République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2019/2020

Estimation Of Continuous Time Semi Markov Process

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar

Discipline : MATHEMATIQUES

Spécialité: Analyse Stochastique Statistique des Processus et Applications

par

Melle. Boubred Maroua[1]

Sous la direction de

Dr. F. Mokhtari

Soutenue le 16/09/2020 devant le jury composé de

Dr. S. Soltani	Université Dr Tahar Moulay - Saïda	Président
Dr. F. Mokhtari	Université Dr Tahar Moulay - Saïda	Encadreur
Dr. S. Rahmani	Université Dr Tahar Moulay - Saïda	Examinatrice
Dr. R. Rouane	Université Dr Tahar Moulay - Saïda	Examinatrice

[1]e-mail: boubred.maroua17@gmail.com.

Dedication

*All praise to **Allah**, today we fold the day's tiredness and the errand summing up between the cover of this humble work.*

I dedicate my work to:

*My great teacher and messenger, **Mohammed-peace and grace from Allah be upon him***, who taught us the purpose of life.

***My parents**, who have been our source of inspiration and gave us strength when we thought of giving up, who continually provide their moral, spiritual, emotional, and financial. God save them.*

*To my friend and my aunt **Fatima.**God bless her.*

*My brothers **Abdelkader**, **Mohamed** and **Aissa**.*

*My best friend and my sister **Bekheita Sadli**.*

*To my friends **Chafiaa Ayhar**, **Ikram Boukhecha**, **Aziza Driss**, **Fatima Medjahed** and **Fatima Amari** and their famillies.*

Last but not least I am dedicating this to my aunts and my cousins.

All those if my pen forget them, my heart will not forgotten them.

All those who are looking glory and pride in Islam and nothing else.

Acknowledgments

In the Name of Allah, the Most Merciful, the Most Compassionate all praise be to Allah and prayers and peace be upon Mohammed His servant and messenger.
First and foremost, I must acknowledge my limitless thanks to Allah, the Ever-Magnificent, the Ever-Thankful, for His help and bless. I am totally sure that this work would have never become truth, without His guidance.

I would like to express my wholehearted thanks to my supervisor Madam **Dr. F. Mokhtari**, who worked hard with me from the beginning till the completion of the present master thesis. I greatly appreciate her support, constructive ideas, invaluable guidance, encouragement and patience throughout this master thesis.

I wish to thank my master thesis committee members: **Dr. S. Soltani**, **Dr. S. Rahmani** and **Dr. R. Rouane**, for generously offering their time, support, guidance and good will throughout the preparation and review of this document.

I owe a deep debt of gratitude to our university, laboratory of stochastic models, statistic and applications and his members for giving us an opportunity to complete this work.

I would like to take this opportunity to say warm thanks to all my beloved friends, who have been so supportive along the way of doing my master thesis.

I owe profound gratitude to my family for their generous support they provided me throughout my entire life and particularly through the process of pursuing the master degree. Because of their unconditional love and prayers, I have the chance to complete this work.

Last but not least, deepest thanks go to all people who took part in making this master thesis real. To all a big thank you.

Contents

Notations

$\mathbb{N}$	**Set of positive natural numbers.**
$\mathbb{R}_+$	**Set of nonnegative real numbers.**
$(\Omega, \mathcal{F}, \mathbb{P})$	**Probability space.**
$\mathbb{E}$	**Expectation with respect to** $\mathbb{P}$.
$\mathrm{E} = \{1, \ldots, s\}$	**Finite state space.**
$\mathcal{M}_{\mathrm{E}}$	**Set of real matrix on** $\mathrm{E} \times \mathrm{E}$.
$\mathcal{M}_{\mathrm{E}}(\mathbb{N})$	**Matrix-valued functions defined on** $\mathbb{N}$**, with values in** $\mathcal{M}_{\mathrm{E}}$.
$Z := (Z_k)_{k \in \mathbb{N}}$	**Semi-Markov chain (SMC).**
$Z := (Z_t)_{t \in \mathbb{R}_+}$	**Semi-Markov process (SMP).**
$(J, S) := (J_n, S_n)_{n \in \mathbb{N}}$	**Markov renewal chain (MRC).**
$J := (J_n)_{n \in \mathbb{N}}$	**Visited states, embedded Markov chain (EMC).**
$S := (S_n)_{n \in \mathbb{N}}$	**Jump times.**
$X := (X_n)_{n \in \mathbb{N}}$	**Sojourn times.**
M	**Fixed censoring time.**
$N(M)$	**Number of jumps of Z in the time interval** $[1, M]$.
$N_i(M)$	**Number of visits to state i of the EMC, up to time M.**
$N_{ij}(M)$	**Number of transitions from state i to state j of the EMC, up to time M.**
$N_{ij}(k, M)$	**Number of transitions from state i to state j of the EMC, up to time M, with sojourn time in state i equal to k.**

$\mathbf{p} := (p_{ij})_{i,j\in\mathrm{E}}$	**Transition matrix of the EMC J.**
$\mathbf{q} := (q_{ij}(k))_{i,j\in\mathrm{E},k\in\mathbb{N}}$	**Semi-Markov kernel.**
$\mathbf{q} := (q_{ij}(t))_{i,j\in\mathrm{E},t\in\mathbb{R}_+}$	**Density of the Markov renewal kernel.**
$\mathbf{Q} := (Q_{ij}(k))_{i,j\in\mathrm{E},k\in\mathbb{N}}$	**Cumulated semi-Markov kernel.**
$\mathbf{Q} := (Q_{ij}(t))_{i,j\in\mathrm{E},t\in\mathbb{R}_+}$	**Markov renewal kernel.**
$\mathbf{f} := (f_{ij}(k))_{i,j\in\mathrm{E},k\in\mathbb{N}}$	**Conditional sojourn time distribution in state i, before visiting state j.**
$\mathbf{F} := (F_{ij}(k))_{i,j\in\mathrm{E},k\in\mathbb{N}}$	**Conditional cumulative sojourn time distribution in state i, before visiting state j.**
$\mathbf{F} := (F_{ij}(t))_{i,j\in\mathrm{E},t\in\mathbb{R}_+}$	**Sojourn time distribution in state i, before visiting state j.**
$\mathbf{h} := (h_i(k))_{i\in\mathrm{E},k\in\mathbb{N}}$	**Sojourn time distribution in state i.**
$\mathbf{H} := (H_i(k))_{i\in\mathrm{E},k\in\mathbb{N}}$	**Cumulative distribution of sojourn time in state i.**
$\mathbf{H} := (H_i(t))_{i\in\mathrm{E},t\in\mathbb{R}_+}$	**Sojourn time distribution in state i.**
$\overline{\mathbf{H}} := (\overline{H}_i(k))_{i\in\mathrm{E},k\in\mathbb{N}}$	**Survival function in state i.**
$\mathbf{P} := (P_{ij})(k)_{i,j\in\mathrm{E},k\in\mathbb{N}}$	**Transition function of the semi-Markov chain Z.**
$\mathbf{P} := (P_{ij})(t)_{i,j\in\mathrm{E},t\in\mathbb{R}_+}$	**Transition function of the semi-Markov process Z.**
$\psi(t) = (\psi_{ij}(t))_{i,j\in\mathrm{E},t\in\mathbb{R}_+}$	**Markov renewal matrix.**
$\lambda(t) := (\lambda_{ij}(t))_{i,j\in\mathrm{E},t\in\mathbb{R}_+}$	**Hazard rate function.**
μ_{ij}	**Mean first passage time from state i to state j, for semi Markov process Z.**
μ^*_{ij}	**Mean first passage time from state i to state j, for embedded Markov chain J.**
$\nu = (\nu(j)_{j\in\mathrm{E}})$	**Stationary distribution of the EMC J.**
$\alpha = (\alpha_i)_{i\in\mathrm{E}}$	**Initial distribution of semi-Markov process Z.**
$A * B$	**Discrete-time matrix convolution product of A, B.**
$Q \star \phi$	**Stieltjes convolution of ϕ, Q.**
$A^{(n)}$	**n-fold convolution of $A \in \mathcal{M}_{\mathrm{E}}(\mathbb{N})$.**

$\xrightarrow{a.s}$	**Almost sure convergence (strong consistency).**
$\xrightarrow{P}$	**Convergence in probability.**
$\xrightarrow{\mathcal{D}}$	**Convergence in distribution.**
δ_{ij}	**Symbole of Kronecker.**
$\mathbb{1}_A$	**Indicatrice function of A.**
$\mathcal{N}(0, \sigma^2)$	**Standard normal random variable (mean $\mu = 0$, variance σ^2).**
$DTMC$	**Discrete-time Markov Chain.**
$CTMC$	**Continuous-time Markov Chain.**
SMC	**Semi-Markov Chain.**
SMP	**Semi-Markov Process.**
RC	**Renewal Chain.**
EMC	**Embedded Markov Chain.**
MLE	**Maximum-Likelihood Estimator.**
$SLLN$	**Strong Law of Large Numbers.**
CLT	**Central Limit Theorem.**
$r.v$	**random variable.**

Introduction

In recent years, the evolution of a system in applications concern queuing theory, reliability and maintenance, survival analysis, performance evaluation, biology, DNA analysis, risk processes, insurance and finance, earthquake modelling, etc, is modelled by a stochastic continuous-time process or discrete. Among the models which are widely used as a standard tool to describe the evolution of a system, we have the Markov models and the semi-Markov models.

Much work has been carried out in the field of Markov processes, and a huge amount of Markov process applications can be found in the literature of the last 50 years. One of the reasons for applying Markov process theory in various fields is that the Markovian hypothesis 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. 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 distribution (continuous case) or geometrically distribution (discrete 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 calendar time in this case, but on the sojourn time in the state.

The semi-Markov processes were introduced independently and almost simultaneously by Levy [18], Smith [30], and Takacs [31] in 1954-1955. The essential developments of semi-Markov processes theory were proposed by Pyke [25, 26], Cinlar [9], Koroluk and Turbin [16, 15], Limnios [19], Takacs [32]. For the semi-Markov processes, the distribution of the sojourn time in a state can be arbitrary, and the future evolution depends on the time spent through the last transition.

A semi-Markov process can also be defined by a two-dimensional process, the first component represents the states successively visited by the process (Markov chain) and the second describes the moments of change of process state. This two- dimensional process is called Markov renewal process.

The problem of statistical inference for semi-Markov processes is of increasing interest in literature. There is a growing literature concerning inference problems for continuous-time semi-Markov processes. For instance, Moore and Pyke (1968)[21] studied empirical and maximum likelihood estimators for semi-Markov kernels; Lagakos et al. (1978)[17] obtained the non-parametric maximum likelihood estimator for the kernel of a finite state semi-Markov process with some absorbing states; Akritas and Roussas (1979)[1] studied the asymptotic local normality: Gill (1980)[11] constructed an estimator for the kernel of a finite state semi-Markov kernel, using counting processes; Ouhbi and Limnios (1999)[23] studied empirical estimators for non-linear functionals of finite semi-Markov kernels.

This master memory falls into four chapters.

In chapter 1, we give some background and some basic concepts, properties, and theorems on homogeneous Markov chains and continuous-time homogeneous Markov processes with a discrete set of states.

In chapter 2, we consider a homogeneous discrete-time finite state space semi-Markov model. We introduce its basic probabilistic properties and we present their empirical estimators for the main characteristics (kernel, sojourn time distributions, transition probabilities, etc.), which proves to be also an approached maximum likelihood estimator. The estimation made by considering a sample path of the discrete-time semi-Markov process (DTSMP) in the time interval $[0, M]$ with M an arbitrarily chosen positive integer. After having obtained these general results, we investigate the asymptotic properties of the estimators, namely, the strong consistency and the asymptotic normality. We continue by giving the Markov renewal equation in the discrete case.

In chapter 3, we develop the theory of continuous-time semi-Markov processes. Results about estimation and the asymptotic behaviors of the empirical estimators of this processes are also transposed here but other specific results about Markov renewal equation and hazard rate function are given as well.

In chapter 4 we present the **R** package **semiMarkov** for parametric estimation in multi-state semi-Markov models and we give a detailed description of the package with an application to asthma control. After that, we apply the nonparametric estimation of the semi-Markov model to the Coronavirus data sets in Tunisia and Algeria. For the Coronavirus data set in Algeria, we use the **R** package **semiMarkov** to determine the hazard rate functions in a parametric way.

Chapter 1

Introduction and preliminaries

In this chapter we introduce some basic concepts, properties, and theorems on homogeneous Markov chains and continuous-time homogeneous Markov processes with a discrete set of states, which will be useful later.

1.1 Definitions and theorems

Consider a finite set $\mathrm{E} = \{1, \ldots, s\}$. We denote by $\mathcal{M}_E$ the set of real matrices on $\mathrm{E} \times \mathrm{E}$ and by $\mathcal{M}_E(\mathbb{N})$ the set of matrix valued functions defined on $\mathbb{N}$, with values in $\mathcal{M}_E$.

Definition 1.1.1. ***(discrete-time matrix convolution product)***

Let $\mathbf{A}, \mathbf{B} \in \mathcal{M}_E(\mathbb{N})$ be two matrix-valued functions. The matrix convolution product $\mathbf{A} * \mathbf{B}$ is the matrix-valued function $\mathbf{C} \in \mathcal{M}_E(\mathbb{N})$ defined by

$$C_{ij}(k) := \sum_{r\in\mathrm{E}} \sum_{l=0}^{k} A_{ir}(k-l)B_{rj}(l), \quad i, j \in \mathrm{E}, \quad k \in \mathbb{N},$$

or, in matrix form,

$$\mathbf{C}(k) := \sum_{l=0}^{k} \mathbf{A}(k-l)\mathbf{B}(l).$$

Lemma 1.1.1. *[5] Let $\delta\mathbf{I} = (\delta_{ij}(k); i, j \in \mathrm{E}) \in \mathcal{M}_E(\mathbb{N})$ be the matrix-valued* *function defined by*

$$\delta_{ij}(k) := \begin{cases} 1, & \textit{if } i = j \textit{ and } k = 0, \\ 0, & \textit{elsewhere.} \end{cases}$$

or, in matrix form,

$$\delta\mathbf{I}(k) := \begin{cases} \mathbf{I}, & \textit{if } k = 0, \\ \mathbf{0}, & \textit{elsewhere.} \end{cases}$$

Then $\delta\mathbf{I}$ *satisfies*

$$\delta\mathbf{I} * \mathbf{A} = \mathbf{A} * \delta\mathbf{I} = \mathbf{A}, \quad \mathbf{A} \in \mathcal{M}_E(\mathbb{N})$$

i.e., $\delta\mathbf{I}$ *is the identity element for the discrete-time matrix convolution product.*

Definition 1.1.2. ***(discrete-time n-fold convolution)*** *Let* $\mathbf{A} \in \mathcal{M}_E(\mathbb{N})$ *be a matrix-valued function and* $n \in \mathbb{N}$. *The n-fold convolution* $\mathbf{A}^{(n)}$ *is the matrix-valued function defined recursively by:*

$$\begin{aligned} A_{ij}^{(0)}(k) &:= \delta_{ij}(k) \begin{cases} 1, & \textit{if } i = j \textit{ and } k = 0, \\ 0, & \textit{elsewhere,} \end{cases} \\ A_{ij}^{(1)}(k) &:= A_{ij}(k), \end{aligned}$$

$$A_{ij}^{(n)}(k) := \sum_{r \in \mathrm{E}} \sum_{l=0}^{k} A_{ir}(l) A_{rj}^{(n-1)}(k-l), \quad n \geq 2, \ k \in \mathbb{N},$$

that is,

$$\mathbf{A}^{(0)} := \delta\mathbf{I}, \qquad \mathbf{A}^{(1)} := \mathbf{A} \quad \textit{and} \qquad \mathbf{A}^{(n)} := \mathbf{A} * \mathbf{A}^{(n-1)}.$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathrm{E}, \varepsilon)$ be a measurable space and let I be a set called a parameter set. Generally, I is a subset of $\mathbb{R}$, usually $\mathbb{N}$ or $\mathbb{R}_+$.

Definition 1.1.3. ***(Stochastic process, state space)***

A stochastic process is a family of random variables $\{X(t), t \in I\}$ *defined on* $(\Omega, \mathcal{F}, \mathbb{P})$ *with values in* E. *For every* $t \in I$, $X(t)$ *is a random variable* $X(t) : \Omega \to \mathrm{E}$, *whose value for the outcome* $\omega \in \Omega$ *is noted* $X(t, \omega)$. *If instead of* t *we fix an* $\omega \in \Omega$, *we obtain the function* $X(., \omega) : I \to \mathrm{E}$ *which is called a trajectory or a path-function or a sample function of the process.*

The set E *is called the state space of the stochastic process* $X = (X(t), t \in I)$. *The stochastic process may be denoted by* X_t *instead of* $X(t)$ *(respectively,* X_n *if* $I = \mathbb{N}$*).*

Theorem 1.1.1. ***(Strong Law of Large Numbers)****[19] Let* $(X_1, X_2, \ldots)$ *is an infinite sequence of i.i.d. Lebesgue integrable random variables with expected value* $\mathbb{E}[X_1] = \mathbb{E}[X_2] = \ldots$, *then we have*

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{a.s} \mathbb{E}[X_1].$$

Theorem 1.1.2. ***(Glivenko-Cantelli theorem)*** *[7] Let* $F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{X_k \leq x\}}$ *be the empirical distribution function of the i.i.d. random sample* $X_1, \ldots, X_n$. *Denote by* F *the common distribution function of* X_i , $i = 1, \ldots, n$. *Thus*

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \xrightarrow[n\to\infty]{a.s} 0.$$

Theorem 1.1.3. *[14] Let* $(Y_n)_{n\in\mathbb{N}}$ *be a sequence of random variables and* $(N_n)_{n\in\mathbb{N}}$ *a positive integer-valued stochastic process. Suppose that*

$$Y_n \xrightarrow[n\to\infty]{a.s} Y \text{ and } N_n \xrightarrow[n\to\infty]{a.s} \infty.$$

Then,

$$Y_{N_n} \xrightarrow[n\to\infty]{a.s} Y.$$

Definition 1.1.4. ***(Martingale)*** *Let* $\mathbf{F} = (\mathcal{F}_n, n \geq 0)$ *be a family of sub-*σ*-algebras of* $\mathcal{F}$ *such that* $\mathcal{F}_n \subset \mathcal{F}_m$, *when* $n < m$. *We say that* $\mathbf{F}$ *is a filtration of* $\mathcal{F}$. *A real-valued* $\mathbf{F}$*-adapted stochastic process* X_n *is (*$\mathcal{F}_n$*-measurable for* $n \geq 0$*) called martingale with respect to a filtration* $\mathbf{F}$ *if, for every* $n = 0, 1, \ldots$

1. $\mathbb{E}|X_n| < \infty$ *; and*

2. $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ *(a.s).*

Theorem 1.1.4. ***(CLT for martingales)[6]***

Let $(X_n)_{n\in\mathbb{N}^\star}$ be a martingale with respect to the filtration $\mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ and define the process $Y_n = X_n - X_{n-1}$, $n \in \mathbb{N}^\star$ (with $Y_1 := X_1$), called a difference martingale. If

1. $\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}[Y_k^2|\mathcal{F}_{k-1}] \xrightarrow[n\to\infty]{P} \sigma^2 > 0$;

2. $\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}[Y_k^2 \mathbf{1}_{\{|Y_k|>\epsilon\sqrt{n}\}}] \xrightarrow[n\to\infty]{} 0$, *For all* $\epsilon > 0$,

then

$$\frac{X_n}{n} \xrightarrow[n\to\infty]{a.s} 0,$$

and

$$\frac{1}{\sqrt{n}}X_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} Y_k \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma^2).$$

Theorem 1.1.5. ***(Anscombe's theorem)[8]***

Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of random variables and $(N_n)_{n\in\mathbb{N}}$ a positive integer-valued stochastic process. Suppose that

$$\frac{1}{\sqrt{n}}\sum_{m=1}^{n} Y_m \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma^2) \quad \text{and} \quad N_n/n \xrightarrow[n\to\infty]{P} \theta,$$

where θ is a constant, $0 < \theta < \infty$. Then,

$$\frac{1}{\sqrt{N_n}}\sum_{m=1}^{N_n} Y_m \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma^2).$$

1.2 Discrete-time Markov chain

Let $(J_n)_{n\geq 0}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in a measurable space $(\mathrm{E}, \varepsilon)$. Unless otherwise stated, we assume that $\mathrm{E} = \{1, 2, \ldots, s\}$ or $\mathrm{E} = \{1, 2, \ldots\}$.

Definition 1.2.1. ***(Discrete-time Markov chain)***

1. *A stochastic process* $(J_n)_{n\geq 0}$ *is called discrete time Markov process or Markov chain with state space* E *if, for any* $n \in \mathbb{N}$ *and any state sequence* $i_1, i_2, \ldots, i, j \in \mathrm{E}$,

$$\mathbb{P}(\underbrace{J_{n+1} = j}_{Future} \,|\, \underbrace{J_1 = i_1, \ldots, J_n = i}_{Past\,and\,present}) = \mathbb{P}(\underbrace{J_{n+1} = j}_{Future}|\underbrace{J_n = i}_{Present}).$$

2. *If, additionally, the probability* $\mathbb{P}(J_{n+1} = j|J_n = i)$ *does not depend on* n, $(J_n)_{n\geq 0}$ *is said to be homogeneous with respect to time.*

Definition 1.2.2. ***(Transition matrix)***

The function $(i, j) \to p_{ij} := \mathbb{P}(J_{n+1} = j|J_n = i)$ *is called transition function of the chain. For any* $i, j \in \mathrm{E}$ *and* $n \geq 0$, *the transition function has the following properties :*

1. $p_{ij} \geq 0$, *for any* $i, j \in \mathrm{E}$,
2. $\sum_{j\in \mathrm{E}} p_{ij} = 1$, *for any* $i \in \mathrm{E}$,
3. $\sum_{k\in \mathrm{E}} p_{ik}p_{kj} = \mathbb{P}(J_{n+2} = j|J_n = i) = p^{(2)}_{ij}$.

If E *finite, we can represent transition function as a square matrix (transition matrix),*

$$\mathbf{p} = (p_{ij})_{i,j\in \mathrm{E}} = \begin{pmatrix} p_{11} & \cdots & p_{1s} \\ \vdots & & \vdots \\ p_{s1} & \cdots & p_{ss} \end{pmatrix}$$

Notation: $p^{(n)}_{ij} := \mathbb{P}(J_n = j|J_0 = i)$ is called the n-step transition function.

Remark 1.2.1. *If* E *finite,* $\mathbf{p^{(n)}}$ *represents the usual n-fold matrix product of* $\mathbf{p}$, *that is*

$$\mathbf{p^{(n)}} = \mathbf{p^n}.$$

In order to define a Markov chain $(J_n)_{n\geq 0}$ we need :

1. transition function (matrix) $\mathbf{p} = (p_{ij})_{i,j\in \mathrm{E}}$.

2. $\alpha = (\alpha_1, \dots, \alpha_s)$, the initial distribution of the chain, that is the distribution of J_0, $\alpha_i = \mathbb{P}(J_0 = i)$ for any state $i \in \mathrm{E}$.

