University Utrecht

Bachelor thesis

Bachelor of Mathematics

Semi-Markov model applied to the estimation of earthquake occurrences

Author:
Patricia de Bruin

Student number:
5979838

Supervisor:
Cristian Spitoni

June 14, 2019

Utrecht University

Contents

Chapter 1

Introduction

The process of earthquake occurrences can be represented with several analytical models (Votsi et al., [1]). Some of them are based on empirical observations of preliminary events, others on physical modelling of the earthquake process and some on statistical analysis of patterns of seismicity.

In this thesis we present a semi-Markov model and a homogeneous Markov model in continuous-time to analyse the process of earthquake occurrences. For the semi-Markov model, we estimate the semi-Markov kernel, Markov renewal functions, transitions probabilities and distributions of sojourn time for every state through a non-parametric method. We apply our model to a dataset of the Northern Aegean Sea region in Greece.

We will examine the expected number of earthquake occurrences and the rate of occurrences of earthquakes, which we will define in the thesis. We formulate 95% confidence intervals for these quantities. Our semi-Markov model will be compared with the homogeneous Markov model based on the hazard rate function of the semi-Markov process. We discuss which model is potentially a better fit to the dataset.

In the last part of the this thesis, we look at an additional dataset of the Northern Aegean Sea region and look at the effect of covariates on the semi-Markov model. In particular, we want investigate the influence of the location where the earthquakes occurred.

1.1 Motivation

We can apply the theory of Markov processes in various fields, because the Markov property is very intuitive: if we know the past and present of a system, then the future development of the system is only determined by its present state (Barbu & Limnios, [11]). So the history of the system does not play a role in its future development. We also call this the memoryless property.

However, the Markov property has its limitations. It enforces restrictions on the distribution of the sojourn time in a state, which is exponentially distributed in the continuous case. This is a disadvantage when we apply Markov processes in real-life applications.

Therefore, we can introduce the semi-Markov process. This process allows us to have arbitrary distributed sojourn time in any state and still provides the Markov property, but in a more flexible way. The memoryless property does not act on the calender time in this case, but on the sojourn time in the state.

Because of the above reasoning, it is convenient to apply the semi-Markov model to the analysis of the process of earthquake occurrences. From historical information, we know when certain earthquakes occurred and for how long there was no seismic activity.

1.2 Summary

Here we give an overview of the subjects that are covered in the chapters of this thesis.

In chapter 2 we introduce definitions about Markov chains in discrete-time and continuous-time. The Markov renewal processes with corresponding renewal kernels are discussed and eventually we define what a semi-Markov process is. This chapter is based on (Ross, [3]) chapters 4, 6 and 7, ([4]) and (Grabski, [5]).

Chapter 3 presents further definitions related to the semi-Markov process. We introduce empirical estimators for some of the quantities that are defined. Also we look at the asymptotic behaviour of the empirical estimators. We follow (Votsi et al., [1]), which is the main article this thesis is based on. This chapter is further based on (Ouhbi & Limnios, [2]), (Grabski, [5]), (Limnios & Oprişan, [7]) chapter 4 and (Limnios & Ouhbi, [8]).

In chapter 4 an introduction of the homogeneous Markov model is given. We discuss the properties of this model with respect to the rate of occurrences of failures. This chapter is based on (Ouhbi & Limnios, [2]) and (Grabski, [5]).

We apply the two models to the dataset of the Northern Aegean Sea region in chapter 5, which we use from (Votsi et al., [1]). We define three states of earthquakes. We look at the hazard rate function of the semi-Markov process, which will be an important quantity if we want to compare the homogeneous Markov model and the semi-Markov model. The R package SemiMarkov (Listwon & Saint-Pierre, [12]) is used to determine the hazard rate functions in a parametric way. Furthermore, we discuss the expected number of earthquake occurrences from any state to the state with magnitude $M \geq 6.1$ and show how to determine the rate of occurrences of earthquakes with $M \geq 6.1$. This chapter is based on (Votsi et al., [1]), (Limnios & Oprişan, [7]) chapter 5 and (Listwon & Saint-Pierre, [12]).

In chapter 6 we compare the homogeneous Markov model and the semi-Markov model based on the hazard rate function of the semi-Markov process. We compute the semi-Markov kernels and look at the Wald test to draw conclusions about the fit of both models. Parametric and non-parametric estimators will also be discussed. This last subject is based on (Rice, [6]).

We apply the semi-Markov model to a new dataset in chapter 7, which we use from (Votsi et al., [1]). This dataset consists of two states of earthquakes. A covariate is added to the dataset, namely the location of the earthquake. The Northern Aegean Sea region is divided into four sub areas. We use the R package SemiMarkov (Listwon & Saint-Pierre, [12]) again to determine the hazard rate functions of the semi-Markov process and look at the Wald test to derive the consequences of the addition of covariates.

Chapter 8 concludes this thesis about the semi-Markov model and its application to the estimation of earthquakes occurrences. In the appendix we find the proofs of certain theses used in the previous chapters and tables and figures.

We mention that in this thesis we will not go into detail about measure theoretic definitions. For the one who is interested in these subjects, see (Cohn, [13]).

Chapter 2

Preliminaries

In this section we introduce some definitions which we will apply in the following sections. We discuss Markov chains, Markov renewal processes and eventually semi-Markov processes.

Definition 2.0.1 (Ross, [3]) (Stochastic process, state space). A *stochastic process* $\{X(t), t \in T\}$ is a collection of random variables. The index $X(t)$ denotes the state of the process at time t and T the index set of the process. When T is countable, the stochastic process is said to be a discrete-time process. When T is an interval, the stochastic process is said to be a continuous-time process. The set I of all possible values that the random variable $X(t)$ can assume, is called the *state space* of the stochastic process.

Definition 2.0.2 (Ross, [3]) (Discrete-time Markov chain). Let $\{X_n, n = 0, 1, 2, \dots\}$ be a stochastic process that takes on a countable number of possible values in a set I. A *discrete-time Markov chain* is a stochastic process where the conditional distribution of a future state X_{n+1} given the past states $X_0, X_1, \dots, X_{n-1}$ and the present state X_n, is independent of the past states and only depends on the present state. The process satisfies the property

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i), \tag{2.0.1}$$

for all $n \geq 0$ and $i_0, \dots, i_{n-1}, i, j \in I$.

If we have a discrete-time Markov chain and in addition it holds true that the right-hand side of (2.0.1) is independent of the time n, then the discrete-time Markov chain is said to have stationary or homogeneous transition probabilities. It follows that there exists a fixed probability p_{ij} that, when the process starts in state i, it will be next in state j. For the probabilities p_{ij}, it holds true that

(i) $p_{ij} \geq 0$, for all $i, j \in I$.

(ii) $\sum_{j=0}^{\infty} p_{ij} = 1$ for all $i = 0, 1, 2, \dots$

Example 2.0.1 (Ross, [3]) A common example for which we can use Markov chains is predicting the weather. Suppose the probability that it rains tomorrow depends only on whether or not it rains today and does not depend on the past weather conditions. We assume that if it rains today, the probability that it rains tomorrow is equal to α. If it does not rain today, the probability that it will rain tomorrow is equal to β. We say that the process is in state 0 if it will rain tomorrow and the process is in state 1 if it does not rain tomorrow. Then the transition probabilities are given by the matrix

$$P = \begin{pmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{pmatrix}. \tag{2.0.2}$$

Definition 2.0.3 (Ross, [3]) (Continuous-time Markov chain). Let $\{X(t), t \geq 0\}$ be a continuous-time stochastic process which takes on values in the set I of non-negative integers. A *continuous-time Markov chain* is a stochastic process with the property that the conditional distribution of the future $X(t+s)$ given the present $X(s)$ and the past $X(u)$, $0 \leq u < s$, depends only on the present and is independent of the past. The process satisfies

$$\mathbb{P}\left(X(t+s) = j \mid X(s) = i, X(u) = x(u), 0 \leq u < s\right) = \mathbb{P}\left(X(t+s) = j \mid X(s) = i\right), \tag{2.0.3}$$

for all $s, t \geq 0$ and non-negative integers $i, j, x(u) \in I$ with $0 \leq u < s$.

If we have a continuous-time Markov chain and in addition it holds true that the right-hand side of (2.0.3) is independent of the time s, then the continuous-time Markov chain is said to have stationary or homogeneous transition probabilities.

Definition 2.0.4 (Ross, [3]) A Markov chain is said to be *irreducible*, if all states commute with each other. In other words, two states i and j are accessible to each other for all $i, j \in I$.

Definition 2.0.5 (Ross, [3]) We call a stochastic process $\{N(t), t \geq 0\}$ a *counting process*, if $N(t)$ represents the number of events that occurred up to time t.

A counting process $N(t)$ must satisfy the following properties:

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

(ii) $N(t) \in \mathbb{Z}$.

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

(iv) For $s < t$, $N(t) - N(s)$ is equal to the number of events that occurred during the interval $(s, t]$.

Definition 2.0.6 (Ross, [3]) (Renewal process). Let $\{N(t), t \geq 0\}$ be a counting process and let X_n be the time between the $(n-1)$th event and the nth event of the process with $n \geq 1$. If the sequence $\{X_1, X_2, \dots\}$ is independent and identically distributed, then we call the counting process $N(t)$ a *renewal process*.

Definition 2.0.7 (Ross, [3], [4]) (Markov renewal process). Let E be the state space. A *Markov renewal process* is a bivariate stochastic process (J_n, S_n), where J_n are the values of the state space E in the Markov chain and S_n are the jump times. We define $X_n := S_n - S_{n-1}$ to be the sojourn time in the state. The process has to satisfy the following properties

$$P\left(J_{n+1} = j, X_n \leq t \mid (J_0, S_0), (J_1, S_1), \dots, (J_n = i, S_n)\right) = P\left(J_{n+1} = j, X_n \leq t \mid J_n = i\right), \tag{2.0.4}$$

$$P\left(J_0 = i, X_0 = 0\right) = P(J_0 = i), \tag{2.0.5}$$

for all $n \geq 0$ and $t \geq 0$ and $i, j \in E$.

Definition 2.0.8 (Grabski, [5]) (Renewal matrix, renewal kernel). Let E be the state space and consider the Markov renewal process (J_n, S_n) as in definition 2.0.7. The matrix defined as

$$Q(t) = \{Q_{ij}(t) : i, j \in E\}, \tag{2.0.6}$$

$$Q_{ij}(t) := \mathbb{P}\left(J_{n+1} = j, X_n \leq t \mid J_n = i\right), \tag{2.0.7}$$

is called a *renewal matrix*. We identify the renewal matrix $Q(t)$ as the *renewal kernel*.

The Markov renewal matrix $Q(t)$ satisfies the following conditions:

(i) For all $t \geq 0$ and $i, j \in E$, it holds true that $Q_{ij}(t) \geq 0$.

(ii) The functions $Q_{ij}(t)$ are right-continuous.

(iii) For all $i, j \in E$, it holds true that $Q_{ij}(0) = 0$ and $Q_{ij}(t) \leq 1$ for all $t \geq 0$.

(iv) For all $i \in E$, it holds that $\lim_{t\to\infty} \sum_{j\in E} Q_{ij}(t) = 1$.

Definition 2.0.9 ([4]) (Semi-Markov process). Consider a Markov renewal process (J_n, S_n) as in definition 2.0.7. Define the stochastic process $Z_t := J_n$ for $t \in [S_n, S_{n+1})$. Then Z_t is called a *semi-Markov process*.

The main difference between a Markov renewal process and a semi-Markov process is that we define the Markov renewal process as a two-tuple of states and times ([4]). The semi-Markov process is the process that evolves over time and all realisations of the process have a defined state for any given time. We can read the semi-Markov process as follows. Suppose that a process can be in one of N states. Each time it enters a state i, it stays there for a random amount of time, and then it makes a transition from state i to state j with transition probability p_{ij}.

If the amount of time spend in state i before it makes a transition to state j is constant, then the semi-Markov process is just a Markov chain. However, the amount of time spend in the states before transition can also depend on a distribution. We consider the following lemma about this concept.

Lemma 2.0.1 (Grabski, [5]) A homogeneous Markov process $\{X(t), t \geq 0\}$ with discrete state space E and right-continuous trajectories keeping constant values on half open intervals, given by the transition rate matrix $\Lambda = (\lambda_{ij})$ for $i, j \in E$ with $0 < -\lambda_{ii} = \lambda_i < \infty$, is a semi-Markov process with kernel $Q(t) = \{Q_{ij}(t) : i, j \in E\}$, where

$$Q_{ij}(t) = p_{ij}\left(1 - e^{-\lambda_i t}\right) \quad \text{for } t \geq 0, \tag{2.0.8}$$

$$p_{ij} = \frac{\lambda_{ij}}{\lambda_i} \quad \text{for } i \neq j, \tag{2.0.9}$$

$$p_{ii} = 0. \tag{2.0.10}$$

Proof. See for example (Grabski, [5]), theorem 1. ■

Chapter 3

Semi-Markov model

This chapter presents the definitions of the semi-Markov model. First, we summarize the definitions for the semi-Markov process which we will apply in the rest of the thesis. After that we introduce empirical estimators for the quantities of a finite state space semi-Markov process. We discuss the asymptotic behaviour of the empirical estimators for some of the quantities.

3.1 Definition of semi-Markov process

We will follow the definitions of Votsi et al., [1]. Consider a Markov-renewal process (J_n, S_n) defined on a complete probability space and with state space $E = \{1, 2, \ldots, s\}$. It holds true that J_n are the values of the state space E in the Markov chain and S_n are the jump times for $n \geq 0$. We assume that S_n take values in $[0, \infty)$. Define $X_n := S_n - S_{n-1}$ to be the sojourn time in the state and we assume that $X_0 = S_0 = 0$. We let $Z_t := J_{N(t)}$ for $t \geq 0$ be the semi-Markov process where $N(t)$ is the counting process of the semi-Markov process up to time t. It holds true that $N(t) := \max\{n \geq 0 : S_n \leq t\}$. In figure 3.1.1 (Barbu & Limnios, [11]) we see an example of a sample path of the semi-Markov chain.

The semi-Markov process depends on its initial law, which we assume is equal to $\pi_i = \mathbb{P}(J_0 = i)$, and on its semi-Markov kernel

$$\begin{aligned} Q_{ij}(t) &:= \mathbb{P}\left(J_{n+1} = j, X_{n+1} \leq t \mid J_0, J_1, \ldots, J_n = i, X_1, X_2, \ldots, X_n\right) \\ &= \mathbb{P}\left(J_{n+1} = j, X_{n+1} \leq t \mid J_n = i\right), \end{aligned} \tag{3.1.1}$$

for all $t \geq 0$ and $i, j \in E$. We consider that $Q_{ii}(t) \neq 0$ for all $i \in E$.

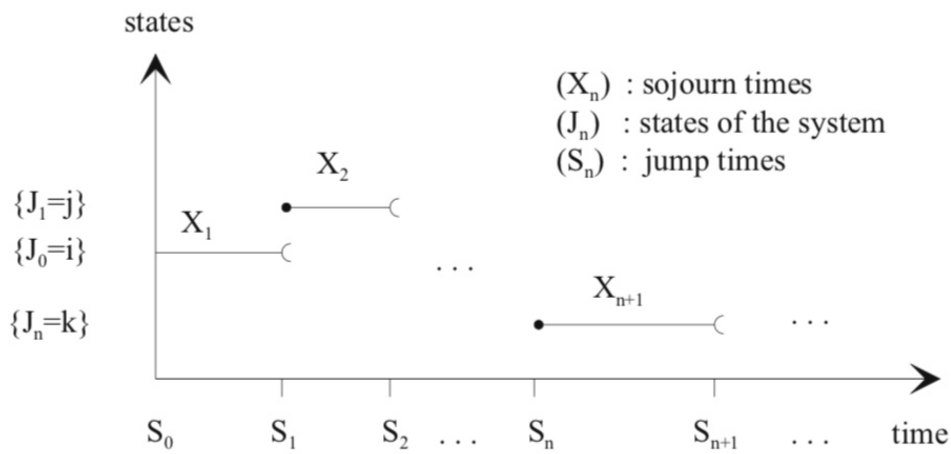

Figure 3.1.1: Sample path of the semi-Markov chain. (Barbu & Limnios, [11])

The probabilities

$$\begin{aligned} p_{ij} &:= \lim_{t\to\infty} Q_{ij}(t) = Q_{ij}(\infty) \\ &= \mathbb{P}\left(J_{n+1} = j \mid J_n = i\right) \end{aligned} \tag{3.1.2}$$

are the transition probabilities from state i to state j of the embedded Markov chain $\{J_n, n = 0, 1, 2, \dots\}$. We assume that the transition probabilities do not depend on the time n.

We define

$$F_{ij}(t) := \mathbb{P}\left(X_{n+1} \le t \mid J_n = i, J_{n+1} = j\right) \tag{3.1.3}$$

to be the distribution function associated with the sojourn time in state i, before going to state j. From this definition we can derive the following result. The proof is given in the appendix.

Theorem 3.1.1 (Grabski, [5]) It holds true that

$$F_{ij}(t) = \frac{Q_{ij}(t)}{p_{ij}}, \tag{3.1.4}$$

for all $t \ge 0$ and $i, j \in E$.

Proof. See appendix A.1. ■

The distribution function of the sojourn time, also called the waiting time, in state i is equal to

$$\begin{aligned} H_i(t) &:= \mathbb{P}\left(X_{n+1} \le t \mid J_n = i\right) \\ &= \sum_{j\in E} Q_{ij}(t). \end{aligned} \tag{3.1.5}$$

Because the semi-Markov process is connected to (J_n, S_n) through $Z_t = J_{N(t)}$ for $t \ge 0$, where $N(t)$ is the counting process of the semi-Markov process up to time t, we define the transition function $P_{ij}(t)$ of the process $\{Z_t, t \ge 0\}$ as (Ouhbi & Limnios, [2])

$$\begin{aligned} P_{ij}(t) &:= \mathbb{P}\left(Z_t = j \mid Z_0 = i\right) \\ &= \mathbb{P}\left(J_{N(t)} = j \mid J_0 = i\right), \end{aligned} \tag{3.1.6}$$

for all $i, j \in E$. Then the unconditional semi-Markov state probability is equal to

$$\begin{aligned} P_j(t) &:= \mathbb{P}\left(Z_t = j\right) = \mathbb{P}\left(J_{N(t)} = j\right) \\ &= \sum_{i=1}^{s} \mathbb{P}\left(J_{N(t)} = j \mid J_0 = i\right) \mathbb{P}(J_0 = i) \\ &= \sum_{i=1}^{s} \pi_i P_{ij}(t), \end{aligned} \tag{3.1.7}$$

where $\pi_i := P(J_0 = i)$ is the initial distribution of the Markov renewal process.

Let $T \ge 0$ be a fixed time and $t \ge 0$ a specific time in the process. We define T as the end time of the process, where it holds true that $T > t$ for all t.

We denote $N_i(T)$ to be the number of visits of $\{J_n, n = 0, 1, 2, \dots\}$ to state i up to time T, and $N_{ij}(T)$ to be the number of transitions from state i to state j up to time T. So

$$N_i(T) := \sum_{n=1}^{N(t)} \mathbb{1}_{\{J_n = i\}} = \sum_{n=1}^{\infty} \mathbb{1}_{\{J_n = i, S_n \le T\}}, \tag{3.1.8}$$

$$N_{ij}(T) := \sum_{n=1}^{N(t)} \mathbb{1}_{\{J_{n-1} = i, J_n = j\}} = \sum_{n=1}^{\infty} \mathbb{1}_{\{J_{n-1} = i, J_n = j, S_n \le T\}}, \tag{3.1.9}$$

where $\mathbb{1}_A$ is the indicator function. The function $\mathbb{1}_A$ is defined as

$$\mathbb{1}_A := \begin{cases} 1 & \text{if } A, \\ 0 & \text{otherwise.} \end{cases}$$

Lastly, we want to define the quantity $Q_{ij}^{(n)}(t) := \mathbb{P}(J_n = j, S_n \leq t \mid J_0 = i)$, which is called the n-fold convolution of the semi-Markov kernel $Q_{ij}(t)$ for $t \geq 0$. This is the probability that, starting from state i, the semi-Markov chain will make its nth transition at time t to state j. Therefore, we need the following definition.

Definition 3.1.1 (Votsi et al., [1]) (Stieltjes convolution). Let $\phi(i,t)$ for $t \geq 0$ and $i \in E$ be a real valued measurable function and Q be a semi-Markov kernel. Then the *Stieltjes convolution* of ϕ by Q is defined as

$$Q * \phi(i,t) := \sum_{k \in E} \int_0^t Q_{ik}(ds)\phi(k, t-s). \tag{3.1.10}$$

We obtain the following recursive formula for $Q_{ij}^{(n)}(t)$:

$$Q_{ij}^{(n)}(t) := \begin{cases} \sum_{k \in E} \int_0^t Q_{ik}(ds) Q_{kj}^{(n-1)}(t-s) & \text{if } n \geq 2, \\ Q_{ij}(t) & \text{if } n = 1, \\ \delta_{ij} \mathbb{1}_{\{t \geq 0\}} & \text{if } n = 0, \end{cases} \tag{3.1.11}$$

where δ_{ij} is Kronecker's delta symbol.

3.2 Empirical estimators

In this section we introduce the empirical estimators for some of the quantities we discussed previously. We mention that we can use the below empirical estimators, because later in chapter 5 and chapter 7, we are working with finite datasets. We discuss the asymptotic behaviour for some of the empirical estimators as well.

Let T be the end time of the process. For the semi-Markov kernel $Q_{ij}(t)$ we have the following empirical estimator

$$\widehat{Q}_{ij}(t,T) := \frac{1}{N_i(T)} \sum_{n=1}^{N(t)} \mathbb{1}_{\{J_{n-1}=i, J_n=j, X_n \leq t\}}. \tag{3.2.1}$$

Because $F_{ij}(t) = Q_{ij}(t)/p_{ij}$, in a similar way we obtain that $\widehat{F}_{ij}(t,T) = \widehat{Q}_{ij}(t,T)/\widehat{p}_{ij}(T)$ with

$$\widehat{F}_{ij}(t,T) := \frac{1}{N_{ij}(T)} \sum_{n=1}^{N(t)} \mathbb{1}_{\{J_{n-1}=i, J_n=j, X_n \leq t\}}, \tag{3.2.2}$$

$$\widehat{p}_{ij}(T) := \frac{N_{ij}(T)}{N_i(T)}. \tag{3.2.3}$$

The quantities $\widehat{F}_{ij}(t,T)$ and $\widehat{p}_{ij}(T)$ are respectively the empirical estimators for the conditional transition functions and the transition probabilities. We see that for $\widehat{p}_{ij}$ we divide the number of transitions from state i to state j by the (total) number of visits to state i.

