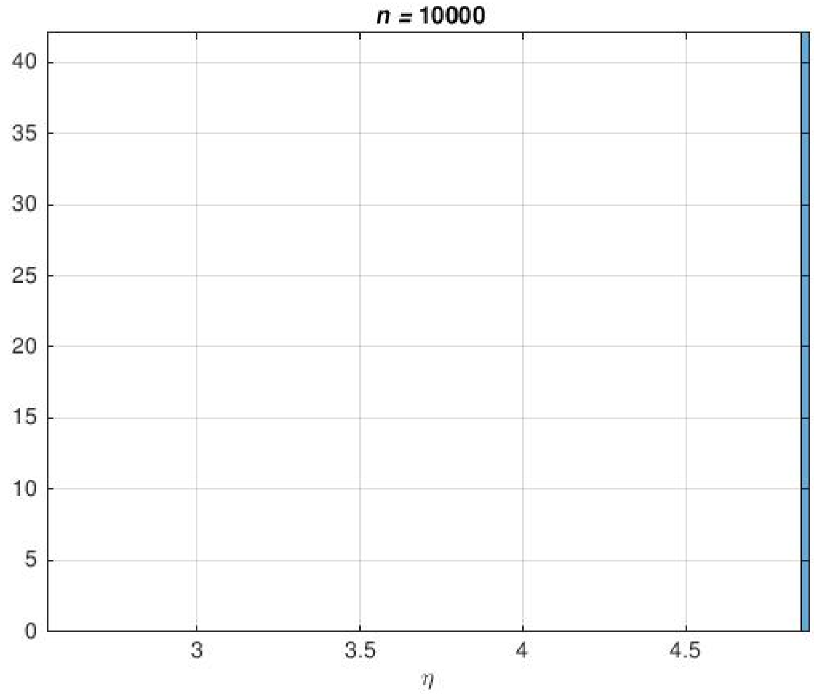

Figure A1.6. Histogram plot of the sample paths of MCMC runs for the scale parameter, where inter-arrival times follow Weibull distribution.

(continues)

Appendix 1. (continued)

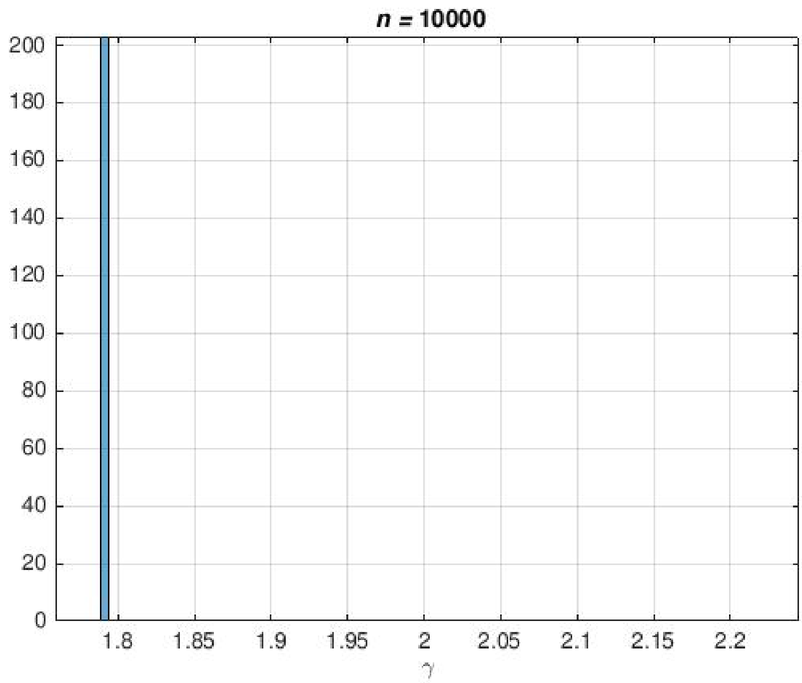

Figure A1.7. Histogram plot of the sample paths of MCMC runs for the shape parameter, where inter-arrival times follow Weibull distribution.

