University of Cyprus

DEPARTMENT OF MATHEMATICS AND STATISTICS

STATISTICAL INFERENCE FOR MULTI-STATE RELIABILITY SYSTEMS

DOCTOR OF PHILOSOPHY DISSERTATION

Andreas Makrides

2016

Andreas Makrides

University of Cyprus

DEPARTMENT OF MATHEMATICS AND STATISTICS

STATISTICAL INFERENCE FOR MULTI-STATE RELIABILITY SYSTEMS

Andreas Makrides

A Dissertation Submitted to the University of Cyprus in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

DECEMBER 2016

Andreas Makrides

Andreas Makrides

© Copyright by Andreas Makrides, 2016

VALIDATION PAGE

Doctoral Candidate: Andreas Makrides

Doctoral Thesis Title: Statistical Inference for Multi-State Reliability Systems

The present Doctoral Dissertation was submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the ***Department of Mathematics and Statistics*** *and was approved on December 16, 2016 by the members of the* ***Examination Committee.***

Examination Committee:

Research Co-Supervisor: Anastassia Baxevani, Assistant Professor,
Department of Mathematics and Statistics, University of Cyprus ____________________

Committee Member: Tasos Christofides, Professor,
Department of Mathematics and Statistics, University of Cyprus ____________________

Research Co-Supervisor: Alexandros Karagrigoriou, Associate Professor retired,
Department of Mathematics and Statistics, University of Cyprus ____________________

Committee Member: Sergios Agapiou, Lecturer,
Department of Mathematics and Statistics, University of Cyprus ____________________

Committee Member: George Tsaklidis, Professor
Department of Mathematics, Aristotle University of Thessaloniki ____________________

Committee Member: Vlad Stefan Barbu, Associate Professor,
Université de Rouen, Laboratoire de Mathématiques ____________________

Andreas Makrides

DECLARATION OF DOCTORAL CANDIDATE

The present doctoral dissertation was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of the University of Cyprus. It is a product of original work of my own, unless otherwise mentioned through references, notes, or any other statements.

Andreas Makrides

Andreas Makrides

To my family

Andreas Makrides

Abstract

A stochastic process is a sequence of random variables defined on a basic probability space $(\Omega, \mathcal{F}, P)$, indexed by a parameter t, where t varies over an index set M. In this work, we deal with stochastic processes where t represents the time up to an event (failure, repair, etc). Consider a process defined on a probability space $(\Omega, \mathcal{F}, P)$ with state space $E = \{1, 2, \ldots, N\}$. Markov processes represent typical tools for modelling such a system.

In this work we focus on multi state systems that we model by means of continuous time Markov processes, which generalize typical Markov jump processes by allowing general distributions (not necessarily Exponential) for sojourn times or residing times or failure times. For this reason, the semi-Markov processes are more adapted to reliability studies (and for applications in general).

The main quantity of interest in such settings is the transition probability of moving from state i to state j. Observe that in the above setting, the time t represents the residing time or sojourn time in state i before moving to state j. Although in the literature it is frequently assumed that this time is exponentially distributed, other more general distributions with heavier tails could be considered. Very frequently, in reliability, in economics, in physics and in engineering, the interest lies in the occurrence of rather exceptional or rare events (natural disasters, total power supply failures, global economic crises, etc) which are associated with the tail part of the distribution. All rare events and the rate at which they occur are related to the shape and the heaviness of the tail of the generating mechanism that produces such events. Since failures may be considered as *rare events*, distributions with heavier tails may be more appropriate for the description of sojourn times. In such cases, appropriate models should be used and the relevant parameter estimators should be determined and analyzed. Then, the problem of transition

Andreas Makrides

probabilities will be addressed together with that of the associated transition rates.

In the first part of this thesis we investigate parameter estimation, transition rates (instantaneous transition probabilities) and transition probabilities using various distributions like Exponential as well as other, heavier tail distributions like Weibull, Pareto, etc for sojourn times. The estimating technique used in this part is the standard method of moments. Note that the distributions considered belong to a general class of distributions. An application of the proposed methodology is presented for illustrative purposes. The application deals with a data set of 113 great earthquakes from the South America region covering the period 1899-2010 with the purpose of making earthquake forecasts.

In the second part of this thesis, we formulate the MLE methodology and provide the associated estimators for MSS reliability indices. The consistency of the estimators is also provided. Two statistical settings will be considered: in the first one we dispose of one sample path of the system; in the second one several sample paths are available. On each situation we take into account two different cases: in the first case, we observe all the sojourn times; in the second one, the sojourn time in the last visited state can be right censored (lost to follow-up, for instance).

The thesis ends with concluding remarks and open problems for future work.