We want to show that the empirical estimator $\widehat{Q}_{ij}(t,T)$ of the semi-Markov kernel is strongly consistent and asymptotically normal. Therefore, we need the following definition and theses.

Definition 3.2.1 ([10]) Let X_i, $i = 1, \ldots, n$, be an independent and identically distributed sequence of random variables with distribution function F on $\mathbb{R}$. Then the *empirical distribution function* is defined by

$$\widehat{F}_n(x) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq x\}}. \tag{3.2.4}$$

Theorem 3.2.1 ([10]) (Glivenko-Cantelli theorem) Let X_i, $i = 1, \ldots, n$, be an independent and identically distributed sequence of random variables with distribution function F on $\mathbb{R}$. Then,

$$\sup_{x \in \mathbb{R}} \left| \widehat{F}_n(x) - F(x) \right| \to 0 \quad \text{(a.s.)} \tag{3.2.5}$$

as $n \to \infty$.

Proof. See for example ([10]), theorem 1.1. ∎

Theorem 3.2.2 (Barbu & Limnios, [11]) The empirical estimator $\widehat{p}_{ij}(T)$ of p_{ij} for all $i, j \in E$ is strongly consistent, i.e.

$$\widehat{p}_{ij}(T) \to p_{ij} \quad \text{(a.s.)} \tag{3.2.6}$$

as $T \to \infty$.

We note that the empirical estimator $\widehat{F}_{ij}(t, T)$ of the distribution function satisfies theorem 3.2.1, because it meets the definition of the empirical distribution function. Now, we give the properties of the empirical estimator $\widehat{Q}_{ij}(t, T)$. The proof of theorem 3.2.3 is given in the appendix.

Theorem 3.2.3 (Limnios & Oprişan, [7]) The empirical estimator $\widehat{Q}_{ij}(t, T)$ of $Q_{ij}(t)$ for all $i, j \in E$ is strongly consistent, i.e.

$$\max_{i,j \in E} \sup_{t \in [0,T)} \left| \widehat{Q}_{ij}(t, T) - Q_{ij}(t) \right| \to 0 \quad \text{(a.s.)} \tag{3.2.7}$$

as $T \to \infty$.

Proof. See appendix A.2. ∎

Theorem 3.2.4 (Limnios & Oprişan, [7]) The empirical estimator $\widehat{Q}_{ij}(t, T)$ is asymptotically normal, i.e. for fixed $t > 0$

$$\sqrt{T} \left| \widehat{Q}_{ij}(t, T) - Q_{ij}(t) \right| \xrightarrow{d} N(0, \sigma_{ij}^2) \tag{3.2.8}$$

as $T \to \infty$. It holds true that

$$\sigma_{ij}^2 = \mu_{ii} Q_{ij}(t) \left(1 - Q_{ij}(t)\right), \tag{3.2.9}$$

where μ_{ii} is the mean time between two visits to state i.

Proof. See for example (Limnios & Oprişan, [7]), theorem 4.26. ∎

3.2.1 Hazard rate function

For all $i, j \in E$ we define the hazard rate function of transition distributions between states, $\lambda_{ij}(t)$ for $t \geq 0$, of the semi-Markov kernel by (Limnios & Ouhbi, [8])

$$\lambda_{ij}(t) := \begin{cases} \dfrac{q_{ij}(t)}{1 - H_i(t)} & \text{if } p_{ij} > 0 \text{ and } H_i(t) < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.10}$$

It holds true that $q_{ij}(t)$ is the derivative of the (i, j)th element of the semi-Markov kernel $Q(t)$. We assume that this derivative exists for all $i, j \in E$. We define the cumulative hazard rate function from state i to state j at time t by $\Lambda_{ij}(t) := \int_0^t \lambda_{ij}(s)\, ds$. The total cumulative hazard rate function of state i at time t is defined as $\Lambda_i(t) = \sum_{j \in E} \Lambda_{ij}(t)$. It holds true that

$$Q_{ij}(t) = \int_0^t e^{-\Lambda_i(s)} \lambda_{ij}(s)\, ds. \tag{3.2.11}$$

The empirical estimator of the hazard rate function of the semi-Markov process is equal to

$$\widehat{\lambda}_{ij}(t,T) := \begin{cases} \dfrac{\widehat{q}_{ij}(t,T)}{1-\widehat{H}_i(t,T)} & \text{if } \widehat{p}_{ij}(T) > 0 \text{ and } \widehat{H}_i(t,T) < 1, \\ 0 & \text{otherwise,} \end{cases} \tag{3.2.12}$$

where $\widehat{p}_{ij}(T)$ is the empirical estimator for the transition probabilities. Furthermore, $\widehat{H}_i(t,T)$ is the empirical estimator of the distribution function of the sojourn time in state i, which is defined as

$$\widehat{H}_i(t,T) := \sum_{j\in E} \widehat{Q}_{ij}(t,T), \tag{3.2.13}$$

where $\widehat{Q}_{ij}(t,T)$ is the empirical estimator of the semi-Markov kernels. The empirical estimator of the derivative function $\widehat{q}_{ij}(t,T)$ of the semi-Markov kernel is (Ouhbi & Limnios, [2])

$$\widehat{q}_{ij}(t,T) := \frac{\widehat{Q}_{ij}(t+\Delta,T) - \widehat{Q}_{ij}(t,T)}{\Delta}. \tag{3.2.14}$$

It holds true that $\Delta = T^{-\alpha}$ for $0 < \alpha < 1$.

3.2.2 Markov renewal matrix

We consider the following counting function of the renewal process (J_n, S_n) (Limnios & Oprişan, [7])

$$N_j(t) := \sum_{n=0}^{\infty} \mathbb{1}_{\{J_n=j, S_n\le t\}}. \tag{3.2.15}$$

This is the total number of visits to state j up to time t. It holds true that, for any $t \geq 0$, $N(t) = \sum_{j\in E} N_j(t)$.

The Markov renewal matrix $\Psi(t) = (\psi_{ij}(t))$ is defined as

$$\psi_{ij}(t) := \mathbb{E}_i[N_j(t)] = \mathbb{E}(N_j(t) \mid J_0 = i) \tag{3.2.16}$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(J_n = j, S_n \le t \mid J_0 = i) = \sum_{n=0}^{\infty} Q_{ij}^{(n)}(t), \tag{3.2.17}$$

for $t \geq 0$ and $i, j \in E$. Here, $\psi_{ij}(t) = \mathbb{E}_i[N_j(t)]$ is the expected number of visits from state i to state j up to time t. As an estimator for the (i,j)th element of the matrix $\Psi(t)$, we use the empirical estimator

$$\widehat{\psi}_{ij}(t,T) := \sum_{n=0}^{\infty} \widehat{Q}_{ij}^{(n)}(t,T), \tag{3.2.18}$$

where $\widehat{Q}_{ij}^{(n)}(t,T)$ is the n-fold convolution of $\widehat{Q}_{ij}(t,T)$ which we defined in section 3.1.

For the empirical estimator $\widehat{Q}_{ij}^{(n)}(t,T)$ of the n-fold convolution of the semi-Markov kernel, the following theorem holds true. The proof is given in the appendix.

Theorem 3.2.5 (Barbu & Limnious, [11]) The empirical estimator $\widehat{Q}_{ij}^{(n)}(t,T)$ of $Q_{ij}^{(n)}(t)$ for all $i, j \in E$ is strongly consistent, i.e. for any fixed $n \in \mathbb{N}$

$$\max_{i,j\in E} \max_{t\in[0,T]} \left| \widehat{Q}_{ij}^{(n)}(t,T) - Q_{ij}^{(n)}(t) \right| \to 0 \quad \text{(a.s.)} \tag{3.2.19}$$

as $T \to \infty$.

Proof. See appendix A.3. ∎

Let us define the finite Markov renewal matrix as

$$\psi_{ij}^{(k)}(t) := \sum_{n=0}^{k} Q_{ij}^{(n)}(t), \tag{3.2.20}$$

where $Q_{ij}^{(n)}(t)$ is the n-fold convolution of the semi-Markov kernel. The corresponding empirical estimator is

$$\widehat{\psi}_{ij}^{(k)}(t,T) := \sum_{n=0}^{k} \widehat{Q}_{ij}^{(n)}(t,T). \tag{3.2.21}$$

Then the following theorem about the finite Markov renewal matrix holds true. The proof is given in the appendix.

Theorem 3.2.6 (Barbu & Limnious, [11]) The empirical estimator $\widehat{\psi}_{ij}^{(k)}(t,T)$ of $\psi_{ij}^{(k)}(t)$ for all $i,j \in E$ is strongly consistent, i.e. for any fixed $k \in \mathbb{N}$

$$\max_{i,j\in E} \left|\widehat{\psi}_{ij}^{(k)}(t,T) - \psi_{ij}^{(k)}(t)\right| \to 0 \quad \text{(a.s.)} \tag{3.2.22}$$

as $T \to \infty$.

Proof. See appendix A.4. ■

The empirical estimator $\widehat{\psi}_{ij}(t,T)$ of the elements of the renewal matrix has the following two properties.

Theorem 3.2.7 (Limnios & Oprişan, [7]) The empirical estimator $\widehat{\psi}_{ij}(t,T)$ of $\psi_{ij}(t)$ for all $i,j \in E$ is strongly consistent, i.e.

$$\max_{i,j\in E} \sup_{t\in[0,T)} \left|\widehat{\psi}_{ij}(t,T) - \psi_{ij}(t)\right| \to 0 \quad \text{(a.s.)} \tag{3.2.23}$$

as $T \to \infty$.

Theorem 3.2.8 (Limnios & Oprişan, [7]) The empirical estimator $\widehat{\psi}_{ij}(t,T)$ is asymptotically normal, i.e. for fixed $t > 0$

$$\sqrt{T}\left|\widehat{\psi}_{ij}(t,T) - \psi_{ij}(t)\right| \xrightarrow{d} N(0, \sigma_{ij}^2(t)) \tag{3.2.24}$$

as $T \to \infty$. It holds true that

$$\sigma_{ij}^2(t) = \sum_{k=1}^{s}\sum_{l=1}^{s} \mu_{kk}\left[(\psi_{ik} * \psi_{lj})^2 * Q_{kl} - (\psi_{ik} * \psi_{lj} * Q_{kl})^2\right](t), \tag{3.2.25}$$

where μ_{kk} is the mean time between two visits to state k.

Proof. See for example (Barbu & Limnios, [11]), theorem 4.5. ■

From theorem 3.2.7 and theorem 3.2.8, we conclude that

$$\mathbb{P}\left(\frac{|\widehat{\psi}_{ij}(t,T) - \psi_{ij}(t)|}{\sigma_{ij}(t)}\right) \xrightarrow{d} N(0,1), \tag{3.2.26}$$

as $T \to \infty$ for all $i,j \in E$.

We can derive $100(1-\alpha)\%$ confidence intervals for the (i,j)th element $\psi_{ij}(t)$ of the Markov renewal matrix for all $t \geq 0$ from the above result. Let $z(\alpha)$ be that number such that the area under the standard normal density function to the right of $z(\alpha)$ is equal to α (Rice, [6]). With use of (3.2.26), we obtain

$$\mathbb{P}\left(-z(\alpha/2) \leq \frac{\widehat{\psi}_{ij}(t,T) - \psi_{ij}(t)}{\sigma_{ij}(t)} \leq z(\alpha/2)\right) \approx 1-\alpha,$$

$$\mathbb{P}\left(\widehat{\psi}_{ij}(t,T) - z(\alpha/2)\sigma_{ij}(t) \leq \psi_{ij}(t) \leq \widehat{\psi}_{ij}(t,T) + z(\alpha/2)\sigma_{ij}(t)\right) \approx 1-\alpha.$$

This yields the $100(1-\alpha)\%$ confidence intervals for $\psi_{ij}(t)$ for all $i,j \in E$:

$$\left[\widehat{\psi}_{ij}(t,T) - z(\alpha/2)\sigma_{ij}(t), \quad \widehat{\psi}_{ij}(t,T) + z(\alpha/2)\sigma_{ij}(t)\right], \quad t \geq 0. \tag{3.2.27}$$

3.2.3 Rate of occurrences of failures

Now we define the rate of occurrences of failures, which we denote by $ro(t)$. Therefore, consider the state space E. We can naturally partition E into two sets, U and D, where U is the set of working states and D is the set of repair states. It holds true that $E = U \cup D$ and $U \cap D = \emptyset$, where $U, D \neq \emptyset$. We let $N_f(t)$ be the counting process of the transitions from U to D, at time t. In other words, this is the number of failures up to time t. We define the rate of occurrences of failures $ro(t)$ as (Ouhbi & Limnios, [2])

$$ro(t) := \lim_{\Delta t \to 0} \frac{\mathbb{E}\left[N_f(t+\Delta t) - N_f(t)\right]}{\Delta t}. \tag{3.2.28}$$

The rate of occurrences of failures is the probability that a failure, which is not necessarily the first, will occur in the next time interval.

To explicitly calculate $ro(t)$, we have to make the following assumptions:

(i) The semi-Markov kernel $Q(t)$ is absolutely continuous with respect to the Lebesgue measure on $[0, \infty)$, and has a derivative function $q(t) = (q_{ij}(t))$ for $i, j \in E$.

(ii) For all $t \geq 0$, it holds true that

$$\Psi'(t) = \sum_{n=0}^{\infty} \left[Q^{(n)}\right]'(t) < \infty.$$

We can prove that the following theorem about the rate of occurrences of failures holds.

Theorem 3.2.9 (Ouhbi & Limnios, [2]) Under the assumptions (i) and (ii), the rate of occurrences of failures for the semi-Markov process is given by

$$ro(t) = \sum_{i \in U} \sum_{j \in D} \sum_{l=1}^{s} \pi_l \psi_{li} * q_{ij}(t). \tag{3.2.29}$$

It holds true that $\pi_l := P(J_0 = l)$.

Proof. See for example (Ouhbi & Limnios, [2]), theorem 1. ∎

The corresponding empirical estimator for the rate of occurrences of failures is

$$\widehat{ro}(t, T) := \sum_{i \in U} \sum_{j \in D} \sum_{l=1}^{s} \pi_l \widehat{\psi}_{li} * \widehat{q}_{ij}(t, T), \tag{3.2.30}$$

where the empirical estimator of the derivative function $\widehat{q}_{ij}$ of the semi-Markov kernel is defined as before.

We look at the asymptotic behaviour of the empirical estimator of the rate of occurrences of failures. It can be shown that the following theses hold true. The proof of theorem 3.2.10 is given in the appendix.

Theorem 3.2.10 (Ouhbi & Limnios, [2]) Under assumptions (i) and (ii), the empirical estimator $\widehat{ro}(t, T)$ of $ro(t)$ is strongly consistent for $ro(t)$, i.e. for all $M \in [0, \infty)$ we have

$$\sup_{t \in [0, M]} \left|\widehat{ro}(t, T) - ro(t)\right| \to 0 \quad \text{(a.s.)} \tag{3.2.31}$$

as $T \to \infty$.

Proof. See appendix A.5. ∎

Theorem 3.2.11 (Ouhbi & Limnios, [2]) Let $q_{ij}(.)$ be twice continuously differentiable at t for all $i \in U$ and $j \in D$. Then $\widehat{ro}(t, T)$ is asymptotically normal with mean $ro(t)$ and variance

$$\sigma^2 = \sum_{i \in U} \sum_{j \in D} \mu_{ii} \frac{\sum_{l=1}^{s} \pi_l \psi_{li} * q_{ij}(t)}{T^{1-\alpha}} + \mathcal{O}(T^{-1}), \tag{3.2.32}$$

where μ_{ii} is the mean time between two visits to state i and $\pi_l := P(J_0 = l)$.

Proof. See for example (Ouhbi & Limnios, [2]), theorem 3. ∎

From theorem 3.2.10 and theorem 3.2.11, we conclude that

$$\mathbb{P}\left(|\widehat{ro}(t,T) - ro(t)|\right) \xrightarrow{d} N(0,\sigma^2), \tag{3.2.33}$$

as $T \to \infty$.

We can derive the $100(1-\alpha)\%$ confidence intervals for the rate of occurrences of failures $ro(t)$ for all $t \geq 0$ from the above result. For $T \to \infty$, it holds true that

$$\mathbb{P}\left(\frac{|\widehat{ro}(t,T) - ro(t)|}{\sigma}\right) \xrightarrow{d} N(0,1). \tag{3.2.34}$$

Let $z(\alpha)$ be that number such that the area under the standard normal density function to the right of $z(\alpha)$ is equal to α (Rice, [6]). With use of (3.2.34), we obtain

$$\mathbb{P}\left(-z(\alpha/2) \leq \frac{\widehat{ro}(t,T) - ro(t)}{\sigma} \leq z(\alpha/2)\right) \approx 1-\alpha,$$
$$\mathbb{P}\left(\widehat{ro}(t,T) - z(\alpha/2)\sigma \leq ro(t) \leq \widehat{ro}(t,T) + z(\alpha/2)\sigma\right) \approx 1-\alpha.$$

This yields the $100(1-\alpha)\%$ confidence interval for $ro(t)$:

$$\left[\widehat{ro}(t,T) - z(\alpha/2)\sigma, \quad \widehat{ro}(t,T) + z(\alpha/2)\sigma\right], \quad t \geq 0. \tag{3.2.35}$$

Chapter 4

Homogeneous Markov model

In this chapter we introduce the homogeneous Markov model. We define a continuous-time Markov chain as in chapter 2.

Definition 4.0.1 (Ross, [3]) (Continuous-time Markov chain). Let $\{X(t), t \geq 0\}$ be a continuous-time stochastic process which takes on values in the set I of non-negative integers. A *continuous-time Markov chain* is a stochastic process with the property that the conditional distribution of the future $X(t+s)$ given the present $X(s)$ and the past $X(u)$, $0 \leq u < s$, depends only on the present and is independent of the past. The process satisfies

$$\mathbb{P}\left(X(t+s)=j \mid X(s)=i, X(u)=x(u), 0 \leq u < s\right) = \mathbb{P}\left(X(t+s)=j \mid X(s)=i\right). \tag{4.0.1}$$

for all $s, t \geq 0$ and non-negative integers $i, j, x(u) \in I$ with $0 \leq u < s$.

If we have a continuous-time stochastic process $\{X(t), t \geq 0\}$ and in addition

$$P\left(X(t+s)=j \mid X(s)=i\right) \tag{4.0.2}$$

is independent of the time s, then the continuous-time Markov chain is said to have homogeneous transition probabilities.

We want to compare the semi-Markov model with the homogeneous Markov model in chapter 6. Therefore, we consider the definition of the semi-Markov kernel for a homogeneous Markov process, which we have already seen in lemma 2.0.1. We mention it here again.

Lemma 4.0.1 (Grabski, [5]) A homogeneous Markov process $\{X(t), t \geq 0\}$ with discrete state space E and right-continuous trajectories keeping constant values on half open intervals, given by the transition rate matrix $\Lambda = (\lambda_{ij})$ for $i, j \in E$ with $0 < -\lambda_{ii} = \lambda_i < \infty$, is a semi-Markov process with kernel $Q(t) = \{Q_{ij}(t) : i, j \in E\}$, where

$$Q_{ij}(t) = p_{ij}\left(1 - e^{-\lambda_i t}\right) \quad \text{for } t \geq 0, \tag{4.0.3}$$

$$p_{ij} = \frac{\lambda_{ij}}{\lambda_i} \quad \text{for } i \neq j, \tag{4.0.4}$$

$$p_{ii} = 0. \tag{4.0.5}$$

With use of lemma 4.0.1, we obtain the following theorem about the rate of occurrences of failures in case of the homogeneous Markov model. The proof is given in the appendix.

Theorem 4.0.1 (Ouhbi & Limnios, [2]) When we have a homogeneous Markov process as in lemma 4.0.1, it holds true that

$$ro(t) = \sum_{i \in U} \sum_{j \in D} P_i(t) \lambda_{ij}, \tag{4.0.6}$$

where λ_{ij} is the (i, j)th entry of the generating matrix of the Markov process.

Proof. See appendix A.6. ∎

Chapter 5

Application to dataset

In this chapter we apply the two models from chapter 3 and chapter 4 to a dataset of the region of the Northern Aegean Sea in Greece. Many researches are interested in this region and its surrounding, because of the high seismic activity (Votsi et al., [1]). The area has experienced several destructive earthquakes, which is known from both instrumental data and historical information.

The Northern Aegean Sea region experienced a decent number of strong earthquakes ($M \geq 6.4$), along with a decent amount of moderate events ($M \geq 5.5$) since 1953. We consider all earthquakes since 1953 with $M \geq 5.5$, which are given in table B.0.1 (Votsi et al., [1]) in the appendix.

Our dataset consist events of magnitude $M \geq 5.5$ that occurred during the period 1953 to 2007. We mention that after shocks of the earthquakes were removed from the dataset. We define three states of earthquakes corresponding to the magnitudes:

State 1: $[5.5, 5.6]$.

State 2: $[5.7, 6.0]$.

State 3: $[6.1, 7.2]$.

These intervals are defined to specify the discrete states of the system. We derive that the state space is equal to the set $E = \{1, 2, 3\}$.

For this application of the semi-Markov model to the region of the Northern Aegean Sea in Greece, we are most interested in the earthquakes of the third state. That is why, in the upcoming sections, we will determine the expected number of earthquake occurrences from any state to the third state and the rate of occurrences of earthquakes with $M \geq 6.1$.

Considering the dataset, we ordered table B.0.1 (Votsi et al., [1]) on date and time of occurrence. The magnitude M and state is mentioned as well, which are the most important part for us.

Let $t = 0$ be the time we observed the first earthquake. This earthquake took place at May 2, 1953. We note that it is unknown how long the first earthquake was in its state. We set the end time T later than the last earthquake we observed. The last earthquake we observed was at November 9, 2007. So we set $T = 55$ years.