Proposition 1.1. *[5] Let $(J_n)_{n\geq 0}$ be a Markov chain of transition matrix $\mathbf{p}$.*

1. *The sojourn time of the chain in a state $i \in \mathrm{E}$ is a geometric random variable of parameter $1 - p_{ii}$.*

2. *The probability that the chain enters state j when it leaves state i is $\frac{p_{ij}}{1-p_{ii}}$ (for $p_{ii} \neq 1$).*

Definition 1.2.3. ***(Stationary distribution)*** *A probability distribution ν on E is said to be stationary or invariant for the Markov chain $(J_n)_{n\geq 0}$ if, for any $j \in \mathrm{E}$*

$$\sum_{j\in \mathrm{E}} \nu(i) p_{ij} = \nu(j),$$

or, in matrix form,

$$\nu \mathbf{p} = \nu,$$

where $\nu = (\nu(1), \dots, \nu(s))$ is a row vector.

1.2.1 State classification

Definition 1.2.4. ***(Accessible state)*** *We say that state j is accessible from state i, written as $i \to j$ if $p_{ij}^{(n)} > 0$. We assume every state is accessible from itself since $p_{ii}^{(0)} = 1$.*

Definition 1.2.5. ***(Communicate state)*** *Two states i and j are said to communicate, written as $i \leftrightarrow j$ if they are accessible from each other. In other words,*

$$i \leftrightarrow j \quad \text{means} \quad i \to j \quad \text{and} \quad j \to i.$$

Definition 1.2.6. ***(Irreducible Markov chain)*** *A Markov chain is said to be irreducible if all states communicate with each other.*

Definition 1.2.7. *(**Recurrent state**) A state is said to be recurrent if, any time that we leave that state, we will return to that state in the future with probability one. On the other hand, if the probability of returning is less than one, the state is called transient. Here, we provide a formal definition: For any state i, we define*

$$G_{ii} = \mathbb{P}(J_n = i, \textit{for some} \quad n \geq 1 | J_0 = i).$$

State i is recurrent if $G_{ii} = 1$, and it is transient if $G_{ii} < 1$.

Definition 1.2.8. *(**Periodic,aperiodic state**) A state $i \in \mathrm{E}$ is said to be periodic of period $d > 1$, or d-periodic, if d is equal to the greatest common divisor of all n such that $\mathbb{P}(J_{n+1} = i | J_1 = i) > 0$. If $d = 1$, then the state i is said to be aperiodic.*

Definition 1.2.9. *(**Ergodic state**) An aperiodic recurrent state is called ergodic. An irreducible Markov chain with one state ergodic (and then all states ergodic) is called ergodic.*

1.3 Continuous-time Markov chain

Definition 1.3.1. *(**Continuous-time Markov chain**) Let $(J(t))_{t \in \mathbb{R}_+}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in a measurable space $(\mathrm{E}, \varepsilon)$. Unless otherwise stated, we assume that*
$\mathrm{E} = \{1, 2, \ldots, s\}$ *or* $\mathrm{E} = \{1, 2, \ldots\}$.

1. *A stochastic process $(J(t))_{t \in \mathbb{R}_+}$ is called continuous-time Markov chain with the state space E if, for any h, $t \geq 0$ and $j \in \mathrm{E}$ we have*

   $$\mathbb{P}(J(h+t) = j | J(h_1) = i_1, \ldots, J(h_n) = i_n, J(h) = i) = \mathbb{P}(J(h+t) = j | J(h) = i)$$

   $0 \leq h_1 < \ldots < h_n < h$, $n \in \mathbb{N}$, $i_1, \ldots, i_n, i, j \in \mathrm{E}$.

2. *If $\mathbb{P}(J(h+t) = j | J(h) = i)$ does not depend on h, then $(J(t))_{t \in \mathbb{R}_+}$ is said to be homogeneous with respect to time.*

Definition 1.3.2. ***(Transition matrix)*** *Let $(J(t))_{t\in\mathbb{R}_+}$ be a continuous-time Markov process with state space* E. *The functions defined on $\mathbb{R}_+$ by*

$$t \to p_{ij}(t) := \mathbb{P}(J(h+t) = j|J(h) = i), \;\; i, j \in \mathrm{E}$$

are called transition functions of the process. The matrix $\mathbf{p}(t) = (p_{ij}(t))_{i,j\in\mathrm{E}}$ is called the transition matrix (possibly infinite) and $(p(t))_{t\in\mathbb{R}_+}$ is called the transition semigroup of the continuous-time Markov process.

Proposition 1.3.1. ***(Properties of the transition function)****[13]*

1. $\mathbf{p}(t)$ *is a stochastic matrix.*
2. $\mathbf{p}(t)$ *verifies the Chapman-Kolmogrov equation :* $\mathbf{p}(t+h) = \mathbf{p}(t)\mathbf{p}(h)$.
3. $\mathbf{p}(0) = I$.

Proposition 1.2. *[12] Let T_i be the waiting time in state i. The Chapman Kolmogorov equation allows that T_i always has an exponential distribution with a parameter $\lambda_i > 0$*

$$G_i(t) = \mathbb{P}(T_i \leq t) = 1 - e^{-\lambda_i t}, \quad t \geq 0, \quad i \in \mathrm{E}.$$

Example 1. *From the definition of the Poisson process it follows that it is the process with stationary independent increments and*

$$\mathbb{P}(J(t+h) - J(h) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k \in \mathrm{E}, \quad \textit{for all} \quad t > 0, \quad h \geq 0.$$

Each process with stationary independent increments is a homogeneous Markov process with transition probabilities:

$$p_{ij}(t) = \mathbb{P}(J(t+h) - J(h) = j - i).$$

Hence, the Poisson process is the homogeneous Markov process with the transition probabilities given by

$$p_{ij}(t) = \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t}, \quad i, j \in \mathrm{E}.$$

Chapter 2

Discrete-time semi-Markov process

Discrete-time semi-Markov processes (DTSMPs) and discrete-time Markov renewal processes (DTMRPs) are a class of stochastic processes which generalize discrete-time Markov chains and discrete-time renewal processes.

For a discrete-time Markov process, the sojourn time in each state is geometrically distributed. In the semi-Markov case, the sojourn time distribution can be any distribution on $\mathbb{N}^*$. This is the reason why the semi-Markov approach is much more suitable for applications than the Markov one.

2.1 Markov renewal chains and semi-Markov chains

Let us consider:

- E the state space. We suppose E to be finite, with $\mid \mathrm{E} \mid = s$.
- The stochastic process $J = (J_n)_{n \geq 0}$ with state space E for the system state at the n^{th} jump.
- The stochastic process $S = (S_n)_{n \geq 0}$ with state space $\mathbb{N}$ for the n^{th} jump. We suppose $S_0 = 0$ and $0 < S_1 < S_2 < \ldots < S_n < S_{n+1} < \ldots$

- The stochastic process $X = (X_n)_{n\geq 0}$ with state space $\mathbb{N}^\star$ for the sojourn time X_n in state J_{n-1} before the n^{th} jump. Thus, $X_n = S_n - S_{n-1}$, for all $n \in \mathbb{N}^\star$.

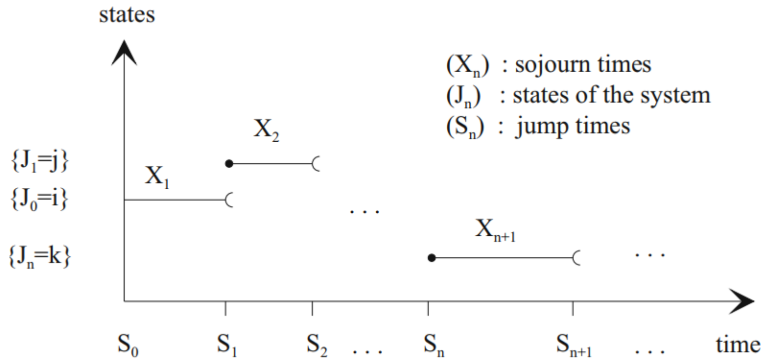

Fig 1.1 : Sample path of a semi-Markov chain.

Definition 2.1.1. *(**Markov renewal chain**) The stochastic process $(J,S) = (J_n, S_n)_{n\in\mathbb{N}}$ is said to be a Markov renewal chain (MRC) if for all $n \in \mathbb{N}$, for all $i, j \in \mathrm{E}$ and for all $k \in \mathbb{N}$ it almost surely satisfies*

$$\mathbb{P}(J_{n+1} = j, S_{n+1} - S_n = k | J_0, \ldots, J_n; S_0, \ldots, S_n) = \mathbb{P}(J_{n+1} = j, S_{n+1} - S_n = k | J_n). \tag{2.1}$$

Moreover, if equation (2.1) is independent of n, (J,S) is said to be homogeneous, with discrete semi-Markov kernel $\mathbf{q} = (q_{ij}(k); i, j \in \mathrm{E}, k \in \mathbb{N})$ defined by

$$q_{ij}(k) = \mathbb{P}(J_{n+1} = j, X_{n+1} = k | J_n = i),\ k > 0, \text{ and } q_{ij}(0) = 0.$$

Let us introduce the cumulated semi-Markov kernel
$\mathbf{Q} = (\mathbf{Q}(k), k \in \mathbb{N}) \in \mathcal{M}_\mathrm{E}(\mathbb{N})$ defined, for all $i, j \in \mathrm{E}$ and for all $k \in \mathbb{N}$, by

$$Q_{ij}(k) = \mathbb{P}(J_{n+1} = j, X_{n+1} \leq k | J_n = i) = \sum_{l=0}^{k} q_{ij}(l).$$

Proposition 2.1. *[5] For all $i, j \in \mathrm{E}$, for all n and $k \in \mathbb{N}$, we have*

$$\mathbb{P}(J_n = j, S_n = k | J_0 = i) = q_{ij}^{(n)}(k).$$

Proof. We prove the result by induction. For $n = 0$, we have

$$\mathbb{P}(J_0 = j, S_0 = k | J_0 = i) = q_{ij}^{(0)}(k).$$

Obviously, for $k \neq 0$ or $i \neq j$, this probability is zero. On the other hand, if $i = j$ and $k = 0$, the probability is one, thus the result follows.
For $n = 1$, the result obviously holds true, using the definition of the semi-Markov kernel $\mathbf{q}$ and of $q_{ij}^{(1)}(k)$. For $n \geq 2$:

$$\begin{aligned}
\mathbb{P}(J_n = j, S_n = k | J_0 = i) &= \sum_{r \in \mathrm{E}} \sum_{l=1}^{k-1} \mathbb{P}(J_n = j, S_n = k, J_1 = r, S_1 = l | J_0 = i) \\
&= \sum_{r \in \mathrm{E}} \sum_{l=1}^{k-1} \mathbb{P}(J_n = j, S_n = k | J_1 = r, S_1 = l, J_0 = i) \\
&\quad \mathbb{P}(J_1 = r, S_1 = l | J_0 = i) \\
&= \sum_{r \in \mathrm{E}} \sum_{l=1}^{k-1} \mathbb{P}(J_{n-1} = j, S_{n-1} = k - l | J_0 = r) \mathbb{P}(J_1 = r, X_1 = l | J_0 = i) \\
&= \sum_{r \in \mathrm{E}} \sum_{l=1}^{k-1} q_{rj}^{(n-1)}(k - l) q_{ir}(l) = q_{ij}^{(n)}(k). \square
\end{aligned}$$

Let us also consider the matrix function $\psi = (\psi(k), k \in \mathbb{N}) \in \mathcal{M}_{\mathrm{E}}(\mathbb{N})$, defined by

$$\psi_{ij}(k) = \sum_{n=0}^{\infty} q^{(n)}(k) = \sum_{n=0}^{k} q^{(n)}(k), \;\; i, j \in \mathrm{E}, \;\; k \in \mathbb{N}.$$

The infinite series which appears in the definition of ψ proves to be a finite series due to the fact that $q_{ij}^{(n)}(k) = 0$ for all n and $k \in \mathbb{N}$ such that $n > k$. Note that this property is specific to a semi-Markov process with discrete-time.

Definition 2.1.2. ***(Discrete-time semi-Markov chain)*** *Let (J, S) be a Markov renewal chain. The chain $Z = (Z_k)_{k \in \mathbb{N}}$ is said to be a semi-Markov chain associated to the MRC (J, S) if*

$$Z_k := J_{N(k)}, \quad k \in \mathbb{N}$$

where

$$N(k) := max\{n \geq 0;\ \ S_n \leq k\},$$

is the discrete-time counting process of the number of jumps in $[1, k] \subset \mathbb{N}$.

Thus Z_k *gives the system state at time* k. *We have also* $J_n = Z_{S_n}$.

Let the row vector $\alpha = (\alpha_1, \ldots, \alpha_s)$ denote the initial distribution of the semi-Markov chain $Z = (Z_k)_{k\in\mathbb{N}}$ i.e $\alpha_i := \mathbb{P}(Z_0 = i) = \mathbb{P}(J_0 = i),\ \ i \in \mathrm{E}$.

Remark 2.1.1. $J = (J_n)_{n\in\mathbb{N}}$ *is a Markov chain, called the embedded Markov chain* (EMC).

Definition 2.1.3. ***(Transition function of the semi-Markov)*** *The transition function of the semi-Markov chain* Z *is the matrix-valued function* $\mathbf{P} = (P_{ij}(k); i, j \in \mathrm{E}, k \in \mathbb{N}) \in \mathcal{M}_{\mathrm{E}}(\mathbb{N})$ *defined by*

$$P_{ij}(k) := \mathbb{P}(Z_k = j|Z_0 = i), i, j \in \mathrm{E}, k \in \mathbb{N}.$$

Definition 2.1.4. ***(Conditional distributions of sojourn times)*** *For all* $i, j \in \mathrm{E}$, *let us define:*

- $f_{ij}(.)$, *the conditional distribution of sojourn time in state i before going to state j:*

$$f_{ij}(k) := \mathbb{P}(X_{n+1} = k|J_n = i, J_{n+1} = j),\ \ \forall k \in \mathbb{N}.$$

- $F_{ij}(.)$, *the conditional cumulative distribution of* X_{n+1}, $n \in \mathbb{N}$*:*

$$F_{ij}(k) := \mathbb{P}(X_{n+1} \leq k|J_n = i, J_{n+1} = j) = \sum_{l=0}^{k} f_{ij}(l),\ \ \forall k \in \mathbb{N}.$$

- $h_i(.)$, *the sojourn time distribution in state i:*

$$h_i(k) := \mathbb{P}(X_{n+1} = k|J_n = i) = \sum_{j\in\mathrm{E}} q_{ij}(k), \forall k \in \mathbb{N}.$$

- *$H_i(.)$, the sojourn time cumulative distribution function in state i:*

$$H_i(k) := \mathbb{P}(X_{n+1} \leq k | J_n = i) = \sum_{l=1}^{k} h_i(l), \forall k \in \mathbb{N}.$$

- *$\overline{H}_i(.)$, the survival function of sojourn time in state i:*

$$\overline{H}_i(k) := \mathbb{P}(X_{n+1} > k | J_n = i), \forall k \in \mathbb{N}.$$

Obviously, for all $i, j \in \mathrm{E}$ and $k \in \mathbb{N}$, we have $q_{ij}(k) = p_{ij} f_{ij}(k)$.

The following assumptions concerning the Markov renewal chain will be needed in the rest of this work.

A1 The Markov chain $(J_n)_{n\in\mathbb{N}}$ is irreducible.

A2 The mean sojourn times are finite, i.e. $\sum_{k=0} k h_i(k) < \infty$ for any state $i \in \mathrm{E}$.

A3 The Markov renewal process $(J_n, S_n)_{n\in\mathbb{N}}$ is aperiodic.

2.2 Elements of statistical estimation

Let us consider a sample path of the DTMRP $(J_n, S_n)_{n\in\mathbb{N}}$, censored at time $M \in \mathbb{N}$ ($X_{N(M)+1}$ is above u_M but it is unknown by how much).

$$\mathcal{H}(M) := (J_0, X_1, \ldots, J_{N(M)-1}, X_{N(M)}, J_{N(M)}, u_M),$$

where $N(M)$ is the number of jumps of the process in $[1, M] \subset \mathbb{N}$ and $u_M := M - S_{N(M)}$ is the censored sojourn time in the last visited state $J_{N(M)}$.

2.2.1 Empirical estimators

Taking a sample path $\mathcal{H}(M)$ of a DTMRP, for all $i, j \in E$ and $k \in \mathbb{N}$, $k \leq M$, we define the empirical estimators of the transition matrix of the embedded Markov chain p_{ij}, of the conditional sojourn time $f_{ij}(k)$ and of the discrete semi-Markov kernel $q_{ij}(k)$ by

$$\widehat{p}_{ij}(M) := \frac{N_{ij}(M)}{N_i(M)}, \ \widehat{f}_{ij}(k, M) := \frac{N_{ij}(k, M)}{N_{ij}(M)}, \ \widehat{q}_{ij}(k, M) := \frac{N_{ij}(k, M)}{N_i(M)}. \quad (2.2)$$

where $N_{ij}(k, M)$, $N_i(M)$ and $N_{ij}(M)$ are given by

- $N_i(M) := \sum_{n=1}^{N(M)} \mathbf{1}_{\{J_n=i\}}$: the number of visits to state i, up to time M;
- $N_{ij}(M) := \sum_{n=1}^{N(M)} \mathbf{1}_{\{J_{n-1}=i, J_n=j\}}$: the number of transitions from i to j, up to time M;
- $N_{ij}(k, M) := \sum_{n=1}^{N(M)} \mathbf{1}_{\{J_{n-1}=i, J_n=j, X_n=k\}}$: the number of transitions from i to j, up to time M, with sojourn time in state i equal to k, $1 \leq k \leq M$.

The likelihood function corresponding to the history $\mathcal{H}(M)$ is

$$\mathrm{L}(M) = \alpha_{J_0} \prod_{k=1}^{N(M)} p_{J_{k-1}J_k} f_{J_{k-1}J_k}(X_k) \overline{H}_{J_{N(M)}}(u_M),$$

where $\overline{H}_{J_{N(M)}}$ is the survival function in state i and α_i is the initial distribution of state i.

Lemma 2.2.1.1. *[5]*

For a semi-Markov chain $Z = (Z_n)_{n\in\mathbb{N}}$ we have

$$u_M/M \xrightarrow[M\to\infty]{a.s} 0.$$

The previous lemma tells us that, for large M, u_M does not add significant information to the likelihood function. For these reason, we will neglect the term $\overline{H}_{J_{N(M)}}(u_M)$ in the expression of the likelihood function $\mathrm{L}(M)$. On the other side, the sample path $\mathcal{H}(M)$ of the MRC $(J_n, S_n)_{n\in\mathbb{N}}$ contains only one observation of the initial distribution α of $(J_n)_{n\in\mathbb{N}}$, so the information on α_{J_0} does not increase with M. As we are interested in large-sample estimation of semi-Markov chains, the term α_{J_0} will be equally neglected in the expression of the likelihood function.

Consequently, we will be concerned with the maximization of the approached likelihood function defined by

$$\mathrm{L}_1(M) = \prod_{k=1}^{N(M)} p_{J_{k-1}J_k} f_{J_{k-1}J_k}(X_k). \tag{2.3}$$

Proposition 2.2. *[5] For a sample path of a DTMRP $(J_n, S_n)_{n\in\mathbb{N}}$, censored at time $M \in \mathbb{N}$, the empirical estimators $\widehat{p}_{ij}(M)$, $\widehat{f}_{ij}(k,M)$ and $\widehat{q}_{ij}(k,M)$, proposed in equation* (2.2)*, are approached non-parametric maximum likelihood estimators i.e. they maximize the approached likelihood function* L_1*, given in equation* (2.3)*.*

Proof. We consider the approached likelihood function $\mathrm{L}_1(M)$ given by equation (2.3). Using the equality

$$\sum_{j=1}^{s} p_{ij} = 1 \tag{2.4}$$

the approached log-likelihood function can be written in the form

$$\log(\mathrm{L}_1(M)) = \sum_{k=1}^{M}\sum_{i,j=1}^{s}[N_{ij}(M)\log(p_{ij}) + N_{ij}(k,M)\log(f_{ij}(k)) + \lambda_i(1-\sum_{j=1}^{s}p_{ij})], \tag{2.5}$$

where the Lagrange multipliers λ_i are arbitrarily chosen constants.
In order to obtain the approached MLE of p_{ij} we maximize equation (2.5) with respect to p_{ij}, and get $p_{ij} = N_{ij}(M)/\lambda_i$. Equation (2.4) becomes

$$1 = \sum_{j=1}^{s} p_{ij} = \sum_{j=1}^{s}\frac{N_{ij}(M)}{\lambda_i} = \frac{N_i(M)}{\lambda_i}.$$

Finally, we infer that the values λ_i which maximize equation (2.5) with respect to p_{ij} are given by $\lambda_i = N_i(M)$ and we obtain

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

The expression of $\widehat{f}_{ij}(k,M)$ can be obtained by the same method. Indeed, using the equality

$$\sum_{k=1}^{\infty} f_{ij}(k) = 1 \tag{2.6}$$

we write the approached log-likelihood function in the form

$$\log(\mathrm{L}_1(M)) = \sum_{k=1}^{M}\sum_{i,j=1}^{s}[N_{ij}(M)\log(p_{ij}) + N_{ij}(k,M)\log(f_{ij}(k)) + \lambda_{ij}(1-\sum_{k=1}^{\infty} f_{ij}(k))], \tag{2.7}$$

where λ_{ij} are arbitrarily chosen constants. Maximizing (2.7) with respect to $f_{ij}(k)$ we obtain $\widehat{f}_{ij}(k,M) := N_{ij}(k,M)/\lambda_{ij}$.

From Equation (2.6) we obtain $\lambda_{ij}(M) = N_{ij}(M)$. Thus $\widehat{f}_{ij}(k,M) := N_{ij}(k,M)/N_{ij}(M)$.

In an analogous way we can prove that the expression of the approached MLE of the kernel $q_{ij}(k)$ is given by equation (2.2). □

Lemma 2.2.1. *[5] For a MRC that satisfies Assumptions A1 and A2, we have:*

1. $\lim_{M\to\infty} S_M = \infty$ *a.s;*
2. $\lim_{M\to\infty} N(M) = \infty$ *a.s.*

Lemma 2.2.2. *[5] For the DTMRP $(J_n, S_n)_{n\in\mathbb{N}}$. We have*

$$\frac{N_i(M)}{M} \xrightarrow[M\to\infty]{a.s} \frac{1}{\mu_{ii}},\quad \frac{N_{ij}(M)}{M} \xrightarrow[M\to\infty]{a.s} \frac{p_{ij}}{\mu_{ii}},\quad \frac{N(M)}{M} \xrightarrow[M\to\infty]{a.s} \frac{1}{\nu(l)\mu_{ll}}.$$

where μ_{ii} is the mean recurrence time of state i for the semi-Markov process $(Z_n)_{n\in\mathbb{N}}$, $(\nu(l); l\in \mathrm{E})$ the stationary distribution and l is an arbitrary fixed state.

2.3 Asymptotic properties of the estimators

In this section, we study the asymptotic properties (consistency and asymptotic normality) of the proposed estimators $\widehat{p}_{ij}(M)$, $\widehat{f}_{ij}(k,M)$ and $\widehat{q}_{ij}(k,M)$.