Andreas Makrides

Περίληψη

Μια στοχαστική διαδικασία είναι μια ακολουθία ορισμένη σε ένα χώρο πιθανοτήτων $(\Omega, \mathcal{F}, P)$, που αντιπροσωπεύεται από μια παράμετρο t, όπου το t παίρνει τιμές σε ένα σύνολο M. Στην παρούσα διατριβή ασχολούμαστε με στοχαστικές διαδικασίες όπου το t αντιπροσωπεύει το χρόνο μέχρι να συμβεί κάποιο γεγονός (αποτυχία, επισκευή, κ.λ.π). Έστω μια στοχαστική διαδικασία ορισμένη σε ένα χώρο πιθανοτήτων $(\Omega, \mathcal{F}, P)$ με χώρο καταστάσεων $E = \{1, 2, \ldots, N\}$. Οι Μαρκοβιανές διαδικασίες αποτελούν τυπικά εργαλεία για την μοντελοποίηση ενός τέτοιου συστήματος.

Η παρούσα δαιτριβή θα επικεντρωθεί σε multi state systems (MSS) τα οποία μοντελοποιούμε με τη βοήθεια των ημιμαρκοβιανών διαδικασιών, οι οποίες γενικεύουν τυπικές Μαρκοβιανές διαδικασίες αλμάτων επιτρέποντας γενικές κατανομές (όχι απαραίτητα Εκθετική) για τους χρόνους παραμονής. Για το λόγο αυτό, οι ημιμαρκοβιανές διαδικασίες είναι καταλληλότερες σε μελέτες αξιοπιστίας (και για εφαρμογές γενικότερα).

Το κύριο αντικείμενο μελέτης στην περίπτωση αυτή είναι η πιθανότητα μετάβασης από την κατάσταση i στην κατάσταση j. Σημειώνεται ότι στην πιο πάνω περίπτωση, ο χρόνος t αντιπροσωπεύει το χρόνο παραμονής στην κατάσταση i πριν από τη μετάβαση στην κατάσταση j. Παρόλο που στη βιβλιογραφία συχνά ο χρόνος αυτός θεωρείται ότι ακολουθεί την Εκθετική κατανομή, εντούτοις μπορούν να θεωρηθούν κι άλλες πιο γενικές κατανομές με πιο βαριές ουρές. Πολύ συχνά, στη θεωρία αξιοπιστίας, στην οικονομία, στη φυσική αλλά και στη μηχανική, το ενδιαφέρον έγκειται στην εμφάνιση ή όχι εξαιρετικών ή σπάνιων γεγονότων (φυσικές καταστροφές, ολική διακοπή παροχής ηλεκτρικού ρεύματος, παγκόσμια οικονομική κρίση κ.λ.π.). Όλα τα σπάνια γεγονότα αλλά και ο ρυθμός με τον οποίο συμβαίνουν συνδέονται τόσο με τη μορφή όσο και με το πόσο βαριά είναι η ουρά της κατανομής. Επειδή οι αποτυχίες μπορούν να θεωρηθούν ως σπάνια γεγονότα, οι κατανομές με πιο βαριές ουρές

Andreas Makrides

μπορούν να θεωρηθούν ως καταλληλότερες για την περιγραφή των χρόνων παραμονής. Σε μια τέτοια περίπτωση, κατάλληλα μοντέλα θα πρέπει να χρησιμοποιηθούν, καθώς επίσης και οι αντίστοιχοι εκτιμητές των παραμέτρων θα πρέπει να προσδιοριστούν και να αναλυθούν. Στη συνέχεια, το πρόβλημα των πιθανοτήτων μετάβασης θα πρέπει να αντιμετωπιστεί παράλληλα με το πρόβλημα των αντίστοιχων ρυθμών μετάβασης.

Στο πρώτο μέρος της διατριβής αυτής, διερευνούμε εκτίμηση παραμέτρων, ρυθμούς μετάβασης (στιγμιαίες πιθανότητες μετάβασης) και πιθανότητες μετάβασης χρησιμοποιώντας διάφορες κατανομές για τους χρόνους παραμονής όπως είναι η Εκθετική, καθώς και άλλες κατανομές με πιο βαριές ουρές όπως η Weibull, η Pareto, κ.λ.π. Η μέθοδος εκτίμησης που χρησιμοποιείται σε αυτή την περίπτωση είναι η μέθοδος των ροπών. Σημειώνεται ότι οι κατανομές που έχουν θεωρηθεί ανήκουν σε μια γενική κλάση κατανομών. Εν συνεχεία, παρουσιάζεται μια εφαρμογή σε δεδομένα σεισμών.