In the upcoming sections we apply the dataset to the semi-Markov model in continuous-time. The number of observed transitions in the dataset from any state i to any state j for $i, j \in E$ are presented as elements in the matrix N. The elements of this matrix are the values $N_{ij}(T)$ for all $i, j \in E$.

$$N = \begin{pmatrix} 6 & 6 & 3 \\ 5 & 2 & 2 \\ 4 & 1 & 3 \end{pmatrix} \tag{5.0.1}$$

We read the matrix N as follows: four times there was a transition from state 3 to state 1. From the matrix N we can conclude that the embedded Markov chain $\{J_n, n = 0, 1, 2, \dots\}$ is irreducible, because we can make a transition from state i to state j for all $i, j \in E$.

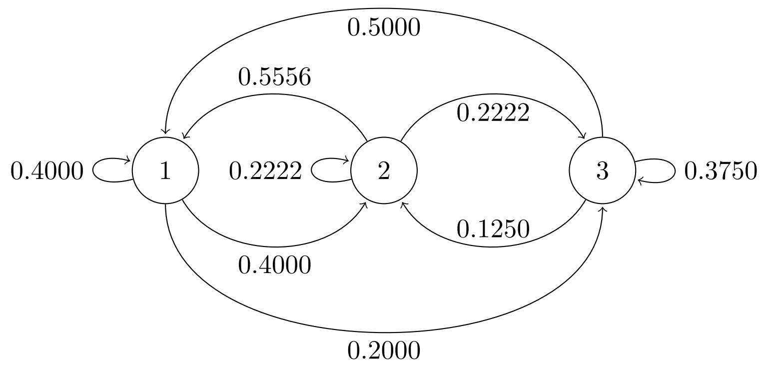

Figure 5.0.1: Transitions from state i to state j with transition probabilities, for all $i, j \in E$.

The values $N_i(T)$ for $i \in E$ are equal to

$$N_1(T) = 15, \qquad N_2(T) = 9, \qquad N_3(T) = 8. \tag{5.0.2}$$

We remember that the empirical estimator of the transition probabilities from any state i to any state j for $i, j \in E$ of the embedded Markov chain are defined as

$$\widehat{p}_{ij}(T) = \frac{N_{ij}(T)}{N_i(T)}. \tag{5.0.3}$$

The estimations of the transition probabilities from any state i to any state j are presented as elements in the matrix $\widehat{P} = (\widehat{p}_{ij})$.

$$\widehat{P} = \begin{pmatrix} 0.4000 & 0.4000 & 0.2000 \\ 0.5556 & 0.2222 & 0.2222 \\ 0.5000 & 0.1250 & 0.3750 \end{pmatrix} \tag{5.0.4}$$

We read the matrix $\widehat{P}$ as follows: the probability that there is a transition from state 3 to state 1 is equal to 0.5000. The transitions from state i to state j with the corresponding transition probabilities are shown in figure 5.0.1 for all transitions from state i to state j, $i, j \in E$.

With use of the definition for $\widehat{Q}_{ij}(t, T)$, we can estimate the semi-Markov kernels for transitions from state i to state j. We remember that

$$\widehat{Q}_{ij}(t, T) := \frac{1}{N_i(T)} \sum_{n=1}^{N(t)} \mathbb{1}_{\{J_{n-1}=i, J_n=j, X_n \leq t\}}. \tag{5.0.5}$$

The semi-Markov kernels are shown in figure 5.0.2 for all transitions from state i to state j, $i, j \in E$ and $t \geq 0$. The sojourn time is measured in years.

The empirical estimators for conditional transition functions $\widehat{F}_{ij}(t, T)$, associated with the sojourn time in each state before transition, are shown in figure 5.0.3 for all transitions from state i to state j, $i, j \in E$ and $t \geq 0$. We measured the sojourn time in years. The conditional transition functions are defined as

$$\widehat{F}_{ij}(t, T) := \frac{1}{N_{ij}(T)} \sum_{n=1}^{N(t)} \mathbb{1}_{\{J_{n-1}=i, J_n=j, X_n \leq t\}}. \tag{5.0.6}$$

Because $\widehat{F}_{ij}(t, T) = \widehat{Q}_{ij}(t, T)/\widehat{p}_{ij}(T)$, note that figure 5.0.3 is just a rescaling of figure 5.0.2.

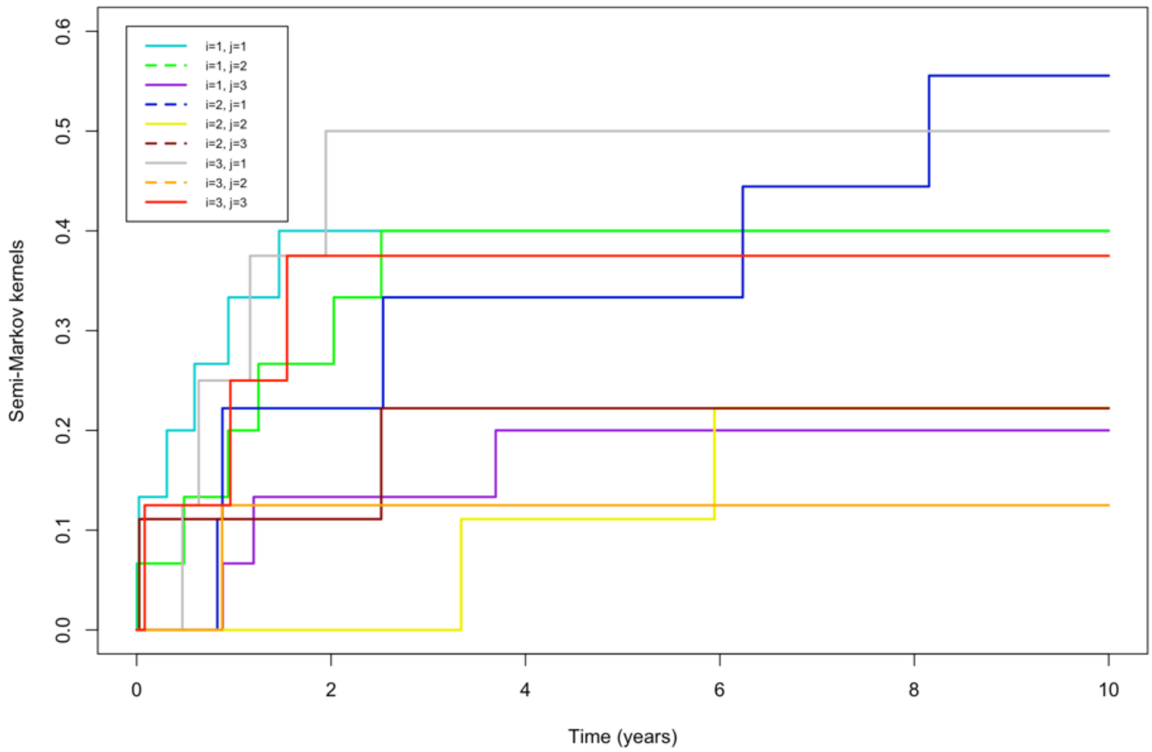

Figure 5.0.2: Empirical estimators for semi-Markov kernels, $\widehat{Q}_{ij}(t, T)$ for all transitions from state i to state j, $i, j \in E$.

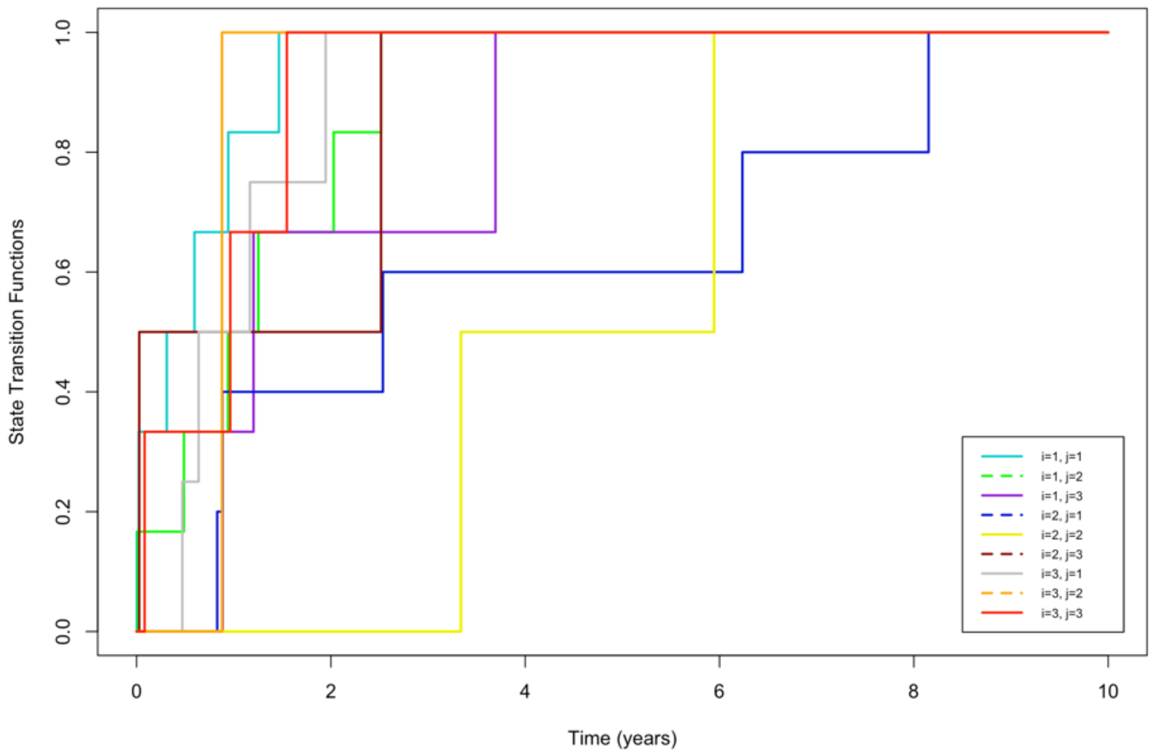

Figure 5.0.3: Empirical estimators for conditional transition functions, $\widehat{F}_{ij}(t, T)$ for all transitions from state i to state j, $i, j \in E$.

5.1 Hazard rate function

If we want to compare the two models that we apply to the dataset of table B.0.1 (Votsi et al., [1]), an important quantity to look at is the hazard rate function of the semi-Markov process $\lambda_{ij}(t)$ for all $i, j \in E$ and $t \geq 0$.

For the semi-Markov model, we can estimate this function using the empirical estimator

$$\widehat{\lambda}_{ij}(t, T) = \frac{\widehat{q}_{ij}(t, T)}{1 - \widehat{H}_i(t, T)}, \tag{5.1.1}$$

for $i, j \in E$ and $t \geq 0$.

With use of the R package SemiMarkov (Listwon & Saint-Pierre, [12]) we are able to find the estimated hazard rate functions of the semi-Markov process for all the transitions from state i to state j, $i \neq j \in E$. We cannot derive the hazard rate functions of the semi-Markov process for the transitions from state i to itself, because by definition of the semi-Markov model the waiting times must be known. If these kind of transitions are observed in the package, the transition is combined with the next transition such that we obtain a transition from state i to state j for $i \neq j \in E$.

For the case $i = j$, we define the hazard rate function as follows

$$\widehat{\lambda}_{ii}(t, T) = -\sum_{j \neq i} \widehat{\lambda}_{ij}(t, T). \tag{5.1.2}$$

The R package SemiMarkov uses the parametric maximum likelihood estimation. We remark that the semi-Markov model we discussed in the previous chapters is non-parametric. We will explain the difference between parametric and non-parametric estimators in the next chapter, when we compare the two models with each other.

5.1.1 Homogeneous Markov model

First, we derive the hazard rate function of the semi-Markov process for the homogeneous Markov model. We changed the original dataset from table B.0.1 (Votsi et al., [1]) to meet the requirements to apply for the package, which we call it *markov*1. We measure the time t in years. From lemma 4.0.1 we know that a semi-Markov process is a homogeneous Markov process if and only if the sojourn time is exponentially distributed. Therefore, we choose the exponential distribution E for the sojourn time. We fit the data with use of the function *semiMarkov*(.). Figure 5.1.1 shows the performed steps in Rstudio.

In figure 5.1.2, we see the estimates of parameters of the waiting time distributions, the standard deviations, the confidence intervals and the Wald test statistics. In the following matrix $\Sigma = (\sigma_{ij})$ we give the values for the parameters of the exponential distribution for all $i \neq j \in E$:

$$\Sigma = \begin{pmatrix} - & 1.535 & 2.386 \\ 5.538 & - & 1.384 \\ 1.598 & 1.302 & - \end{pmatrix}. \tag{5.1.3}$$

With use of the R package SemiMarkov we can determine two hazard rate functions, namely the hazard rate function of the waiting time $\alpha_{ij}(t)$ and the hazard rate function of the semi-Markov process $\lambda_{ij}(t)$. First, we give the hazard rate function of the waiting time and after that the hazard rate function of the semi-Markov process.

When we choose an exponential distribution for the sojourn time, the hazard rate functions of the waiting time with scale parameter σ_{ij} are defined as (Listwon & Saint-Pierre, [12])

$$\alpha_{ij}(t) = \frac{1}{\sigma_{ij}}, \tag{5.1.4}$$

for all $i, j \in E$ and $t \geq 0$. Then we obtain the following estimated hazard rate functions of the waiting time $\alpha_{ij}(t)$ for the homogeneous Markov model for all $i \neq j \in E$:

$$\alpha_{12}(t) = 0.6515, \qquad \alpha_{13}(t) = 0.4191,$$
$$\alpha_{21}(t) = 0.1806, \qquad \alpha_{23}(t) = 0.7225,$$
$$\alpha_{31}(t) = 0.6258, \qquad \alpha_{32}(t) = 0.7680.$$

The plots of the hazard rate functions of the waiting time are shown in figure B.0.1 in the appendix.

When we choose an exponential distribution for the sojourn time, the density functions of the sojourn time with scale parameter σ_{ij} are defined as

$$f_{ij}(t) = \frac{1}{\sigma_{ij}} e^{-t/\sigma_{ij}}, \tag{5.1.5}$$

for all $i, j \in E$ and $t \geq 0$. We obtain the following estimated density functions $f_{ij}(t)$ in case of the homogeneous Markov model for all $i \neq j \in E$:

$$\begin{aligned} f_{12}(t) &= 0.6515e^{-0.6515t}, & f_{13}(t) &= 0.4191e^{-0.4191t}, \\ f_{21}(t) &= 0.1806e^{-0.1806t}, & f_{23}(t) &= 0.7225e^{-0.7225t}, \\ f_{31}(t) &= 0.6258e^{-0.6258t}, & f_{32}(t) &= 0.7680e^{-0.7680t}. \end{aligned}$$

The probability distribution functions of the sojourn time with scale parameter σ_{ij} are defined as

$$F_{ij}(t) = 1 - e^{-t/\sigma_{ij}}, \tag{5.1.6}$$

for all $i, j \in E$ and $t \geq 0$. The estimated probability distribution functions of the sojourn time for all $i \neq j \in E$ are

$$\begin{aligned} F_{12}(t) &= 1 - e^{-0.6515t}, & F_{13}(t) &= 1 - e^{-0.4191t}, \\ F_{21}(t) &= 1 - e^{-0.1806t}, & F_{23}(t) &= 1 - e^{-0.7225t}, \\ F_{31}(t) &= 1 - e^{-0.6258t}, & F_{32}(t) &= 1 - e^{-0.7680t}. \end{aligned}$$

We know from the beginning of this chapter that the transition probability matrix is given by

$$\widehat{P} = \begin{pmatrix} 0.4000 & 0.4000 & 0.2000 \\ 0.5556 & 0.2222 & 0.2222 \\ 0.5000 & 0.1250 & 0.3750 \end{pmatrix}. \tag{5.1.7}$$

Define $G_i(t) = 1 - H_i(t) = \sum_{j \in E} p_{ij}(1 - F_{ij}(t))$ as the survival function of the sojourn time in state i. Here, $F_{ij}(t)$ is the probability distribution function of the sojourn time and p_{ij} the transition probability of the embedded Markov chain. For the derivative of the semi-Markov kernel, we know that $q_{ij}(t) = p_{ij} f_{ij}(t)$. With use of this information, we can determine the hazard rate function of the semi-Markov process $\lambda_{ij}(t)$.

If we plug the expressions for the density functions of the sojourn time, the probability distribution functions of the sojourn time and the transition probabilities in case of the homogeneous Markov model into the hazard rate function of the semi-Markov process

$$\lambda_{ij}(t) = \frac{q_{ij}(t)}{1 - \widehat{H}_i(t)} = \frac{p_{ij} f_{ij}(t)}{\sum_{j \in E} p_{ij}(1 - F_{ij}(t))}, \tag{5.1.8}$$

we obtain the estimated hazard rate functions of the semi-Markov process for all $i \neq j \in E$. It holds true that $\lambda_{ij}(t) = -\sum_{j \neq i} \lambda_{ij}(t)$ for all $i \in E$.

The plots of the hazard rate functions of the semi-Markov process are shown in figure B.0.2 in the appendix.

5.1.2 Semi-Markov model

For the semi-Markov model we use the same dataset as before, and call it *semimarkov*1. We measure the time t in years. Because we want an estimator for the semi-Markov process, we choose the (non-Markov) Weibull distribution W for the sojourn time for all transitions except for the transition from state 3 to state 2.[1] Here, we choose the exponential distribution. We fit the data with use of the function $semiMarkov(.)$ as before. Figure 5.1.3 shows the performed steps in Rstudio.

From figure 5.1.4, we can derive the estimates of parameters of the waiting time distributions, the standard deviations, the confidence intervals and the Wald test statistics. In the following two matrices $\Sigma = (\sigma_{ij})$ and $V = (\nu_{ij})$ we give the values for the parameters of the Weibull and exponential distribution for all $i \neq j \in E$:

$$\Sigma = \begin{pmatrix} - & 1.493 & 2.398 \\ 5.381 & - & 1.152 \\ 1.461 & 1.311 & - \end{pmatrix}, \quad V = \begin{pmatrix} - & 1.061 & 1.854 \\ 1.263 & - & 0.495 \\ 2.083 & - & - \end{pmatrix}. \tag{5.1.9}$$

As we said before, we can determine two hazard rate functions, one for the waiting time $\alpha_{ij}(t)$ and one for the semi-Markov process $\lambda_{ij}(t)$. First, we give the hazard rate function of the waiting time and then the hazard rate function of the semi-Markov process.

When we choose a Weibull distribution for the sojourn time, the hazard rate functions of the waiting time with scale parameter σ_{ij} and shape parameter ν_{ij} are defined as (Listwon & Saint-Pierre, [12])

$$\alpha_{ij}(t) = \frac{\nu_{ij}}{\sigma_{ij}} \left(\frac{t}{\sigma_{ij}}\right)^{\nu_{ij}-1}, \tag{5.1.10}$$

for all $i, j \in E$ and $t \geq 0$. Then we obtain the estimated hazard rate functions $\lambda_{ij}(t)$ for the semi-Markov process

$$\begin{aligned} \alpha_{12}(t) &= 0.7106(0.6698t)^{0.061}, & \alpha_{13}(t) &= 0.7731(0.4170t)^{0.854}, \\ \alpha_{21}(t) &= 0.2347(0.1858t)^{0.263}, & \alpha_{23}(t) &= 0.4297(0.8681t)^{-0.505}, \\ \alpha_{31}(t) &= 1.4257(0.6845t)^{1.083}, & \alpha_{32}(t) &= 0.7628. \end{aligned}$$

The plots of the hazard rate functions of the waiting time are shown in figure B.0.3 in the appendix.

When we choose a Weibull distribution for the sojourn time, the density functions of the sojourn time with scale parameter σ_{ij} and shape parameter ν_{ij} are defined as

$$f_{ij}(t) = \frac{\nu_{ij}}{\sigma_{ij}} \left(\frac{t}{\sigma_{ij}}\right)^{\nu_{ij}-1} e^{-(t/\sigma_{ij})^{\nu_{ij}}}, \tag{5.1.11}$$

for all $i, j \in E$ and $t \geq 0$. We obtain the following estimated density functions $f_{ij}(t)$ in case of the semi-Markov model for all $i \neq j \in E$:

$$\begin{aligned} f_{12}(t) &= 0.7106(0.6698t)^{0.061} e^{-(0.6698t)^{1.061}}, & f_{13}(t) &= 0.7731(0.4170t)^{0.854} e^{-(0.4170t)^{1.854}}, \\ f_{21}(t) &= 0.2347(0.1858t)^{0.263} e^{-(0.1858t)^{1.263}}, & f_{23}(t) &= 0.4297(0.8681t)^{-0.505} e^{-(0.8681t)^{0.495}}, \\ f_{31}(t) &= 1.4257(0.6845t)^{1.083} e^{-(0.6845t)^{2.083}}, & f_{32}(t) &= 0.7628 e^{-0.7628t}. \end{aligned}$$

The probability distributions of the sojourn time with scale parameter σ_{ij} and shape parameter ν_{ij} are defined as

$$F_{ij}(t) = 1 - e^{-(t/\sigma_{ij})^{\nu_{ij}}}, \tag{5.1.12}$$

for all $i, j \in E$ and $t \geq 0$. The estimated probability distribution functions of the sojourn time for all $i \neq j \in E$ are

$$\begin{aligned} F_{12}(t) &= 1 - e^{-(0.6698t)^{1.061}}, & F_{13}(t) &= 1 - e^{-(0.4170t)^{1.854}}, \\ F_{12}(t) &= 1 - e^{-(0.1858t)^{1.263}}, & F_{13}(t) &= 1 - e^{-(0.8681t)^{0.495}}, \\ F_{12}(t) &= 1 - e^{-(0.6845t)^{2.083}}, & F_{13}(t) &= 1 - e^{-0.7628t}. \end{aligned}$$

[1] If we choose the Weibull distribution for the transition from state 3 to state 2, we get an infinite estimator for the both hazard rate functions. Maybe, this is due to the fact we have only one observation of a transition from state 3 to state 2. Therefore, we choose to distribute the sojourn time exponentially in this case.