2.3.1 Strong consistency

Corollary 2.3.1. *[5] For any $i,j\in \mathrm{E}$, under A1, we have*

$$\widehat{p}_{ij}(M) = \frac{N_{ij}(M)}{N_i(M)} \xrightarrow[M\to\infty]{a.s} p_{ij}.$$

For $i,j\in\mathrm{E}$ two fixed states, let us also define the empirical estimator of the conditional cumulative distribution of $(X_n)_{n\in\mathbb{N}^*}$

$$\widehat{F}_{ij}(k, M) := \sum_{l=0}^{k} \widehat{f}_{ij}(l, M) = \sum_{l=0}^{k} \frac{N_{ij}(l, M)}{N_{ij}(M)}. \tag{2.8}$$

The following result concerns the convergence of $\widehat{f}_{ij}(k, M)$ and $\widehat{F}_{ij}(k, M)$.

Proposition 2.3. *[5] For any fixed arbitrary states $i, j \in \mathrm{E}$, the empirical estimators $\widehat{f}_{ij}(k, M)$ and $\widehat{F}_{ij}(k, M)$ proposed in equations* (2.2) *and* (2.8)*, are uniformly strongly consistent, i.e.*

1. $\max\limits_{i,j\in \mathrm{E}} \max\limits_{0\le k\le M} |\widehat{F}_{ij}(k, M) - F_{ij}(k)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0.$

2. $\max\limits_{i,j\in \mathrm{E}} \max\limits_{0\le k\le M} |\widehat{f}_{ij}(k, M) - f_{ij}(k)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0.$

Proof. We first prove the strong consistency of the estimators using the SLLN theorem 1.1.1 . Second, we show the uniform consistency, i.e., that the convergence does not depend on the chosen k, $0 \le k \le M$. This second part is done by means of the Glivenko-Cantelli theorem 1.1.2.

Obviously, the strong consistency can be directly obtained using Glivenko-Cantelli theorem 1.1.2. Anyway, we prefer to derive separately the consistency result because it is easy and constructive.

Let us denote by $\{n_1, n_2, \ldots, n_{N_{ij}(M)}\}$ the transition times from state i to state j, up to time M. Note that we have

$$\widehat{F}_{ij}(k, M) = \frac{1}{N_{ij}(M)} \sum_{l=1}^{N_{ij}(M)} \mathbf{1}_{\{X_{n_l} \le k\}},$$

and

$$\widehat{f}_{ij}(k, M) = \frac{1}{N_{ij}(M)} \sum_{l=1}^{N_{ij}(M)} \mathbf{1}_{\{X_{n_l} = k\}}.$$

For any $l \in \{1, 2, \ldots, N_{ij}(M)\}$ we have

$$\mathbb{E}[\mathbf{1}_{\{X_{n_l} \le k\}}] = \mathbb{P}(X_{n_l} \le k) = F_{ij}(k),$$

and

$$\mathbb{E}[\mathbf{1}_{\{X_{n_l} = k\}}] = \mathbb{P}(X_{n_l} = k) = f_{ij}(k).$$

Since $N_{ij}(M) \xrightarrow[M\to\infty]{a.s} \infty$, applying the SLLN theorem 1.1.1 to the sequences of i.i.d. random variables $\{\mathbf{1}_{\{X_{n_l}\leq k\}}\}_{l\in\{1,2,\ldots,N_{ij}(M)\}}$ and $\{\mathbf{1}_{\{X_{n_l}=k\}}\}_{l\in\{1,2,\ldots,N_{ij}(M)\}}$, and using Theorem 1.1.3, we get

$$\widehat{F}_{ij}(k,M) = \frac{1}{N_{ij}(M)} \sum_{l=1}^{N_{ij}(M)} \mathbf{1}_{\{X_{n_l}\leq k\}} \xrightarrow[M\to\infty]{a.s} \mathbb{E}[\mathbf{1}_{\{X_{n_l}\leq k\}}] = F_{ij}(k),$$

and

$$\widehat{f}_{ij}(k,M) = \frac{1}{N_{ij}(M)} \sum_{l=1}^{N_{ij}(M)} \mathbf{1}_{\{X_{n_l}=k\}} \xrightarrow[M\to\infty]{a.s} \mathbb{E}[\mathbf{1}_{\{X_{n_l}=k\}}] = f_{ij}(k).$$

In order to obtain uniform consistency, from the Glivenko-Cantelli theorem 1.1.2, we have

$$\max_{0\leq k\leq m} |\frac{1}{m}\sum_{l=1}^{m}\mathbf{1}_{\{X_{n_l}\leq k\}} - F_{ij}(k)| \xrightarrow[M\to\infty]{a.s} 0.$$

Let us define $\xi_m := \max_{0\leq k\leq m} |\frac{1}{m}\sum_{l=1}^{m}\mathbf{1}_{\{X_{n_l}\leq k\}} - F_{ij}(k)|$. The previous convergence tells us that $\xi_m \xrightarrow[m\to\infty]{a.s} 0$. As $N(M) \xrightarrow[M\to\infty]{a.s} \infty$ (2.2.1) applying Theorem 1.1.3 we obtain $\xi_{N(M)} \xrightarrow[M\to\infty]{a.s} 0$ which reads

$$\max_{0\leq k\leq M} |\widehat{F}_{ij}(k,M) - F_{ij}(k)| \xrightarrow[M\to\infty]{a.s} 0.$$

As the state space E is finite, we take the maximum with respect to $i, j \in \mathrm{E}$ and the desired result for $\widehat{F}_{ij}(k,M)$ follows.

Concerning the uniform consistency of $\widehat{f}_{ij}(k,M)$, note that we have

$$\max_{i,j\in\mathrm{E}} \max_{0\leq k\leq M} |\widehat{f}_{ij}(k,M) - f_{ij}(k)| = \max_{i,j\in\mathrm{E}} \max_{0\leq k\leq M} |\widehat{F}_{ij}(k,M) - \widehat{F}_{ij}(k-1,M) - F_{ij}(k) + F_{ij}(k-1)|$$

$$\leq \max_{i,j\in\mathrm{E}} \max_{0\leq k\leq M} |\widehat{F}_{ij}(k,M) - F_{ij}(k)| + \max_{i,j\in\mathrm{E}} \max_{0\leq k\leq M} |\widehat{F}_{ij}(k-1,M) - F_{ij}(k-1)|$$

and the result follows from the uniform strong consistency of $\widehat{F}_{ij}(k,M)$.□

Proposition 2.4. *[5] The empirical estimator of the semi-Markov kernel proposed in equation* (2.2) *is uniformly strongly consistent, i.e.*

$$\max_{i,j\in\mathrm{E}} \max_{0\leq k\leq M} |\widehat{q}_{ij}(k,M) - q_{ij}(k)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0.$$

Proof. Firstly, from Corollary 2.3.1, we immediately obtain the almost sure convergence of $\widehat{p}_{ij}(M)$. The uniform strong consistency of $\widehat{q}_{ij}(k, M)$ follows from the consistency of the estimators $\widehat{p}_{ij}(M)$, $\widehat{f}_{ij}(k, M)$ (Proposition 2.2) and from the following inequality

$$\begin{aligned}
\max_{i,j\in E}\max_{0\le k\le M}|\widehat{q}_{ij}(k,M) - q_{ij}(k)| &= \max_{i,j\in E}\max_{0\le k\le M}|\widehat{p}_{ij}(M)\widehat{f}_{ij}(k,M) - \widehat{p}_{ij}(M)f_{ij}(k) \\
&\quad +\widehat{p}_{ij}(M)f_{ij}(k) - p_{ij}f_{ij}(k)| \\
&\le \max_{i,j\in E}\widehat{p}_{ij}(M)\max_{i,j\in E}\max_{0\le k\le M}|\widehat{f}_{ij}(k,M) - f_{ij}(k)| \\
&\quad +\max_{i,j\in E}\max_{0\le k\le M} f_{ij}(k)\max_{i,j\in E}|\widehat{p}_{ij}(M) - p_{ij}| \\
&\le \max_{i,j\in E}|\widehat{p}_{ij}(M) - p_{ij}| + \max_{i,j\in E}\max_{0\le k\le M}|\widehat{f}_{ij}(k,M) - f_{ij}(k)|.
\end{aligned}$$

The conclusion follows from the consistency of $\widehat{p}_{ij}(M)$ and $\widehat{f}_{ij}(k, M)$ □

2.3.2 Asymptotic normality

We present further theorem CLT for additive functionals of Markov renewal chains. Let f be a real function defined on $E \times E \times \mathbb{N}$. Define, for each $M \in \mathbb{N}$, the functional $W_f(M)$ as

$$W_f(M) := \sum_{n=1}^{N(M)} f(J_{n-1}, J_n, X_n),$$

or, equivalently,

$$W_f(M) := \sum_{i,j=1}^{s}\sum_{n=1}^{N_{ij}(M)} f(i, j, X_{ijn}),$$

where X_{ijn} is the n^{th} sojourn time of the chain in state i, before going to state j. Set

$$\begin{aligned}
A_{ij} &:= \sum_{x=1}^{\infty} f(i,j,x)q_{ij}(x), \qquad A_i := \sum_{j=1}^{s} A_{ij}, \\
B_{ij} &:= \sum_{x=1}^{\infty} f^2(i,j,x)q_{ij}(x), \qquad B_i := \sum_{j=1}^{s} B_{ij},
\end{aligned}$$

if the sums exist. Define

$$
\begin{aligned}
r_i &:= \sum_{j=1}^{s} A_j \frac{\mu_{ii}^*}{\mu_{jj}^*}, \qquad m_f := \frac{r_i}{\mu_{ii}} \\
\sigma_i^2 &:= -r_i^2 + \sum_{j=1}^{s} B_j \frac{\mu_{ii}^*}{\mu_{jj}^*} + 2\sum_{r=1}^{s}\sum_{l\neq i}\sum_{k\neq i} A_{rl}A_k\mu_{ii}^* \frac{\mu_{li}^* + \mu_{ik}^* - \mu_{lk}^*}{\mu_{rr}^*\mu_{kk}^*}, B_f := \frac{\sigma_i^2}{\mu_{ii}}
\end{aligned}
$$

Where μ_{ii}^* is the mean recurrence time of state i for the Markov chain $(J_n)_{n\geq 0}$.

Theorem 2.3.1. ***(Central Limit Theorem)*** *[21]*
For an aperiodic Markov renewal chain that satisfies Assumptions A1 and A2 we have

$$\sqrt{M}\left[\frac{W_f(M)}{M} - m_f\right] \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}(0, B_f).$$

Theorem 2.3.2. *[5] For $i, j \in \mathrm{E}$, and $k \in \mathbb{N}$,*
$\sqrt{M}[\widehat{q}_{ij}(k, M) - q_{ij}(k)]$ converges in distribution, as $M \to \infty$, to a zero mean normal random variable with variance $\mu_{ii}q_{ij}(k)[1 - q_{ij}(k)]$.

Proof. We present two different proofs of the theorem. The first one is based on the CLT for Markov renewal chains (Theorem 2.3.1). The second one relies on the Lindeberg-Lévy CLT for martingales (Theorem 1.1.4).

Method 1.

$$\sqrt{M}[\widehat{q}_{ij}(k, M) - q_{ij}(k)] = \frac{M}{N_i(M)}\frac{1}{\sqrt{M}}\sum_{n=1}^{N(M)}[\mathbf{1}_{\{J_n=j,X_n=k\}} - q_{ij}(k)]\mathbf{1}_{\{J_{n-1}=i\}} = \sum_{n=1}^{N(M)} f(J_{n-1}, J_n, X_n).$$

Let us consider the function

$$f(m, l, u) := \mathbf{1}_{\{m=i,l=j,u=k\}} - q_{ij}(k)\mathbf{1}_{\{m=i\}}.$$

Using the notation from the Pyke and Schaufele's CLT, we have

$$W_f(M) = \sum_{n=1}^{N(M)} f(J_{n-1}, J_n, X_n) = \sum_{n=1}^{N(M)} [\mathbf{1}_{\{J_n=j,X_n=k\}} - q_{ij}(k)]\mathbf{1}_{\{J_{n-1}=i\}}.$$

In order to apply Pyke and Schaufeles'central limit theorem for Markov renewal processes (Theorem 2.3.1), we need to compute $A_{ml}, A_m, B_{ml}, B_m, m_f$ and B_f for $m, l \in$ E.

$$
\begin{aligned}
A_{ml} &:= \sum_{u=1}^{\infty} f(m,l,u) q_{ml}(u), \\
&:= \sum_{u=1}^{\infty} \mathbf{1}_{\{m=i,l=j,u=k\}} q_{ml}(u) - \sum_{u=1}^{\infty} \mathbf{1}_{\{m=i\}} q_{ij}(k) q_{ml}(u) \\
&:= \delta_{mi}\delta_{lj} \sum_{u=1}^{\infty} \mathbf{1}_{\{u=k\}} q_{ij}(u) - \delta_{mi} q_{ij}(k) \sum_{u=1}^{\infty} q_{il}(u) = q_{ij}(k)\delta_{mi}(\delta_{lj} - p_{il}) \\
A_m &:= \sum_{l=1}^{s} A_{ml} = q_{ij}(k)\delta_{mi}[\sum_{l=1}^{s} \delta_{lj} - \sum_{l=1}^{s} p_{il}] = 0. \\
B_{ml} &:= \sum_{u=1}^{\infty} f^2(m,l,u) q_{ml}(u) \\
&:= \sum_{u=1}^{\infty} \mathbf{1}_{\{m=i,l=j,u=k\}} q_{ml}(u) + \sum_{u=1}^{\infty} \mathbf{1}_{\{m=i\}} q_{ij}^2(k) q_{ml}(u) \\
&\quad -2 \sum_{u=1}^{\infty} \mathbf{1}_{\{m=i,l=j,u=k\}} q_{ij}(k) q_{ml}(u) \\
&:= q_{ij}(k)\delta_{mi}\delta_{lj} + q_{ij}^2(k)\delta_{mi} p_{il} - 2q_{ij}^2(k)\delta_{mi}\delta_{lj} \\
B_m &:= \sum_{l=1}^{s} B_{ml} = \delta_{mi} q_{ij}(k)[1 - q_{ij}(k)].
\end{aligned}
$$

Finally, we obtain

$$
r_i := \sum_{m=1}^{s} A_m \frac{\mu_{ii}^*}{\mu_{mm}^*} = 0, \qquad m_f := \frac{r_i}{\mu_{ii}} = 0,
$$

$$
\sigma_i^2 := \sum_{m=1}^{s} B_m \frac{\mu_{ii}^*}{\mu_{mm}^*} = q_{ij}(k)[1 - q_{ij}(k)], \qquad B_f := \frac{\sigma_i^2}{\mu_{ii}} = \frac{q_{ij}(k)[1 - q_{ij}(k)]}{\mu_{ii}}.
$$

Since $N_i(M)/M \xrightarrow[M\to\infty]{a.s} 1/\mu_{ii}$ (see Lemma 2.2.2), we conclude as follows:

$$
\sqrt{M}[\widehat{q}_{ij}(k,M) - q_{ij}(k)] \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}(0, \mu_{ii} q_{ij}(k)[1 - q_{ij}(k)]).
$$

Method 2.

For $i, j \in E$ arbitrarily fixed states and $k \in \mathbb{N}$ arbitrarily fixed positive integer, we write the random variable $\sqrt{M}[\widehat{q}_{ij}(k,M) - q_{ij}(k)]$ as

$$\sqrt{M}[\widehat{q}_{ij}(k,M) - q_{ij}(k)] = \frac{M}{N_i(M)} \frac{1}{\sqrt{M}} \sum_{n=1}^{N(M)} \left[\mathbf{1}_{\{J_n=j, X_n=k\}} - q_{ij}(k)\right] \mathbf{1}_{\{J_{n-1}=i\}}.$$

Let $\mathcal{F}_n$ be the σ-algebra defined by $\mathcal{F}_n := \sigma(J_l, X_l; l \le n), n \ge 0$, and let Y_n be the random variable

$$Y_n = \mathbf{1}_{\{J_{n-1}=i, J_n=j, X_n=k\}} - q_{ij}(k)\mathbf{1}_{\{J_{n-1}=i\}}.$$

Obviously, Y_n is $\mathcal{F}_n$-measurable and $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, for all $n \in \mathbb{N}$. Moreover, we have

$$\begin{aligned} \mathbb{E}(Y_n \mid \mathcal{F}_{n-1}) &= \mathbb{P}(J_{n-1} = i, J_n = j, X_n = k \mid \mathcal{F}_{n-1}) - q_{ij}(k)\mathbb{P}(J_{n-1} = i \mid \mathcal{F}_{n-1}) \\ &= \mathbf{1}_{\{J_{n-1}=i\}}\mathbb{P}(J_n = j, X_n = k \mid J_{n-1} = i) - q_{ij}(k)\mathbf{1}_{\{J_{n-1}=i\}} \\ &= 0. \end{aligned}$$

Therefore, $(Y_n)_{n\in\mathbb{N}}$ is an $\mathcal{F}_n$-martingale difference and $(\sum_{l=1}^{n} Y_l)_{l\in\mathbb{N}}$ is an $\mathcal{F}_n$-martingale. Note also that, as Y_l is bounded for all $l \in \mathbb{N}$, we have

$$\frac{1}{\sqrt{n}} \sum_{l=1}^{n} \mathbb{E}(Y_l^2 \mathbf{1}_{\{|Y_l| > \epsilon\sqrt{n}\}}) \underset{n\to\infty}{\longrightarrow} 0.$$

For any $\epsilon > 0$. Using the CLT for martingales (Theorem 1.1.4) we obtain

$$\frac{1}{\sqrt{n}} \sum_{l=1}^{n} Y_l \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \sigma^2), \tag{2.9}$$

where $\sigma^2 > 0$ is given by

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{l=1}^{n} \mathbb{E}(Y_l^2 \mid \mathcal{F}_{l-1}) > 0.$$

As $N(M)/M \xrightarrow[M\to\infty]{a.s} 1/\nu(l)\mu_{ll}$ applying Anscombe's theorem (Theorem 1.1.5) we obtain

$$\frac{1}{\sqrt{N(M)}} \sum_{l=1}^{N(M)} Y_l \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}(0, \sigma^2). \tag{2.10}$$

To obtain σ^2, we need to compute Y_l^2 and $\mathbb{E}(Y_l^2 \mid \mathcal{F}_{l-1})$. First,

$$Y_l^2 = \mathbf{1}_{\{J_{l-1}=i, J_l=j, X_l=k\}} + (q_{ij}(k))^2 \mathbf{1}_{\{J_{l-1}=i\}} - 2q_{ij}(k)\mathbf{1}_{\{J_{l-1}=i, J_l=j, X_l=k\}}.$$

Second,

$$\begin{aligned}
\mathbb{E}(Y_l^2 \mid \mathcal{F}_{l-1}) &= \mathbf{1}_{\{J_{l-1}=i\}}\mathbb{P}(J_l=j, X_l=k \mid J_{l-1}=i) \\
&\quad + (q_{ij}(k))^2\mathbf{1}_{\{J_{l-1}=i\}} - 2\mathbf{1}_{\{J_{l-1}=i\}}q_{ij}(k)\mathbb{P}(J_l=j, X_l=k \mid J_{l-1}=i) \\
&= \mathbf{1}_{\{J_{l-1}=i\}}q_{ij}(k) + (q_{ij}(k))^2\mathbf{1}_{\{J_{l-1}=i\}} - 2(q_{ij}(k))^2\mathbf{1}_{\{J_{l-1}=i\}} \\
&= \mathbf{1}_{\{J_{l-1}=i\}}q_{ij}(k)[1-q_{ij}(k)].
\end{aligned}$$

Thus, σ^2 given by

$$\sigma^2 = \lim_{n\to\infty}\left(\frac{1}{\sqrt{n}}\sum_{l=1}^{n}\mathbf{1}_{\{J_{l-1}=i\}}\right)q_{ij}(k)[1-q_{ij}(k)] = \nu(i)q_{ij}(k)[1-q_{ij}(k)],$$

where ν is the stationary distribution of the embedded Markov chain $(J_n)_{n\in\mathbb{N}}$. The random variable of interest $\sqrt{M}[\widehat{q}_{ij}(k,M) - q_{ij}(k)]$ can be written as

$$\begin{aligned}
\sqrt{M}[\widehat{q}_{ij}(k,M) - q_{ij}(k)] &= \frac{M}{N_i(M)}\frac{1}{\sqrt{M}}\sqrt{N(M)}\frac{1}{\sqrt{N(M)}}\sum_{l=1}^{N(M)}Y_l \\
&= \frac{M}{N_i(M)}\sqrt{\frac{N(M)}{M}}\frac{1}{\sqrt{N(M)}}\sum_{l=1}^{N(M)}Y_l.
\end{aligned}$$

Note that we have

$$\begin{aligned}
\frac{N_i(M)}{M} &\xrightarrow[M\to\infty]{a.s} \frac{1}{\mu_{ii}}, \\
\frac{N(M)}{M} &\xrightarrow[M\to\infty]{a.s} \frac{1}{\nu(i)\mu_{ii}}.
\end{aligned}$$

Using these results and convergence (2.10), we obtain that $\sqrt{M}[\widehat{q}_{ij}(k,M) - q_{ij}(k)]$ converges in distribution, as M tends to infinity, to a zero-mean normal random variable, of variance

$$\begin{aligned}
\sigma_q^2(i,j,k) &= (\mu_{ii}\sqrt{1/\mu_{ii}\nu(i)})^2\nu(i)q_{ij}(k)[1-q_{ij}(k)] \\
&= \mu_{ii}q_{ij}(k)[1-q_{ij}(k)],
\end{aligned}$$

which is the desired result. □

2.4 Markov renewal equation

Definition 2.4.1. ***(Discrete-time Markov renewal equation).***

Let $\mathbf{L} = (L_{ij}(k); i, j \in \mathrm{E}, k \in \mathbb{N}) \in \mathcal{M}_{\mathrm{E}}(\mathbb{N})$ *be an unknown matrix-valued function and* $\mathbf{U} = (U_{ij}(k); i, j \in \mathrm{E}, k \in \mathbb{N}) \in \mathcal{M}_{\mathrm{E}}(\mathbb{N})$ *be a known one. The equation*

$$\mathbf{L}(k) = \mathbf{U}(k) + \mathbf{Q} * \mathbf{L}(k), \quad k \in \mathbb{N},$$

is called a discrete-time Markov renewal equation (DTMRE).

The following result consists in a recursive formula for computing the transition function $\mathbf{P}$ of the semi-Markov chain Z, which is a first example of a Markov renewal equation.