Στο δεύτερο μέρος της διατριβής αυτής, εφαρμόζουμε τη μέθοδο μέγιστης πιθανοφάνειας και παραθέτουμε τους αντίστοιχους εκτιμητές των δεικτών αξιοπιστίας ενός MSS, καθώς επίσης και η συνέπεια των εκτιμητών αυτών. Για το σκοπό αυτό εξετάζονται δύο περιπτώσεις: στην πρώτη περίπτωση θεωρούμε ένα μόνο δείγμα διαδρομής του συστήματος ενώ στη δεύτερη περίπτωση υπάρχουν διαθέσιμες πολλές διαδρομές. Σε κάθε περίπτωση λαμβάνουμε υπόψη δύο περιπτώσεις: αρχικά ότι όλοι οι χρόνοι παραμονής είναι διαθέσιμοι, και στη δεύτερη περίπτωση, ότι ο χρόνος παραμονής στην τελευταία κατάσταση είναι λογοκριμένος (για παράδειγμα έχει χαθεί από την παρακολούθηση).

Andreas Makrides

Acknowledgments

First I would like to express my deepest gratitude to my supervisor, Dr. Alex Karagrigoriou for his great support during my PhD studies and always being very positive and helpful. It is certain that without his valuable help I could not manage to complete this work.

I am also thankful to Dr. Anatoly Lisnianski and Dr. Ilia Frenkel who presented to me the problem of modeling the multi state systems by means of continuous time Markov processes.

I owe a big thanks to Dr. Vlad Barbu for his welcome and for his hospitality at the University of Rouen. His guidance was the catalyst in completion of this thesis. I am also thankful to Dr. Anastassia Baxevani, Dr. Tasos Christofides, Dr. Sergios Agapiou and Dr. George Tsaklidis for their kind advice and useful suggestions.

I am grateful to Dr. Stavros Vakeroudis and Dr. Irene Votsi for their constructive comments which improve the quality of this work.

I finally would like also to thank Dr. Theodoros Tsapanos for his constructive collaboration in the illustrative part of my thesis and for all the information concerning Geosciences.

Andreas Makrides

Contents

Andreas Makrides

Andreas Makrides

Andreas Makrides

List of Figures

Andreas Makrides

List of Tables

Andreas Makrides

Chapter 1

Introduction

Andreas Makrides

All systems are designed to perform their intended tasks in a given environment. Some systems can perform their tasks with various distinctive levels of efficiency usually referred to as performance rates. A system that can have a finite number of performance rates is called a multi-state system (MSS). Usually a MSS is composed of elements that in their turn can be multi-state. A binary system is the simplest case of a MSS having two distinctive states, namely *perfect functioning* and *complete failure*. An inherent weakness of traditional reliability theory is that the system was considered as a binary one and its components are always described as being in either of the above two distinctive states. Early attempts to replace this by a theory of multi-state systems with multi-state components were made in the late 70s. The basic concepts of MSS reliability were introduced by Murchland (1975), El-Neveihi et al. (1978), Barlow and Wu (1978), Ross (1979) and Aven and Jensen (1999). Extensions and generalizations of the above results were obtained by Natvig (1982), Block and Savits (1982) and Hudson and Kapur (1982). Since that time MSS reliability began intensive development. Essential achievements that were attained up to the mid 1980's are reflected in Natvig (1985) and in El-Neveihi and Prochan (1984) where one can find the state of the art in the field of MSS reliability at this stage.

For the history of ideas in MSS reliability theory in recent years one is referred to Sahner et al. (1996), Lisnianski and Levitin (2003), Lisnianski et al. (2010) and Natvig (2011). The simplest examples of such a situation that are briefly discussed below, are the well-known 1-out-of-n, k-out-of-n, $(n-1)$-out-of-n and n-out-of-n systems.

1.1 Basic Concepts

1.1.1 Simple Multi-state Systems

A typical engineering system consists of n units. A binary system is the simplest case of a MSS. At any given moment of time t each unit of a binary system can be in one of two distinctive states, namely ”failure state” (”*off* mode” or ”complete failure” or ”total failure”) and ”functioning state” (”*on* mode” or ”perfect functioning” or ”nominal performance”).

For describing the state of the i^{th} binary unit, $i = 1, 2, \ldots, n$, we often use an indicator function x_i taking the values 0 or 1 depending on whether the unit i is in the ”failure state” or the ”functioning state”. The vector of states for all n units of the system is denoted by $x = (x_1, x_2, \ldots, x_n)$.

The state of the entire system depends on the state of the units and it can also be in one of two distinctive states: the failure state or the functioning state. For describing the state of the system we use the indicator function $\phi(x)$ that takes on two values, 0 and 1.

Definition 1.1.1 *Consider a binary system consisting of* n *units. The function* $\phi : \{0,1\}^n \to \{0,1\}$ *which for each vector* x *of states for all units of the system, describes the state* $\phi(x)$ *of the system, is called the structure function of the system.*

Example 1.1.1 *SS - Serial System.*

A serial system consisting of n units fails when at least one of the n units fail, or equivalently, it functions when all n units function. A typical serial system can be represented as follows:

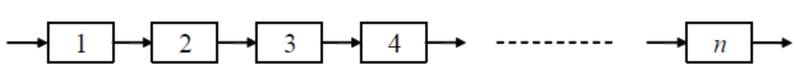

Figure 1.1: Serial system.