As before, the transition probability matrix is given by

$$\widehat{P} = \begin{pmatrix} 0.4000 & 0.4000 & 0.2000 \\ 0.5556 & 0.2222 & 0.2222 \\ 0.5000 & 0.1250 & 0.3750 \end{pmatrix}. \tag{5.1.13}$$

If we plug the expressions for the density functions of the sojourn time, the probability distribution functions of the sojourn time and the transition probabilities in case of the semi-Markov model into the hazard rate function of the semi-Markov process

$$\lambda_{ij}(t) = \frac{q_{ij}(t)}{1 - \widehat{H}_i(t)} = \frac{p_{ij} f_{ij}(t)}{\sum_{j \in E} p_{ij}(1 - F_{ij}(t))}, \tag{5.1.14}$$

we obtain the estimated hazard rate function of the semi-Markov process for all $i \neq j \in E$. It holds true that $\lambda_{ij}(t) = -\sum_{j \neq i} \lambda_{ij}(t)$ for all $i \in E$.

The plots of the hazard rate functions of the semi-Markov process are shown in figure B.0.4 in the appendix.

```
> library(numDeriv)
> library(MASS)
> library(Rsolnp)
> library(SemiMarkov)
> markov1 <- read.delim(file.choose(), header = TRUE)
> View(markov1)
> states <- c("1", "2", "3")
> mtrans <- matrix(FALSE, nrow = 3, ncol = 3)
> mtrans[1, 2:3] <- c("E", "E")
> mtrans[2, c(1,3)] <- c("E", "E")
> mtrans[3, c(1,2)] <- c("E", "E")
> fit <- semiMarkov(data = markov1, states = states, mtrans = mtrans)

Iter: 1 fn: 50.3817      Pars:  1.53458 2.38574 5.53841 1.38388 1.59834 1.30158 0.64633 0.77133
 0.81071
Iter: 2 fn: 50.3817      Pars:  1.53460 2.38573 5.53844 1.38385 1.59832 1.30158 0.64633 0.77133
 0.81071
solnp--> Completed in 2 iterations
```

Figure 5.1.1: The setting of the hazard rate function in case of the homogeneous Markov model.

```
> fit$table.dist
$Sigma
  Type Index Transition Estimation   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2      1.535 0.66     0.25     2.82    1.00      0.66  0.4166
2 dist     2     1 -> 3      2.386 1.43    -0.42     5.19    1.00      0.94  0.3323
3 dist     3     2 -> 1      5.538 2.49     0.66    10.41    1.00      3.33  0.0680
4 dist     4     2 -> 3      1.384 1.13    -0.82     3.59    1.00      0.12  0.7290
5 dist     5     3 -> 1      1.598 0.84    -0.05     3.24    1.00      0.51  0.4751
6 dist     6     3 -> 2      1.302 1.65    -1.94     4.54    1.00      0.03  0.8625
```

Figure 5.1.2: Estimates of parameters of the waiting time distribution in case of the homogeneous Markov model.

```
> library(numDeriv)
> library(MASS)
> library(Rsolnp)
> library(SemiMarkov)
> semimarkov1 <- read.delim(file.choose(), header = TRUE)
> View(semimarkov1)
> states <- c("1", "2", "3")
> mtrans <- matrix(FALSE, nrow = 3, ncol = 3)
> mtrans[1, 2:3] <- c("W", "W")
> mtrans[2, c(1,3)] <- c("W", "W")
> mtrans[3, c(1,2)] <- c("W", "E")
> fit <- semiMarkov(data = semimarkov1, states = states, mtrans = mtrans)

Iter: 1 fn: 47.3272      Pars:  1.49337 2.39772 5.38077 1.15168 1.46124 1.31125 1.06096 1.85406
 1.26277 0.49500 2.08268 0.62623 0.75694 0.81211
Iter: 2 fn: 47.3272      Pars:  1.49338 2.39774 5.38079 1.15169 1.46124 1.31124 1.06095 1.85403
 1.26275 0.49500 2.08269 0.62623 0.75694 0.81211
solnp--> Completed in 2 iterations
```

Figure 5.1.3: The setting of the hazard rate function in case of the semi-Markov model.

```
> fit$table.dist
$Sigma
  Type Index Transition Sigma   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2 1.493 0.60     0.31     2.67    1.00      0.67  0.4131
2 dist     2     1 -> 3 2.398 0.77     0.89     3.90    1.00      3.31  0.0689
3 dist     3     2 -> 1 5.381 1.96     1.54     9.22    1.00      5.01  0.0252
4 dist     4     2 -> 3 1.152 2.10    -2.97     5.27    1.00      0.01  0.9203
5 dist     5     3 -> 1 1.461 0.37     0.73     2.19    1.00      1.55  0.2131
6 dist     6     3 -> 2 1.311 1.68    -1.98     4.61    1.00      0.03  0.8625

$Nu
  Type Index Transition    Nu   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     7     1 -> 2 1.061 0.37     0.33     1.79    1.00      0.03  0.8625
2 dist     8     1 -> 3 1.854 0.76     0.36     3.35    1.00      1.25  0.2636
3 dist     9     2 -> 1 1.263 0.47     0.33     2.19    1.00      0.31  0.5777
4 dist    10     2 -> 3 0.495 0.31    -0.11     1.10    1.00      2.64  0.1042
5 dist    11     3 -> 1 2.083 0.84     0.44     3.73    1.00      1.67  0.1963
```

Figure 5.1.4: Estimates of parameters of the waiting time distribution in case of the semi-Markov model.

5.2 Expected number of earthquake occurrences

In this section, we derive 95% confidence intervals for the expected number of earthquake occurrences. Therefore, we need the Markov renewal matrix $\Psi(t)$ we introduced in section 3.2, because the renewal process is determined by this matrix. The entries of the renewal matrix are defined as follows

$$\psi_{ij}(t) = \sum_{n=0}^{\infty} Q_{ij}^{(n)}(t). \tag{5.2.1}$$

To determine these quantities, we need the n-fold convolution of the semi-Markov kernel $Q_{ij}(t)$. As we see in figure 5.0.2, in our case the semi-Markov kernels are step functions. We will derive how we can compute the n-fold convolution of the semi-Markov kernel when these function are step functions (Limnios & Oprişan, [7]).

Let us consider a partition of the time interval $[0, t]$, so $\{0 = t_0 < t_1 < \cdots < t_n = t\}$. We define the step function $\widetilde{Q}_{ij}(.)$ for $0 \leq u < t$ as follows

$$\widetilde{Q}_{ij}(u) := \sum_{r=0}^{n-1} Q_{ij}(t_r) \mathbb{1}_{\{t_r \leq u < t_{r+1}\}}. \tag{5.2.2}$$

Then the following theorem for step functions $\widetilde{Q}_{ij}(.)$ holds true. The proof is given in the appendix.

Theorem 5.2.1 (Limnios & Oprişan, [7]) By the recursive formula for the Stieltjes convolution, the n-fold convolution of $\widetilde{Q}_{ij}(.)$ for a fixed time t, is equal to

$$\widetilde{Q}_{ij}^{(n)}(t) = \sum_{k_{n-1} \in E} \cdots \sum_{k_1 \in E} \sum_{r_{n-1}} \cdots \sum_{r_1} \widetilde{Q}_{ik_1}(t - t_{r_{n-1}} - \cdots - t_{r_1}) \prod_{s=1}^{n-1} \Delta \widetilde{Q}_{k_s, k_{s+1}}(t_{r_s}), \tag{5.2.3}$$

where $\Delta \widetilde{Q}_{kj}(t_n) := \widetilde{Q}_{kj}(t_n) - \widetilde{Q}_{kj}(t_{n-1})$ and $k_n = j$ and

$$\begin{gathered} k_1 \in E, \ldots, k_{n-1} \in E, \\ r_{n-1} : 0 < t_{r_{n-1}} \leq t, \\ r_{n-2} : 0 < t_{r_{n-2}} \leq t - t_{r_{n-1}}, \\ \vdots \\ r_1 : 0 < t_{r_1} \leq t - t_{r_{n-1}} - \cdots - t_{r_2}. \end{gathered}$$

Proof. See appendix A.7. ∎

Now, we look at our dataset again. We are interested in the expected number of earthquake occurrence from any state i to the third state. This, because earthquakes of $M \geq 6.1$ are the most disastrous ones. The quantity of our interest is equal to

$$\psi_{i3}(t) = \sum_{n=0}^{\infty} Q_{i3}^{(n)}(t). \tag{5.2.4}$$

The corresponding empirical estimator is

$$\widehat{\psi}_{i3}(t, T) = \sum_{n=0}^{\infty} \widehat{Q}_{i3}^{(n)}(t, T). \tag{5.2.5}$$

The 95% confidence intervals for $\widehat{\psi}_{i3}(t, T)$ are defined as

$$\left[\widehat{\psi}_{i3}(t, T) - z(\alpha/2)\sigma_{i3}(t), \quad \widehat{\psi}_{i3}(t, T) + z(\alpha/2)\sigma_{i3}(t)\right], \quad t \geq 0, \tag{5.2.6}$$

with

$$\sigma_{i3}^2(t) = \sum_{k=1}^{3} \sum_{l=1}^{3} \mu_{kk} \left[(\psi_{ik} * \psi_{l3})^2 * Q_{kl} - (\psi_{ik} * \psi_{l3} * Q_{kl})^2\right](t). \tag{5.2.7}$$

In figure 5.3.1 (Votsi et al., [1]) we find the expected number of earthquake occurrences from any state $i \in E$ to state 3 with its 95% confidence intervals. The quantity $\widehat{\psi}_{i3}(t, T)$ is plotted against the time in months. We observe that the expected number of earthquake of the third state is the largest over time, when the initial state is equal to state 3. However, we can conclude that the expected number of earthquake occurrences do not significantly differ from each other when we change the initial state.

5.3 Rate of occurrences of earthquakes

We determine the rate of occurrences of earthquakes with $M \geq 6.1$. This is the probability that an earthquake of the third state, which is not necessarily the first, will occur in the next time interval. For the semi-Markov model we assume that the assumptions for which theorem 3.2.9 is true, hold.

We remember that the state space E can be partitioned into the two sets U and D, where U is the set of working states and D is the set of repair states. In our case, it holds true that $E = \{1, 2, 3\}$, $U = \{1, 2\}$ and $D = \{3\}$, because we are interested in the transitions of earthquakes from state 1 and state 2 to state 3. It is trivial that U and D form a partition of E.

It follows that the rate of occurrences of earthquakes of state 3 is equal to

$$ro(t) = \sum_{i \in \{1,2\}} \sum_{l=1}^{3} \pi_l \psi_{li} * q_{i3}(t). \tag{5.3.1}$$

The corresponding empirical estimator is equal to

$$\widehat{ro}(t, T) = \sum_{i \in \{1,2\}} \sum_{l=1}^{3} \pi_l \widehat{\psi}_{li} * \widehat{q}_{i3}(t, T). \tag{5.3.2}$$

In the section 3.2 we derived the $100(1-\alpha)\%$ confidence intervals for $\widehat{ro}(t)$ in general. In this particular case, the 95% confidence intervals for the rate of occurrences of earthquakes with $M \geq 6.1$ are defined as

$$\left[\widehat{ro}(t, T) - 1.96\sigma, \quad \widehat{ro}(t, T) + 1.96\sigma\right], \quad t \geq 0, \tag{5.3.3}$$

with

$$\sigma^2 = \sum_{i \in \{1,2\}} \mu_{ii} \frac{\sum_{l=1}^{3} \pi_l \psi_{li} * q_{i3}(t)}{T^{1-\alpha}} + \mathcal{O}(T^{-1}). \tag{5.3.4}$$

In figure 5.3.2 (Votsi et al., [1]) the rate of occurrences of earthquakes with $M \geq 6.1$ with its 95% confidence intervals is shown. The quantity $\widehat{ro}(t, T)$ is plotted against the time in months. We conclude that the rate increases as time goes by.

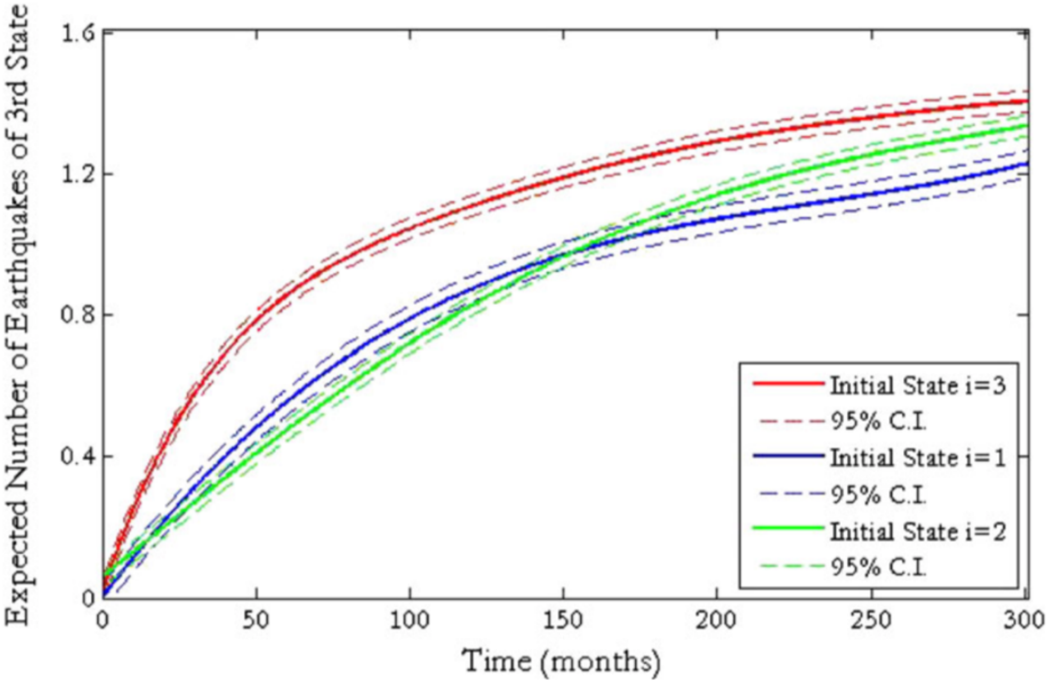

Figure 5.3.1: The 95% confidence intervals of the expected number of earthquake occurrences from any state $i \in E$ to state 3, $\widehat{\psi}_{i3}(t, T)$ for all $i \in E$. (Votsi et al., [1])

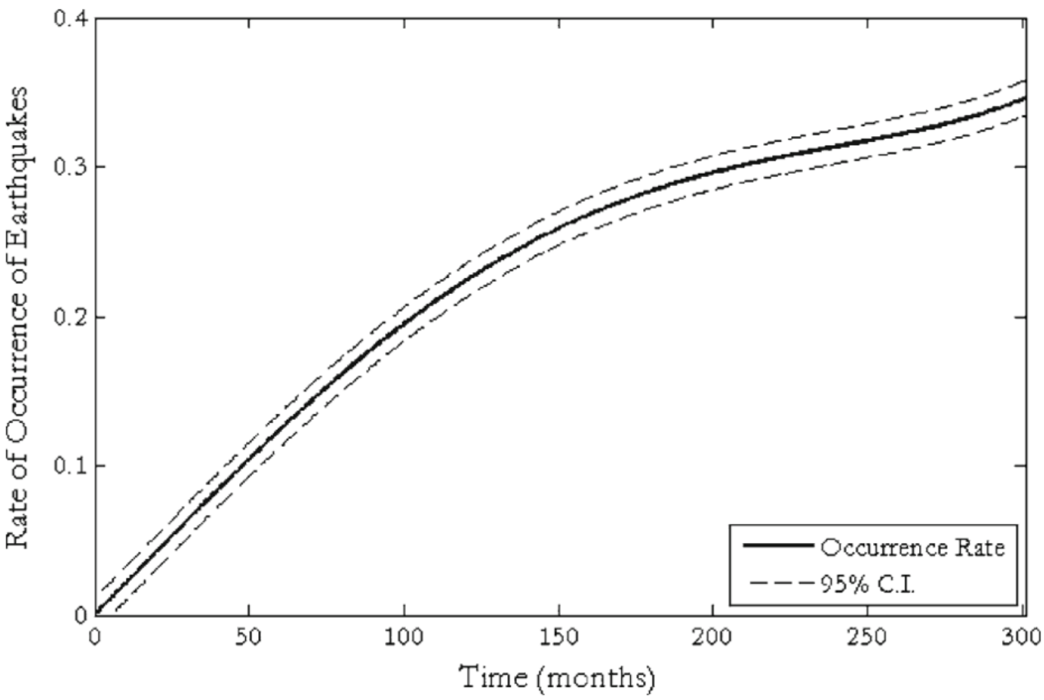

Figure 5.3.2: Rate of occurrences of earthquakes of state 3 ($M \geq 6.1$), $\widehat{ro}(t, T)$. (Votsi et al., [1])

Chapter 6

Comparison

In this chapter we make a comparison between the homogeneous Markov model and the semi-Markov model. We compute the semi-Markov kernels with use of sojourn time distributions and look at the Wald test for the hazard rate function of the semi-Markov process. We will also discuss parametric and non-parametric estimators.

We note that the main difference between the homogeneous Markov model and the semi-Markov model is that the homogeneous Markov model depends on the calendar time and the semi-Markov model depends on the sojourn time in the state. So we can also ask ourselves how well we can compare the two models with each other based on this fact.

6.1 Semi-Markov kernels

We determine the semi-Markov kernels for both the homogeneous Markov model and the semi-Markov model. It holds true that the semi-Markov kernels are defined as

$$Q_{ij}(t) = p_{ij}F_{ij}(t), \tag{6.1.1}$$

for all $i, j \in E$ and $t \geq 0$. Here, p_{ij} are the transition probabilities of the embedded Markov chain and $F_{ij}(t)$ are the probability distribution functions of the sojourn time.

In case of the homogeneous Markov model, we choose the exponential distribution as the sojourn time distribution. In this case, the probability distribution functions of the sojourn time are defined as

$$F_{ij}(t) = 1 - e^{-(t/\sigma_{ij})} \tag{6.1.2}$$

for all $i, j \in E$ and $t \geq 0$. Here, σ_{ij} is the scale parameter of the distribution.

For the semi-Markov model, we choose the Weibull distribution as the sojourn time distribution. In this case, the probability distribution functions of the sojourn time are defined as

$$F_{ij}(t) = 1 - e^{-(t/\sigma_{ij})^{\nu_{ij}}}, \tag{6.1.3}$$

for all $i, j \in E$ and $t \geq 0$. Here, σ_{ij} is the scale parameter and ν_{ij} is the shape parameter of the distribution.

From chapter 5 that the transition probability matrix is equal to

$$\widehat{P} = \begin{pmatrix} 0.4000 & 0.4000 & 0.2000 \\ 0.5556 & 0.2222 & 0.2222 \\ 0.5000 & 0.1250 & 0.3750 \end{pmatrix}. \tag{6.1.4}$$

In the following matrix we find the parameters σ_{ij} of the exponential distribution functions for the homogeneous Markov model.

$$\Sigma = \begin{pmatrix} - & 1.535 & 2.386 \\ 5.538 & - & 1.384 \\ 1.598 & 1.302 & - \end{pmatrix} \tag{6.1.5}$$

Then for all $i \neq j \in E$, the probability distribution functions of the sojourn time for the homogeneous Markov model are equal to

$$
\begin{aligned}
F_{12}(t) &= 1 - e^{-0.6515t}, & F_{13}(t) &= 1 - e^{-0.4191t}, \\
F_{21}(t) &= 1 - e^{-0.1806t}, & F_{23}(t) &= 1 - e^{-0.7225t}, \\
F_{31}(t) &= 1 - e^{-0.6258t}, & F_{32}(t) &= 1 - e^{-0.7680t}.
\end{aligned}
$$

We obtain the following semi-Markov kernels for the homogeneous Markov model with $i \neq j \in E$:

$$
\begin{aligned}
Q_{12}(t) &= 0.4000\left(1 - e^{-0.6515t}\right), & Q_{13}(t) &= 0.2000\left(1 - e^{-0.4191t}\right), \\
Q_{21}(t) &= 0.5556\left(1 - e^{-0.1806t}\right), & Q_{23}(t) &= 0.2222\left(1 - e^{-0.7225t}\right), \\
Q_{31}(t) &= 0.5000\left(1 - e^{-0.6258t}\right), & Q_{32}(t) &= 0.1250\left(1 - e^{-0.7680t}\right).
\end{aligned}
$$

We define $Q_{ii}(t) = 0$ for all $i \in E$ and $t \geq 0$, because we have no information about the sojourn time distributions of the transitions from state i to itself.

In figure B.0.5 in the appendix, the plots of the semi-Markov kernels are shown for the homogeneous Markov model for all transitions from state i to state j, $i \neq j \in E$.

For the semi-Markov model, in the following matrices we find the parameters σ_{ij} and ν_{ij} of the exponential and Weibull distribution functions.[1]

$$
\Sigma = \begin{pmatrix} - & 1.493 & 2.398 \\ 5.381 & - & 1.152 \\ 1.461 & 1.311 & - \end{pmatrix}, \quad V = \begin{pmatrix} - & 1.061 & 1.854 \\ 1.263 & - & 0.495 \\ 2.083 & - & - \end{pmatrix} \tag{6.1.6}
$$

The distribution functions of the sojourn time for the semi-Markov model, for all $i \neq j \in E$, are equal to

$$
\begin{aligned}
F_{12}(t) &= 1 - e^{-(0.6698t)^{1.061}}, & F_{13}(t) &= 1 - e^{-(0.4170t)^{1.854}}, \\
F_{12}(t) &= 1 - e^{-(0.1858t)^{1.263}}, & F_{13}(t) &= 1 - e^{-(0.8681t)^{0.495}}, \\
F_{12}(t) &= 1 - e^{-(0.6845t)^{2.083}}, & F_{13}(t) &= 1 - e^{-0.7628t}.
\end{aligned}
$$

Then, for the semi-Markov model, we obtain the following semi-Markov kernels for $i \neq j \in E$:

$$
\begin{aligned}
Q_{12}(t) &= 0.4000\left(1 - e^{-(0.6698t)^{1.061}}\right), & Q_{13}(t) &= 0.2000\left(1 - e^{-(0.4170t)^{1.854}}\right), \\
Q_{21}(t) &= 0.5556\left(1 - e^{-(0.1858t)^{1.263}}\right), & Q_{23}(t) &= 0.2222\left(1 - e^{-(0.8681t)^{0.495}}\right), \\
Q_{31}(t) &= 0.5000\left(1 - e^{-(0.6845t)^{2.083}}\right), & Q_{32}(t) &= 0.1250\left(1 - e^{-0.7628t}\right).
\end{aligned}
$$

We define $Q_{ii}(t) = 0$ for all $i \in E$ and $t \geq 0$, because we have no information about the sojourn time distributions of the transitions from state i to itself.