Proposition 2.5. ***(Markov renewal equation of the semi-Markov)****[5]*
For all $i, j \in \mathrm{E}$ *and* $k \in \mathbb{N}$*, we have:*

$$P_{ij}(k) = \delta_{ij}[1 - H_i(k)] + \sum_{r \in \mathrm{E}} \sum_{l=0}^{k} q_{ir}(l) P_{rj}(k - l), \tag{2.11}$$

For all $k \in \mathbb{N}$*, let us define* $\mathbf{H}(k) := diag(H_i(k); i \in \mathrm{E})$*,* $\mathbf{H} := (\mathbf{H}(k); k \in \mathbb{N})$*. In matrix-valued function notation, equation (2.11) becomes*

$$\mathbf{P}(k) = (\mathbf{I} - \mathbf{H})(k) + \mathbf{q} * \mathbf{P}(k), \quad k \in \mathbb{N}.$$

Proof. For all $i, j \in \mathrm{E}$ and $k \in \mathbb{N}$, we have

$$\begin{aligned}
P_{ij}(k) &= \mathbb{P}(Z_k = j | Z_0 = i) \\
&= \mathbb{P}(Z_k = j, S_1 \leq k | Z_0 = i) + \mathbb{P}(Z_k = j, S_1 > k | Z_0 = i) \\
&= \sum_{r \in \mathrm{E}} \sum_{l=0}^{k} \mathbb{P}(Z_k = j, Z_{S_1} = r, S_1 = l | Z_0 = i) + \delta_{ij}(1 - H_i(k)) \\
&= \sum_{r \in \mathrm{E}} \sum_{l=0}^{k} \mathbb{P}(Z_k = j | Z_{S_1} = r, S_1 = l, Z_0 = i) \mathbb{P}(J_1 = r, S_1 = l | J_0 = i) + \delta_{ij}(1 - H_i(k)) \\
&= \sum_{r \in \mathrm{E}} \sum_{l=0}^{k} \mathbb{P}(Z_{k-l} = j | Z_0 = r) \mathbb{P}(J_1 = r, X_1 = l | J_0 = i) + \delta_{ij}(1 - H_i(k)) \\
&= \delta_{ij}(1 - H_i(k)) + \sum_{r \in \mathrm{E}} \sum_{l=0}^{k} P_{rj}(k - l) q_{ir}(l),
\end{aligned}$$

and we obtain the desired result.□

Solving the Markov renewal equation for the semi-Markov transition function $\mathbf{P}$ [3] we obtain that the unique solution is

$$\mathbf{P}(k) = (\delta\mathbf{I} - \mathbf{q})^{(-1)} * (\mathbf{I} - \mathbf{H})(k) = (\psi * (\mathbf{I} - diag(\mathbf{Q} \cdot \mathbf{1})))(k),$$

where $(\delta\mathbf{I} - \mathbf{q})^{(-1)}$ denotes the left convolution inverse of the matrix function $(\delta\mathbf{I} - \mathbf{q})$, $\mathbf{H}(k) := diag(\mathbf{H}_i(k))_{i\in \mathrm{E}}$ and $\mathbf{1}$ denotes the s-column vector whose all elements equal 1.

We propose the following estimator for $\mathbf{P}(\cdot)$:

$$\begin{aligned}\widehat{\mathbf{P}}(k, M) &:= [(\delta\mathbf{I} - \widehat{\mathbf{q}}(\cdot, M))^{(-1)} * (\mathbf{I} - diag(\widehat{\mathbf{Q}}(\cdot, M) \cdot \mathbf{1}))](k) \\ &= [\widehat{\psi}(\cdot, M) * (\mathbf{I} - diag(\widehat{\mathbf{Q}}(\cdot, M) \cdot \mathbf{1}))](k).\end{aligned}$$

Theorem 2.4.1. *[4] The estimator of the semi-Markov transition matrix is strongly consistent and, for any fixed $k \in \mathbb{N}$, $k \leq M$, and $i, j \in \mathrm{E}$, we have*

$$\sqrt{M}(\widehat{P}_{ij}(k, M) - P_{ij}(k)) \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}(0, \sigma_{ij}^2(k)),$$

where

$$\begin{aligned}\sigma_{ij}^2(k) = & \sum_{m=1}^{s} \mu_{mm} \left\{ \sum_{r=1}^{s} [\delta_{mj}\Psi_{ij} - (1 - \mathbf{H}_j) * \psi_{im}\psi_{rj}]^2 * q_{mr}(k) \right\} \\ & - \left[\delta_{mj}\psi_{ij} * \mathbf{H}_m(k) - \sum_{r=1}^{s} (1 - \mathbf{H}_j) * \psi_{im}\psi_{rj} * q_{mr} \right]^2 (k),\end{aligned}$$

and $\Psi = (\Psi(k), k \in \mathbb{N} \in \mathcal{M}_{\mathrm{E}}(\mathbb{N})$ is the matrix renewal function of the DTMRP given by

$$\Psi_{ij}(k) := \mathbb{E}_i[N_j(k)] = \sum_{n=0}^{k} Q_{ij}^{(n)}(k) = \sum_{l=0}^{k} \psi_{ij}(l), \;\; i, j \in \mathrm{E} \;\; and \;\; k \in \mathbb{N}.$$

Chapter 3

Continuous-time semi-Markov process

This chapter provides the definitions and basic properties related to Continuous-time semi-Markov process (CTSMP). The semi Markov process (SMP) is constructed by the so-called Markov renewal process (MRP) that is a special case of the two-dimensional Markov sequence. The MRP is defined by the transition probabilities matrix, called the renewal kernel and an initial distribution, or by other characteristics that are equivalent to the renewal kernel. The counting process corresponding to the SMP allows us to determine the concept of process regularity. The process is said to be regular if the corresponding counting process has a finite number of jumps in a finite period.

3.1 Definitions and properties

Definition 3.1.1. ***(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* $X_{n+1} = S_{n+1} - S_n$ *to be the sojourn time in the state. The process has to satisfy the following equality*

$$\mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_0, J_1, \ldots, J_n, S_0, S_1, \ldots, S_n) = \mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n), \tag{3.1}$$

for all $j \in \mathrm{E}$*, all* $t \in \mathbb{R}_+$ *and all* $n \in \mathbb{N}$*.*

Definition 3.1.2. ***(Renewal matrix, renewal kernel)*** *Let* E *be the state space and consider the Markov renewal process* (J_n, S_n)*, we define* $X_{n+1} = S_{n+1} - S_n$ *to be the sojourn time in the state. The matrix defined as*

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

$$Q_{ij}(t) := \mathbb{P}(J_{n+1} = j, X_{n+1} \leq t | J_n = i),$$

is called a renewal matrix. We identify the renewal matrix $\mathbf{Q}$ *as the renewal kernel.*

Proposition 3.1. *[12] The Markov renewal matrix* $\mathbf{Q}$ *satisfies the following conditions:*

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

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

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

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

Definition 3.1.3. *The probabilities*

$$\begin{aligned} p_{ij} &= \lim_{t\to\infty} Q_{ij}(t) = Q_{ij}(\infty) \\ &= \mathbb{P}(J_{n+1} = j | J_n = i), \end{aligned}$$

are the transition probabilities from state i *to state* j *of the embedded Markov chain* $\{J_n; n \in \mathbb{N}\}$*.*

We assume that the transition probabilities do not depend on the time n.

Proposition 3.2. *[12] For a Markov renewal process with a renewal kernel* $\mathbf{Q}(t), t \geq 0$ *a following equality is satisfied*

$$\mathbb{P}(J_0 = i_0, J_1 = i_1, X_1 \leq t_1, \ldots, J_n = i_n, X_n \leq t_n) = \alpha_{i_0} Q_{i_0 i_1}(t_1) Q_{i_1 i_2}(t_2) \ldots Q_{i_{n-1} i_n}(t_n),$$

where $\alpha_{i_0} := \mathbb{P}(J_0 = i_0)$ *is the initial distribution of the Markov renewal process.*

For $t_1 \to \infty, \ldots, t_n \to \infty$*, we obtain*

$$\mathbb{P}(J_0 = i_0, J_1 = i_1, \ldots, J_n = i_n) = \alpha_{i_0} p_{i_0 i_1} p_{i_1 i_2} \ldots p_{i_{n-1} i_n}.$$

Definition 3.1.4. *(**Continuous-time semi-Markov process**) Consider a Markov-renewal process* $\{(J_n, S_n) : n \in \mathbb{N}\}$ *defined on a complete probability space and with state space* E*. The stochastic process* $\{Z_t; t \in \mathbb{R}_+\}$ *defined by*

$$Z_t = J_{N(t)},$$

is called a Semi-Markov Process (SMP) where $N(t) = \max\{n \in \mathbb{N} : S_n \leq t\}$ *is the counting process of the semi-Markov process up to time t. we can also define the semi-Markov Process by*

$$Z_t = J_n \text{ For } t \in [S_n, S_{n+1}),\ n \in \mathbb{N}.$$

Definition 3.1.5. *we define the transition matrix of the process* $\{Z_t; t > 0\}$ *as*

$$\begin{aligned} \mathbf{P}(t) &= \{P_{ij}(t) : i, j \in \mathrm{E}\}, \\ P_{ij}(t) &= \mathbb{P}(Z_t = j | Z_0 = i), \\ &= \mathbb{P}(J_{N(t)} = j | J_0 = i). \end{aligned}$$

For all $i, j \in \mathrm{E}$*.*

Then the unconditional semi-Markov state probability is equal to

$$\begin{aligned} P_j(t) &= \mathbb{P}(Z_t = j) = \mathbb{P}(J_{N(t)} = j) \\ &= \sum_{i=1}^{s} \mathbb{P}(J_{N(t)} = j | J_0 = i)\mathbb{P}(J_0 = i) \\ &= \sum_{i=1}^{s} \alpha_i P_{ij}(t). \end{aligned}$$

Where $\alpha_i = \mathbb{P}(J_0 = i)$ is the initial distribution of the Markov renewal process.

Definition 3.1.6. ***(Regularity of SMP)*** *A SMP* $\{Z_t; t \in \mathbb{R}_+\}$ *is said to be regular if the corresponding counting process* $\{N(t); t > 0\}$ *has a finite number of jumps in a finite period with probability* 1*:*

$$\forall t \in \mathbb{R}_+, \qquad \mathbb{P}(N(t) < \infty) = 1. \tag{3.2}$$

The equality (3.2) *is equivalent to a relation*

$$\forall t \in \mathbb{R}_+, \qquad \mathbb{P}(N(t) = \infty) = 0.$$

Definition 3.1.7. ***(Distribution functions of sojourn time)*** *for all* $i, j \in \mathrm{E}$, $\forall t \in \mathbb{R}_+$.

1. $F_{ij}(.)$*, the distribution function associated with the sojourn time in state* i*, before going to state* j*:*

$$F_{ij}(t) := \mathbb{P}(X_{n+1} \leq t | J_n = i, J_{n+1} = j).$$

2. $H_i(.)$*, the distribution function of the sojourn time, also called the waiting time, in state* i*:*

$$H_i(t) := \mathbb{P}(X_{n+1} \leq t | J_n = i) = \sum_{j \in \mathrm{E}} Q_{ij}(t).$$

From the definition before we can derive the following result.

Proposition 3.3. *[12] It holds true that*

$$F_{ij}(t) = \tfrac{Q_{ij}(t)}{p_{ij}}.$$

For all $t \geq 0$ *and* $i, j \in \mathrm{E}$

Proof. From the definition of conditional probabilities, it follows that

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

3.2 Elements of statistical estimation

Estimators for semi Markov kernel $Q_{ij}(t)$ are defined on sample functions of the MRP over $[0, M]$. These sample functions of the MRP are equivalent to the sample functions $(J_0, J_1, \dots, J_{N(M)}, X_0, X_1, \dots, X_{N(M)})$.

3.2.1 Empirical estimators

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

$$\widehat{Q}_{ij}(t, M) = \tfrac{1}{N_i(M)} \sum_{n=1}^{N(t)} \mathbf{1}_{\{J_{n-1}=i, J_n=j, X_n \le t\}},$$

where

$$N_i(M) := \sum_{n=1}^{N(t)} \mathbf{1}_{\{J_n=i\}} = \sum_{n=1}^{\infty} \mathbf{1}_{\{J_n=i, S_n \le M\}}.$$

We define the empirical estimator of the transition matrix of the embedded Markov chain p_{ij} by

$$\widehat{p}_{ij} = \frac{N_{ij}(M)}{N_i(M)},$$

where

$$N_{ij}(M) := \sum_{n=1}^{N(t)} \mathbf{1}_{\{J_{n-1}=i, J_n=j\}} = \sum_{n=1}^{\infty} \mathbf{1}_{\{J_{n-1}=i, J_n=i, S_n \le M\}}.$$

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

$$\widehat{F}_{ij}(t, M) = \tfrac{1}{N_{ij}(M)} \sum_{n=1}^{N(t)} \mathbf{1}_{\{J_{n-1}=i, J_n=j, X_n \le t\}}.$$

The quantities $\widehat{F}_{ij}(t, M)$ and $\widehat{p}_{ij}$ 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.

3.2.2 Asymptotic properties of the estimators

Strong consistency

From corollary (2.3.1) the empirical estimator $\widehat{p}_{ij}(M)$ of p_{ij} for all $i, j \in \mathrm{E}$ is strongly consistent, i.e.

$$\widehat{p}_{ij}(M) \xrightarrow[M\to\infty]{a.s} p_{ij}.$$

Theorem 3.2.1. *[10] The empirical estimator $\widehat{Q}_{ij}(t, M)$ of $Q_{ij}(t)$ for all $i, j \in \mathrm{E}$ is strongly consistent, i.e.*

$$\max_{i,j\in\mathrm{E}} \sup_{t\in[0,M]} |\widehat{Q}_{ij}(t, M) \, - \, Q_{ij}(t)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0$$

Proof. It holds true that $Q_{ij}(t) = F_{ij}(t)p_{ij}$ and therefore $\widehat{Q}_{ij}(t, M) = \widehat{F}_{ij}(t, M)\widehat{p}_{ij}(M)$ as well. Then it follows that

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

The first term converges to 0 (a.s.). The second converges to 0 (a.s.) as well by theorem (Glivenko-Cantelli theorem 1.1.2). □

Asymptotic normality

It is assumed throughout that the MRP is irreducible, recurrent, and that $F_{ij} = H_i$ for $1 \leq j \leq s$. This last assumption incurs no loss of generality as is pointed out in [27]. Consider the estimator defined by:

$$\widehat{Q}_{ij}(t, M) = \widehat{H}_i(t, M)\widehat{p}_{ij}(M) \tag{3.3}$$

$$\widehat{H}_i(t, M) = N_i(M)^{-1} \sum_{k=1}^{N_i(M)} \epsilon(t - X_{ik}) \tag{3.4}$$

and where $\epsilon(u)$ equals one if $u \geq 0$ and zero otherwise. $\widehat{H}_i(t, M)$ is therefore the ordinary empirical distribution function determined from the sample, of random size $N_i(M)$, of the holding times in state i.

The limiting distributions of the quantities in (3.3) may be obtained as consequences of the central limit theorem for MRP (3.2.2).

For a real measurable function f, defined on $\mathrm{E} \times \mathrm{E} \times \mathbb{R}$. Define, for each $M > 0$, the functional $W_f(M)$ as

$$W_f(M) := \sum_i \sum_{n=1}^{N_i(M)} f(i, X_{in}), \tag{3.5}$$

where X_{in} is the n^{th} sojourn time of the chain in state i ie $X_{in} = S^i_{n+1} - S^i_n$. The functional $W_f(M)$ can be defined only if the series in (3.5) converges.

Set

$$A_{ij} := \int_0^\infty f(i, j, x) dQ_{ij}(x), A_i := \sum_{j=1}^{s} A_{ij},$$

$$B_{ij} := \int_0^\infty (f(i, j, x))^2 dQ_{ij}(x), B_i := \sum_{j=1}^{s} B_{ij},$$

Let μ_{ij} and μ^*_{ij} denote the mean first passage times from state i to j in the MRP and in the corresponding Markov Chain,$\{J_n; n \geq 0\}$, respectively.

Write

$$
\begin{aligned}
r_i &:= \sum_{j=1}^{s} A_j \frac{\mu^*_{ii}}{\mu^*_{jj}} \\
\sigma_i^2 &:= -r_i^2 + \sum_{j=1}^{s} B_j \frac{\mu^*_{ii}}{\mu^*_{jj}} + 2\sum_{r=1}^{s}\sum_{l\neq i}\sum_{k\neq i} A_{rl} A_k \mu^*_{ii} \frac{\mu^*_{li}+\mu^*_{ik}-\mu^*_{ii}}{\mu^*_{lk}\mu^*_{kk}}
\end{aligned}
$$

Finally, put

$$
\begin{aligned}
m_f &:= \frac{r_i}{\mu_{ii}} \\
B_f &:= \frac{\sigma_i^2}{\mu_{ii}}
\end{aligned}
$$

Theorem 3.2.2. ***(Central Limit Theorem)****[21]*

For an irreducible recurrent MRP and if the above moments are finite, we have,

$$
M^{-1/2}\left[W_f(M) - M\cdot m_f\right] \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}(0, B_f).
$$

To apply this result in the proofs of the theorems of this section it will only be necessary to produce the appropriate function f and to compute the corresponding moments.

Theorem 3.2.3. *[21] For fixed i, j, t,* $(M^{1/2}[\widehat{p}_{ij}(M) - p_{ij}], M^{1/2}[\widehat{H}_i(t,M) - H_i(t)])$ *converges in law as* $M\to\infty$ *to a bivariate normal r.v. with means zero and covariance matrix* (σ_{ij}) *given by*

$$
\sigma_{11} = \mu_{ii}p_{ij}(1-p_{ij}), \quad \sigma_{22} = \mu_{ii}H_i(t)(1-H_i(t)), \quad \sigma_{12} = \sigma_{21} = 0.
$$

Proof. Let ω_1 and ω_2 be arbitrary constants. To prove the asymptotic joint normality it suffices to show that

$$
\omega_1 M^{1/2}[\widehat{p}_{ij}(M) - p_{ij}] + \omega_2 M^{1/2}[\widehat{H}_i(t,M) - H_i(t)], \tag{3.6}
$$

converges in law to a normal r.v. for all ω_1 and ω_2. We rewrite (3.6) as the product of $[M/N_i(M)]M^{-1/2}$ and a sum of the form (3.5) by using the function f defined by

$$
f(r,s,y) = \{\omega_1[\delta_{sj} - p_{ij}] + \omega_2[\epsilon(t-y) - H_i(t)]\}\delta_{ri}. \tag{3.7}
$$

For this function

$$A_r = \omega_1 \delta_{ri}[p_{rj} - p_{ij}] + \omega_2 \delta_{ri}[H_r(t) - H_i(t)] = 0,$$

and

$$B_r = \{\omega_1^2[p_{rj} + p_{ij}^2 - 2p_{rj}p_{ij}] + \omega_2^2[H_r(t) + H_i^2(t) - 2H_r(t)H_i(t)]\}\delta_{ri},$$

for $l \leq r \leq s$; hence r_i and the third sum in (3.6) is zero. Then

$$\sigma_i^2 = \sum_{r=1}^{s} B_r \frac{\mu_{ii}^*}{\mu_{rr}^*} = \omega_1^2 p_{ij}[1 - p_{ij}] + \omega_2 H_i(t)[1 - H_i(t)].$$

The variance σ_i^2 is finite, so from Lemma 7.1 of [27] the limiting distribution of $M^{-1/2}W_f(M)$ for the f given in (3.7) is normal with zero mean and variance σ_i^2/μ_{ii}. But $M/N_i(M) \to \mu_{ii}$ (a.s.) so the limiting distribution of (3.6) is normal with zero mean and variance $\sigma_i^2\mu_{ii}$ as required. □

The zero correlation between $\widehat{p}_{ij}$ and $\widehat{H}_i(t, M)$ yields the following results.

Corollary 3.2.1. *[21] For fixed i, j, t, $\widehat{p}_{ij}$ and $\widehat{H}_i(t, M)$ are asymptotically independent.*

Theorem 3.2.4. *[21] The empirical estimator $\widehat{Q}_{ij}(t, M)$ is asymptotically normal, i.e. for fixed $t > 0$*

$$\sqrt{M}[\widehat{Q}_{ij}(t, M) \ - \ Q_{ij}(t)] \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}(0, \sigma^2).$$

With

$$\sigma^2 = \mu_{ii} H_i(t) p_{ij}[H_i(t) - 2H_i(t)p_{ij} + p_{ij}].$$

Proof. We rewrite $\sqrt{M}[\widehat{Q}_{ij}(t, M) \ - \ Q_{ij}(t)]$ as

$$M^{1/2}\widehat{H}_i(t, M)[\widehat{p}_{ij}(M) - p_{ij}] + M^{1/2}p_{ij}[\widehat{H}_i(t, M) - H_i(t)]). \qquad (3.8)$$

By a well-known convergence theorem [10] the limiting distribution of (3.8) is the same as that of

$$M^{1/2}H_i(t)[\widehat{p}_{ij}(M) - p_{ij}] + M^{1/2}p_{ij}[\widehat{H}_i(t, M) - H_i(t)]) \qquad (3.9)$$

With the particular choice $\omega_1 = H_i(t)$ and $\omega_2 = p_{ij}$, (3.9) is just the same as (3.6) and the proof is complete.□

3.3 Markov renewal matrix

Definition 3.3.1. ***(Stieltjes convolution)*** *Let $\phi(i,t)$ for $t \geq 0$ and $i \in \mathrm{E}$ be a real valued measurable function and Q be a semi-Markov kernel. Then the Stieltjes convolution of ϕ by Q is defined as*

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

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

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

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

$$Q_{ij}^{(n)}(t) = \mathbb{P}(J_n = j, S_n \leq t | J_0 = i),$$

and therefore an MRP is regular if and only if $\sum_{j\in \mathrm{E}} Q_{ij}^{(n)}(t) \to 0$, as $n \to \infty$, for all i.

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

$$\begin{aligned} \psi_{ij}(t) &:= \mathbb{E}_i[N_j(t)] \\ &= \mathbb{E}(N_j(t)|J_0 = i) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(J_n = j, S_n \leq t|J_0 = i) = \sum_{n=0}^{\infty} Q_{ij}^{(n)}(t), \end{aligned}$$

for $t \geq 0$, $i, j \in \mathrm{E}$. 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) element of the matrix $\psi(t)$, we use the empirical estimator

$$\widehat{\psi}_{ij}(t,M) = \sum_{n=0}^{\infty} \widehat{Q}_{ij}^{(n)}(t,M),$$

where $\widehat{Q}_{ij}^{(n)}(t,M)$ is the n-fold convolution of $\widehat{Q}_{ij}(t,M)$. For the empirical estimator $\widehat{Q}_{ij}^{(n)}(t,M)$ of the n-fold convolution of the semi-Markov kernel, the following theorem holds true.