Andreas Makrides

Fig. 1.1 shows that for a "signal" to move from left to right, it will have to go through all n units. Therefore, the "signal" passes through and the system functions (i.e. $\phi = 1$) when all n units function so that $x_1 = x_2 = \cdots = x_n = 1$. On the other hand, if even a single unit fails i.e. $x_i = 0$ for some i, $i = 1, 2, \ldots, n$ then the system fails and $\phi = 0$. Thus, the structure function of the serial system can be written as:

$$\phi(x) = \min\{x_1, x_2, \ldots, x_n\} = \prod_{i=1}^{n} x_i. \tag{1.1}$$

Obviously the above function is equal to 1 if and only if all x_i's are equal to 1.

Example 1.1.2 *PS - Parallel system.*

A parallel system consisting of n units fails when all n units fail, or equivalently it functions when at least one unit functions. The graphical representation of a parallel system is given below:

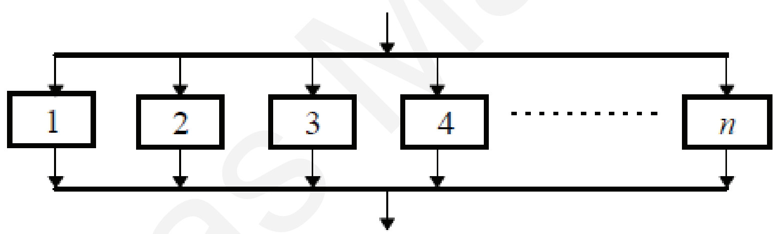

Figure 1.2: Parallel system.

A "signal" passes from top to bottom if at least one of the n units functions. Therefore $\phi = 1$ if at least one of the x_i's is equal to 1. Thus, the structure function of the parallel system can be written as

$$\phi(x) = \max\{x_1, x_2, \ldots, x_n\} = 1 - \prod_{i=1}^{n}(1 - x_i). \tag{1.2}$$

Example 1.1.3 $S(n, k) : G$, *k-out-of-n:Good.*

An engineering system consisting of n components is said to be a k-out-of-n system if the system functions if and only if at least k-out-of-n components function with $1 \leq k \leq n$. Suppose that all components function independently of each other. Hence, the structure

Andreas Makrides

function of this system is

$$\phi(x) = \begin{cases} 1, & \sum_{i=1}^{n} x_i \geq k, \\ 0, & \sum_{i=1}^{n} x_i < k, \end{cases} \tag{1.3}$$

which is equivalent to the function

$$\phi(x) = x_{(n-k+1)},$$

where $x_{(n-k+1)}$ is the $(n-k+1)^{th}$ order statistic.
Note that the special case $k = n$ corresponds to the serial system, whereas $k = 1$ corresponds to the parallel system.
Another interesting case of practical importance is the $(n-1)$-out-of-n system for which $k = n - 1$. According to this system, which sometimes is referred to as "fail-safe" structure, failure of a single component is not sufficient to cause system failure, but the failure of two components does cause system failure (Barlow and Proschan, 1965, p. 218).

MSS behavior is characterized by its evolution in the space of states. The entire set of possible system states can be divided into two disjoint subsets corresponding to acceptable and unacceptable system functioning. The system entrance into the subset of unacceptable states constitutes a failure. MSS reliability can be defined as the system's ability to remain in acceptable states during the operation period or alternatively the system's ability to operate without failure for a specified period of time.

1.1.2 MSS Examples

Example 1.1.4 *Power Supply Unit.*

In a power supply system, each generating unit can function at different levels of capacity. Generating units are complex assemblies of many parts. The failures of different parts may lead to situations in which the generating unit continues to operate, but at a reduced capacity. For example, Billinton and Allan (1996) describe a three-state 50 MW generating unit. The performance rates (generating capacity) corresponding to these states and

Andreas Makrides

probabilities of the states are presented in Table 1.1.

State	Generating Capacity (MW)	Probability
1	50	0.960
2	30	0.033
3	0	0.007

Table 1.1: Capacity distribution of a power supply unit

Example 1.1.5 *Refrigeration System.*

The most commonly used refrigeration system for supermarkets today is the multiplex direct expansion system (Baxter, 2002). All display cases and cold storerooms use direct-expansion air-refrigerant coils that are connected to the system compressors in a remote machine room located at the back or on the roof of the store. Heat rejection is usually done with aircooled condensers with simultaneously working axial blowers mounted outside.

An example of such a refrigeration system consists of 2 subsystems:(a) the 5-level/state compressor and (b) the condenser with 3-level/state axial blowers (see Figure 1.3). The performance of the elements is measured by their produced cold capacity (BTU per year).