(continues)

Appendix 1. (continued)

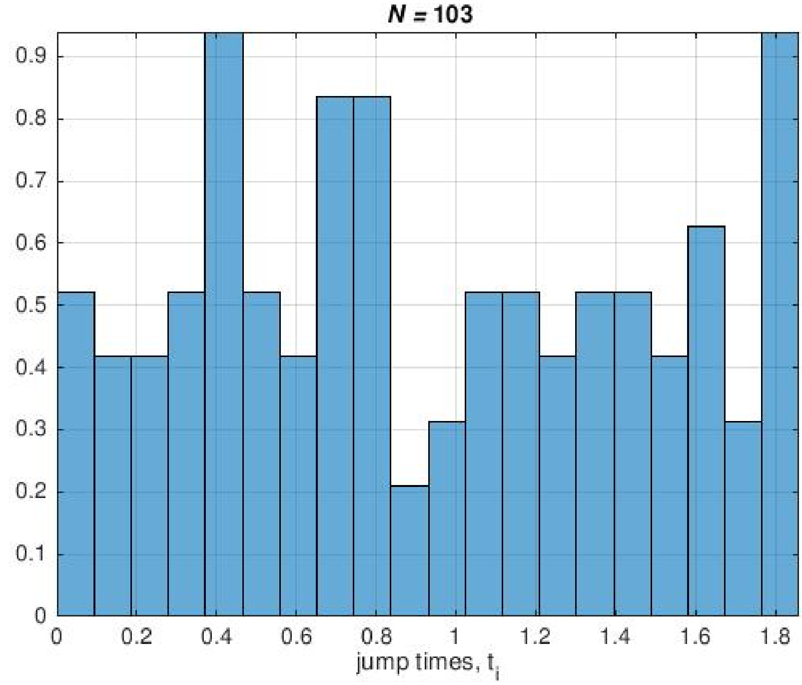

Figure A1.8. Histogram plot of the sample paths of the predicted jump times generated using maximum likelihood estimation method, where inter-arrival times follow exponential distribution.

(continues)

Appendix 1. (continued)

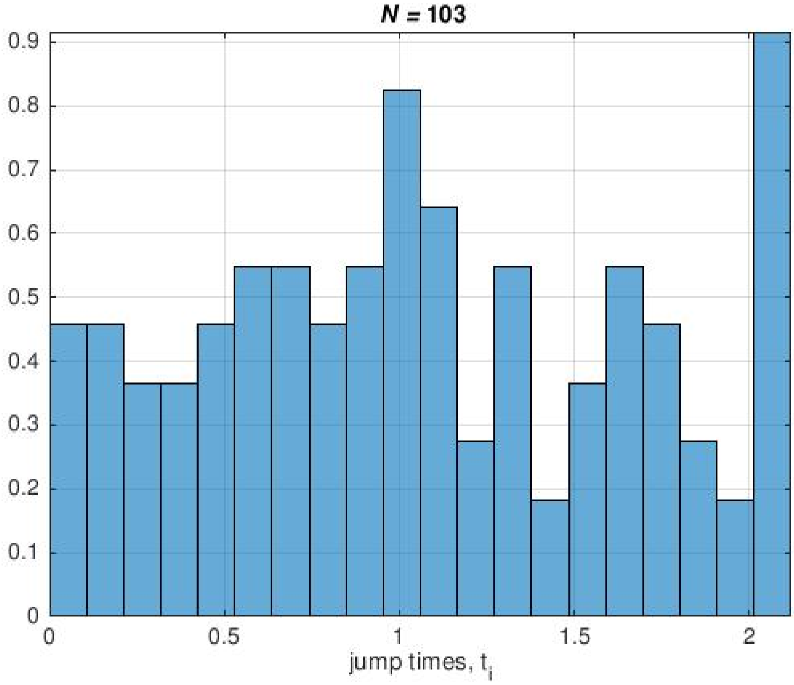

Figure A1.9. Histogram plot of the sample paths of the predicted jump times generated using Bayesian estimation method, where inter-arrival times follow exponential distribution.