Theorem 3.3.1. *[22] The empirical estimator $\widehat{Q}_{ij}^{(n)}(t,M)$ of $Q_{ij}^{(n)}(t)$ for all $i,j \in \mathrm{E}$ is strongly consistent, i.e. for any fixed $n \in \mathbb{N}$*

$$\max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} |\widehat{Q}_{ij}^{(n)}(t,M) - Q_{ij}^{(n)}(t)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0.$$

Proof. By induction. For the case $n=1$, the result follows from theorem 3.2.1. Assume that it holds true for n = m. So

$$\max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} |\widehat{Q}_{ij}^{(m)}(t,M) - Q_{ij}^{(m)}(t)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0.$$

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

$$\begin{aligned}
\max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} |\widehat{Q}_{ij}^{(m+1)}(t,M) - Q_{ij}^{(m+1)}(t)| &= \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} |\sum_{k\in \mathrm{E}} \widehat{Q}_{ik}(t,M) \star \widehat{Q}_{ij}^{(m)}(t,M) - \sum_{k\in \mathrm{E}} Q_{ik}(t) \star Q_{ik}^{(m)}(t)| \\
&= \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} |\sum_{k\in \mathrm{E}} (\widehat{Q}_{ik}(t,M) \star \widehat{Q}_{ij}^{(m)}(t,M) - Q_{ik}(t) \star Q_{ik}^{(m)}(t))| \\
&= \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} |\widehat{Q}_{ik}(t,M) \star \widehat{Q}_{ij}^{(m)}(t,M) - Q_{ik}(t) \star Q_{ik}^{(m)}(t)| \\
&= \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} |\widehat{Q}_{ik}(t,M) \star \widehat{Q}_{ij}^{(m)}(t,M) - Q_{ik}(t) \star Q_{ik}^{(m)}(t,M) \\
&\quad + Q_{ik}(t) \star Q_{ik}^{(m)}(t,M) - Q_{ik}(t) \star Q_{ik}^{(m)}(t)| \\
&\leq \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} |\widehat{Q}_{ik}(t,M) \star \widehat{Q}_{ij}^{(m)}(t,M) - Q_{ik}(t) \star Q_{ik}^{(m)}(t,M)| \\
&\quad + \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} |Q_{ik}(t) \star Q_{ik}^{(m)}(t,M) - Q_{ik}(t) \star Q_{ik}^{(m)}(t)| \\
&= \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} |[\widehat{Q}_{ik}(t,M) - Q_{ik}(t)] \star \widehat{Q}_{ij}^{(m)}(t,M)| \\
&\quad + \max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} |Q_{ik}(t) \star [Q_{ik}^{(m)}(t,M) - Q_{ik}^{(m)}(t)]| \\
&\leq \max_{i,k\in \mathrm{E}} \sup_{t\in[0,M]} |\widehat{Q}_{ik}(t,M) - Q_{ik}(t)| \max_{k,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} \widehat{Q}_{ij}^{(m)}(t,M) \\
&\quad + \max_{i,k\in \mathrm{E}} \sup_{t\in[0,M]} |Q_{ik}^{(m)}(t,M) - Q_{ik}^{(m)}(t)| \max_{k,j\in \mathrm{E}} \sup_{t\in[0,M]} \sum_{k\in \mathrm{E}} Q_{ik}(t) \\
&\leq s \max_{i,k\in \mathrm{E}} \sup_{t\in[0,M]} |\widehat{Q}_{ik}(t,M) - Q_{ik}(t)| \\
&\quad + \max_{i,k\in \mathrm{E}} \sup_{t\in[0,M]} |Q_{ik}^{(m)}(t,M) - Q_{ik}^{(m)}(t)|
\end{aligned}$$

The last step holds true, because $\mathrm{E} = \{1, 2, \ldots, s\}$. By theorem 3.2.1 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.□

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

Theorem 3.3.2. *[19] The empirical estimator $\widehat{\psi}_{ij}(t,M)$ of the Markov renewal function $\psi_{ij}(t)$ is uniformly strongly consistent, for all $i,j \in \mathrm{E}$ is* *strongly consistent, i.e.*

$$\max_{i,j\in \mathrm{E}} \sup_{t\in[0,M]} |\widehat{\psi}_{ij}(t,M) - \psi_{ij}(t)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0.$$

Theorem 3.3.3. *[19] The empirical estimator $\widehat{\psi}_{ij}(t,M)$ of the Markov renewal function $\psi_{ij}(t)$ converges in distribution, for any fixed $t > 0$, as $M \to \infty$, to a normal random variable, i. e.,*

$$\sqrt{M}|\widehat{\psi}_{ij}(t,M) - \psi_{ij}(t)| \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2_{ij}(t)).$$

It holds true that

$$\sigma^2_{ij}(t) = \sum_{k=1}^{s}\sum_{l=1}^{s} \mu_{kk}[(\psi_{ik} \star \psi_{lj})^2 \star Q_{kl} - (\psi_{ik} \star \psi_{lj} \star Q_{kl})^2](t).$$

Proof. See for example (Ouhbi and Limnios, [22]), theorem 3.

3.4 Markov renewal equation

Definition 3.4.1. ***(Continuous-time Markov renewal equation)*** *Let $\mathbf{L} = (L_{ij}(t); i,j \in \mathrm{E}, t \in \mathbb{R}_+) \in \mathcal{M}_{\mathrm{E}}(\mathbb{R}_+)$ be an unknown matrix-valued function and $\mathbf{U} = (U_{ij}(t); i,j \in \mathrm{E}, t \in \mathbb{R}_+) \in \mathcal{M}_{\mathrm{E}}(\mathbb{R}_+)$ be a known one. The equation*

$$\mathbf{L}(t) = \mathbf{U}(t) + \mathbf{Q} \star \mathbf{L}(t), \tag{3.10}$$

is called a Markov renewal equation (MRE).

Note that Equation (3.10) is equivalent to equation

$$(\mathbf{I} - \mathbf{Q}) \star \mathbf{L}(t) = \mathbf{U}(t).$$

And has a unique solution [29] $\mathbf{L}(t) = \psi \star \mathbf{U}(t)$.

Proposition 3.4. *[19] For all $i,j \in \mathrm{E}$ and $t \in \mathbb{R}_+$, we have:*

$$P_{ij}(t) = \delta_{ij}[1 - H_i(t)] + \sum_{k\in\mathrm{E}} \int_0^t Q_{ik}(ds)P_{kj}(t-s). \tag{3.11}$$

For all $t \in \mathbb{R}_+$, let us define $\mathbf{H}(t) := diag(H_i(t); i \in \mathrm{E})$, $\mathbf{H} := (\mathbf{H}(t); t \in \mathbb{R}_+)$. In matrix-valued function notation, equation (3.11) becomes

$$\mathbf{P}(t) = (\mathbf{I} - \mathbf{H})(t) + \mathbf{Q} \star \mathbf{P}(t), \quad t \in \mathbb{R}_+.$$

Whose unique solution is

$$\mathbf{P}(t) = \psi \star (\mathbf{I} - \mathbf{H})(t).$$

Let us define now the empirical estimator of the transition function of the semi-Markov process $P_{ij}(t)$, $i, j \in \mathrm{E}$ and $t \in \mathbb{R}_+$. In a matrix form, we have

$$\widehat{\mathbf{P}}(t, M) = \widehat{\psi} \star (\mathbf{I} - diag(\widehat{Q}(t, M)).$$

Then, the following results hold:

Theorem 3.4.1. *[19] For any fixed $C > 0$ and $i, j \in \mathrm{E}$, we have*

$$\lim_{M \to \infty} \max_{i,j} \sup_{t \in [0,C]} |\widehat{P}_{ij}(t, M) - P_{ij}(t)| = 0. \quad (a.s)$$

Theorem 3.4.2. *[19] For $i, j \in \mathrm{E}$, we have*

$$M^{1/2}(\widehat{P}_{ij}(t, M) - P_{ij}(t)) \xrightarrow[T \to \infty]{\mathcal{D}} \mathcal{N}(0, \sigma_{ij}^2(t)),$$

where

$$\sigma_{ij}^2(t) = \sum_{r \in \mathrm{E}} \sum_{k \in \mathrm{E}} \mu_{rr} \left[(1 - H_i) \star B_{irkj} - \psi_{ij} \mathbf{1}_{\{r=j\}}\right]^2 \star Q_{rk}(t) - \{[(1-G_i) \star B_{irkj} - \psi_{ij} \mathbf{1}_{\{r=j\}}] \star Q_{rk}(t)\}^2,$$

and

$$B_{irkj}(t) = \sum_{n=1}^{\infty} \sum_{l=1}^{n} Q_{ir}^{(l-1)} \star Q_{kj}^{(n-l)}(t).$$

3.5 Hazard rate function

We define the instantaneous transition rate function, $\lambda_{ij}(t)$ for $t \geq 0$, $i, j \in \mathrm{E}$ of a semi-Markov kernel by

$$\begin{aligned} \lambda_{ij}(t) &= \lim_{\Delta t \downarrow 0} \frac{1}{\Delta t} \mathbb{P}\left[J_{n+1} = j, t < X_{n+1} \leq t + \Delta t | J_n = i, X_{n+1} > t\right] \\ &= \begin{cases} \frac{q_{ij}(t)}{1 - H_i(t)}, & \text{if } p_{ij} > 0 \text{ and } H_i(t) < 1, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

The quantity $\lambda_{ij}(t)\Delta t + o(\Delta t)$, $i \neq j$, is the probability that the process has spent t units of time in state i and will transit to state j in $(t, t+\Delta t]$.

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

$$\widehat{\lambda}_{ij}(t, M) := \begin{cases} \frac{\widehat{q}_{ij}(t,M)}{1-\widehat{H}_i(t,M)}, & \text{if } \widehat{p}_{ij}(M) > 0 \text{ and } \widehat{H}_i(t, M) < 1, \\ 0, & \text{otherwise} \end{cases} \tag{3.12}$$

Where the empirical estimator of the derivative function $\widehat{q}_{ij}(t, M)$ of the semi-Markov kernel is

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

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

Theorem 3.5.1. *[23] For $0 < \alpha < 1/2$, The empirical estimator $\widehat{\lambda}_{ij}(t, M)$ of $\lambda_{ij}(t)$ is uniformly strongly consistent,in all compacts $[0, C]$, $C \in \mathbb{R}_+$, in the sense that, ie*

$$\max_{i,j\in \mathrm{E}} \sup_{t\in[0,C]} |\widehat{\lambda}_{ij}(t, M) - \lambda_{ij}(t)| \xrightarrow[M\longrightarrow\infty]{a.s.} 0.$$

Definition 3.5.1. *We define the hazard rate function of the waiting time $\alpha_{ij}(t)$ by:*

$$\alpha_{ij}(t) = \lim_{\Delta t\downarrow 0} \frac{1}{\Delta t}\mathbb{P}[X_{n+1} \in (t, t + \Delta t)|J_{n+1} = j, J_n = i, X_{n+1} > t].$$

We know that the survival function of the sojourn time $S_{ij}(t)$ is

$$S_{ij}(t) = \mathbb{P}[X_{n+1} > t|J_{n+1} = j, J_n = i] = 1 - F_{ij}(t)$$

(Note that $S_{ij}(t)$ is a decreasing function, that is $S_{ij}(0) = 1$ and $\lim_{t\to\infty} S_{ij}(t) = 0$).

Now, by the definition of conditional probability we have

$$\alpha_{ij}(t) = \frac{f_{ij}(t)}{S_{ij}(t)}.$$

The relation between the hazard rate function of the semi-Markov and hazard rate function of the waiting time is

$$\lambda_{ij}(t) = \frac{p_{ij} S_{ij}(t)}{S_j(t)} \alpha_{ij}(t).$$

Where $S_j(t) = \mathbb{P}[X_{n+1} > t | J_{n+1} = j]$. *We call this the survival function of the waiting time in state* j.

Remark 3.5.1. *The hazard rate of waiting time at time t represents the conditional probability that a transition from state* i *to state* j *is observed given that no event occurs until time* t *and The hazard rate of the semi-Markov process at time* t *represents the conditional probability that a transition into state* j *is observed given that the subject is in state* i *and that no event occurs until time* t. *The hazard rate of the semi-Markov process can be interpreted as the subject's risk of passing from state* i *to state* j.

Chapter 4

Applications

4.1 Application to asthma control data

As an illustrative example, we revisit the analysis of severe asthmatic patients which was conducted in France between 1997 and 2001 by ARIA (Association pour la Recherche en Intelligence Artificielle). Adult asthmatics were prospectively enrolled over a 4-year period by a number of French chest physicians. The data reflects the real follow-up of patients consulting at varied times according to their perceived needs. At each visit, several covariates were recorded and asthma was evaluated

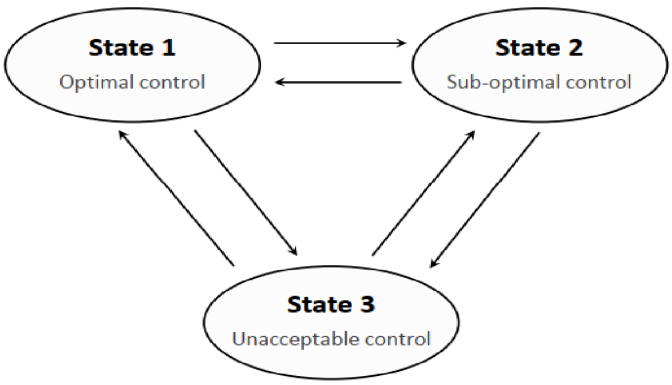

Figure 4.1: The three states model used for asthma control evolution.

The considered model to study the evolution of asthma consists of three transient states Figure 4.1: the optimal control (State 1), the sub-optimal control (State 2), the unacceptable control (State 3) and three covariates are included in the data: Severity (disease severity : coded 1 if severe, 0 if mild-moderate asthma), BMI (Body Mass Index : 1 if BMI ≥ 25, 0 if BMI$<$25) and Sex (1 if men, 0 if women). A random selection of 371 patients with at least two visits (data `asthma`) is included in the package SemiMarkov.

4.1.1 The SemiMarkov R package

Package description

The **SemiMarkov** package was developed to analyze longitudinal data using multi-state semi-Markov models. The main function `semiMarkov` of the package computes the parametric maximum likelihood estimation in multi-state semi-Markov models in continuous-time. The effect of time varying or fixed covariates can be studied using a proportional intensities model for the hazard of the sojourn time.

Format of data.

The data frame to be used in the function `semiMarkov` must be similar to the `asthma` data : a table in long format (one row per transition and possibly several rows by individual) that must contain the following informations.

1. `id:` the individual identification number
2. `state.h:` state left by the process
3. `state.j:` state entered by the process
4. `time:` sojourn time in `state.h`

The data set may also include additional explanatory variables (for instance, some individual's characteristics). The values of these covariates must be given for each individual and for each transition in order to take fixed or time-dependent covariates into account (one value for each row of the data frame data).

Functions description.

Following is a brief description of the package functions.

- `table.state:` Computes a frequency table counting the number of observed transitions in the data set.
- `param.init:` Defines default or specified initial values of the parameters.
- `semiMarkov:` Computes the parametric maximum likelihood estimation of multi-state semi-Markov models.
- `hazard:` For any object of classes `semiMarkov` and `param.init`, the function computes the values of the hazard rate of sojourn times or the values of the hazard rate of the semi-Markov process for a given vector of times.
- `summary.semiMarkov, summary.hazard, print.semiMarkov, print.hazard:` Summary and printing methods for objects of classes `semiMarkov` and `hazard`.
- `plot.hazard:` Plot method for objects of class `hazard`.

Sojourn times distribution.

The parametric estimation in homogeneous semi-Markov models is based on the specification of the sojourn times distribution. The simplest model is obtained using the exponential distribution $\mathcal{E}(\sigma_{ij})$, for which the hazard rate is constant over time (corresponding to the Markov case) and is related to a single positive parameter σ,

$$\alpha_{ij}(t) = \frac{1}{\sigma_{ij}}, \qquad \forall t \geq 0.$$

The Weibull distribution which generalizes the exponential one, is often used in practical applications. Indeed, the Weibull distribution with two parameters $\mathcal{W}(\sigma_{ij}, \nu_{ij})$ is well adapted to deal with various shapes of monotone hazards,

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

where $\sigma_{ij} > 0$ is a scale parameter and ν_{ij} is a shape parameter. The exponentiated Weibull distribution $\mathcal{EW}(\sigma_{ij}, \nu_{ij}, \theta_{ij})$ with an additional shape parameter $\theta_{ij} > 0$ is very useful to fit $\cap$ and $\cup$ shapes of hazard rates

$$\alpha_{ij}(t) = \frac{\theta_{ij}\frac{\nu_{ij}}{\sigma_{ij}}(\frac{t}{\sigma_{ij}})^{\nu_{ij}-1}\exp(-(\frac{t}{\sigma_{ij}})^{\nu_{ij}})[1-\exp(-(\frac{t}{\sigma_{ij}})^{\nu_{ij}})]^{\theta_{ij}}}{1-[1-\exp(-(\frac{t}{\sigma_{ij}})^{\nu_{ij}})]}.$$

These three distributions are available in the package **SemiMarkov** : exponential (`"E"`, `"Exp"` or `"Exponential"`), Weibull (`"W"` or `"Weibull"`) and exponentiated Weibull (`"EW"`, `"EWeibull"` or `"Exponentiated Weibull"`). which allow to fit various shapes of the hazard ratio are nested: a $\mathcal{EW}(\sigma_{ij}, 1, 1)$ is equivalent to $\mathcal{W}(\sigma_{ij}, 1)$ which is equivalent to a $\mathcal{E}(\sigma_{ij})$. The estimations of the distribution parameters are given with standard deviations and p-values of the Wald test [1] ($H_0 : \theta_{ij} = 1$). One can then evaluate, for instance, the relevance of the exponentiated Weibull distribution in comparison to the Weibull or the exponential distribution.

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}

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

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

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

[1] The Wald test is an econometric parametric test whose name comes from the American mathematician of Hungarian origin Abraham Wald (October 31, 1902- December 13, 1950) with a wide variety of uses. Whenever we have a relationship within or between data items that can be expressed as a statistical model with parameters to be estimated, and all from a sample, the Wald test can be used to "test the true value of the parameter" based on the sample estimate.

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. If $p \leq 0.05$, we reject the null-hypothesis H_0. If $p > 0.05$, we fail to reject H_0.

Multi-state model definition.

The multi-state approach requires to define the states of the process and to specify the structure of the model (the number of states and the possible transitions between them). In case of the three-state model described in Figure 4.1 where the sojourn times associated to each transition are Weibull distributed, the matrix `mtrans` of possible transitions will be defined as follows

```
R> mtrans
       [,1]      [,2]      [,3]
[1,]   "FALSE"   "W"       "W"
[2,]   "W"       "FALSE"   "W"
[3,]   "W"       "W"       "FALSE"
```

The argument `states` is a character vector used to define the names of states, possible values are those included in the data's columns `state.h` and `state.j`.

Covariates

The effect of covariates on the process evolution can be investigated considering a Cox proportional hazard model [2] for the hazard rates of waiting times. 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

$$\lambda_{ij}(t|Z) = \lambda_{ij}^{(\theta_{ij})}(t)e^{\beta_{ij}^T Z}.$$

The interpretation of the regression coefficients in terms of relative risks (as in the Cox model) can help to quantify the effect of covariates and to

[2]The Cox proportional-hazards model (Cox, 1972) is essentially a regression model commonly used statistical in medical research for investigating the association between the survival time of patients and one or more predictor variables.

understand the process evolution. For each estimation of regression coefficients, standard deviation and p-value of the Wald test ($H_0 : \beta = 0$) are given.

Initial values.

The optimization procedure used in the maximum likelihood estimation requires definition of initial values of the parameters: the distribution parameters, the transition probabilities and the regression coefficients associated to the covariates.

Parametric maximum likelihood estimation

The semiMarkov function.

In a parametric framework, distributions of sojourn times are supposed to belong to a class of parametric functions. For each transition, the distribution (which depends on a finite number of parameters) can be specified using either the hazard rate λ_{ij}, the density f_{ij} or the cumulative distribution function F_{ij}.

The main function `semiMarkov` estimates the parameters of a multi-state homogeneous semi-Markov model using the parametric maximum likelihood estimation. The following arguments are used in the function `semiMarkov`: arguments related to the data `(data, cov)`, arguments related to the model `(states, mtrans, cov_tra, cens)` and initial values `(dist_init, proba_init, coef_init)`.

This function gives informations on the optimization method and provides the parameters estimation together with their standard deviations. For each regression coefficient β, the p-value of the Wald test when testing the absence of effect ($H_0 : \beta = 0$) is also provided whereas for each distribution parameter σ(or ν or θ) the p-value of the Wald test when testing ($H_0 : \sigma = 1$) is given.

The hazard function.

The hazard rate of sojourn time and the hazard rate of the semi-Markov process can be deduced from the parameters and the distributions of sojourn times. The function `hazard` computes vectors of hazard rates values using either the estimations included in an object of class `semiMarkov` or the specific values defined by an object of class `param.init`. The argument `type` is used to choose the type of hazard rate: `alpha` for the hazard rates of waiting times and `lambda` for the hazard rates of the semi-Markov process. If covariates are used in the model, the hazard rates can be obtained for given values of the covariates using the argument `cov`.