In figure B.0.6 in the appendix, the plots of the semi-Markov kernels are shown for the semi-Markov model for all transitions from state i to state j, $i \neq j \in E$.

[1] We note that the exponential distribution for the sojourn time distribution is chosen for the transition from state 3 to state 2.

6.2 Wald test and p-value

For each of the parameters of the hazard rate functions of the semi-Markov process, the R package SemiMarkov performed the Wald test. The Wald test gives us the relevance of the given distribution. In our case we test the distribution parameters σ_{ij} for $i, j \in E$ for the exponential distribution and σ_{ij}, ν_{ij} for $i, j \in E$ for the Weibull distribution. We have the following hypothesis test for the scale parameter σ_{ij} (Listwon & Saint-Pierre, [12]):

$$\begin{cases} H_0 : \sigma_{ij} = 1, \\ H_1 : \sigma_{ij} \neq 1. \end{cases} \tag{6.2.1}$$

Similarly, we have the hypothesis test for the shape parameter ν_{ij}:

$$\begin{cases} H_0 : \nu_{ij} = 1, \\ H_1 : \nu_{ij} \neq 1. \end{cases} \tag{6.2.2}$$

The p-value illustrates when we can reject the null-hypothesis H_0. It is defined to be the smallest significance level at which the null hypothesis is rejected (Rice, [6]). If $p \leq 0.05$, we reject the null-hypothesis H_0. If $p > 0.05$, we fail to reject H_0.

First we look at the homogeneous Markov model. Here, we choose the exponential distribution for the sojourn time of the process. In figure 6.2.1 the results of the Wald test are shown.

$Sigma

	Type	Index	Transition	Estimation	SD	Lower_CI	Upper_CI	Wald_H0	Wald_test	p_value
1	dist	1	1 -> 2	1.535	0.66	0.25	2.82	1.00	0.66	0.4166
2	dist	2	1 -> 3	2.386	1.43	-0.42	5.19	1.00	0.94	0.3323
3	dist	3	2 -> 1	5.538	2.49	0.66	10.41	1.00	3.33	0.0680
4	dist	4	2 -> 3	1.384	1.13	-0.82	3.59	1.00	0.12	0.7290
5	dist	5	3 -> 1	1.598	0.84	-0.05	3.24	1.00	0.51	0.4751
6	dist	6	3 -> 2	1.302	1.65	-1.94	4.54	1.00	0.03	0.8625

Figure 6.2.1: Wald test p-values for the homogeneous Markov model.

We derive that for all σ_{ij}, $i \neq j \in E$, we fail to reject the null-hypothesis. For the semi-Markov model the results of the Wald test are shown in figure 6.2.2. We remember that for the transitions from state 3 to state 2 we chose an exponential distribution for the sojourn time instead of the Weibull distribution. So we exclude this transition from our conclusions.

$Sigma

	Type	Index	Transition	Sigma	SD	Lower_CI	Upper_CI	Wald_H0	Wald_test	p_value
1	dist	1	1 -> 2	1.493	0.60	0.31	2.67	1.00	0.67	0.4131
2	dist	2	1 -> 3	2.398	0.77	0.89	3.90	1.00	3.31	0.0689
3	dist	3	2 -> 1	5.381	1.96	1.54	9.22	1.00	5.01	0.0252
4	dist	4	2 -> 3	1.152	2.10	-2.97	5.27	1.00	0.01	0.9203
5	dist	5	3 -> 1	1.461	0.37	0.73	2.19	1.00	1.55	0.2131
6	dist	6	3 -> 2	1.311	1.68	-1.98	4.61	1.00	0.03	0.8625

$Nu

	Type	Index	Transition	Nu	SD	Lower_CI	Upper_CI	Wald_H0	Wald_test	p_value
1	dist	7	1 -> 2	1.061	0.37	0.33	1.79	1.00	0.03	0.8625
2	dist	8	1 -> 3	1.854	0.76	0.36	3.35	1.00	1.25	0.2636
3	dist	9	2 -> 1	1.263	0.47	0.33	2.19	1.00	0.31	0.5777
4	dist	10	2 -> 3	0.495	0.31	-0.11	1.10	1.00	2.64	0.1042
5	dist	11	3 -> 1	2.083	0.84	0.44	3.73	1.00	1.67	0.1963

Figure 6.2.2: Wald test p-values for the semi-Markov model.

In this case we only reject the null-hypothesis for the scale parameter σ_{12} which is associated with the transition from state 2 to state 1. In the other case we fail to reject the null-hypothesis for σ_{ij} and ν_{ij}. We observe that ν_{12} and ν_{21} do not significantly differ from the value 1. For these transitions we can use the exponential distribution instead of the Weibull distribution.

We conclude that for some of the hazard rate functions of the semi-Markov process $\lambda_{ij}(t)$ for $i, j \in E$ the homogeneous Markov model maybe a better fit. For the rest of the hazard rate functions of the semi-Markov process, we cannot conclude a preference for a certain model based on the p-values of the Wald test.

6.3 Parametric and non-parametric estimators

As we promised in section 5.1, we will discuss the difference between parametric and non-parametric estimators in this section (Rice, [6]). Suppose that we want to fit a probability law into the given data. Then usually we have to estimate certain parameters associated with the probability law. Based on some scientific theory, one may suggest the form of a probability distribution and the parameters that are of interest to the scientific investigation. If we assume that the given data follows this probability law, then we use parametric estimation. It holds true that non-parametric methods of estimation do not assume that the given data follows a particular distribution.

We have used the parametric method to estimate the hazard rate functions of the semi-Markov process in the last chapter. We assumed that the sojourn time in the states of the process were exponentially distributed or Weibull distributed. The empirical estimators we defined in section 3.2 are non-parametric.

Chapter 7

Application to dataset with sub areas

In this chapter we apply the semi-Markov model to a new dataset of the region of the Northern Aegean Sea in Greece. This time, we are interested in the addition of covariates to the semi-Markov model. A covariate is an independent observed variable that can have an influence to the accuracy of a model and its results. In our case, we want to examine the influence of the location where earthquakes took place on the semi-Markov model.

We use the dataset given in table B.0.2 (Votsi et al., [1]), see the appendix. This time, the after shocks of the earthquakes are not filtered out of the dataset. It consists of 67 earthquakes of magnitude $M \geq 5.2$ that occurred since 1964. We define two states of earthquakes corresponding to the magnitude:

State 1: $[5.2, 5.5]$.

State 2: $[5.6, 7.2]$.

We derive that the state space is equal to the set $E = \{1, 2\}$.

The covariate that we add to the model is the location where the earthquakes occurred, as we said before. We divided the Northern Aegean region into four sub areas, which are shown in figure B.0.7 (Votsi et al., [1]). For each earthquake that occurred, we denoted the date and time of occurrence and also the magnitude M, state and sub area.

We let $t = 0$ be the time that we observed the first earthquake. This earthquake took place at February 23, 1964. It is unknown how long the first earthquake was in its state. We set the end time T later than the last earthquake we observed, which was at December 21, 2006. So we set $T = 33$ years.

Now we are going to apply the dataset to the semi-Markov model in continuous-time. The number of observed transitions in the dataset from any state i to any state j are presented as elements in the matrix N. The elements of this matrix are the values $N_{ij}(T)$ for all $i, j \in E$.

$$N = \begin{pmatrix} 31 & 13 \\ 13 & 9 \end{pmatrix} \tag{7.0.1}$$

The values $N_i(T)$ for $i \in E$ are equal to

$$N_1(T) = 44, \qquad N_2(T) = 22. \tag{7.0.2}$$

In the following matrix $\widehat{P} = (\widehat{p}_{ij})$ the estimations of the transition probabilities are presented as elements.

$$\widehat{P} = \begin{pmatrix} 0.7045 & 0.2955 \\ 0.5909 & 0.4091 \end{pmatrix} \tag{7.0.3}$$

In figure 7.0.1 the transitions from state i to state j with the corresponding transition probabilities are shown for all $i, j \in E$.

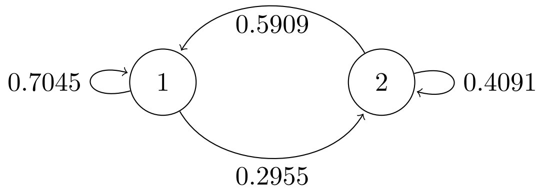

Figure 7.0.1: Transitions from state i to state j with transition probabilities, for all $i, j \in E$.

7.1 Hazard rate function

We examine the influence of covariates on the semi-Markov model, with use of the hazard rate function of the semi-Markov process. We remember that we can estimate the hazard rate function of the semi-Markov process with the following empirical estimator

$$\widehat{\lambda}_{ij}(t,T) = \begin{cases} \dfrac{\widehat{q}_{ij}(t,T)}{1-\widehat{H}_i(t,T)} & \text{if } i \neq j, \\ -\displaystyle\sum_{j\neq i} \widehat{\lambda}_{ij}(t,T) & \text{if } i = j. \end{cases} \tag{7.1.1}$$

With use of the R package SemiMarkov (Listwon & Saint-Pierre, [12]) we estimate the hazard rate functions of the semi-Markov process for all transitions from state i to state j, $i \neq j \in E$. We cannot derive the hazard rate functions for the transition from state i to itself, because by definition of the semi-Markov model the waiting times must be known. If these kind of transitions are observed in the package, the transition is combined with the next transition such that we obtain a transition from state i to state j for $i \neq j \in E$.

We determine the hazard rate functions of the semi-Markov process. We used the data from table B.0.2 (Votsi et al., [1]) and called it *semimarkov*2. We choose the Weibull distribution W for the sojourn time for all transitions. For the covariate, we choose the sub areas from the dataset which column is called *subareas*. We fit the data with use of the function *semiMarkov*(.). Figure 7.1.1 shows the performed steps in Rstudio.

In figure 7.1.2, we find the estimates of parameters of the waiting time distributions, the standard deviations, the confidence intervals and the Wald test statistics. In the following two matrices $\Sigma = (\sigma_{ij})$ and $V = (\nu_{ij})$ we give the values for the parameters of the Weibull distribution for all $i \neq j \in E$:

$$\Sigma = \begin{pmatrix} - & 4.772 \\ 8.028 & - \end{pmatrix}, \quad V = \begin{pmatrix} - & 0.494 \\ 0.259 & - \end{pmatrix}. \tag{7.1.2}$$

The R package SemiMarkov can determine two hazard rate functions, namely the hazard rate function of the waiting time $\alpha_{ij}(t)$ and the hazard rate function of the semi-Markov process $\lambda_{ij}(t)$. First, we give the hazard rate function of the waiting time and after that the hazard rate function of the semi-Markov process.

In the case that the sojourn time is Weibull distributed with scale parameter σ_{ij} and shape parameter ν_{ij}, the hazard rate functions of the waiting time are defined as (Listwon & Saint-Pierre, [12])

$$\alpha_{ij}(t) = \frac{\nu_{ij}}{\sigma_{ij}} \left(\frac{t}{\sigma_{ij}}\right)^{\nu_{ij}-1}, \tag{7.1.3}$$

for all $i, j \in E$ and $t \geq 0$.

The estimated hazard rate functions of the waiting time $\alpha_{ij}(t)$ without covariates for all $i \neq j \in E$ are equal to

$$\alpha_{12}(t) = 0.1035(0.2096t)^{-0.506}, \qquad \alpha_{21}(t) = 0.0323(0.1246t)^{-0.741}.$$

The density functions of the sojourn time with scale parameter σ_{ij} and shape parameter ν_{ij} are defined as

$$f_{ij}(t) = \frac{\nu_{ij}}{\sigma_{ij}} \left(\frac{t}{\sigma_{ij}}\right)^{\nu_{ij}-1} e^{-(t/\sigma_{ij})^{\nu_{ij}}}, \tag{7.1.4}$$

for all $i, j \in E$ and $t \geq 0$. We obtain the following estimated density functions $f_{ij}(t)$ for all $i \neq j \in E$:

$$f_{12}(t) = 0.1035(0.2096t)^{-0.506} e^{-(0.2096t)^{0.494}}, \qquad f_{21}(t) = 0.0323(0.1246t)^{-0.741} e^{-(0.1246t)^{0.259}}.$$

The probability distribution of the sojourn time with scale parameter σ_{ij} and shape parameter ν_{ij} are defined as

$$F_{ij}(t) = 1 - e^{-(t/\sigma_{ij})^{\nu_{ij}}}, \tag{7.1.5}$$

for all $i, j \in E$ and $t \geq 0$. The estimated probability distribution functions of the sojourn time for all $i \neq j \in E$ are

$$F_{12}(t) = 1 - e^{-(0.2096t)^{0.494}}, \qquad F_{21}(t) = 1 - e^{-(0.1246t)^{0.259}}.$$

We know from the beginning of this chapter that the transition probability matrix is given by

$$\widehat{P} = \begin{pmatrix} 0.7045 & 0.2955 \\ 0.5909 & 0.4091 \end{pmatrix}. \tag{7.1.6}$$

Define $G_i(t) = 1 - H_i(t) = \sum_{j\in E} p_{ij}(1 - F_{ij}(t))$ as the survival function of the sojourn time in state i. Here, $F_{ij}(t)$ is the probability distribution function of the sojourn time and p_{ij} the transition probability of the embedded Markov chain. For the derivative of the semi-Markov kernel, we know that $q_{ij}(t) = p_{ij} f_{ij}(t)$. With use of this information, we can determine the hazard rate function of the semi-Markov process $\lambda_{ij}(t)$.

If we plug the expressions for the density functions of the sojourn time, the probability distribution functions of the sojourn time and the transition probabilities into the hazard rate function of the semi-Markov process

$$\lambda_{ij}(t) = \frac{q_{ij}(t)}{1 - \widehat{H}_i(t)} = \frac{p_{ij} f_{ij}(t)}{\sum_{j\in E} p_{ij}(1 - F_{ij}(t))}, \tag{7.1.7}$$

we obtain the estimated hazard rate functions of the semi-Markov process for all $i \neq j \in E$. It holds true that $\lambda_{ij}(t) = -\sum_{j\neq i} \lambda_{ij}(t)$ for all $i \in E$.

We can study the influence of covariates on the sojourn time distributions of the semi-Markov process with use of a Cox proportional regression model (Listwon & Saint-Pierre, [12]). Let Z_{ij} be a vector of explanatory variables and β_{ij} a vector of regression parameters associated with the transition from state i to state j. Then the hazard rate is defined as

$$\alpha_{ij}(t \mid Z_{ij}) = \alpha_{ij}(t) e^{\beta_{ij}^T Z_{ij}}, \tag{7.1.8}$$

for all $t \geq 0$ and $i \neq j \in E$. The regression coefficients can be interpreted in terms of relative risk. It holds true that $\alpha_{ij}(t)$ is the baseline hazard as defined in (7.1.3).

We look back at the hazard rate function of the waiting time. Now, we add the covariate *subareas* to the model. Let x be a explanatory variable and β_{ij} the regression parameter associated with the transition from state i to state j. Then the hazard rate functions of the waiting time are defined as

$$\alpha_{ij}(t \mid x) = \alpha_{ij}(t) e^{\beta_{ij} x}. \tag{7.1.9}$$

The explanatory variable x takes on the values of the defined sub areas, so $x \in \{1, 2, 3, 4\}$. From figure 7.1.2, we derive that

$$\beta_{12} = 0.4500, \quad \beta_{21} = 0.2962. \tag{7.1.10}$$

Then the estimated hazard rate functions of the waiting time $\alpha_{ij}(t \mid x)$ with the covariate *subareas* are equal to

$$\alpha_{12}(t \mid x) = 0.1035(0.2096t)^{-0.506} e^{0.4500x},$$

$$\alpha_{21}(t \mid x) = 0.0323(0.1246t)^{-0.741} e^{0.2962x},$$

with $x \in \{1, 2, 3, 4\}$.

The plots of the hazard functions of the waiting time with and without the covariate are shown in figure B.0.8 in the appendix.

For the hazard rate functions of the semi-Markov process, according to the Cox proportional model, it holds true that the functions with the covariate *subareas* for all $t \geq 0$ and $i \neq j \in E$ are defined as

$$\lambda_{ij}(t \mid x) = \lambda_{ij}(t) e^{\beta_{ij} x}, \tag{7.1.11}$$

where x is the explanatory variable and β_{ij} the regression parameter associated with the transition from state i to state j. Because we defined four sub areas, it holds true that $x \in \{1, 2, 3, 4\}$. With use of the hazard rate functions of the semi-Markov process defined in (7.1.7) and the regression parameters from (7.1.10), we can obtain the hazard rate functions of the semi-Markov process with covariates.

The plots of the hazard rate functions of the semi-Markov process with and without the covariate are shown in figure B.0.9 in the appendix. We see that for all values of the covariate, sub areas 1 to 4, the hazard rate functions of the semi-Markov process for all transitions from state i to state j, $i \neq j \in E$, are not significantly different.

```
> library(numDeriv)
> library(MASS)
> library(Rsolnp)
> library(SemiMarkov)
> semimarkov2 <- read.delim(file.choose(), header = TRUE)
> View(semimarkov2)
> states <- c("1", "2")
> mtrans <- matrix(FALSE, nrow = 2, ncol = 2)
> mtrans[1, 2] <- c("W")
> mtrans[2, 1] <- c("W")
> fit <- semiMarkov(data = semimarkov2, states = states, mtrans = mtrans, cov = as.data.frame(s
emimarkov2$subarea))

Iter: 1 fn: -9.3729      Pars:  4.77173 8.02757 0.49412 0.25913 0.45006 0.29615
Iter: 2 fn: -9.3729      Pars:  4.77174 8.02755 0.49412 0.25913 0.45006 0.29616
solnp--> Completed in 2 iterations
```

Figure 7.1.1: The setting of the hazard rate function.

```
> fit$table.dist
$Sigma
  Type Index Transition Sigma   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2 4.772 2.33     0.20     9.35    1.00      2.61  0.1062
2 dist     2     2 -> 1 8.028 6.46    -4.63    20.68    1.00      1.18  0.2774

$Nu
  Type Index Transition    Nu   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     3     1 -> 2 0.494 0.12     0.26     0.73    1.00     17.27 <0.0001
2 dist     4     2 -> 1 0.259 0.06     0.14     0.38    1.00    155.05 <0.0001

> fit$table.coef
  Type Index Transition Covariates Estimation   SD Lower_CI Upper_CI Wald_H0 Wald_test
1 coef     1     1 -> 2      Beta1   0.450060 0.17     0.12     0.78    0.00      7.06
2 coef     2     2 -> 1      Beta1   0.296155 0.11     0.08     0.51    0.00      7.23
  p_value
1  0.0079
2  0.0072
```

Figure 7.1.2: Estimates of parameters of the waiting time distribution.

7.2 Semi-Markov kernels

We determine the semi-Markov kernels for the semi-Markov process. Because we do not have a definition about the influence of covariates on the semi-Markov kernel, we will only compute the semi-Markov kernels without covariates. It holds true that the semi-Markov kernels are defined as

$$Q_{ij}(t) = p_{ij} F_{ij}(t), \tag{7.2.1}$$

for all $i, j \in E$ and $t \geq 0$. Here, p_{ij} are the transition probabilities of the embedded Markov chain and $F_{ij}(t)$ are the probabilities distribution functions of the sojourn time.

For the semi-Markov model, we choose the Weibull distribution as the sojourn time distribution. In this case, the probability distribution functions of the sojourn time are defined as

$$F_{ij}(t) = 1 - e^{-(t/\sigma_{ij})^{\nu_{ij}}}, \tag{7.2.2}$$

for all $i, j \in E$ and $t \geq 0$. Here, σ_{ij} is the scale parameter and ν_{ij} is the shape parameter of the distribution.

In the beginning of this chapter, we obtained the following transition probability matrix

$$\widehat{P} = \begin{pmatrix} 0.7045 & 0.2955 \\ 0.5909 & 0.4091 \end{pmatrix}. \tag{7.2.3}$$

In the following two matrices we find the parameters σ_{ij} and ν_{ij} of the Weibull distribution functions.

$$\Sigma = \begin{pmatrix} - & 4.772 \\ 8.028 & - \end{pmatrix}, \quad V = \begin{pmatrix} - & 0.494 \\ 0.259 & - \end{pmatrix} \tag{7.2.4}$$

Then for all $i \neq j \in E$, the probability distribution functions of the sojourn time are equal to

$$F_{12}(t) = 1 - e^{-(0.2096t)^{0.494}}, \qquad F_{21}(t) = 1 - e^{-(0.1246t)^{0.259}}.$$

Because the transition probabilities from state i to state j for all $i \neq j \in E$ are equal to 1, we obtain the following semi-Markov kernels for $i \neq j \in E$:

$$Q_{12}(t) = 0.2955 \left(1 - e^{-(0.2096t)^{0.494}}\right), \qquad Q_{21}(t) = 0.5909 \left(1 - e^{-(0.1246t)^{0.259}}\right).$$

We define $Q_{ii}(t) = 0$ for all $i \in E$ and $t \geq 0$, because we have no information about the sojourn time distributions of the transitions from state i to itself.

In figure B.0.10 the plots of the semi-Markov kernels are shown for all transitions from state i to state j, $i \neq j \in E$.

7.3 Wald test and p-value

For each of the parameters of the hazard rate functions of the semi-Markov process, the R package SemiMarkov performed the Wald test. We test the distribution parameters σ_{ij}, ν_{ij} for $i, j \in E$ for the Weibull distribution. We have the following hypothesis test for the scale parameter σ_{ij}:

$$\begin{cases} H_0 : \sigma_{ij} = 1, \\ H_1 : \sigma_{ij} \neq 1. \end{cases} \tag{7.3.1}$$

Similarly, we have the hypothesis test for the shape parameter ν_{ij}:

$$\begin{cases} H_0 : \nu_{ij} = 1, \\ H_1 : \nu_{ij} \neq 1. \end{cases} \tag{7.3.2}$$

The p-value illustrates when we can reject the null-hypothesis H_0. It is defined to be the smallest significance level at which the null hypothesis is rejected (Rice, [6]). If $p \leq 0.05$, we reject the null-hypothesis H_0. If $p > 0.05$, we fail to reject H_0.