(continues)

Appendix 1. (continued)

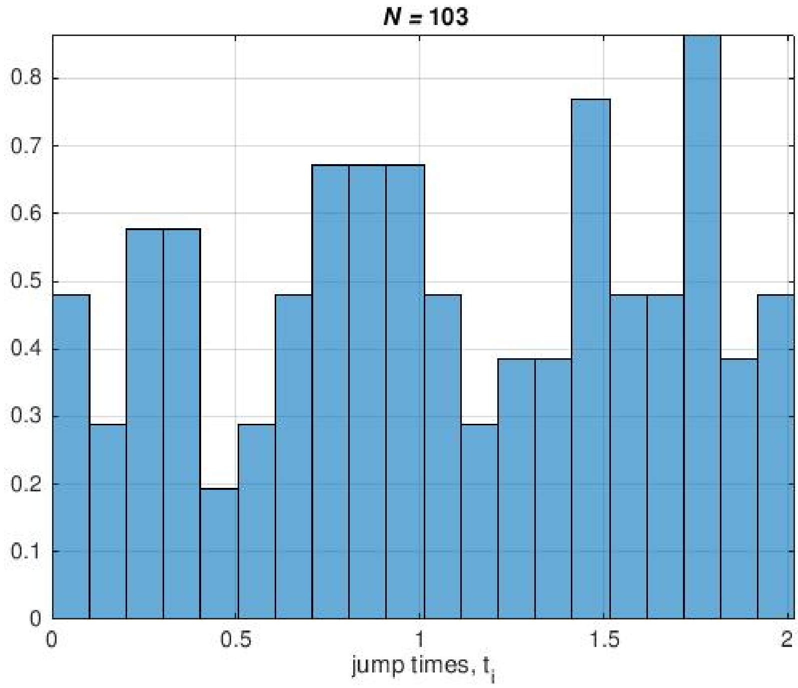

Figure A1.10. Histogram plot of the sample paths of the predicted jump times generated using maximum likelihood estimation method, where inter-arrival times follow gamma distribution.

(continues)

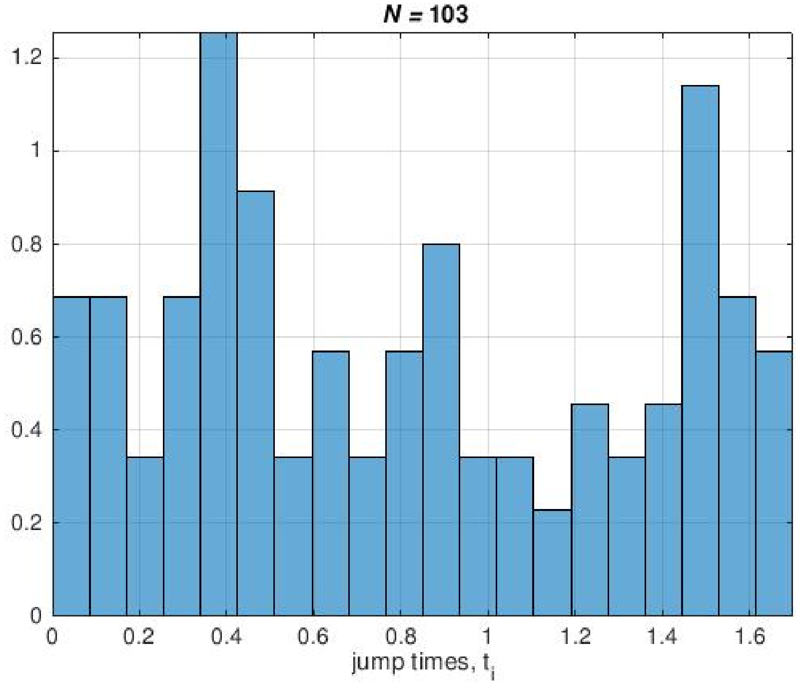

Figure A1.11. Histogram plot of the sample paths of the predicted jump times generated using Bayesian estimation method, where inter-arrival times follow gamma distribution.