The compressor can be in one of five states: a state of total failure corresponding to capacity 0, states of partial failures corresponding to capacities of $3 \cdot 10^9$, $6 \cdot 10^9$, $9 \cdot 10^9$ BTU per year and a fully operational state with a capacity of $12 \cdot 10^9$ BTU per year. The blowers can be in one of three states: a state of total failure corresponding to a capacity of 0, state of partial failure corresponding to capacity $6 \cdot 10^9$ of BTU per year and a fully operational state with a capacity of $12 \cdot 10^9$ BTU per year.

Example 1.1.6 *Survival Analysis.*

Recently, multi-state models have been implemented in (bio)medicine (Giard et al., 2002; Van den Hout and Matthews, 2008; Marshall and Jones, 2007), etc. For instance, in Van den Hout and Matthews (2008) a cognitive ability during old age is considered. An illness-death model is presented for describing the progression of an illness. The model

Andreas Makrides

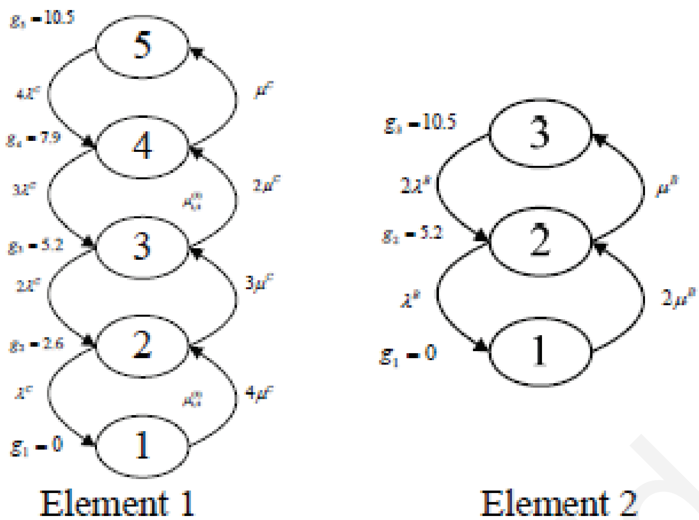

Figure 1.3: State-space diagram of the multi-state elements.

considers three states: the health state, an illness state, and the death state and it is used to derive the probability of transition from one state to another within a specified time interval.

To numerically characterize MSS behavior from a reliability point of view, one has to determine the MSS reliability indices. The time to failure T or TTF, the time between failures TBF, and the number of failures NOF are typical variables of interest. The reliability function usually denoted by $R(\cdot)$ and the meantime to failure (MTTF), namely the mean time up to the instant when the system enters the subset of unacceptable states (e.g. the failure state in a binary system) for the first time are standard reliability indices. The expected number of failures and the probability that the number of failures does not exceed a fixed value are the reliability indices associated with the number of failures. The definitions for all these concepts are given in Section 2.1.

The MSS could be investigated and analyzed with the use of Markov processes. Markov processes are widely used for reliability analysis because the number of failures in arbitrary time intervals in many practical cases can be described as a Poisson process and time T or TTF up to the failure and repair (and maintenance) time TTR are often Exponentially distributed.

Andreas Makrides

1.2 Aims of the Thesis

In general a stochastic process is a sequence of random variables defined on a basic probability space $(\Omega, \mathcal{F}, P)$, indexed by a parameter say t continuous, where t varies over an index set M. In this work, we deal with stochastic processes where t represents the time up to an event (failure, repair, etc). Consider a process defined on a probability space $(\Omega, \mathcal{F}, P)$ with state space $E = \{1, 2, \ldots, N\}$. For example, state "1" is associated with nominal performance of the system and state "N" is associated with total failure. Markov processes represent typical tools for modelling such a system.

In this work we focus on multi state systems that we model by means of semi-Markov processes, which generalize typical Markov jump processes by allowing general distributions (not necessarily Exponential) for sojourn times or residing times or failure times (Limnios and Oprişan (2001)). For this reason, the semi-Markov processes are more adapted to reliability studies (and for applications in general). Chapter 2 is devoted to the proposed methodology. For semi-Markov processes in discrete time with application in reliability one may refer, among others, to Barbu et al. (2004), Barbu and Limnios (2006), Chryssaphinou et al. (2010) and McClean et al. (2004).