4.1.2 Script and concluding remarques

```
R>library("SemiMarkov")
## Asthma control data
R>data("asthma")
R>head(asthma)
```

	id	state.h	state.j	time	Severity	BMI	Sex
1	2	3	2	0.15331964	1	1	0
2	2	2	2	4.12320329	1	1	0
3	3	3	1	0.09582478	1	1	1
4	3	1	3	0.22997947	1	1	1
5	3	3	1	0.26557153	1	1	1
6	3	1	1	5.40725530	1	1	1

There are no absorbing states in the considered model Figure 4.1. The last sojourn time is then right-censored. Its value is the time between the last visit and the date of the end of the study. A censored observation is identified by a transition into the same state. In such case, the value of `state.h` is equal to the value of `state.j` and the value of `time` is the censored sojourn time.

```
R>table.state(asthma)
$table.state
     1   2   3
 1 152  95  44
 2 112 116  71
 3 115 120 103
$Ncens
[1] 371
```

```
## Definition of the model:  states, names
# possible transitions and waiting times distributions
R> states <- c("1","2","3")
R> mtrans <- matrix(FALSE, nrow=3, ncol=3)
R> mtrans[1,2:3] <- c("W","W")
R> mtrans[2,c(1,3)] <- c("W","W")
R> mtrans[3,c(1,2)] <- c("W","W")
```

	[,1]	[,2]	[,3]
[1,]	"FALSE"	"W"	"W"
[2,]	"W"	"FALSE"	"W"
[3,]	"W"	"W"	"FALSE"

```
## Semi-Markov model without covariates
fit1 <- semiMarkov(data=asthma, states=states, mtrans=mtrans)
## Hazard rates of waiting time
alpha1 <- hazard(fit1)
plot(alpha1)
```

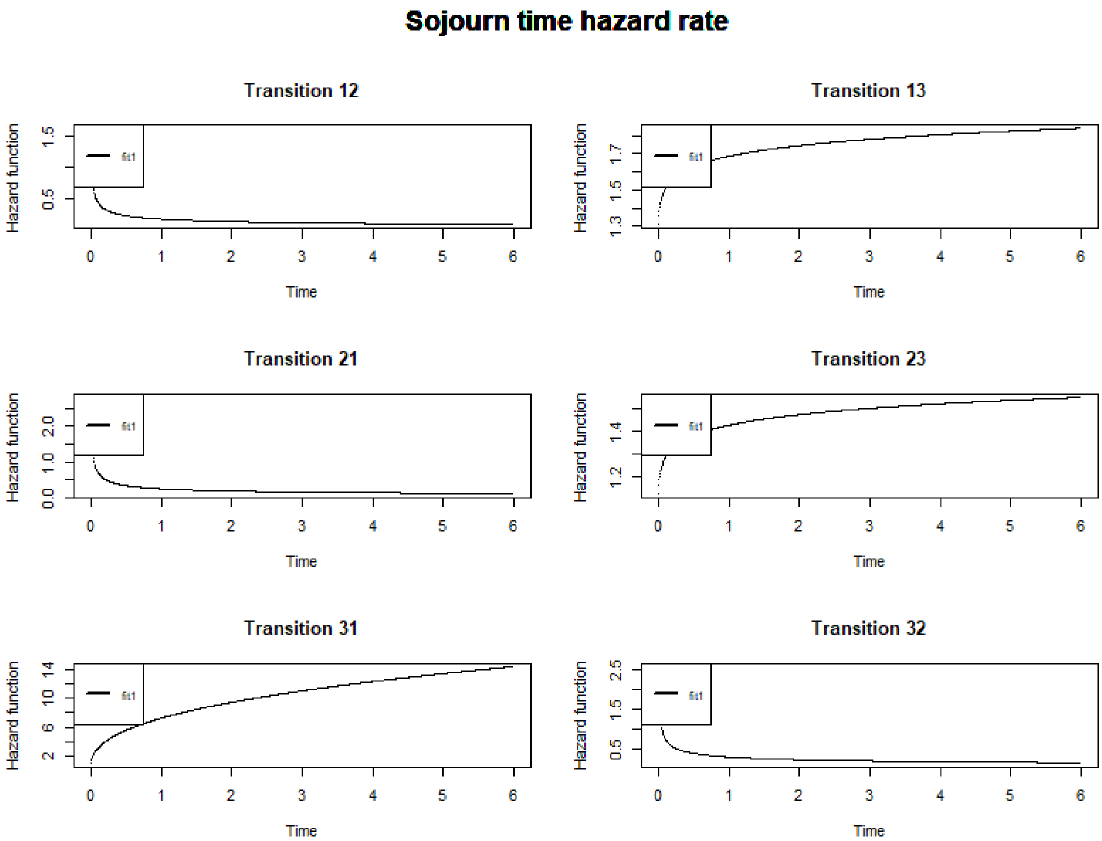

Figure 4.2: The hazard rate of sojourn time.

```
## Hazard rates of the semi-Markov process
lambda1 <- hazard(fit1, type = "lambda")
plot(lambda1)
```

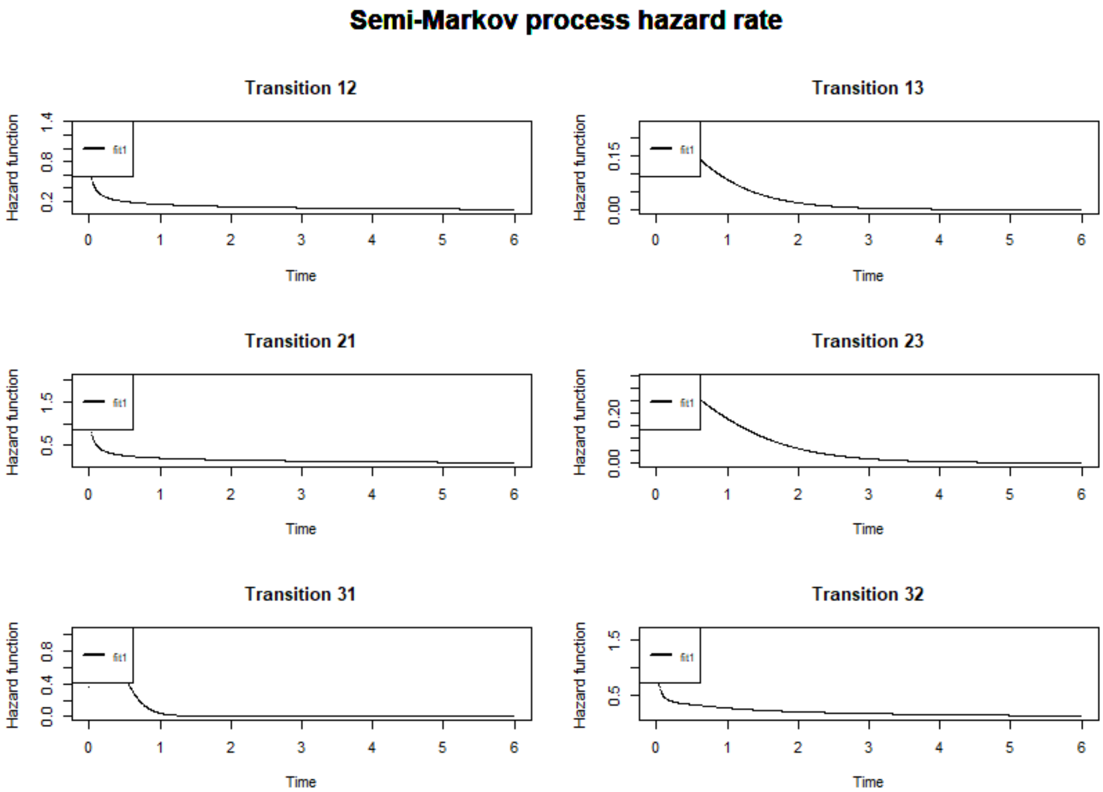

Figure 4.3: The hazard rate of the semi-Markov process.

The effect of covariates and the proportional hazard assumption can be evaluated by representing the hazard rates in each stratum. In a second step, a proportional model can be considered to study the effect of covariates. For instance, one can consider a model with BMI as covariate and the Weibull distribution for the waiting times.

Semi-Markov model with a covariate "BMI"

```
R> BMI <- as.data.frame(asthma$BMI)
## Semi-Markov model with a covariate "BMI"
R> fit2 <- semiMarkov(data=asthma, states=states,
mtrans=mtrans, cov=BMI)
## Estimations of parameters of the waiting times distributions
R> fit2$table.dist
```

```
$Sigma
```

	Transition	Sigma	SD	Lower_CI	Upper_CI	Wald_test	p_value
1	1 -> 2	9.384	2.42	4.64	14.13	12.01	0.0005
2	1 -> 3	0.418	0.08	0.26	0.58	51.54	<0.0001
3	2 -> 1	5.014	1.25	2.57	7.46	10.36	0.0013
4	2 -> 3	0.714	0.12	0.49	0.94	6.06	0.0138
5	3 -> 1	2.233	0.53	1.20	3.26	5.51	0.0189
6	3 -> 2	0.498	0.08	0.34	0.65	41.05	<0.0001

```
$Nu
```

	Transition	Nu	SD	Lower_CI	Upper_CI	Wald_test	p_value
1	1 -> 2	0.531	0.05	0.44	0.63	95.85	<0.0001
2	1 -> 3	1.18	0.14	0.90	1.46	1.65	0.1990
3	2 -> 1	0.51	0.04	0.43	0.59	141.80	<0.0001
4	2 -> 3	1.048	0.10	0.86	1.24	0.25	0.6171
5	3 -> 1	0.499	0.04	0.42	0.58	161.12	<0.0001
6	3 -> 2	0.931	0.06	0.81	1.06	1.14	0.2857

The `semiMarkov` function provides estimations of parameters of the waiting times distributions, the standard deviations, the confidence intervals and the Wald test statistics ($H_0 : \theta_{hj} = 1$). One can observe that the coefficient ν_{23} et ν_{32} associated to the transition from state 2 to state 3 and from 3 to 2 is not significantly different from 1. The exponential distribution can then be used instead of the Weibull distribution for this transitions.

```
## The estimation of the coefficient
R> fit2$table.coef
```

	Transition	Covariates	Estimation	SD	Lower_CI	Upper_CI	Wald_test	p_value
1	1 -> 2	Beta1	-0.27808202	0.22	-0.72	0.16	1.55	0.2131
2	1 -> 3	Beta1	-0.87827455	0.35	-1.57	-0.19	6.27	0.0123
3	2 -> 1	Beta1	0.03216316	0.19	-0.35	0.41	0.03	0.8625
4	2 -> 3	Beta1	-0.11151384	0.27	-0.64	0.41	0.17	0.6801
5	3 -> 1	Beta1	-0.61127842	0.20	-1.00	-0.22	9.43	0.0021
6	3 -> 2	Beta1	-0.23912936	0.21	-0.65	0.17	1.32	0.2506

For this new model, BMI regression coefficients remain significant for transitions from 1 to 3 and from 3 to 1 with $\beta_{13} = -0.88$ and $\beta_{31} = -0.61$, and respective p-values 0.012, and 0.002. The fact that hazard rate of the sojourn time associated with these covariates is less than unity (estimated coefficients are negative), indicates that BMI$\geq$ 25 generally lengthens the duration of the sojourn time in state 1 when making a $1 \rightarrow 3$ transition and generally lengthens the duration of the sojourn time in state 3 when making a $3 \rightarrow 1$ transition. This can also be interpreted as a decrease of the risk of leaving "optimal control" state to "unacceptable control" as well as a decrease of the risk of leaving "unacceptable control" state to "optimal control".

```
## Time fixed covariate
## Covariate equal to 0 and 1 for each transition
R> alpha2 <- hazard(fit2, cov=0)
R> alpha3 <- hazard(fit2, cov=1)
R> plot(alpha2,alpha3)
```

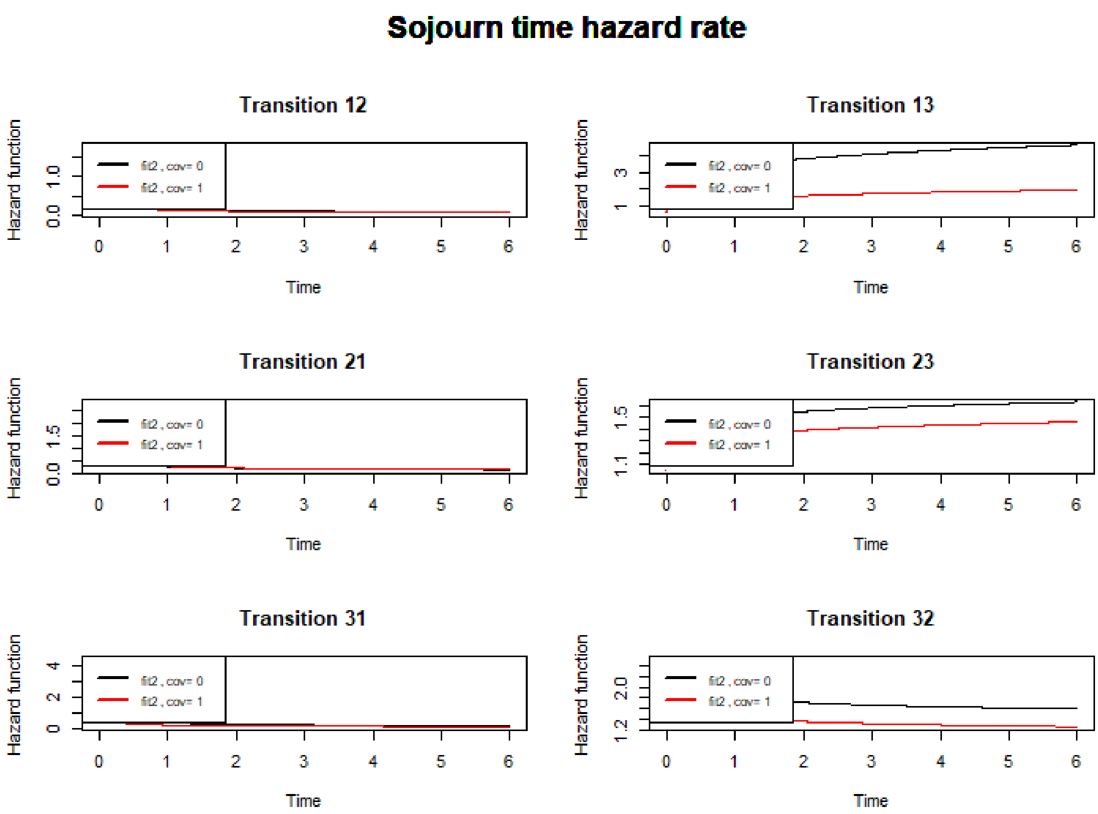

Figure 4.4: The hazard rate of sojourn time

4.2 Application to Covid-19 pandemic

The COVID-19 pandemic, also known as the Coronavirus pandemic is an ongoing global pandemic of Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome Coronavirus 2 (SARS-CoV-2). The outbreak was first identified in Wuhan, China, in December 2019.

The World Health Organization declared the outbreak a Public Health Emergency of International Concern on 30 January 2020, and a pandemic on 11 March. As of 3 July 2020, more than 10.8 million cases of COVID-19 have been reported in more than 188 countries and territories, resulting in more than 521,000 deaths, more than 5.76 million people have recovered.

In this section we apply the semi-Markov model to the data set of the COVID-19 pandemic in Algeria and Tunisia.

4.2.1 Application for Tunisia Coronavirus data

Firstly, we apply the semi-Markov model in continuous-time for Coronavirus cases in Tunisia. So we consider this cases between March and June, which are given in table 4.6.

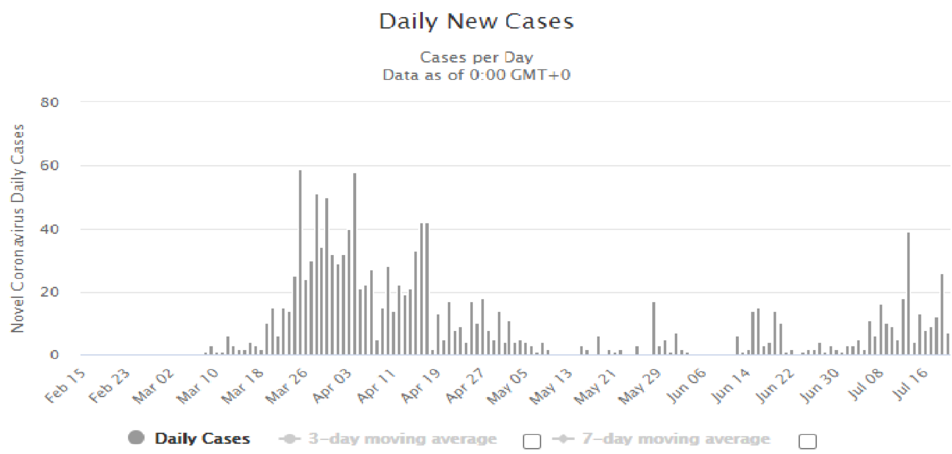

Figure 4.5: Daily new cases in Tunisia.

From the table 4.6 we can define three states corresponding to the number of Coronavirus cases in Tunisia:

- State 1:[0 : 24]
- State 2:[25 : 50]
- State 3:[51 : 75].

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

We can resume the table 4.6 to the following table:

date	state
21/03/2020	3
22/03/2020	1
23/03/2020	2
24/03/2020	3
25/03/2020	1
26/03/2020	2
27/03/2020	3
28/03/2020	2
04/04/2020	3
05/04/2020	1
07/04/2020	2
08/04/2020	1
10/04/2020	2
12/04/2020	1
14/04/2020	2
17/04/2020	1
22/04/2020	2
23/04/2020	1

Table 4.1: Table of the states corresponding to the number of Coronavirus cases in Tunisia.

Let $t = 0$ be the time we observed the first cases which happened in March 21, 2020. We set the end time $M = 96$ days.

The number of observed transitions in the data set 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 $N_{ij}(M)$ for all $i, j \in$ E.

$$N = \begin{pmatrix} 0 & 6 & 0 \\ 4 & 0 & 3 \\ 3 & 1 & 0 \end{pmatrix}.$$

We read the matrix N as follows: four times there was a transition from state 2 to state 1.

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

$$N_1(M) = 6, \qquad N_2(M) = 7, \qquad N_3(M) = 4.$$

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

$$\widehat{P} = \begin{pmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.5714 & 0.0000 & 0.4286 \\ 0.7500 & 0.2500 & 0.0000 \end{pmatrix}.$$

We read the matrix $\widehat{P}$ as follows: the probability that there is a transition from state 3 to state 1 is equal to 0.75.

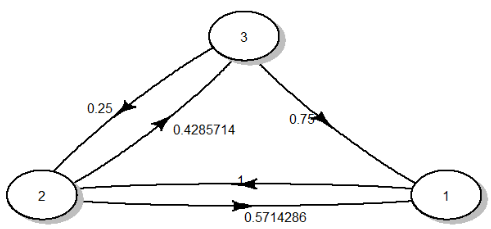

Figure 4.6: Transitions from state i to state j with transition probabilities, for all $i, j \in$ E.

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

The semi-Markov kernels are shown in figure 4.7 for all transitions from state i to state j, $i, j \in \mathrm{E}$ and $t \geq 0$. The sojourn time is measured in weeks.

The empirical estimators for conditional transition functions $\widehat{F}_{ij}(t, M)$, associated with the sojourn time in each state before transition, are shown in figure 4.8 for all transitions from state i to state j, $i, j \in \mathrm{E}$ and $t \geq 0$.

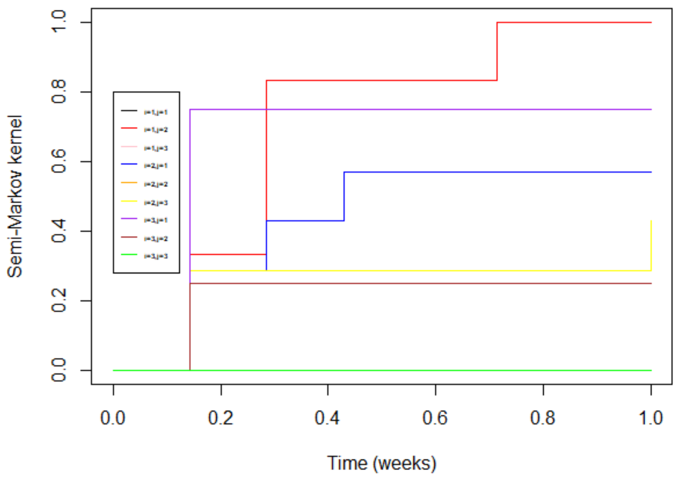

Figure 4.7: Empirical estimators for semi-Markov kernels, $\widehat{Q}_{ij}(t, M)$ for all transitions from state i to state j, $i, j \in \mathrm{E}$.

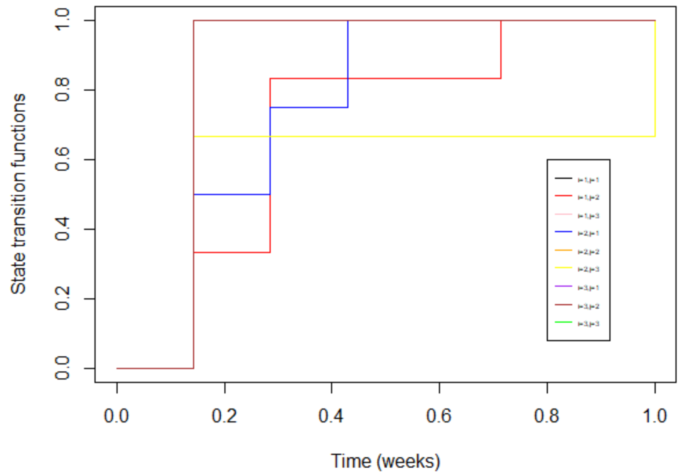

Figure 4.8: Empirical estimators for conditional transition functions, $\widehat{F}_{ij}(t, M)$ for all transitions from state i to state j, $i, j \in$ E.

Given that the last Coronavirus occurrence was in state i and at least a time interval of length t has already elapsed, the probability of an Coronavirus occurrence of state j in the next time interval of length Δ is denoted by $\lambda_{ij}(t)\Delta$. The term *Instantaneous Coronavirus Occurrence Rate* at state j in the next step conditional on the starting state i is used for the description of the probability $\lambda_{ij}(t)\Delta$, which is expressed by means of the semi-Markov kernels via the formula

$$\lambda_{ij}(t)\Delta = \frac{Q_{ij}(t+\Delta) - Q_{ij}(t)}{\overline{H}_i(t)} + o(\Delta).$$

Table 4.2 shows the estimated instantaneous Coronavirus occurrence rate for each type of transitions.

				t=0.1					
Δ	$\lambda_{11}(t)$	$\lambda_{12}(t)$	$\lambda_{13}(t)$	$\lambda_{21}(t)$	$\lambda_{22}(t)$	$\lambda_{23}(t)$	$\lambda_{31}(t)$	$\lambda_{32}(t)$	$\lambda_{31}(t)$
1/2	0	1.667	0	1.143	0	0.571	1.5	0.5	0
1	0	1	0	0.571	0	0.429	0.75	0.25	0
2	0	0.5	0	0.286	0	0.214	0.375	0.125	0

Table 4.2: Estimated instantaneous Coronavirus occurrence rates

4.2.2 Application for Algeria Coronavirus data

Now, we apply the semi-Markov model in continuous-time for Coronavirus cases in Algeria. So we consider this cases between March and June, which are given in table 4.5.

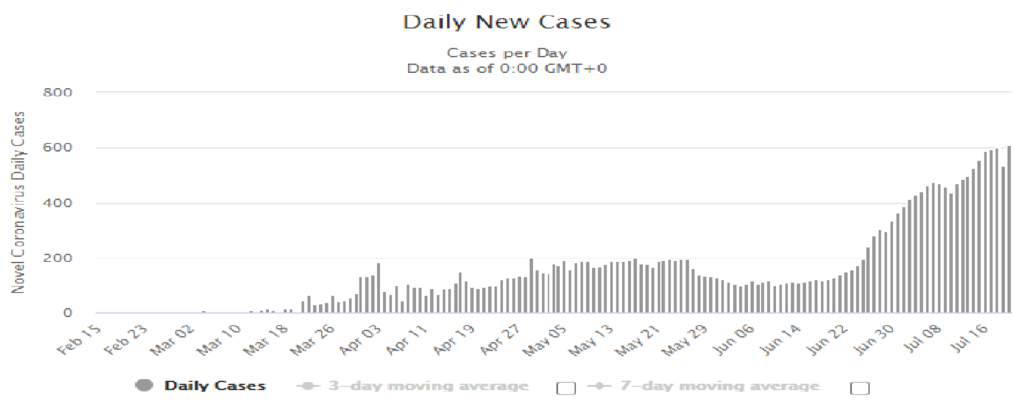

Figure 4.9: Daily new cases in Algeria.

From the table 4.5 we can define three states corresponding to the number of Coronavirus cases in Algeria:

- State 1: [22 : 99]
- State 2: [100 : 150]
- State 3: [151 : 199].

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

We can resume the table 4.5 to the following table:

Date	state
21-03-2020	2
22-03-2020	1
31-03-2020	2
03-04-2020	3
04-04-2020	1
06-04-2020	2
07-04-2020	1
08-04-2020	2
09-04-2020	1
16-04-2020	2
19-04-2020	1
24-04-2020	2
29-04-2020	3
01-05-2020	2
03-05-2020	3
28-05-2020	2
04-06-2020	1
05-06-2020	2
23-06-2020	3

Table 4.3: Table of the states corresponding to the number of Coronavirus cases in Algeria.

Non parametric estimation

Let $t = 0$ be the time we observed the first cases which happened in March 21, 2020, and we set the end time $M = 96$ days.

The number of observed transitions in the data set 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 $N_{ij}(M)$ for all $i, j \in$ E.

$$N = \begin{pmatrix} 0 & 6 & 0 \\ 5 & 0 & 4 \\ 1 & 2 & 0 \end{pmatrix}.$$

We read the matrix N as follows: four times there was a transition from state 2 to state 3.

The values $N_i(M)$ for $i \in \mathrm{E}$ are equal to

$$N_1(M) = 6, \qquad N_2(M) = 9, \qquad N_3(M) = 3.$$

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

$$\widehat{P} = \begin{pmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.5556 & 0.0000 & 0.4444 \\ 0.3333 & 0.6667 & 0.0000 \end{pmatrix}. \tag{4.1}$$

We read the matrix $\widehat{P}$ as follows: the probability that there is a transition from state 3 to state 1 is equal to 0.3333.

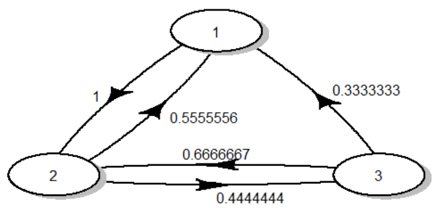

Figure 4.10: Transitions from state i to state j with transition probabilities, for all $i, j \in \mathrm{E}$.