For the covariates in the model, the R package SemiMarkov performs a Wald test as well. The Wald test examines if the addition of a covariate is an extension to the semi-Markov model. The hypothesis test for the regression coefficient β_{ij} is equal to (Listwon & Saint-Pierre, [12])

$$\begin{cases} H_0 : \beta_{ij} = 0, \\ H_1 : \beta_{ij} \neq 0. \end{cases} \tag{7.3.3}$$

In figure 7.3.1 the results of the Wald test are shown.

```
> fit$table.dist
$Sigma
  Type Index Transition Sigma   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2 4.772 2.33     0.20     9.35    1.00      2.61  0.1062
2 dist     2     2 -> 1 8.028 6.46    -4.63    20.68    1.00      1.18  0.2774

$Nu
  Type Index Transition    Nu   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     3     1 -> 2 0.494 0.12     0.26     0.73    1.00     17.27 <0.0001
2 dist     4     2 -> 1 0.259 0.06     0.14     0.38    1.00    155.05 <0.0001

> fit$table.coef
  Type Index Transition Covariates Estimation   SD Lower_CI Upper_CI Wald_H0 Wald_test
1 coef     1     1 -> 2      Beta1   0.450060 0.17     0.12     0.78    0.00      7.06
2 coef     2     2 -> 1      Beta1   0.296155 0.11     0.08     0.51    0.00      7.23
  p_value
1  0.0079
2  0.0072
```

Figure 7.3.1: Wald test p-values for the semi-Markov model with covariates.

We derive that for all σ_{ij}, $i \neq j \in E$, we fail to reject the null-hypothesis. However for all ν_{ij}, $i \neq j \in E$, we do reject the null hypothesis. From this we can conclude that we cannot choose an exponential distribution for the sojourn time instead of the Weibull distribution. So the semi-Markov model is a good fit in this case.

Now, we look at the p-values of the regression parameters β_{ij} associated with the transition from state i to state j for all $i \neq j \in E$. We can read that for all β_{ij}, we reject the null-hypothesis. We conclude that the covariate *subareas* is an extension to the semi-Markov model which leads to a better fit according to the Wald test.

Chapter 8

Conclusions

As we mentioned before, the process of earthquake occurrences can be represented with several statistical models. We can achieve forecasting results with use of the semi-Markov model in the high seismically active region of the Northern Aegean Sea in Greece.

In this thesis we explained the semi-Markov model. We defined empirical estimators of important quantities to the semi-Markov model and proved some theses about the asymptotic behaviour of these quantities. We estimated the relevant quantities to the semi-Markov model, such as the semi-Markov kernels and the distribution functions associated with the sojourn time. Furthermore, we have shown how to determine the expected number of earthquake occurrences and the rate of occurrences of earthquakes with $M \geq 6.1$. We computed the transition probabilities of the embedded Markov chain as well.

We determined the hazard rate functions of the semi-Markov process in a parametric way and compared the semi-Markov model with the homogeneous Markov model. We concluded that it is difficult to derive a preference for a certain model based on the Wald test.

Lastly, we applied the semi-Markov model to a new dataset of the Northern Aegean Sea region to derive the consequence of adding covariates. We derived that according to the Wald test, the addition of sub areas is an extension to the model.

Appendix A

Proofs

In this chapter we give proofs of some theses that we introduced in this thesis.

A.1 Theorem 3.1.1

Proof. (Grabski, [5]) From the definition of conditional probabilities, it follows that

$$
\begin{aligned}
F_{ij}(t) &= P\left(X_{n+1} \leq t \mid J_n = i, J_{n+1} = j\right) \\
&= \frac{P\left(X_{n+1} \leq t, J_n = i, J_{n+1} = j\right)}{P\left(J_n = i, J_{n+1} = j\right)} \\
&= \frac{P\left(X_{n+1} \leq t, J_n = i, J_{n+1} = j\right)}{P\left(J_n = i\right)} \cdot \frac{P\left(J_n = i\right)}{P\left(J_n = i, J_{n+1} = j\right)} \\
&= \frac{P\left(J_{n+1} = j, X_{n+1} \leq t \mid J_n = i\right)}{P\left(J_{n+1} = j \mid J_n = i\right)} \\
&= \frac{Q_{ij}(t)}{p_{ij}}.
\end{aligned}
$$

∎

A.2 Theorem 3.2.3

Proof. (Limnios & Oprişan, [7]) It holds true that $Q_{ij}(t) = F_{ij}(t)p_{ij}$ and therefore $\widehat{Q}_{ij}(t,T) = \widehat{F}_{ij}(t,T)\widehat{p}_{ij}(T)$ as well. Then it follows that

$$\begin{aligned}
\max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{Q}_{ij}(t,T) - Q_{ij}(t)\right| &= \max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{F}_{ij}(t,T)\widehat{p}_{ij}(T) - F_{ij}(t)p_{ij}\right| \\
&= \max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{F}_{ij}(t,T)\widehat{p}_{ij}(T) - \widehat{F}_{ij}(t,T)p_{ij} + \widehat{F}_{ij}(t,T)p_{ij} - F_{ij}(t)p_{ij}\right| \\
&\leq \max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{F}_{ij}(t,T)\widehat{p}_{ij}(T) - \widehat{F}_{ij}(t,T)p_{ij}\right| + \max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{F}_{ij}(t,T)p_{ij} - F_{ij}(t)p_{ij}\right| \\
&= \max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{F}_{ij}(t,T)\left(\widehat{p}_{ij}(T) - p_{ij}\right)\right| + \max_{i,j\in E}\sup_{t\in[0,T)}\left|\left(\widehat{F}_{ij}(t,T) - F_{ij}(t)\right)p_{ij}\right| \\
&= \max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{p}_{ij}(T) - p_{ij}\right|\widehat{F}_{ij}(t,T) + \max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{F}_{ij}(t,T) - F_{ij}(t)\right|p_{ij}.
\end{aligned}$$

By theorem 3.2.2, the first term converges to 0 (a.s.). By theorem 3.2.1 (Glivenko-Cantelli theorem), the second converges to 0 (a.s.) as well. ∎

A.3 Theorem 3.2.5

Proof. (Barbu & Limnios, [11]) By induction. For the case $n = 1$, the result follows from theorem 3.2.3.

Assume that it holds true for $n = m$. So

$$\max_{i,j\in E} \max_{t\in[0,T]} \left| \widehat{Q}_{ij}^{(m)}(t,T) - Q_{ij}^{(m)}(t) \right| \to 0 \quad \text{(a.s.)}$$

as $T \to \infty$.

Now, let $n = m + 1$. It follows that

$$\begin{aligned}
\max_{i,j\in E} \max_{t\in[0,T]} \left| \widehat{Q}_{ij}^{(m+1)}(t,T) - Q_{ij}^{(m+1)}(t) \right| &= \max_{i,j\in E} \max_{t\in[0,T]} \left| \sum_{k\in E} \widehat{Q}_{ik}(t,T) * \widehat{Q}_{kj}^{(m)}(t,T) - \sum_{k\in E} Q_{ik}(t) * Q_{ik}^{(m)}(t) \right| \\
&= \max_{i,j\in E} \max_{t\in[0,T]} \left| \sum_{k\in E} \widehat{Q}_{ik}(t,T) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * Q_{ik}^{(m)}(t) \right| \\
&\leq \max_{i,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \left| \widehat{Q}_{ik}(t,T) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * Q_{ik}^{(m)}(t) \right| \\
&= \max_{i,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \Big| \widehat{Q}_{ik}(t,T) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * \widehat{Q}_{kj}^{(m)}(t,T) \\
&\qquad + Q_{ik}(t) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * Q_{ik}^{(m)}(t) \Big| \\
&= \max_{i,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \Big| \widehat{Q}_{ik}(t,T) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * \widehat{Q}_{kj}^{(m)}(t,T) \\
&\qquad + Q_{ik}(t) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * Q_{ik}^{(m)}(t) \Big| \\
&\leq \max_{i,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \left| \widehat{Q}_{ik}(t,T) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * \widehat{Q}_{kj}^{(m)}(t,T) \right| \\
&+ \max_{i,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \left| Q_{ik}(t) * \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}(t) * Q_{ik}^{(m)}(t) \right| \\
&= \max_{i,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \left| \left[\widehat{Q}_{ik}(t,T) - Q_{ik}(t) \right] * \widehat{Q}_{kj}^{(m)}(t,T) \right| \\
&+ \max_{i,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \left| Q_{ik}(t) * \left[\widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}^{(m)}(t) \right] \right| \\
&\leq \max_{i,k\in E} \max_{t\in[0,T]} \left| \widehat{Q}_{ik}(t,T) - Q_{ik}(t) \right| \max_{k,j\in E} \max_{t\in[0,T]} \sum_{k\in E} \widehat{Q}_{kj}^{(m)}(t,T) \\
&+ \max_{i,k\in E} \max_{t\in[0,T]} \left| \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}^{(m)}(t) \right| \max_{k,j\in E} \max_{t\in[0,T]} \sum_{k\in E} Q_{ik}(t) \\
&\leq s \max_{i,k\in E} \max_{t\in[0,T]} \left| \widehat{Q}_{ik}(t,T) - Q_{ik}(t) \right| \\
&+ \max_{i,k\in E} \max_{t\in[0,T]} \left| \widehat{Q}_{kj}^{(m)}(t,T) - Q_{ik}^{(m)}(t) \right|.
\end{aligned}$$

The last step holds true, because $E = \{1, 2, \ldots, s\}$. By theorem 3.2.3, the first converges to 0 (a.s.). By the induction hypothesis, the second term converges to 0 (a.s.) as well. The result follows from the principle of mathematical induction. ■

A.4 Theorem 3.2.6

Proof. (Barbu & Limnios, [11]) It holds true that

$$\begin{aligned}\max_{i,j\in E}\sup_{t\in[0,T)}\left|\widehat{\psi}_{ij}^{(k)}(t,T)-\psi_{ij}^{(k)}(t)\right| &= \max_{i,j\in E}\sup_{t\in[0,T)}\left|\sum_{n=0}^{k}\widehat{Q}_{ij}^{(n)}(t,T)-\sum_{n=0}^{k}Q_{ij}^{(n)}(t)\right| \\ &= \max_{i,j\in E}\sup_{t\in[0,T)}\left|\sum_{n=0}^{k}\widehat{Q}_{ij}^{(n)}(t,T)-Q_{ij}^{(n)}(t)\right| \\ &\leq \max_{i,j\in E}\sup_{t\in[0,T)}\sum_{n=0}^{k}\left|\widehat{Q}_{ij}^{(n)}(t,T)-Q_{ij}^{(n)}(t)\right|.\end{aligned}$$

By theorem 3.2.5, this converges to 0 (a.s.). ■

A.5 Theorem 3.2.10

Proof. (Ouhbi & Limnios, [2]) It holds true that

$$
\begin{aligned}
\sup_{t\in[0,M]} |\widehat{ro}(t,T) - ro(t)| &= \sup_{t\in[0,M]} \left| \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \widehat{\psi}_{li}(t,T) * \widehat{q}_{ij}(t,T) - \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \psi_{li}(t) * q_{ij}(t) \right| \\
&= \sup_{t\in[0,M]} \left| \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left[\widehat{\psi}_{li}(t,T) * \widehat{q}_{ij}(t,T) - \psi_{li}(t) * q_{ij}(t) \right] \right| \\
&\leq \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{\psi}_{li}(t,T) * \widehat{q}_{ij}(t,T) - \psi_{li}(t) * q_{ij}(t) \right| \\
&= \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{\psi}_{li}(t,T) * \widehat{q}_{ij}(t,T) - \widehat{\psi}_{li}(t,T) * q_{ij}(t) \right. \\
&\qquad\qquad \left. + \widehat{\psi}_{li}(t,T) * q_{ij}(t) - \psi_{li}(t) * q_{ij}(t) \right| \\
&\leq \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{\psi}_{li}(t,T) * \widehat{q}_{ij}(t,T) - \widehat{\psi}_{li}(t,T) * q_{ij}(t) \right| \\
&+ \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{\psi}_{li}(t,T) * q_{ij}(t) - \psi_{li}(t) * q_{ij}(t) \right| \\
&= \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{\psi}_{li}(t,T) * [\widehat{q}_{ij}(t,T) - q_{ij}(t)] \right| \\
&+ \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \left[\widehat{\psi}_{li}(t,T) - \psi_{li}(t) \right] * q_{ij}(t) \right|.
\end{aligned}
$$

Define $\psi_T := \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D} \widehat{\psi}_{ij}(t,T) < \infty$. Then we obtain

$$
\begin{aligned}
\sup_{t\in[0,M]} |\widehat{ro}(t,T) - ro(t)| &\leq \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{q}_{ij}(t,T) - q_{ij}(t) \right| \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D} \widehat{\psi}_{li}(t,T) \\
&+ \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{\psi}_{li}(t,T) - \psi_{li}(t)) \right| \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D} q_{ij}(t) \\
&\leq \psi_T \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D} \left| \widehat{q}_{ij}(t,T) - q_{ij}(t) \right| \\
&+ \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D}\sum_{l=1}^{s} \pi_l \left| \widehat{\psi}_{li}(t,T) - \psi_{li}(t)) \right| \sup_{t\in[0,M]} \sum_{i\in U}\sum_{j\in D} q_{ij}(t).
\end{aligned}
$$

By theorem 3.2.3, it follows that the first term converges to 0 (a.s.). By theorem 3.2.7 the second term converges to 0 (a.s.) as well. ∎

A.6 Theorem 4.0.1

Proof. (Ouhbi & Limnios, [2]) Let $\Lambda = (\lambda_{ij})$ be the generating matrix of the Markov process. It follows from lemma 4.0.1 that the semi-Markov kernel is equal to $Q_{ij}(t) = p_{ij}\left(1 - e^{-\lambda_i t}\right)$, where $\lambda_i = -\lambda_{ii}$ and $p_{ij} = \lambda_{ij}/\lambda_i$. For the distribution function of the sojourn time in state i, it holds true that

$$\begin{aligned} H_i(t) &= \sum_{j \in E} Q_{ij}(t) \\ &= 1 - e^{-\lambda_i t}, \end{aligned}$$

because $\sum_{j \in E} p_{ij} = 1$. We derive that

$$\begin{aligned} q_{ij}(t) &= Q'_{ij}(t) \\ &= -p_{ij} \cdot -\lambda_i \cdot e^{-\lambda_i t} \\ &= \lambda_{ij} e^{-\lambda_i t} \\ &= \lambda_{ij}\left(1 - H_i(t)\right). \end{aligned}$$

We obtain

$$\begin{aligned} ro(t) &= \sum_{i \in U} \sum_{j \in D} \lambda_{ij} \int_0^t \sum_{l=1}^{s} \pi_l \psi_{li}(du)\left(1 - H_i(t-u)\right) \\ &= \sum_{i \in U} \sum_{j \in D} \lambda_{ij} P_i(t). \end{aligned}$$

The last step follows from (Yeh, [9]), theorem 1. ■

A.7 Theorem 5.2.1

Proof. (Limnios & Oprişan, [7]) By induction. For $n = 2$, we can write the Stieltjes convolution we defined in definition 3.1.1 as

$$\begin{aligned}\widetilde{Q}_{ij}^{(2)}(t) &= \widetilde{Q}_{ij}(t) * \widetilde{Q}_{ij}(t)\\ &= \sum_{k\in E}\int_0^t \widetilde{Q}_{ik}(ds)\widetilde{Q}_{kj}(t-s)\\ &= \sum_{k\in E}\sum_{\{n:t_n\in(0,t]\}} \widetilde{Q}_{ik}(t-t_n)\Delta\widetilde{Q}_{kj}(t_n),\end{aligned}$$

where $\Delta\widetilde{Q}_{kj}(t_n) := \widetilde{Q}_{kj}(t_n) - \widetilde{Q}_{kj}(t_{n-1})$.

Now, we consider the case that $n = 3$. Then it follows that

$$\begin{aligned}\widetilde{Q}_{ij}^{(3)}(t) &= \widetilde{Q}_{ij}(t) * \widetilde{Q}_{ij}^{(2)}(t)\\ &= \sum_{k\in E}\int_0^t \widetilde{Q}_{ik}(ds)\widetilde{Q}_{kj}^{(2)}(t-s)\\ &= \sum_{k\in E}\int_0^t \widetilde{Q}_{ik}^{(2)}(ds)\widetilde{Q}_{kj}(t-s)\\ &= \sum_{k\in E}\sum_{\{n:t_n\in(0,t]\}} \widetilde{Q}_{ik}^{(2)}(t-t_n)\Delta\widetilde{Q}_{kj}(t_n).\end{aligned}$$

For $\widetilde{Q}_{ik}^{(2)}(t-t_n)$, we know that

$$\widetilde{Q}_{ik}^{(2)}(t-t_n) = \sum_{l\in E}\sum_{\{m:t_m\in(0,t-t_n]\}} \widetilde{Q}_{il}(t-t_n-t_m)\Delta\widetilde{Q}_{lk}(t_m).$$

If we plug this into our expression for the 3-fold convolution, we obtain

$$\begin{aligned}\widetilde{Q}_{ij}^{(3)}(t) &= \sum_{k\in E}\sum_{\{n:t_n\in(0,t]\}} \widetilde{Q}_{ik}^{(2)}(t-t_n)\Delta\widetilde{Q}_{kj}(t_n)\\ &= \sum_{k\in E}\sum_{l\in E}\sum_{\{n:t_n\in(0,t]\}}\sum_{\{m:t_m\in(0,t-t_n]\}} \widetilde{Q}_{il}(t-t_n-t_m)\Delta\widetilde{Q}_{lk}(t_m)\Delta\widetilde{Q}_{kj}(t_n).\end{aligned}$$

Thus it holds true for $n = 3$.

Now, assume that the expression given in theorem 5.2.1 holds true for $n = m$. Then for $n = m + 1$, we derive that

$$\begin{aligned}\widetilde{Q}_{ij}^{(m+1)}(t) &= \widetilde{Q}_{ij}(t) * \widetilde{Q}_{ij}^{(m)}(t)\\ &= \sum_{k_m}\int_0^t \widetilde{Q}_{ik_m}(ds)\widetilde{Q}_{k_mj}^{(m)}(t-s)\\ &= \sum_{k_m}\int_0^t \widetilde{Q}_{ik_m}^{(m)}(ds)\widetilde{Q}_{k_mj}(t-s)\\ &= \sum_{k_m}\sum_{r_m} \widetilde{Q}_{ik_m}^{(m)}(t-t_{r_m})\Delta\widetilde{Q}_{k_mj}(t_{r_m}),\end{aligned}$$

with $k_m \in E$ and $r_m : 0 < t_{r_m} \leq t$. From the induction hypothesis, it follows that

$$\widetilde{Q}_{ik_m}^{(m)}(t-t_{r_m}) = \sum_{k_{m-1}}\cdots\sum_{k_1}\sum_{r_{m-1}}\cdots\sum_{r_1}\widetilde{Q}_{ik_1}(t-t_{r_m}-t_{r_{m-1}}-\cdots-t_{r_1})\prod_{s=1}^{m-1}\Delta\widetilde{Q}_{k_s,k_{s+1}}(t_{r_s}),$$

where we used the following notation

$$\begin{gathered}
k_1 \in E, \ldots, k_{m-1} \in E, \\
r_{m-1} : 0 < t_{r_{m-1}} \leq t - t_{r_m}, \\
r_{m-2} : 0 < t_{r_{m-2}} \leq t - t_{r_m} - t_{r_{m-1}}, \\
\vdots \\
r_1 : 0 < t_{r_1} \leq t - t_{r_m} - t_{r_{m-1}} - \cdots - t_{r_2}.
\end{gathered}$$

If we plug this into our expression for $\widetilde{Q}_{ij}^{(m+1)}(t)$, we obtain

$$\begin{aligned}
\widetilde{Q}_{ij}^{(m+1)}(t) &= \sum_{k_m} \sum_{r_m} \widetilde{Q}_{ik_m}^{(m)}(t - t_{r_m}) \Delta \widetilde{Q}_{k_m j}(t_{r_m}) \\
&= \sum_{k_m} \sum_{k_{m-1}} \cdots \sum_{k_1} \sum_{r_m} \sum_{r_{m-1}} \cdots \sum_{r_1} \widetilde{Q}_{ik_1}(t - t_{r_m} - t_{r_{m-1}} - \cdots - t_{r_1}) \prod_{s=1}^{m-1} \Delta \widetilde{Q}_{k_s,k_{s+1}}(t_{r_s}) \Delta \widetilde{Q}_{k_m j}(t_{r_m}) \\
&= \sum_{k_m} \sum_{k_{m-1}} \cdots \sum_{k_1} \sum_{r_m} \sum_{r_{m-1}} \cdots \sum_{r_1} \widetilde{Q}_{ik_1}(t - t_{r_m} - t_{r_{m-1}} - \cdots - t_{r_1}) \prod_{s=1}^{m} \Delta \widetilde{Q}_{k_s,k_{s+1}}(t_{r_s}).
\end{aligned}$$

with $j = k_{m+1}$. By the principle of mathematical induction, the desired result follows. ∎

Appendix B

Tables and figures

Table B.0.1: Earthquakes with $M \geq 5.5$, 1953 to 2007. (Votsi et al., [1])