The main quantity of interest in such settings is the transition probability of moving from state i to state j. Observe that in the above setting, the time t represents the residing time or sojourn time in state i before moving to state j. Although it is natural to expect that this time is Exponentially distributed, other more general distributions with heavier tails could be considered. Very frequently, in reliability, in economics, in physics and in engineering, the interest lies in the occurrence of rather exceptional or rare events (natural disasters, total power supply failures, global economic crises, etc) which are associated with the tail part of the distribution. All rare events and the rate at which they occur are related to the shape and the heaviness of the tail of the generating mechanism that produces such events. Since failures may be considered as *rare events*, distributions with heavier tails may be more appropriate for the description of sojourn times. In such cases, appropriate models should be used and the relevant parameter estimators should be determined and analyzed. Then, the problem of transition probabilities will be addressed together with that of the associated transition rates. A non parametric approach for the analysis of the semi-Markov processes can be found in Votsi (PhD, 2013).

In the first part of this work, in Chapter 4, we investigate

(a) Parameter estimation

(b) Transition rates (instantaneous transition probabilities) and

(c) Transition probabilities

using various distributions like Exponential as well as other, heavier tail distributions like Weibull, Pareto, etc. for sojourn times. The estimating technique used in this part is the standard method of moments. Note that the distributions considered belong to a general class of distributions discussed in Chapter 3. An application of the proposed methodology is presented in Chapter 6. The application deals with a data set of 113 great earthquakes from the South America region covering the period 1899-2010 with the purpose of making earthquake forecasts. Seismic hazard assessment from a parametric as well as a non-parametric point of view can be found, among others, in Alvarez (2005), Masala (2012) and Votsi et al. (2014).

In the second part of this work in Chapter 5, we formulate the MLE methodology and provide the associated estimators for MSS reliability indices. The consistency of the estimators is provided in Subsection 5.6. Two statistical settings will be considered: in the first one, presented in Subsection 5.1, we dispose of one sample path of the system; in the second one, described in Subsection 5.2, several sample paths are available. On each situation we take into account two different cases: in the first case, we observe all the sojourn times; in the second one, the sojourn time in the last visited state can be right censored (lost to follow-up, for instance).

The originality of the work lies on new statistical approaches like the use of a general class of distributions for the sojourn times (closed under minima) for making inferences for semi-Markov processes, the application of the framework of multistate systems to earth sciences for the modeling of earthquake occurrence and the use of mean transition rates for predictive purposes. These approaches result in a novel methodology for the analysis of the class of semi-Markov processes with the distribution of sojourn times belonging to the class of generalized Gamma distributions, in a novel setting that allows for self-transitions

Andreas Makrides

of the embedded Markov chain, in a single step/move and finally in good predictive ability since time-homogeneity allows for predictions to be based on as fresh knowledge as possible by choosing as the time origin, the time instant of the most recent available event.

Andreas Makrides

Chapter 2

Multi state system methodology and semi-Markov processes

2.1 Semi-Markov Processes

Consider a discrete-state continuous-time Markov process. We assume that the random system has finite state space with N states: $E = \{1, \ldots, N\}$, $N < \infty$. Assume that its time evolution is governed by a stochastic process $Z = (Z_t)_{t\in\mathbb{R}_+}$. Let us denote by $S = (S_n)_{n\in\mathbb{N}}$ the successive time points when state changes in $(Z_t)_{t\in\mathbb{R}_+}$ occur and by $J = (J_n)_{n\in\mathbb{N}}$ the successive visited states at these time points. Set also $X = (X_n)_{n\in\mathbb{N}}$ for the successive sojourn times in the visited states. Thus,

$$X_n = S_n - S_{n-1},\ n \in \mathbb{N}^*,$$

and, by convention, we set $X_0 = S_0 = 0$.

Let us recall now the definition of a Markov renewal and semi-Markov process (Limnios and Oprişan, 2001). If $(J, S) = (J_n, S_n)_{n\in\mathbb{N}}$ satisfies the relation

$$\begin{aligned}&\mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_0, \ldots, J_n; S_1, \ldots, S_n)\\&= \mathbb{P}(J_{n+1} = j, S_{n+1} - S_n \leq t | J_n), j \in E, t \in \mathbb{R}_+,\end{aligned}$$

Andreas Makrides

then

- (J, S) is called a *Markov renewal process* (MRP) ;
- $Z = (Z_t)_{t\in\mathbb{R}_+}$ is called a *semi-Markov process* (SMP) associated to (J, S), where

  $$Z_t := J_{N(t)} \quad \Leftrightarrow \quad J_n = Z_{S_n},$$

  with

  $$N(t) := \max\{n \in \mathbb{N} \mid S_n \le t\},\ t \in \mathbb{R}_+, \tag{2.1}$$

  is the counting process of the number of jumps in the time interval $(0, t]$. Thus, Z_t gives the state of the system at time t.

If $(J_n, S_n)_{n\in\mathbb{N}}$ is a MRP, it can be immediately checked that $(J_n)_{n\in\mathbb{N}}$ is a Markov chain, called the *embedded Markov chain*.