The semi-Markov kernels are shown in figure 4.11 for all transitions from state i to state j, $i, j \in \mathrm{E}$ and $t = 0$. The sojourn time is measured in weeks.

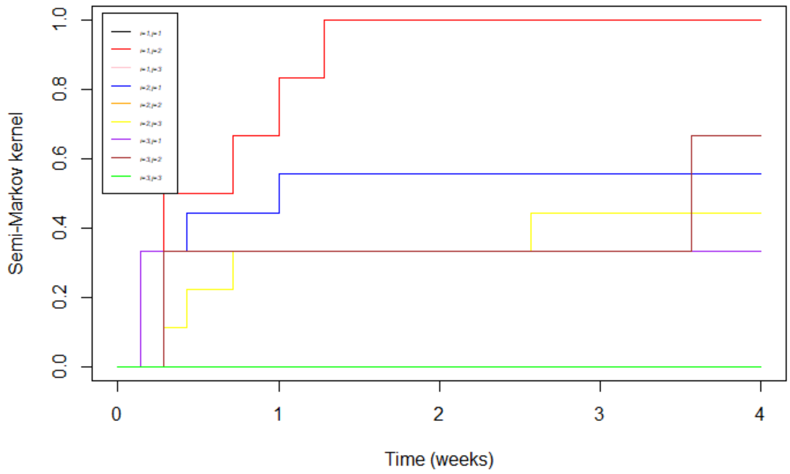

Figure 4.11: Empirical estimators for semi-Markov kernels, $\widehat{Q}_{ij}(t, M)$ for all transitions from state i to state j, $i, j \in$ E.

The empirical estimators for conditional transition functions $\widehat{F}_{ij}(t, M)$, associated with the sojourn time in each state before transition, are shown in figure 4.12 for all transitions from state i to state j, $i, j \in$ E and $t = 0$.

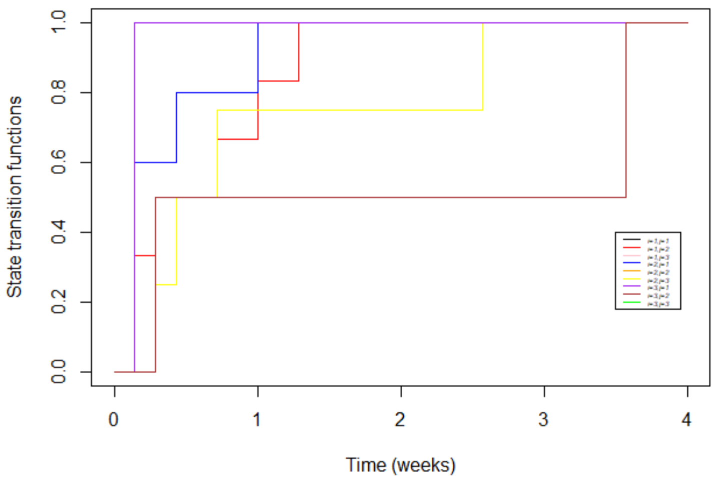

Figure 4.12: Empirical estimators for conditional transition functions, $\widehat{F}_{ij}(t, M)$ for all transitions from state i to state j, $i, j \in$ E.

Table 4.4 shows the estimated instantaneous Coronavirus occurrence rate for each type of transitions.

				t=1/2					
Δ	$\lambda_{11}(t)$	$\lambda_{12}(t)$	$\lambda_{13}(t)$	$\lambda_{21}(t)$	$\lambda_{22}(t)$	$\lambda_{23}(t)$	$\lambda_{31}(t)$	$\lambda_{32}(t)$	$\lambda_{31}(t)$
1	0	1	0	0.333	0	0.333	0	0	0
2	0	0.5	0	0.167	0	0.167	0	0	0
5	0	1	0	0.333	0	0.333	0	0	0

Table 4.4: Estimated instantaneous Coronavirus occurrence rates.

Parametric estimation

With use of the **R** package `SemiMarkov` 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.

Homogeneous Markov model

First, we derive the hazard rate function of the semi-Markov process for the homogeneous Markov model. We changed the original data set from table 4.3 to meet the requirements to apply for the package, which we call it `markov`. We measure the time t in weeks and 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 $\mathcal{E}$ for the sojourn time. We fit the data with use of the function `semiMarkov(.)`. Figure 4.13 shows the performed steps in **R**studio.

```
> library(numDeriv)
> library(MASS)
> library(Rsolnp)
> library(SemiMarkov)
> markov=data.frame(id=rep(1,18),state.h=c(2,1,2,3,1,2,1,2,1,2,1,2,3,2,3,2,1,2),state.j=c(1,2,3,1,2,1,2,
1,2,1,2,3,2,3,2,1,2,3),time=c(0.1428571, 1.2857143, 0.4285714, 0.1428571, 0.2857143, 0.1428571, 0.142857
1, 0.1428571, 1.0000000,0.4285714, 0.7142857, 0.7142857, 0.2857143 ,0.2857143, 3.5714286, 1.0000000, 0.14
28571, 2.5714286))
> states=c("1","2","3")
> mtrans=matrix(FALSE,nrow =3 ,ncol=3)
> mtrans[1,2]=c("E")
> mtrans[2,c(1,3)]=c("E","E")
> mtrans[3,c(1,2)]=c("E","E")
> fit=semiMarkov(data = markov,states = states,mtrans = mtrans)

Iter: 1 fn: 17.3951      Pars:  0.59524 0.37143 1.00002 0.14286 1.92857 0.55556 0.33333
Iter: 2 fn: 17.3951      Pars:  0.59524 0.37143 1.00000 0.14286 1.92857 0.55556 0.33333
solnp--> Completed in 2 iterations
```

Figure 4.13: The setting of the hazard rate function in case of the homogeneous Markov model.

In figure 4.14, we see the estimates of parameters of the waiting time distributions, the standard deviations, the confidence intervals and the Wald test statistics.

```
> fit$table.dist
$Sigma
  Type Index Transition Estimation   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2      0.595 0.24     0.12     1.07    1.00      2.77  0.0960
2 dist     2     2 -> 1      0.371 0.17     0.05     0.70    1.00     14.32  0.0002
3 dist     3     2 -> 3          1 0.50     0.02     1.98    1.00      0.00  1.0000
4 dist     4     3 -> 1      0.143 0.14    -0.14     0.42    1.00     36.00 <0.0001
5 dist     5     3 -> 2      1.929 1.36    -0.74     4.60    1.00      0.46  0.4976
```

Figure 4.14: Estimates of parameters of the waiting time distribution in case of the homogeneous Markov model.

In the following matrix $\Sigma = (\sigma_{ij})$ we give the values for the parameters of the exponential distribution for all $i \neq j \in \mathrm{E}$:

$$\Sigma = \begin{pmatrix} - & 0.595 & - \\ 0.371 & - & 1 \\ 0.143 & 1.929 & - \end{pmatrix}. \tag{4.2}$$

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

$$\begin{aligned} \alpha_{12}(t) &= 1.6807, & \alpha_{31}(t) &= 6.9930 \\ \alpha_{21}(t) &= 2.6954, & \alpha_{32}(t) &= 0.8383 \\ \alpha_{23}(t) &= 1. \end{aligned}$$

The plots of the hazard rate functions of the waiting time are shown in figure 4.15.

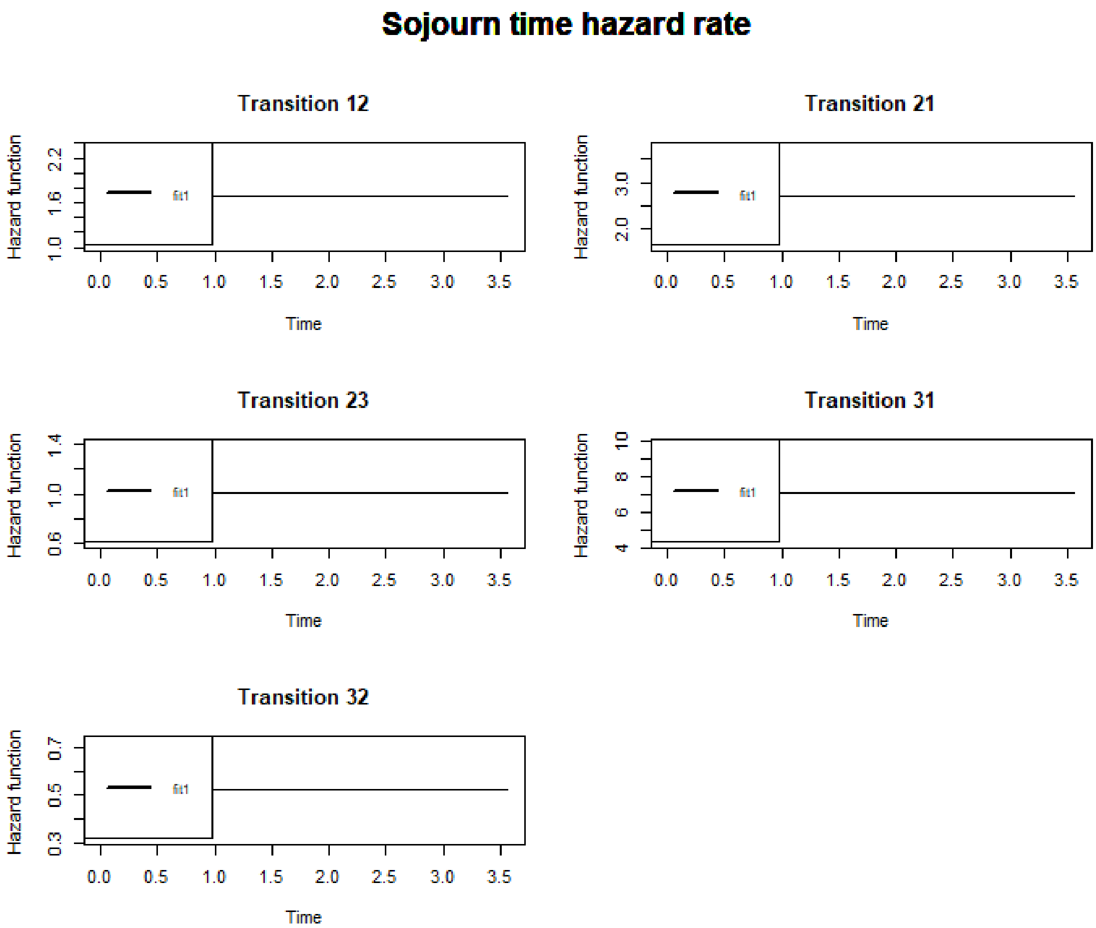

Figure 4.15: Hazard rate of waiting time for the homogeneous Markov model for transitions from state i to state j, $i \neq j \in$ E.

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

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

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:

$$\begin{array}{rclrcl} f_{12}(t) & = & 1.6807e^{-1.6807t}, & f_{31}(t) & = & 6.9930e^{-6.9930t} \\ f_{21}(t) & = & 2.6954e^{-2.6954t}, & f_{32}(t) & = & 0.8383e^{-0.8383t} \\ f_{23}(t) & = & 1e^{-1t}. & & & \end{array}$$

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

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

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

$$\begin{array}{rclrcl} F_{12}(t) & = & 1 - e^{-1.6807t}, & F_{31}(t) & = & 1 - e^{-6.9930t} \\ F_{21}(t) & = & 1 - e^{-2.6954t}, & F_{32}(t) & = & 1 - e^{-0.8383t} \\ F_{23}(t) & = & 1 - e^{-1t}. & & & \end{array}$$

We know that the transition probability matrix is given by

$$\widehat{P} = \begin{pmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.5556 & 0.0000 & 0.4444 \\ 0.3333 & 0.6667 & 0.0000 \end{pmatrix}.$$

We can write that $1 - H_i(t) = \sum_{j \in \mathrm{E}} p_{ij}(1 - F_{ij}(t))$ 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

$$\lambda_{ij}(t) = \frac{q_{ij}(t)}{1 - H_i(t)} = \frac{p_{ij} f_{ij}(t)}{\sum_{j \in \mathrm{E}} p_{ij}(1 - F_{ij}(t))}.$$

We obtain the estimated hazard rate functions of the semi-Markov process for all $i \neq j \in$ E. The plots of the hazard rate function of the semi-Markov process are shown in figure 4.16.

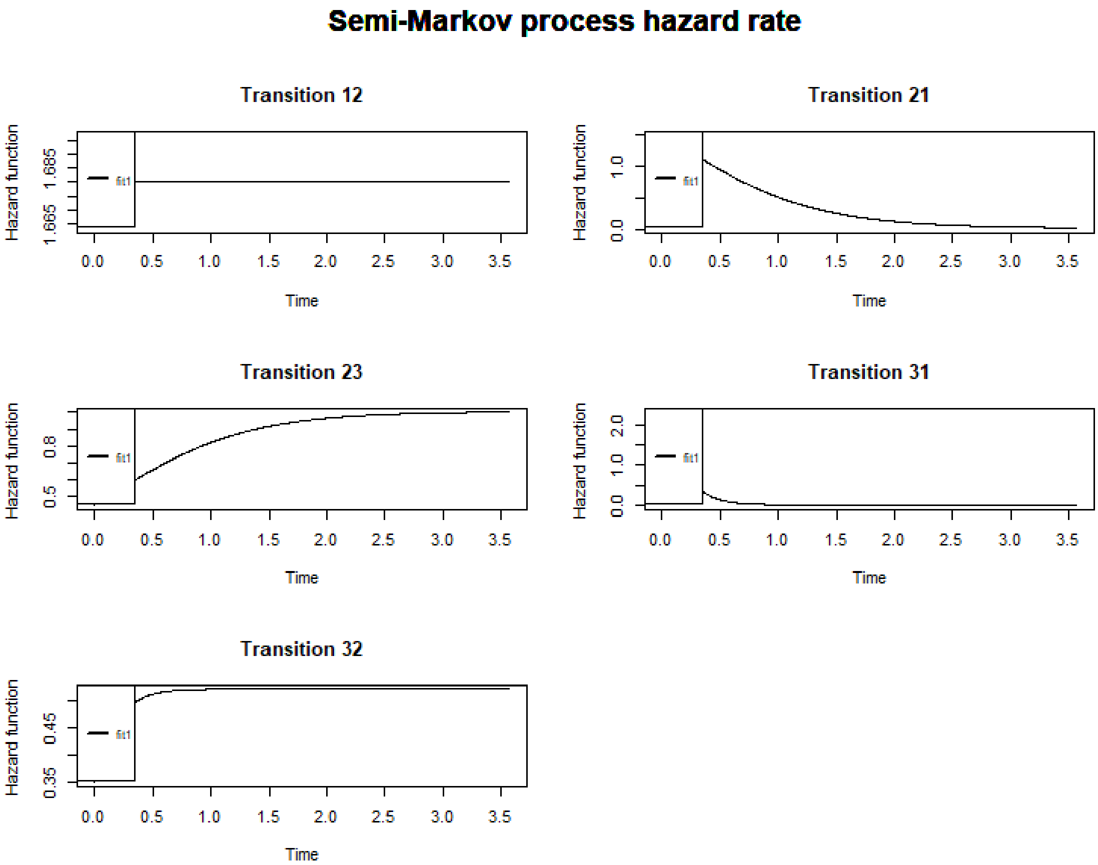

Figure 4.16: Hazard rate of semi-Markov process for the homogeneous Markov model for transitions from state from state i to state j, $i \neq j \in$ E.

Now, we determine the semi-Markov kernels for the homogeneous Markov model.

From 4.1 and 4.2, we obtain the following semi-Markov kernels for the homogeneous Markov model with $i \neq j \in$ E.

$$
\begin{aligned}
Q_{12}(t) &= 1(1-e^{-1.6807t}), & Q_{31}(t) &= 0.3333(1-e^{-6.9930t}) \\
Q_{21}(t) &= 0.5556(1-e^{-2.6954t}), & Q_{32}(t) &= 0.6667(1-e^{-0.8383t}) \\
Q_{23}(t) &= 0.4444(1-e^{-1t}).
\end{aligned}
$$

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

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.

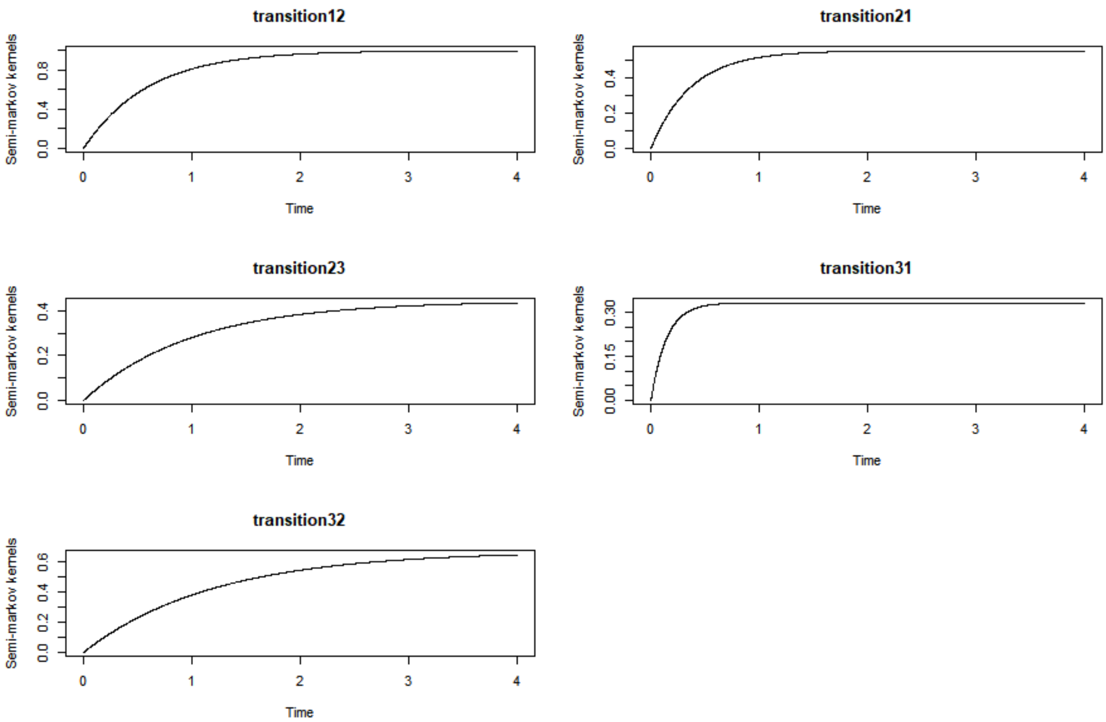

Figure 4.17: Semi-Markov kernels for the homogenous Markov model for all transitions from state i to state j, $i \neq j \in$ E.

Wald test and p-value.

For the homogeneous Markov model, we choose the exponential distribution for the sojourn time of the process. In figure 4.18 the results of the Wald test are shown.

```
> fit$table.dist
$Sigma
  Type Index Transition Estimation    SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2      0.595 0.24     0.12     1.07    1.00      2.77  0.0960
2 dist     2     2 -> 1      0.371 0.17     0.05     0.70    1.00     14.32  0.0002
3 dist     3     2 -> 3          1 0.50     0.02     1.98    1.00      0.00  1.0000
4 dist     4     3 -> 1      0.143 0.14    -0.14     0.42    1.00     36.00 <0.0001
5 dist     5     3 -> 2      1.929 1.36    -0.74     4.60    1.00      0.46  0.4976
```

Figure 4.18: Wald test p-values for the homogeneous Markov model.

We derive that for all σ_{ij}, $i \neq j \in$ E, we only reject the null-hypothesis for the scale parameters σ_{21} and σ_{31}, the rest we fail to reject the null-hypothesis for σ_{ij}.

Semi-Markov model

For the semi-Markov model we use the same data set as before, and call it `semimarkov`. We measure the time t in weeks. Because we want an estimator for the semi-Markov process, we choose the (non-Markov) Weibull distribution $\mathcal{W}$ for the sojourn time for all transitions except for the transition from state 3 to state 1 Here, we choose the exponential distribution. We fit the data with use of the function `semiMarkov(.)` as before.
Figure 4.19 shows the performed steps in **R**studio.

```
> library(numDeriv)
> library(MASS)
> library(Rsolnp)
> library(SemiMarkov)
> markov=data.frame(id=rep(1,18),state.h=c(2,1,2,3,1,2,1,2,1,2,1,2,3,2,3,2,1,2),state.j=c(1,2,3,1,2,1,2,
1,2,1,2,3,2,3,2,1,2,3),time=c(0.1428571, 1.2857143, 0.4285714, 0.1428571, 0.2857143, 0.1428571, 0.142857
1, 0.1428571, 1.0000000,0.4285714, 0.7142857, 0.7142857, 0.2857143 ,0.2857143, 3.5714286, 1.0000000, 0.14
28571, 2.5714286))
> states=c("1","2","3")
> mtrans=matrix(FALSE,nrow =3 ,ncol=3)
> mtrans[1,2]=c("W")
> mtrans[2,c(1,3)]=c("W","W")
> mtrans[3,c(1,2)]=c("E","W")
> fit2=semiMarkov(data = markov,states = states,mtrans = mtrans)

Iter: 1 fn: 16.8176      Pars:  0.64831 0.39995 1.06613 0.14287 1.88666 1.32623 1.21954 1.17870 0.94999
 0.55556 0.33335
Iter: 2 fn: 16.8176      Pars:  0.64831 0.39996 1.06614 0.14286 1.88666 1.32623 1.21954 1.17871 0.94998
 0.55556 0.33333
solnp--> Completed in 2 iterations
```

Figure 4.19: The setting of the hazard rate function in case of the semi-Markov model.

From figure 4.20, we can derive the estimates of parameters of the waiting time distributions, the standard deviations, the confidence intervals and the Wald test statistics.

```
> fit2$table.dist
$Sigma
  Type Index Transition Sigma   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2 0.648 0.21     0.24     1.06    1.00      2.78  0.0954
2 dist     2     2 -> 1   0.4 0.16     0.09     0.71    1.00     14.79  0.0001
3 dist     3     2 -> 3 1.066 0.48     0.12     2.01    1.00      0.02  0.8875
4 dist     4     3 -> 1 0.143 0.14    -0.14     0.42    1.00     36.00 <0.0001
5 dist     5     3 -> 2 1.887 1.48    -1.02     4.79    1.00      0.36  0.5485

$Nu
  Type Index Transition    Nu   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     6     1 -> 2 1.326 0.44     0.46     2.19    1.00      0.55  0.4583
2 dist     7     2 -> 1  1.22 0.41     0.42     2.02    1.00      0.29  0.5902
3 dist     8     2 -> 3 1.179 0.44     0.31     2.04    1.00      0.16  0.6892
4 dist     9     3 -> 2  0.95 0.56    -0.15     2.05    1.00      0.01  0.9203
```

Figure 4.20: Estimates of parameters of the waiting time distribution in case of the semi-Markov model.