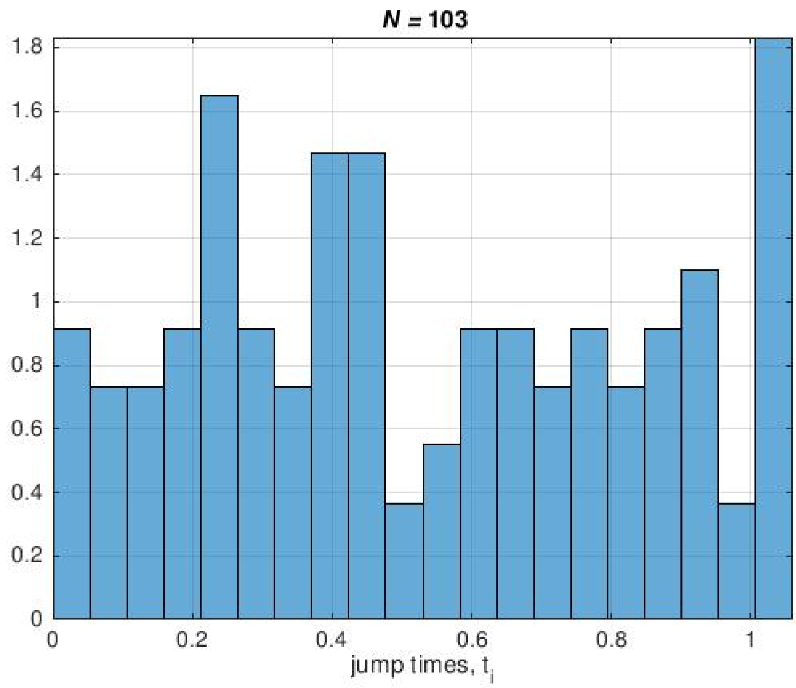

Figure A1.12. Histogram plot of the sample paths of the predicted jump times generated using maximum likelihood estimation method, where inter-arrival times follow Weibull distribution.

(continues)

Appendix 1. (continued)

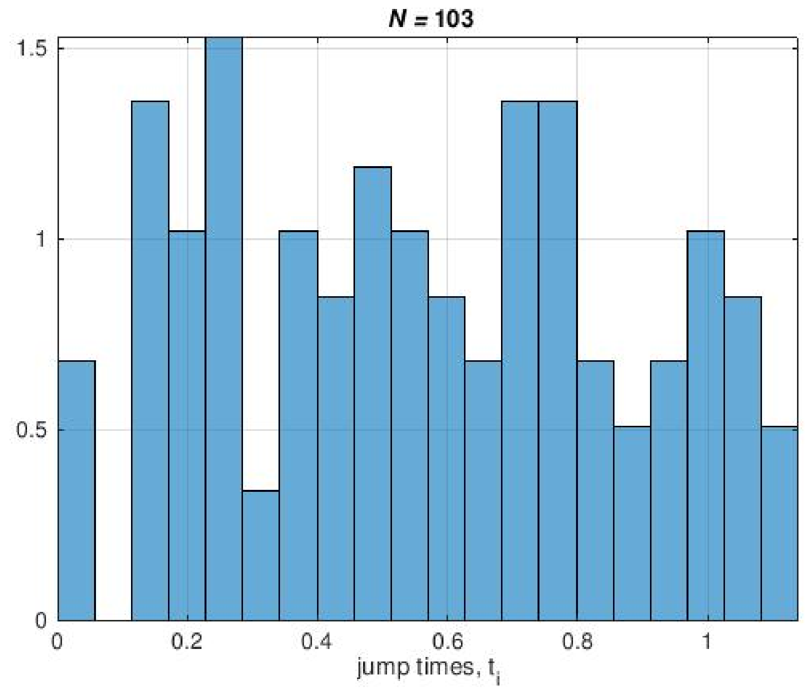

Figure A1.13. Histogram plot of the sample paths of the predicted jump times generated using Bayesian estimation method, where inter-arrival times follow Weibull distribution.

(continues)