Throughout this work we assume that the SMP (or equivalently, the MRP) is regular, that is

$$\mathbb{P}_i(N(t) < \infty) = 1, \text{ for all } \ t > 0 \ \text{ and } \ i \in E,$$

where $\mathbb{P}_i(\cdot)$ is the probability that the process started at state $i \in E$ at initial time $t = 0$. Under this condition,

$$S_n < S_{n+1}, n \in \mathbb{N},$$

$$S_n \xrightarrow[n\to\infty]{a.s.} \infty,$$

$$N(t) \xrightarrow[t\to\infty]{a.s.} \infty.$$

We also assume that the SMP (or equivalently, the MRP) is irreducible and positive-recurrent (see Limnios and Oprisan, 2001, Fox, 1968, Kovalenko et al., 1997 or Levy, 1954).

A particular case of the semi-Markov process is the continuous time Markov jump process which is defined bellow.

Definition 2.1.1 *The stochastic process* $(X_t)_{t\in\mathbb{R}_+}$ *is called a* continuous time Markov process *if for all* $i, j, i_1, \ldots, i_k \in E$*, all* $t, s \ge 0$*, and all* $s_1, \ldots, s_k \ge 0$ *with* $s_l \le s$ *for all*

Andreas Makrides

$l \in [1, k]$,

$$\mathbb{P}(X_{t+s} = j \mid X_t = i,\ X_{s_1} = i_1, \ldots, X_{s_k} = i_k) = \mathbb{P}(X_{t+s} = j \mid X_t = i).$$

For the sake of completeness we close this Section by providing in terms of the stochastic process Z_t, defined above, the definitions of the main reliability measures which will be fully discussed and investigated at Section 5.5.

Let us assume that the state space E is divided into two subsets, U (containing the functioning states, the so-called up-states) and D (containing the failure states, the so-called down-states), such that $E = U \cup D$ and $U \cap D = \varnothing$, where we assume that $U = \{1, \ldots, n\}$ and $D = \{n+1, \ldots, N\}$.

Definition 2.1.2 *The reliability or survival function of the system at time t, $R(t)$, is defined as the probability of being in acceptable states for $s \leq t$, i.e.,*

$$R(t) = \mathbb{P}(T_D > t) = \mathbb{P}(Z_s \in U, s \leq t),$$

where $T_D := \inf\{t \mid Z_t \in D\}$ is the lifetime of the system.

Definition 2.1.3 *The pointwise or instantaneous availability of the system, $A(t)$, is defined as the probability of being in acceptable state at time instant t,i.e.,*

$$A(t) = \mathbb{P}(Z_t \in U).$$

Definition 2.1.4 *The maintainability of the system, $M(t)$, is the probability that the system is repaired up to time t, given that it has failed at time $t = 0$, i.e.,*

$$M(t) = \mathbb{P}(T_U \leq t) = 1 - \mathbb{P}(Z_s \in D, s \leq t),$$

Andreas Makrides

where $T_U := \inf\{t \mid Z_t \in U\}$ *is the duration of repair.*

Definition 2.1.5 *The mean time to failure (MTTF) is defined as the mean lifetime, i.e., the expectation of the hitting time to down set D,*

$$MTTF := \mathbb{E}\,(T_D)\,.$$

2.2 Semi-Markov characteristics

A semi-Markov model is characterized by its *initial distribution* $\alpha = (\alpha_1, \ldots, \alpha_N)$

$$\alpha_j := \mathbb{P}(J_0 = j),\ j \in E,$$

and by the *semi-Markov kernel*

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

Let us also introduce the *transition probabilities* of the embedded Markov chain $(J_n)_{n\in\mathbb{N}}$,

$$p_{ij} := \mathbb{P}(J_n = j | J_{n-1} = i) = \lim_{t\to\infty} Q_{ij}(t), \tag{2.2}$$

which can be viewed as the fraction of time the system stays in state i before moving to j and the *conditional sojourn time distribution functions*

$$\begin{aligned} W_{ij}(t) &:= \mathbb{P}(S_n - S_{n-1} \leq t | J_{n-1} = i, J_n = j) \\ &= \mathbb{P}(X_n \leq t | J_{n-1} = i, J_n = j). \end{aligned}$$

Andreas Makrides

Observe that

$$Q_{ij}(t) = p_{ij} W_{ij}(t).$$

For details the interested readers may see Moore and Pyke (1968).

2.3 The problem setting - Notations

In discrete-state continuous-time Markov process with N possible states the quantity of interest is the transition probability p_{ij} of an event *moving* from state i, $i = 1, 2, \ldots, N$ to state j, $j = 1, \ldots, i, \ldots, N$. Since there is often a distinction between technical and nontechnical systems, for practical purposes the following clarification is necessary:

In technical/reliability systems, starting from state i, the system is not allowed (not possible) to return to the same state in a single step/move so that in such cases the system jumps to state j with $j \in \{1, \ldots, N\} - \{i\}$). In non technical systems the system is allowed to return to the same state in a single step/move so that in these cases the system jumps to j with $j \in \{1, \ldots, i, \ldots, N\}$.
Note that the double subscript $\{ij\}$ represents a transition from state i to state j with j taking the appropriate values according to the situation under investigation.