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:

$$\Sigma = \begin{pmatrix} - & 0.648 & - \\ 0.4 & - & 1.066 \\ 0.143 & 1.887 & - \end{pmatrix}, \qquad V = \begin{pmatrix} - & 1.326 & - \\ 1.22 & - & 1.179 \\ - & 0.95 & - \end{pmatrix}. \tag{4.3}$$

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

$$\begin{aligned} \alpha_{12}(t) &= 2.0463(1.5432t)^{0.326}, & \alpha_{31}(t) &= 6.9930 \\ \alpha_{21}(t) &= 3.05(2.5t)^{0.22}, & \alpha_{32}(t) &= 0.5034(1.0526t)^{-0.05} \\ \alpha_{23}(t) &= 1.106(0.938t)^{0.179}. \end{aligned}$$

The plots of the hazard rate functions of the waiting time are shown in figure 4.21.

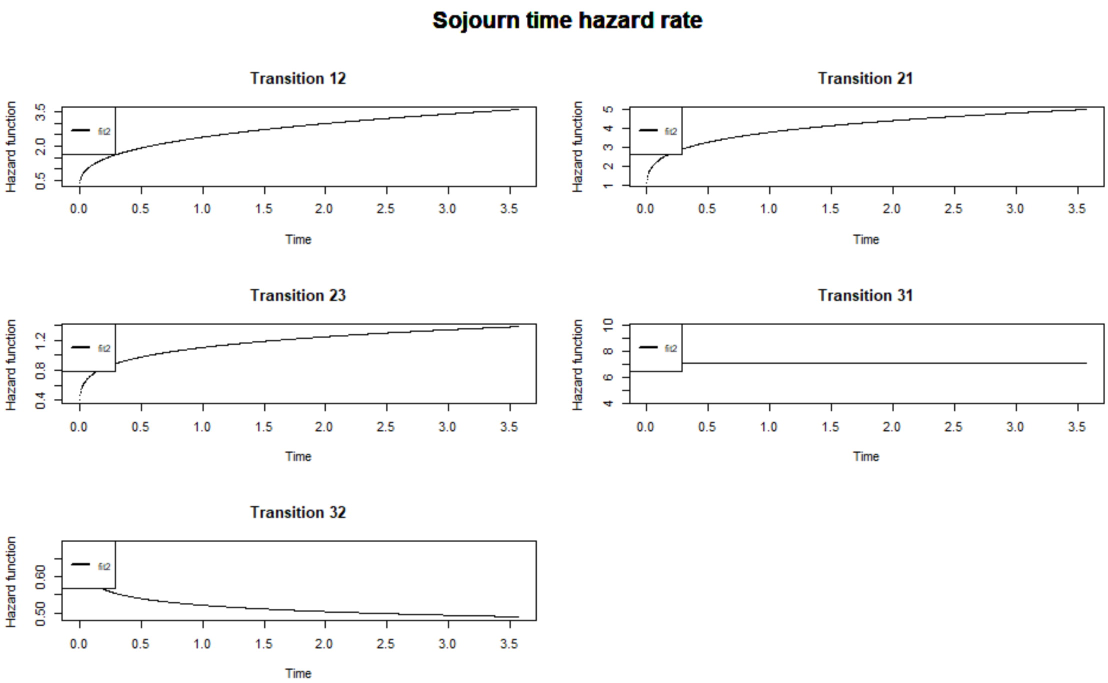

Figure 4.21: Hazard rate of waiting time for the semi-Markov model for transitions from state i to state j, $i \neq j \in$ E.

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

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

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:

$$\begin{aligned} f_{12}(t) &= 2.0463(1.5432t)^{0.326}e^{-(1.5432t)^{1.326}}, & f_{31}(t) &= 6.9930e^{-6.9930t} \\ f_{21}(t) &= 3.05(2.5t)^{0.22}e^{-(2.5t)^{1.22}}, & f_{32}(t) &= 0.5034(1.0526t)^{-0.05}e^{(1.0526t)^{0.95}} \\ f_{23}(t) &= 1.106(0.938t)^{0.179}e^{-(0.938t)^{1.179}}. \end{aligned}$$

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

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

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

$$\begin{aligned} F_{12}(t) &= 1 - e^{-(1.5432t)^{1.326}}, & F_{31}(t) &= 1 - e^{-6.9930t} \\ F_{21}(t) &= 1 - e^{-(2.5t)^{1.22}}, & F_{32}(t) &= 1 - e^{(1.0526t)^{0.95}} \\ F_{23}(t) &= 1 - e^{-(0.938t)^{1.179}}. \end{aligned}$$

As before, the transition probability matrix is given by

$$\widehat{P} = \begin{pmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.5556 & 0.0000 & 0.4444 \\ 0.3333 & 0.6667 & 0.0000 \end{pmatrix}.$$

The plots of the hazard rate functions of the semi-Markov process are shown in figure 4.22.

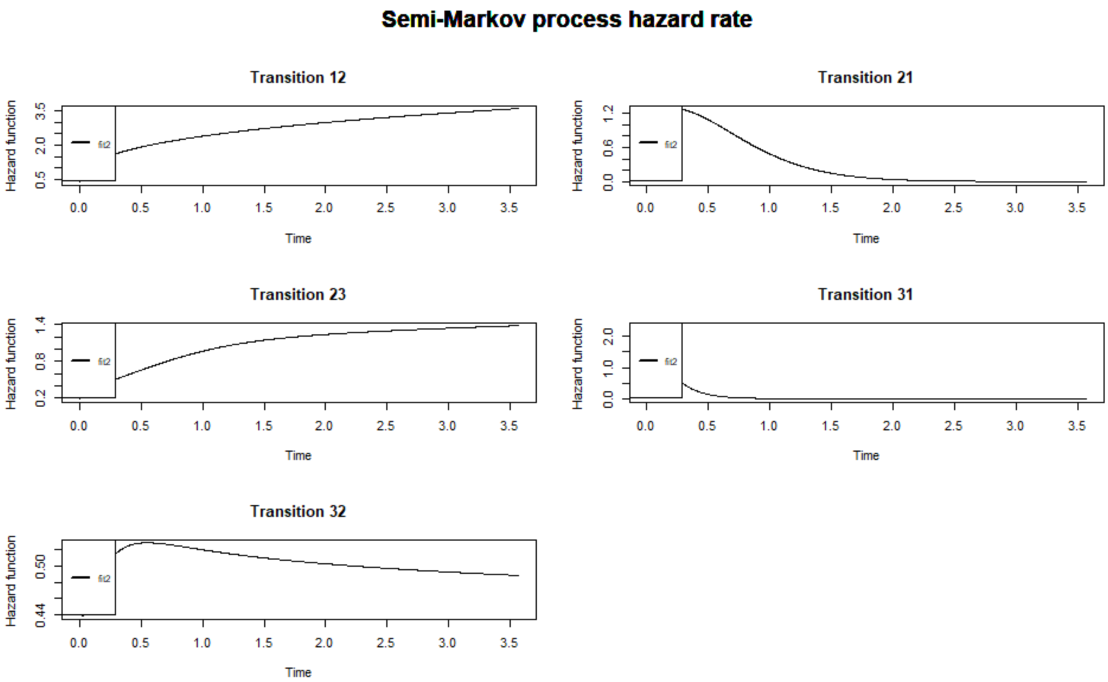

Figure 4.22: Hazard rate of semi-Markov process for the semi-Markov model for transitions from state i to state j, $i \neq j \in \mathrm{E}$.

Now, using 4.1 and 4.3 for $i \neq j \in$ E, we obtain the following semi-Markov kernels for the semi-Markov model.

$$\begin{aligned} Q_{12}(t) &= 1(1-e^{-(1.5432t)^{1.326}}), & Q_{31}(t) &= 0.3333(1-e^{-6.9930t}) \\ Q_{21}(t) &= 0.5556(1-e^{-(2.5t)^{1.22}}), & Q_{32}(t) &= 0.6667(1-e^{(1.0526t)^{0.95}}) \\ Q_{23}(t) &= 0.4444(1-e^{-(0.938t)^{1.179}}). \end{aligned}$$

We define also $Q_{ii}(t) = 0$ for all $i, j \in$ E and $t \geq 0$.

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.

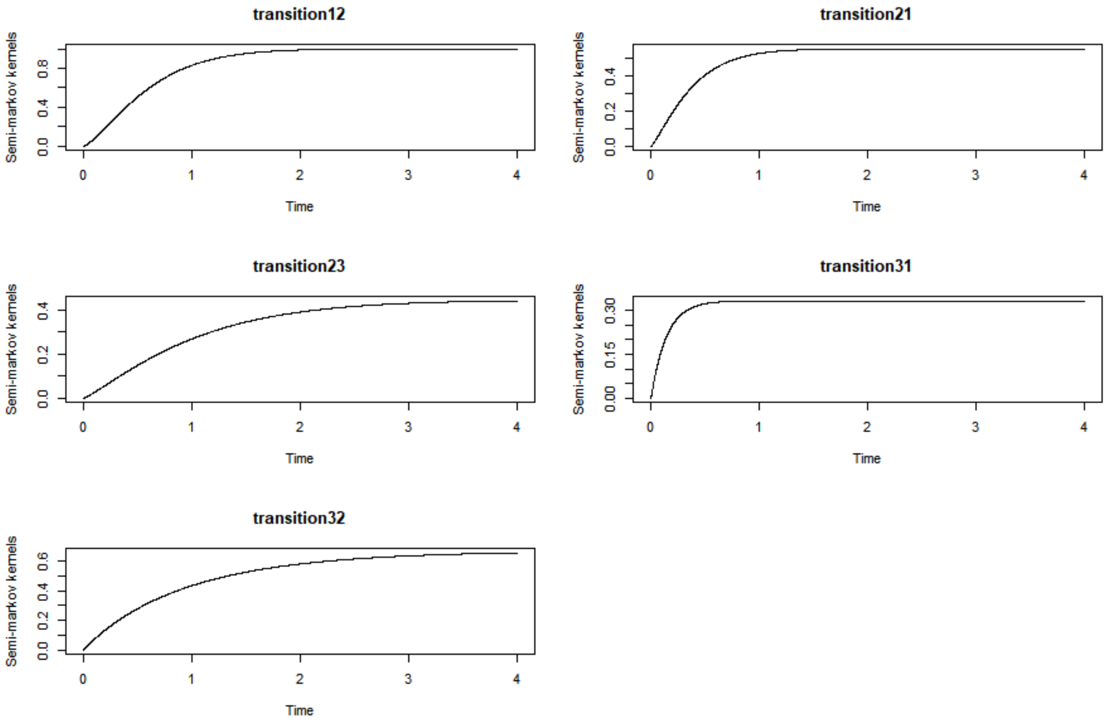

Figure 4.23: Semi-Markov kernels for the semi-Markov model for all transitions from state i to state j, $i \neq j \in$ E.

Wald test and p-value.

For the semi-Markov model the results of the Wald test are shown in figure 4.24. We remember that for the transitions from state 3 to state 1 we chose an exponential distribution for the sojourn time instead of the Weibull distribution. So we exclude this transition from our conclusions.

```
> fit2$table.dist
$Sigma
  Type Index Transition Sigma   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     1     1 -> 2 0.648 0.21     0.24     1.06    1.00      2.78  0.0954
2 dist     2     2 -> 1   0.4 0.16     0.09     0.71    1.00     14.79  0.0001
3 dist     3     2 -> 3 1.066 0.48     0.12     2.01    1.00      0.02  0.8875
4 dist     4     3 -> 1 0.143 0.14    -0.14     0.42    1.00     36.00 <0.0001
5 dist     5     3 -> 2 1.887 1.48    -1.02     4.79    1.00      0.36  0.5485

$Nu
  Type Index Transition    Nu   SD Lower_CI Upper_CI Wald_H0 Wald_test p_value
1 dist     6     1 -> 2 1.326 0.44     0.46     2.19    1.00      0.55  0.4583
2 dist     7     2 -> 1  1.22 0.41     0.42     2.02    1.00      0.29  0.5902
3 dist     8     2 -> 3 1.179 0.44     0.31     2.04    1.00      0.16  0.6892
4 dist     9     3 -> 2  0.95 0.56    -0.15     2.05    1.00      0.01  0.9203
```

Figure 4.24: Wald test p-values for the semi-Markov model.

In this case we only reject the null-hypothesis for the scale parameter σ_{21} 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}. 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.

Conclusion

In this work, we explained the continuous-time semi-Markov model with a discrete set of states. We defined empirical estimators of important quantities such as semi-Markov kernel, sojourn time distributions, transition probabilities, and hazard rate function. We gave results about their asymptotic properties.

The present work aims at the introduction of the continuous-time semi-Markov model as a candidate model for the description of asthma control, Tunisia, and Algeria Coronavirus data. For asthma control, it was very important to study this data with covariate variable (BMI), using the Wald test, we can conclude the decreasing or increasing effects of this variable.

The process of Algeria Coronavirus data was represented with two statistical models Markov and semi Markov model and with the parametric and nonparametric methods. Semi-Markov package in R Language was used for the implementation of the parametric-method however, for the nonparametric one, we had developed our functions. Parametric methods provide estimators with several attractive asymptotic properties; however, these estimators present inconvenience when the sample size is small. Since applications of parametric methods presuppose certain conditions concerning the sample size, this difficulty could be affected through the application of nonparametric methods. For the hazard rate functions, the semi-Markov process maybe a better fit for the previous model.

For providing more accurate forecasting results for Algeria Coronavirus data one more ways the accessibility into instantaneous results about Coronavirus cases and the inclusion of different covariate variables like age, chronic diseases,$\cdots$.

Appendix

Date	Number of cases	state
21-03-2020	139	2
22-03-2020	62	1
23-03-2020	29	1
24-03-2020	34	1
25-03-2020	38	1
26-03-2020	65	1
27-03-2020	42	1
28-03-2020	45	1
29-03-2020	57	1
30-03-2020	73	1
31-03-2020	132	2
01-04-2020	131	2
02-04-2020	139	2
03-04-2020	185	3
04-04-2020	80	1
05-04-2020	69	1
06-04-2020	103	2
07-04-2020	45	1
08-04-2020	104	2
09-04-2020	94	1
10-04-2020	95	1
11-04-2020	64	1
12-04-2020	89	1
13-04-2020	69	1
14-04-2020	87	1
15-04-2020	90	1
16-04-2020	108	2
17-04-2020	150	2
18-04-2020	116	2
19-04-2020	95	1
20-04-2020	89	1
21-04-2020	93	1
22-04-2020	99	1
23-04-2020	97	1
24-04-2020	120	2
25-04-2020	129	2
26-04-2020	126	2
27-04-2020	135	2
28-04-2020	132	2
29-04-2020	199	3
30-04-2020	158	3
01-05-2020	148	2
02-05-2020	141	2
03-05-2020	179	3
04-05-2020	174	3
05-05-2020	190	3
06-05-2020	159	3
07-05-2020	185	3
08-05-2020	187	3

Date	Number of cases	state
09-05-2020	189	3
10-05-2020	165	3
11-05-2020	168	3
12-05-2020	176	3
13-05-2020	186	3
14-05-2020	189	3
15-05-2020	187	3
16-05-2020	192	3
17-05-2020	198	3
18-05-2020	182	3
19-05-2020	176	3
20-05-2020	165	3
21-05-2020	186	3
22-05-2020	190	3
23-05-2020	195	3
24-05-2020	193	3
25-05-2020	197	3
26-05-2020	194	3
27-05-2020	160	3
28-05-2020	140	2
29-05-2020	137	2
30-05-2020	133	2
31-05-2020	127	2
01-06-2020	119	2
02-06-2020	113	2
03-06-2020	107	2
04-06-2020	98	1
05-06-2020	104	2
06-06-2020	115	2
07-06-2020	104	2
08-06-2020	111	2
09-06-2020	117	2
10-06-2020	102	2
11-06-2020	105	2
12-06-2020	109	2
13-06-2020	112	2
14-06-2020	109	2
15-06-2020	112	2
16-06-2020	116	2
17-06-2020	121	2
18-06-2020	117	2
19-06-2020	119	2
20-06-2020	127	2
21-06-2020	140	2
22-06-2020	149	2
23-06-2020	156	3
24-06-2020	176	3
25-06-2020	197	3

Table 4.5: Table of the number cases (COVID-19) in Algeria.

Date	Number of cases	state
21-03-2020	75	3
22-03-2020	14	1
23-03-2020	25	2
24-03-2020	59	3
25-03-2020	24	1
26-03-2020	30	2
27-03-2020	52	3
28-03-2020	33	2
29-03-2020	50	2
30-03-2020	32	2
31-03-2020	29	2
01-04-2020	32	2
03-04-2020	40	2
04-04-2020	58	3
05-04-2020	21	1
06-04-2020	22	1
07-04-2020	27	2
08-04-2020	5	1
09-04-2020	15	1
10-04-2020	28	2
11-04-2020	36	2
12-04-2020	19	1
13-04-2020	21	1
14-04-2020	33	2
15-04-2020	42	2
16-04-2020	42	2
17-04-2020	2	1
18-04-2020	13	1
19-04-2020	5	1
20-04-2020	17	1
21-04-2020	8	1
22-04-2020	30	2
23-04-2020	0	1
24-04-2020	0	1
25-04-2020	10	1
26-04-2020	18	1
27-04-2020	8	1
28-04-2020	5	1
29-04-2020	14	1
30-04-2020	4	1
01-05-2020	11	1
02-05-2020	4	1
03-05-2020	5	1
04-05-2020	4	1
05-05-2020	3	1
06-05-2020	1	1
07-05-2020	4	1
08-05-2020	2	1
09-05-2020	0	1
10-05-2020	0	1
11-05-2020	0	1
12-05-2020	0	1
13-05-2020	0	1
14-05-2020	3	1

Date	Number of cases	state
15-05-2020	2	1
16-05-2020	0	1
17-05-2020	6	1
18-05-2020	1	1
19-05-2020	1	1
20-05-2020	1	1
21-05-2020	2	1
22-05-2020	0	1
23-05-2020	3	1
24-05-2020	0	1
25-05-2020	0	1
26-05-2020	17	1
27-05-2020	3	1
28-05-2020	5	1
29-05-2020	1	1
30-05-2020	7	1
31-05-2020	2	1
01-06-2020	1	1
02-06-2020	0	1
03-06-2020	0	1
04-06-2020	0	1
05-06-2020	0	1
06-06-2020	0	1
07-06-2020	0	1
08-06-2020	0	1
09-06-2020	0	1
10-06-2020	6	1
11-06-2020	1	1
12-06-2020	2	1
13-06-2020	14	1
14-06-2020	15	1
15-06-2020	3	1
16-06-2020	4	1
17-06-2020	14	1
18-06-2020	10	1
19-06-2020	1	1
20-06-2020	2	1
21-06-2020	2	1
22-06-2020	1	1
23-06-2020	2	1
24-06-2020	2	1
25-06-2020	4	1

Table 4.6: Table of the number cases (COVID-19) in Tunisia.

Bibliography

[1] M.G. Akritas Roussas GG Asymptotic expansion of the log-likelihood function based on stopping times defined on a Markov process. Ann Inst Stat Math 31 (1979).

[2] A. Asanjarani, B. Liquet and Y. Nazarathy Estimation of Semi-Markov Multi-state Models: A Comparison of the Sojourn Times and Transition Intensities Approaches.

[3] V. Barbu and N. Limnios. Discrete time semi-Markov processes for reliability andsurvival analysis, Communications in statistic theory and methods. Taylor and Francis publisher 2004.

[4] V. Barbu and N. Limnios. Empirical estimation for discrete time semi-Markov processes with applications in reliability. J. Nonparametr.Statist. 2006a.

[5] V. Barbu and N. Limnios. Semi-Markov chains and Hidden Semi Markov Models toward Applications vol.191, Spring, New York.(2008a).

[6] P. Billingsley. The Lindeberg-L'evy theorem for martingales. Proc. Amer. Math. Soc. 1961b.

[7] P. Billingsley. Probability and Measure. Wiley, New York, 3rd edition, 1995.

[8] P. Billingsley. Convergence of Probability Measures. Wiley, New York, 2nd edition, 1999.

[9] E. Cinlar, Markov renewal theory, Adv. Appl. Prob. 1 (1969).

[10] H. Cramér, Mathematical Methods of Statistics, Princeton University Press, 1946.

[11] R.D. Gill Nonparametric estimation based on censored observations of Markov renewal process. Z Wahrscheinlichkeitstheor Verw Geb 53(1980).

[12] F. Grabski, Semi-Markov Processes: Applications in System Reliability and Maintenance,Polish Naval University Gdynia, Poland, 2015

[13] F. Grabski, Concept of Semi-Markov Process, De Gruyter, 2016. D. A. Darling, " The Kolmogorov-Smirnov, Cram r-von mises tests," Ann.Math. Statist., 28 (1957).

[14] A. Gut. Stopped Random Walks. Limit Theorems and Applications, volume 5 ofApplied Probability. A Series of the Applied Probability Trust. Springer, New York, 1988.

[15] V.S. Korolyuk, A.F. Turbin, Semi-Markov Processes and Their Applications, Naukova Dumka, Kiev, 1976 (in Russian)

[16] V.S. Korolyuk, A.F. Turbin, Decompositions of Large Scale Systems, Kluwer Academic, Singapore, 1993.

[17] S.W. Lagakos, Sommer CJ, Zelen M Semi-Markov models for partially censored data. Biometrika 65 (1978).

[18] P. Levy, Proceesus semi-Markoviens, in: Proc. Int. Cong. Math., Amsterdam, 1954, pp.

[19] N. Limnios and Oprisan, G, 2001, Semi-Markov Processes and Reliability (Boston: Birkhäuser).

[20] A. Listwon and P. Saint-Pierre, SemiMarkov: An R Package for Parametric Estimation in MultiState Semi-Markov Models, Journal of Statistical Software, 2015.

[21] E.H. Moore, R. Pyke, Estimation of the transition distributions of a Markov renewal process, Ann. Inst. Stat. Math. 20 (1968).

[22] B. Ouhbi and N. Limnios. Comportement asymptotique de la matrice de renouvellement markovien. . R. Acad. Sci. Paris, 1997.

[23] B. Ouhbi and N. Limnios. Nonparametric estimation for semi-Markov processes based on its hazard rate functions. Stat. Inference Stoch. Process., 2(2), 1999.

[24] B. Ouhbi and N. Limnios, The rate of occurence of failures for semi-Markov processes and estimation, Elsevier Science B.V, 2002.

[25] R. Pyke, Markov renewal processes: definitions and preliminary properties, Ann. Math. Stat. 32 (1961).

[26] R. Pyke, Markov renewal processes with finitely many states, Ann. Math. Stat. 32 (1961).

[27] R. Pyke, R. Schaufele, "Limit theorems for Markov renewal processes," Ann. Math. Stat. 35 (1964).

[28] S. M. Ross, Introduction to Probability Models, Elsevier, 9th edition, 2007.

[29] V.M. Shurenkov. On the theory of Markov renewal. Theory Probab Appl, 1984.

[30] W.L. Smith, Regenerative stochastic processes, Proc. R. Soc. Lond. A 232 (1955).

[31] L. Takács, Some investigations concerning recurrent stochastic processes of a certain type, Magyar Tud. Akad. Mat. Kutato Int. Kzl. 3 (1954).

[32] L. Takács, On a sojourn time problem in the theory of stochastic processes, Trans. Am. Math. Soc. 93 (1959).

[33] The Coronavirus data. Available at https://www.coronavirus-statistiques.com.

[34] I. Votsi, N. Limnios, George Tsaklidis and Eleftheria Papadimitiou, Estimation of the Expected Number of Earthquake Occurences Based on Semi-Markov Models, Springer, 2011.