Date	Time	M	State
02-05-1953	18:37	5.6	1
03-08-1954	18:18	5.9	2
02-06-1955	23:34	5.5	1
06-01-1956	12:15	5.5	1
16-01-1958	04:18	5.7	2
11-04-1964	16:00	5.5	1
20-04-1964	04:21	5.6	1
09-03-1965	17:57	6.1	3
28-08-1965	14:08	5.6	1
20-12-1965	00:08	5.6	1
04-03-1967	17:58	6.6	3
19-02-1968	22:45	7.1	3
10-10-1968	07:10	5.5	1
06-04-1969	03:49	5.9	2
17-03-1975	05:35	5.8	2
27-03-1975	05:15	6.6	3
11-02-1976	07:35	5.7	2
14-06-1979	11:44	5.9	2
19-12-1981	14:10	7.2	3
18-01-1982	19:27	7.0	3
06-08-1983	15:43	6.8	3
05-10-1984	20:58	5.6	1
25-03-1986	01:41	5.5	1
29-03-1986	18:36	5.8	2
24-05-1994	02:05	5.5	1
25-05-1994	02:18	5.5	1
04-05-1995	00:34	5.8	2
14-11-1997	21:38	5.6	1
26-07-2001	00:21	6.3	3
06-07-2003	19:01	5.5	1
15-06-2004	12:02	5.5	1
21-12-2006	19:30	5.7	2
09-11-2007	01:43	5.5	1

Table B.0.2: Earthquakes with $M \geq 5.2$, 1964 to 2006. (Votsi et al., [1])

Date	Time	M	State	Subarea
23-02-1964	22:41	5.4	1	3
11-04-1964	16:00	5.5	1	1
29-04-1964	04:21	5.6	2	3
29-04-1964	17:00	5.2	1	3
09-03-1965	17:57	6.1	2	3
09-03-1965	17:59	5.7	2	3
09-03-1965	18:37	5.2	1	3
09-03-1965	19:46	5.2	1	3
13-03-1965	04:08	5.3	1	3
13-03-1965	04:09	5.5	1	3
28-08-1965	12:08	5.6	2	1
20-12-1965	00:08	5.6	2	1
04-03-1967	17:58	6.6	2	2
19-02-1968	22:45	7.1	2	2
20-02-1968	02:21	5.2	1	2
10-03-1968	07:10	5.5	1	3
24-04-1968	08:18	5.5	1	2
06-04-1969	03:49	5.9	2	4
17-03-1975	05:11	5.3	1	1
17-03-1975	05:17	5.4	1	1
17-03-1975	05:35	5.8	2	1
27-03-1975	05:15	6.6	2	1
29-04-1975	02:06	5.7	2	1
14-06-1979	11:44	5.9	2	4
12-11-1980	16:04	5.3	1	3
19-12-1981	14:10	7.2	2	2
21-12-1981	14:13	5.2	1	2
27-12-1981	17:39	6.5	2	2
29-12-1981	08:00	5.4	1	2
18-01-1982	19:27	7.0	2	1
18-01-1982	19:31	5.6	2	1
10-04-1982	04:50	5.2	1	2
06-08-1983	15:43	6.8	2	1
10-10-1983	10:17	5.4	1	1
06-05-1984	09:12	5.4	1	4
29-07-1984	01:53	5.2	1	1
05-10-1984	20:58	5.6	2	2
25-03-1986	01:41	5.5	1	3
29-03-1986	18:36	5.8	2	3
03-04-1986	23:32	5.2	1	3
03-06-1986	06:16	5.3	1	3
17-06-1986	17:54	5.4	1	3
06-08-1987	06:21	5.2	1	4
08-08-1987	22:15	5.3	1	1
27-08-1987	16:46	5.2	1	3
30-05-1988	16:47	5.2	1	1
19-03-1989	05:36	5.4	1	3
05-10-1989	06:52	5.4	1	1
23-07-1992	20:12	5.4	1	1
24-05-1994	02:05	5.5	1	4
16-04-1997	13:06	5.2	1	2
14-11-1997	21:38	5.8	2	4
11-04-1998	09:29	5.2	1	1
22-08-2000	03:35	5.2	1	3
10-06-2001	13:11	5.6	2	4
26-07-2001	00:21	6.4	2	3

Table B.0.2: Earthquakes with $M \geq 5.2$, 1964 to 2006. (Votsi et al., [1])

Date	Time	M	State	Subarea
26-07-2001	00:34	5.3	1	3
26-07-2001	02:06	5.2	1	3
26-07-2001	02:09	5.3	1	3
30-07-2001	15:24	5.4	1	3
29-10-2001	20:21	5.4	1	3
06-07-2003	19:10	5.5	1	1
06-07-2003	20:10	5.2	1	1
15-06-2004	12:02	5.2	1	1
22-11-2004	19:13	5.2	1	4
24-08-2005	03:06	5.2	1	2
21-12-2006	18:30	5.3	1	3

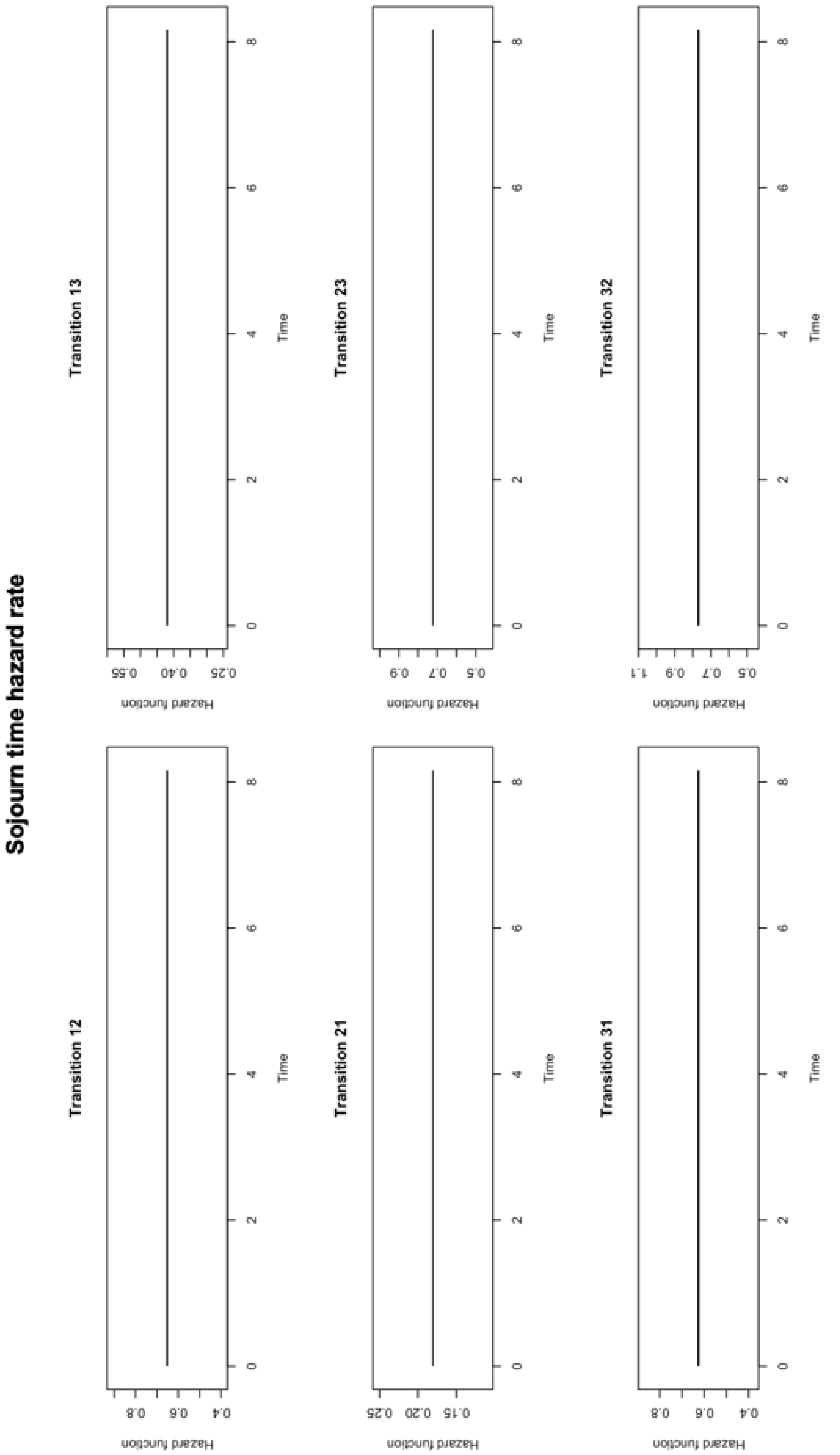

Figure B.0.1: Hazard rate of waiting time for the homogeneous Markov model for transitions from state i to state j, $i \neq j \in E$.

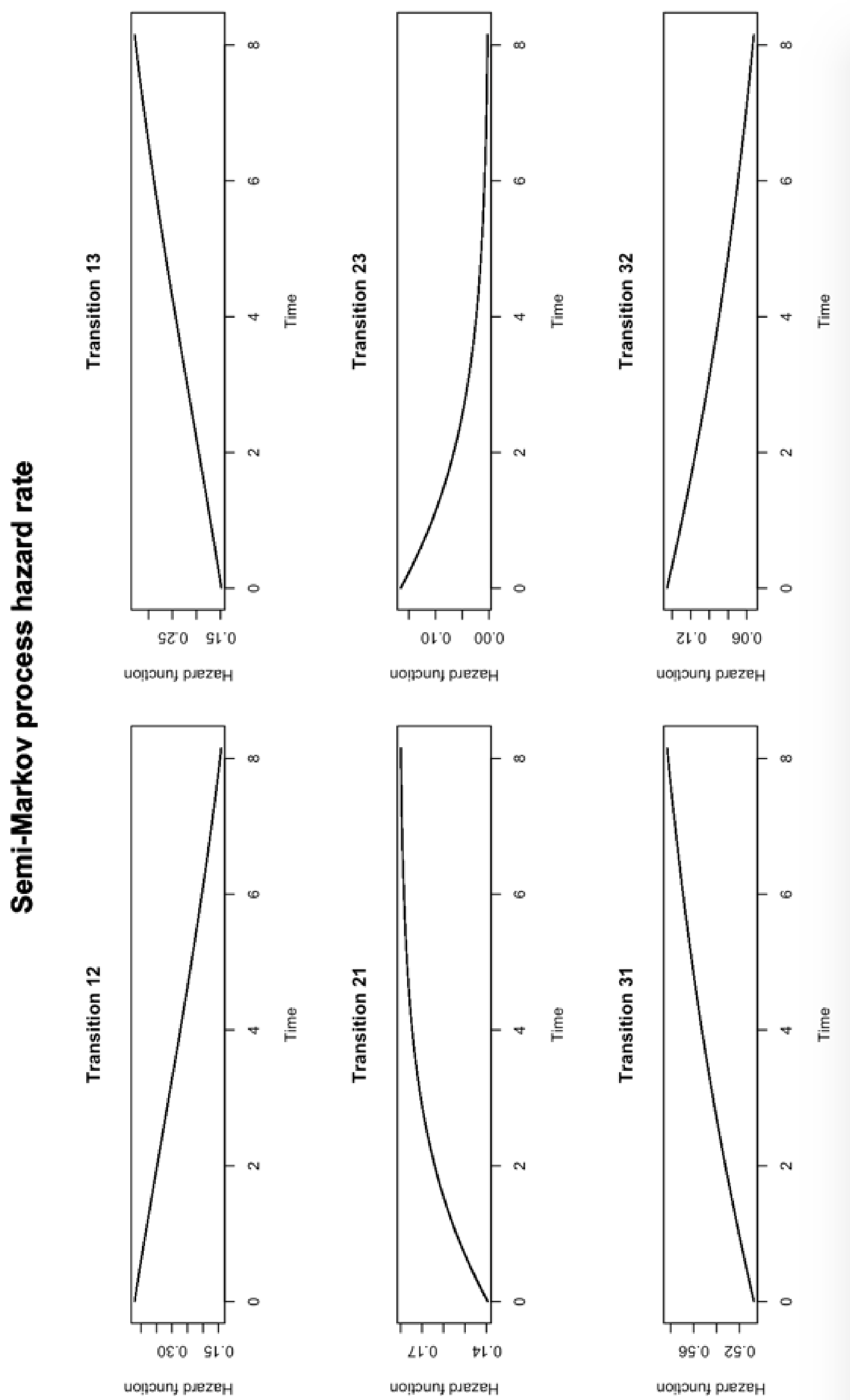

Figure B.0.2: Hazard rate of semi-Markov process for the homogeneous Markov model for transitions from state i to state j, $i \neq j \in E$.

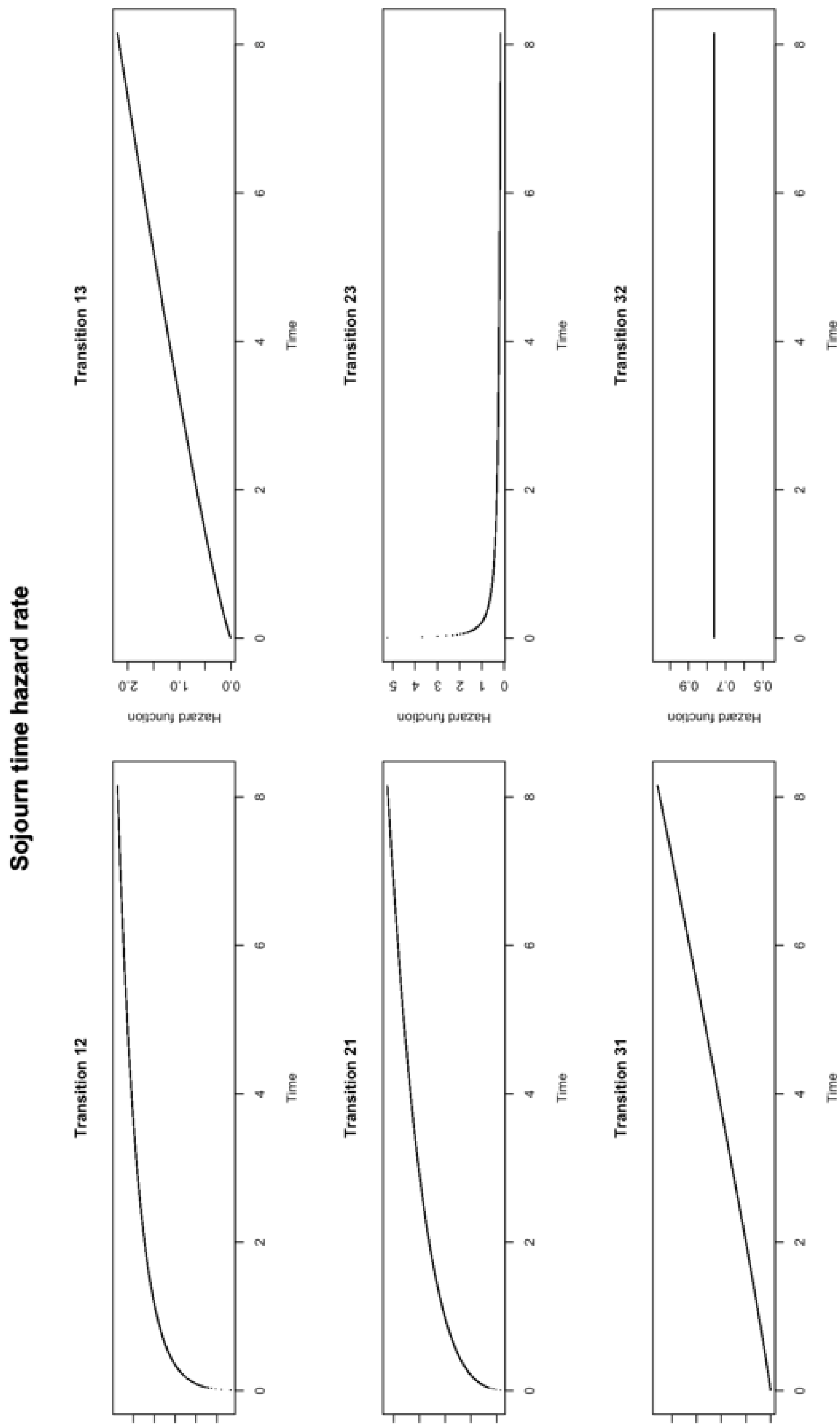

Figure B.0.3: Hazard rate of waiting time for the semi-Markov model for transitions from state i to state j, $i \neq j \in E$.

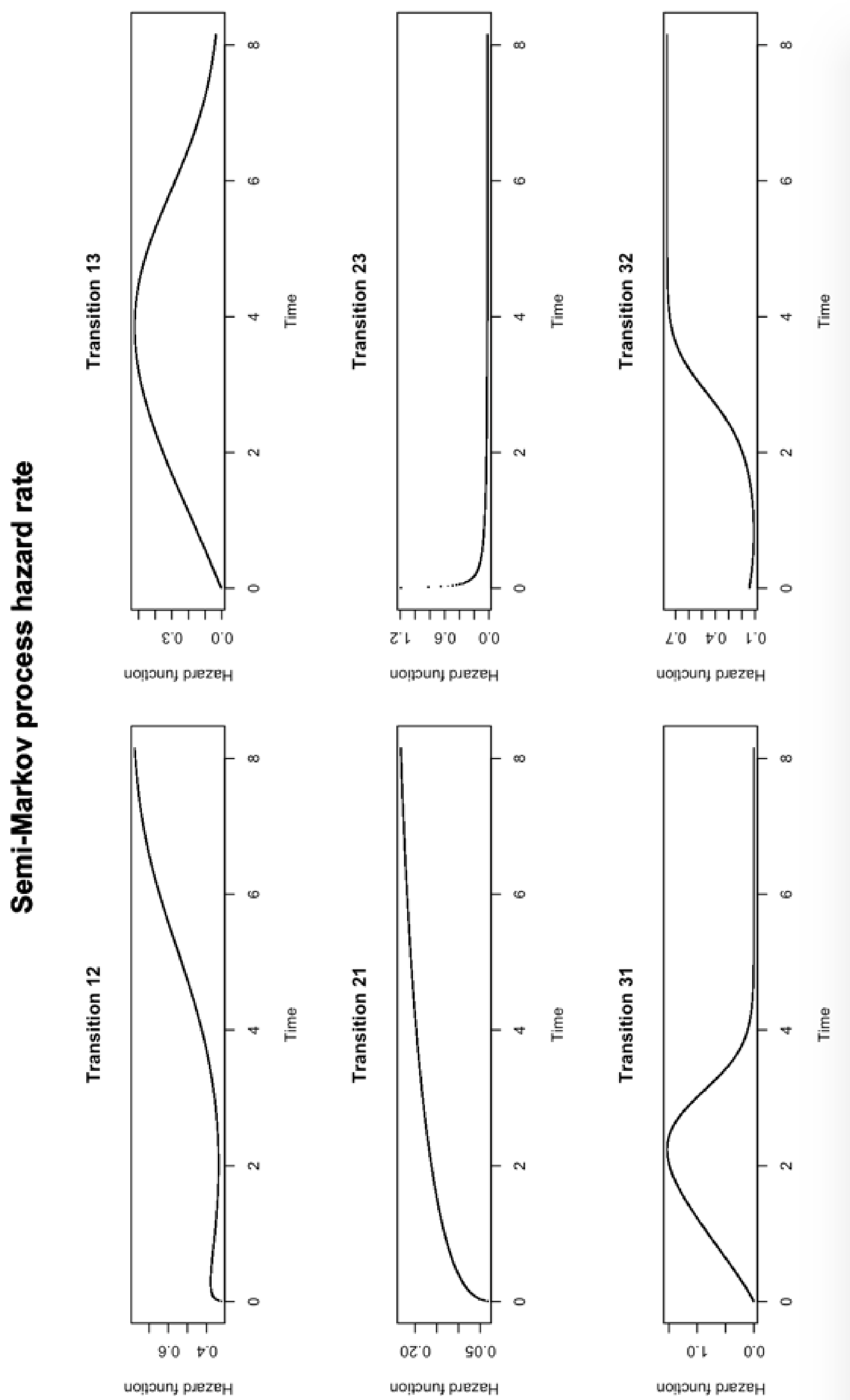

Figure B.0.4: Hazard rate of semi-Markov process for the semi-Markov model for transitions from state i to state j, $i \neq j \in E$.

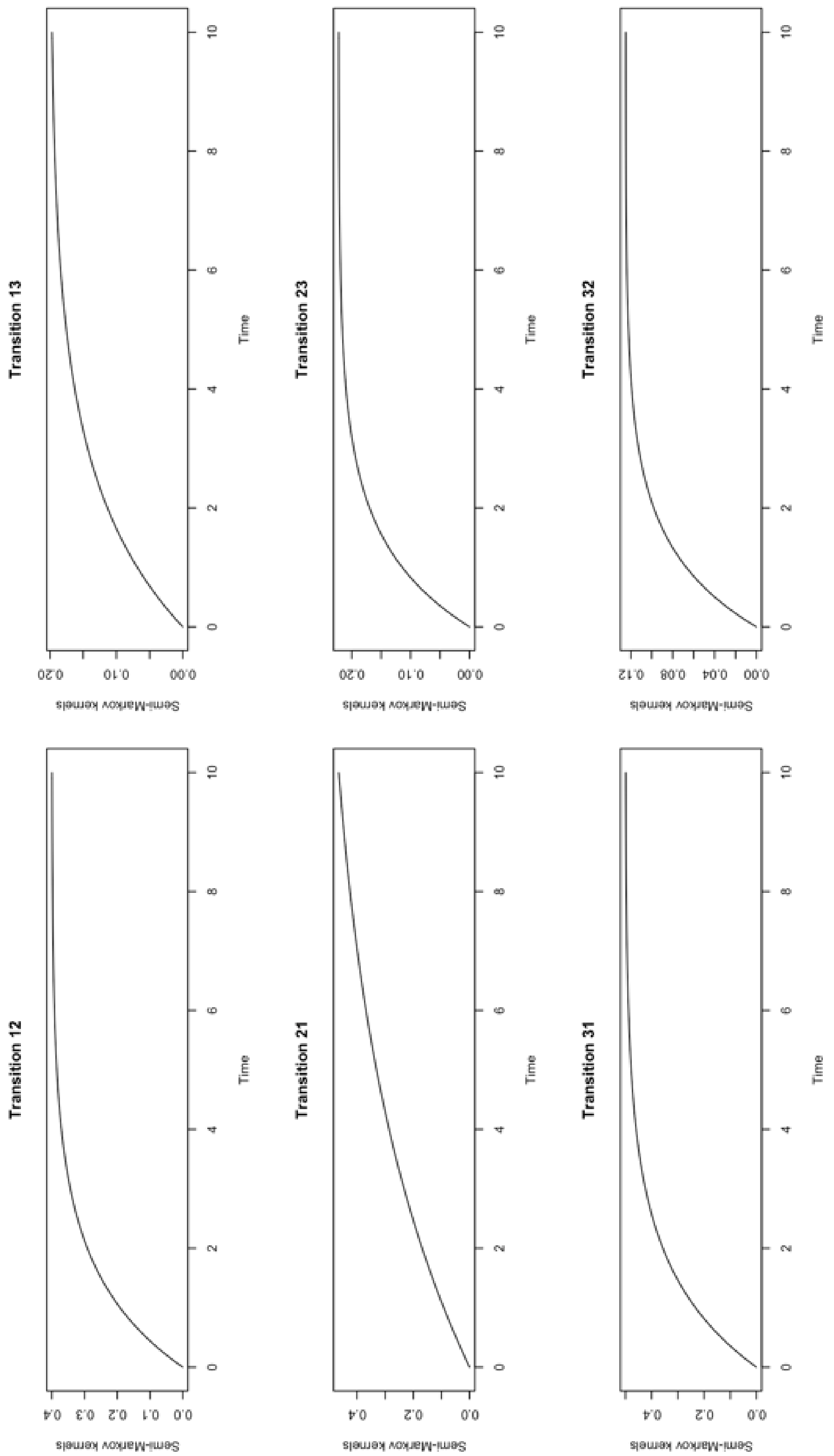

Figure B.0.5: Semi-Markov kernels for the homogeneous Markov model for all transitions from state i to state j, $i \neq j \in E$.

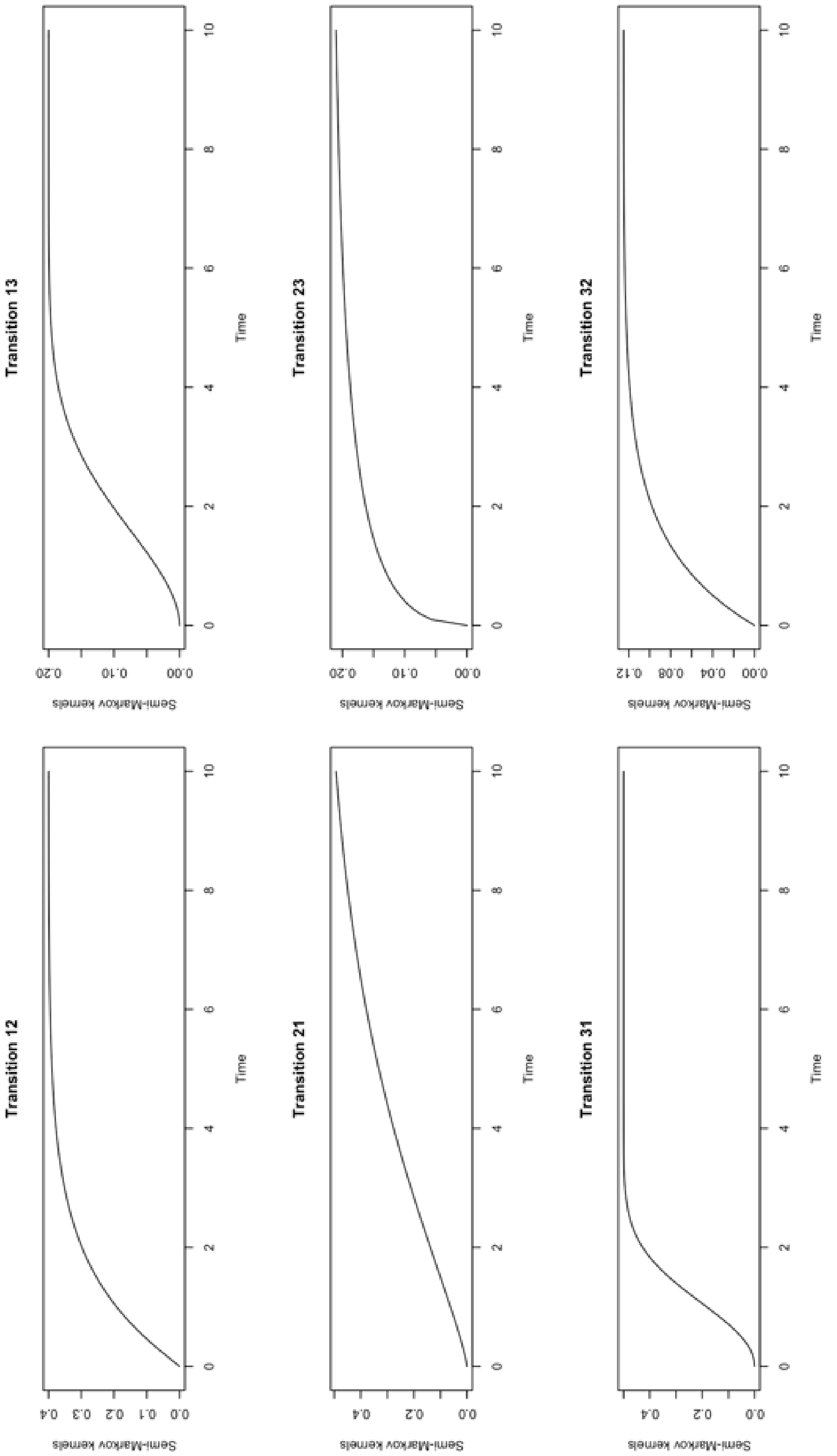

Figure B.0.6: Semi-Markov kernels for the semi-Markov model for all transitions from state i to state j, $i \neq j \in E$.

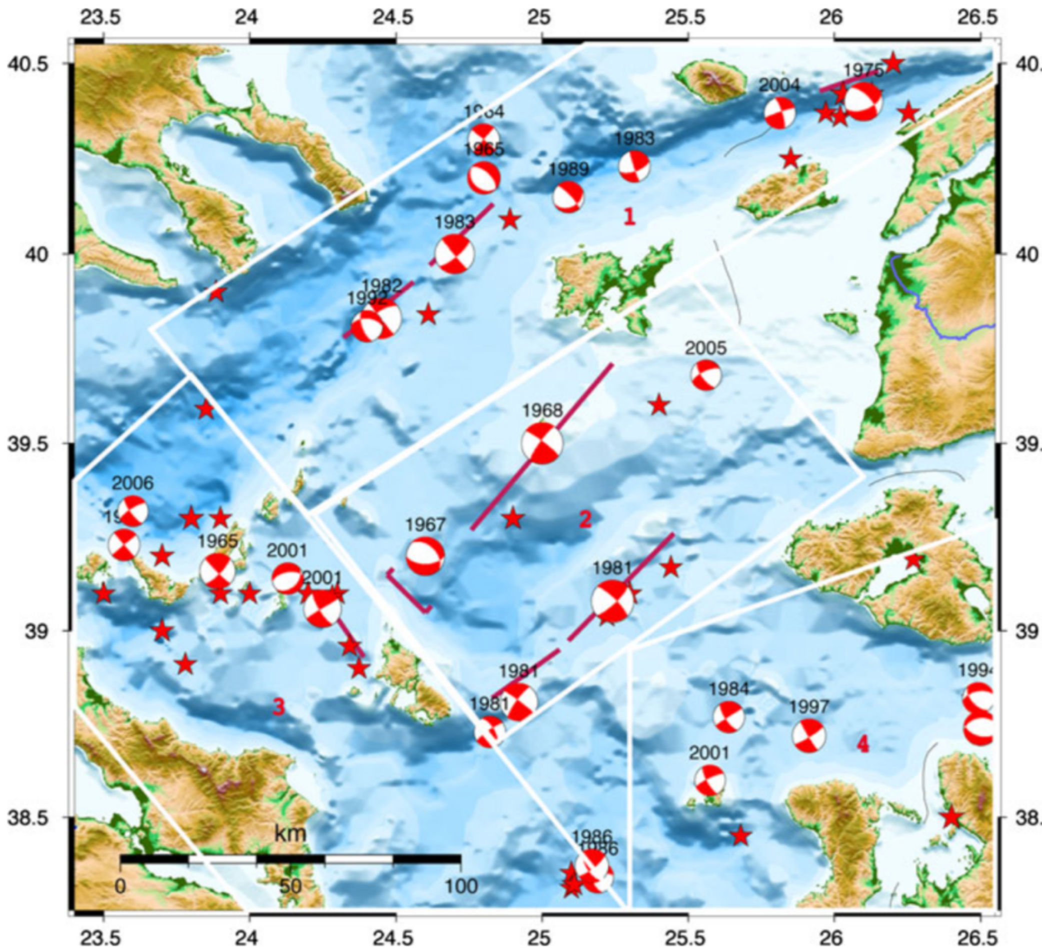

Figure B.0.7: Map of the Northern Aegean region in Greece, showing locations of 67 earthquakes with $M \geq 5.2$ since 1964. (Votsi et al., [1])

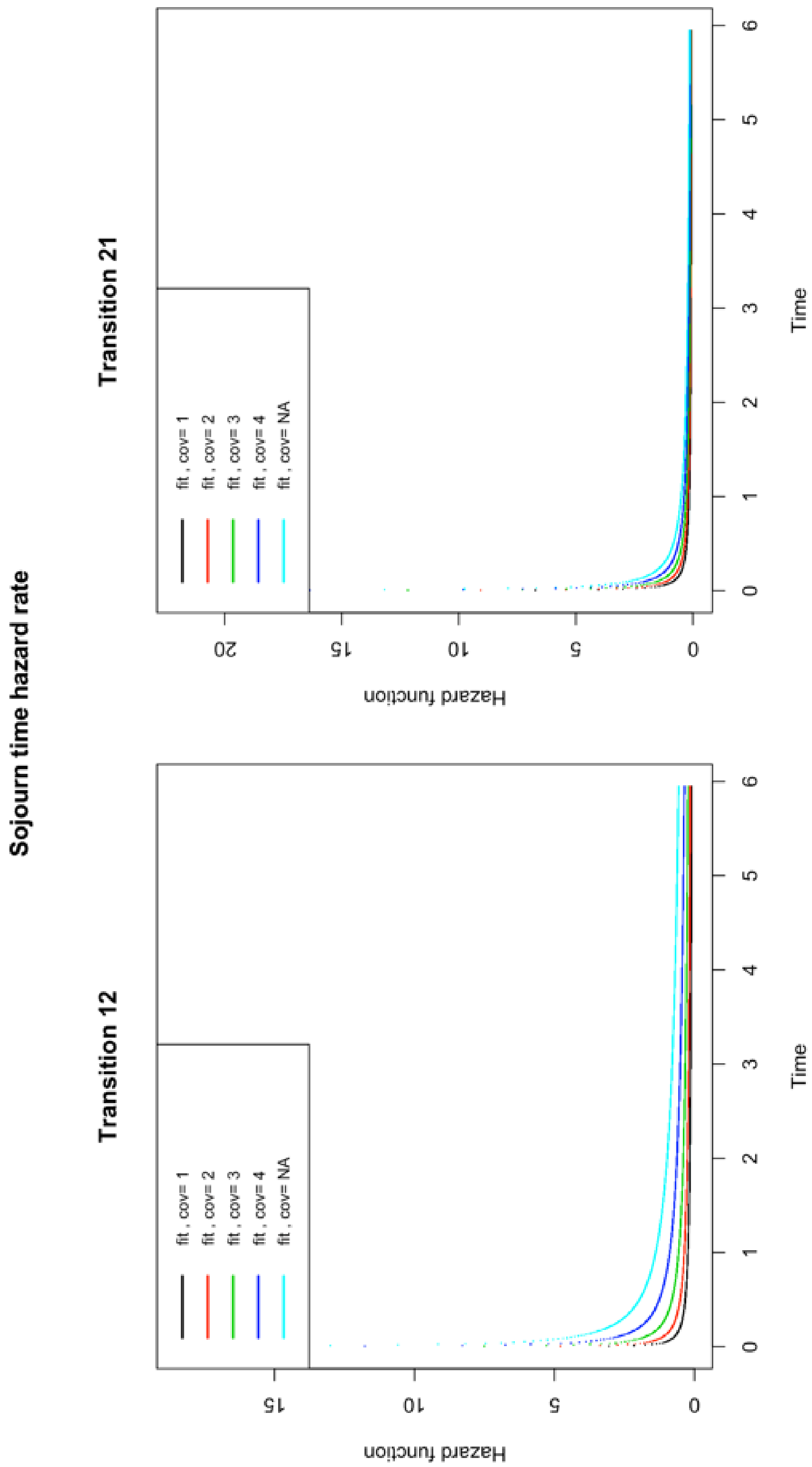

Figure B.0.8: Hazard rate of waiting time with covariates for transitions from state i to state j, $i \neq j \in E$.

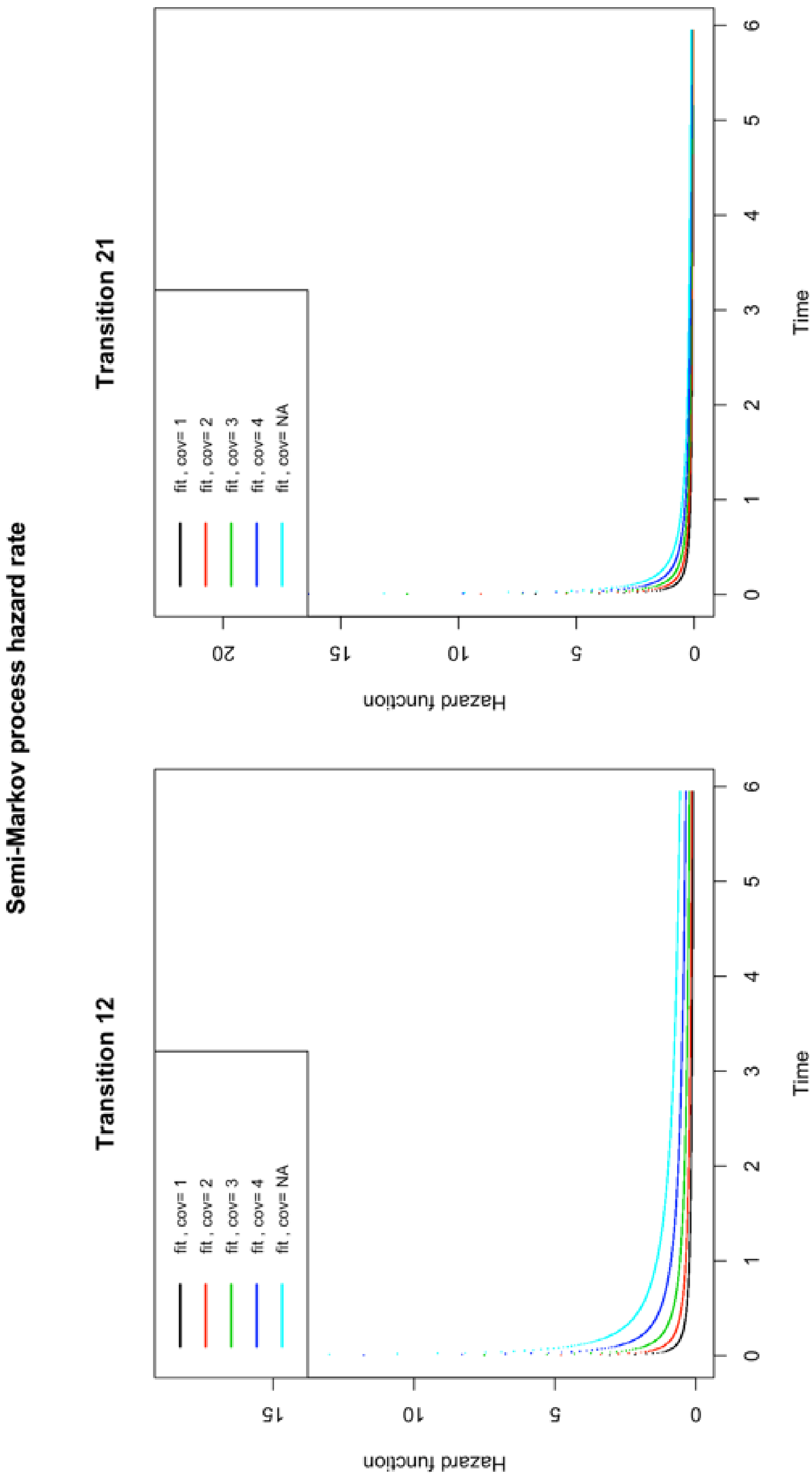

Figure B.0.9: Hazard rate of the semi-Markov process with covariates for transitions from state i to state j, $i \neq j \in E$.

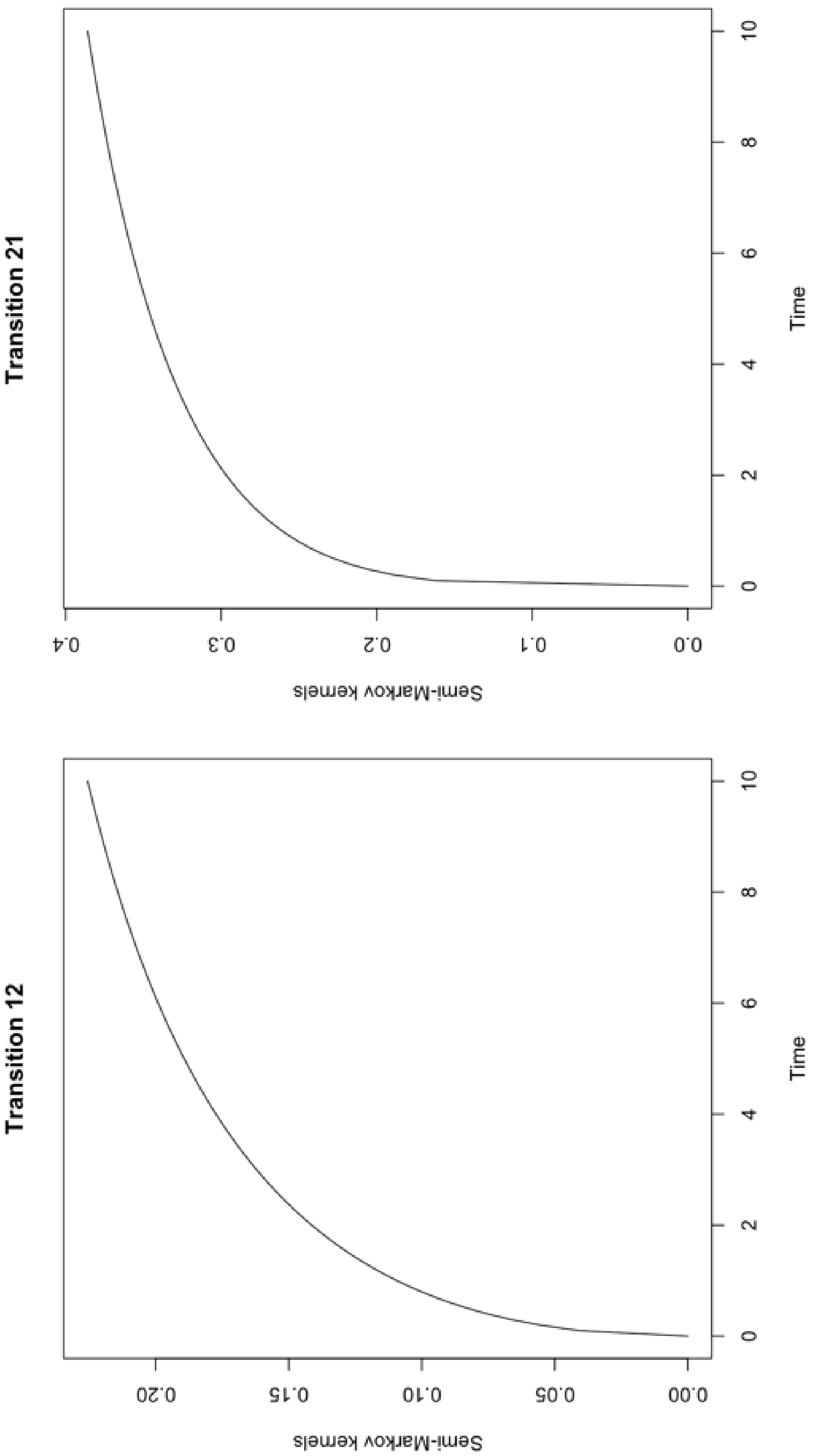

Figure B.0.10: Semi-Markov kernels of new dataset for all transitions from state i to state j, $i \neq j \in E$.

Appendix C

Bibliography

[1] Irene Votsi, Nikolaos Limnios, George Tsaklidis & Eleftheria Papadimitiou, *Estimation of the Expected Number of Earthquake Occurences Based on Semi-Markov Models*, Springer, 2011.

[2] Brahim Ouhbi & Nikolaos Limnios, *The rate of occurence of failures for semi-Markov processes and estimation*, Elsevier Science B.V., 2002.

[3] Sheldon M. Ross, *Introduction to Probability Models*, Elsevier, *9th edition*, 2007.

[4] *Markov renewal process.* Available at `https://en.wikipedia.org/wiki/Markov_renewal_process`. Referred at March 11, 2019.

[5] Franciszek Grabski, *Concept of Semi-Markov Process*, De Gruyter, 2016.

[6] John A. Rice, *Mathematical Statistics and Data Analysis*, Brooks/Cole Cengage Learning, *3rd edition*, 2007.

[7] N. Limnios & G. Oprişan, *Semi-Markov Processes and Reliability*, Springer, 2001.

[8] Nikolaos Limnios & Brahim Ouhbi, *Nonparametric Estimation for Semi-Markov Processes on K-Sample Paths with Application to Reliability*, (n.d.)

[9] Lam Yeh, *The Rate of Occurrence of Failures*, Applied Probability Trust, 1997.

[10] *The Glivenko-Cantelli Theorem.* Available at `http://home.uchicago.edu/~amshaikh/webfiles/glivenko-cantelli_topics.pdf`. Referred at May 11, 2019.

[11] Vlad Stefan Barbu & Nikolaos Limnios, *Semi-Markov Chains and Hidden Semi-Markov Models Toward Applications*, 2008.

[12] Agnieszka Listwon & Philippe Saint-Pierre, *SemiMarkov: An R Package for Parametric Estimation in Multi-State Semi-Markov Models*, Journal of Statistical Software, 2015.

[13] Donald L. Cohn, *Measure Theory*, Birkhauser Boston Inc., *2nd edition*, 2013.