In order to be precise with our notation, let us denote by T_{ij} the time elapsed between two consecutive events the first of which is the time when we arrive at/hit the state i and the second one the time arriving at state j. This time-period can be viewed as the time the process spends at state i before "moving" (directly) to state j. The time is also called *sojourn time* or *residing time*.

Let $F_{ij}(t)$ be the cumulative distribution function (cdf) of sojourn (residing) time T_{ij} in state i when the unit transits from state i to state j and it is assumed to be absolutely continuous with respect to the Lebesgue measure. Thus the associated probability distribution function (pdf) is well defined and it is denoted by $f_{ij}(t)$. Let also $f_i(t)$ be the pdf of the time till the first transit from state i to any state including the state i itself. In

Andreas Makrides

terms of $F_{ij}(t)$ and $f_{ij}(t)$ the reliability function $R_{ij}(t)$ is defined by

$$R_{ij}(t) = 1 - F_{ij}(t)$$

while the intensity rate, $\lambda_{ij}(t)$, from state i to state j in reference to the non-homogeneity is defined by

$$\lambda_{ij}(t) = \frac{f_{ij}(t)}{R_{ij}(t)}. \tag{2.3}$$

Finally, the MTTF is given by

$$MTTF = \int_0^\infty x f_{ij}(x)dx = \int_0^\infty R_{ij}(x)dx.$$

Recall that the intensity (transition) rate given in (2.3) is constant and therefore independent of time if the underlying distribution is Exponential. In all other cases, like the Weibull, Pareto etc. it is a polynomial function of time.

2.4 Variable transition rates

Note that the case of constant intensity rates (i.e. Exponential distribution), is associated with time-homogeneity. Hence, the origin $t = 0$ of the time coordinate can be located at any instant, and the equations remain valid for any time t so that

$$\lambda_{ij}(t) = \lambda_{ij}, \forall\ t.$$

On the other hand, if a transition rate is non-constant function of time, the system equation will be valid only for one particular choice of the origin $t = 0$, which must be the time instant when all the components of a system are new, so that the time parameter t can be identified with the age of the components. We may disallow the use of the time parameter itself as a state variable, but it is still possible to contrive homogeneous Markov models that simulate systems with variable transition rates.

Consider a continuous time Markov process where the failure rate is an arbitrary function

Andreas Makrides

$\lambda(t)$ of time. Now, we cannot simply represent this model as explained previously, because state zero will contain a mixture of systems of different ages, depending on how recently each system exited and returned to this state. The single binary state variable (healthy or functioning state and failure state) is not sufficient to fully determine the causal state of the system. To remedy this, suppose we split up the "healthy" or "functioning" state into a sequence of several states, and allow the system to transition from one to the next at rates that ensure uniform progression. From any one of these "healthy" or "functioning" states, the system can fail and be replaced or refurbished immediately back to the initial healthy state (see Figure 2.1).

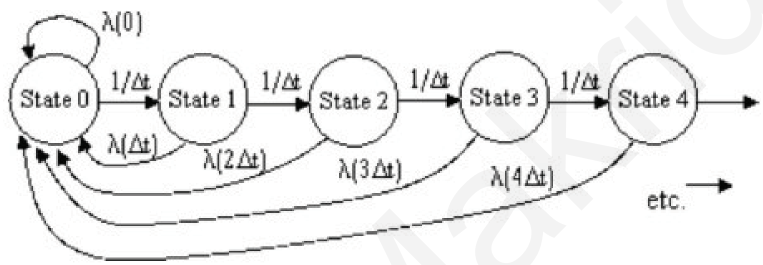

Figure 2.1: Healthy state splitting.

We have assigned a rate of $1/\Delta t$ to each transition rate from one state to the next, so that the mean time from one state to the next is Δt. Then we assign a constant failure rate of $\lambda(n\Delta t)$ to the n^{th} state in this sequence. This arrangement simulates the variation in the failure rate as a function of the elapsed time since the system was most recently in state 0. The steady-state equation of the n^{th} state (except for state 0) is

$$\frac{1}{\Delta t} P_{n-1} = \left(\frac{1}{\Delta t} + \lambda(n\Delta t) \right) P_n,$$

where P_n and P_{n-1} denote the probability of being in state n or $n-1$ respectively. Thus we can define the ratios q_n of the probabilities of consecutive state

$$q_n = \frac{P_n}{P_{n-1}} = \frac{1}{1 + \Delta t \, \lambda(n\Delta t)}.$$

Consistent with this, we will define $q_0 = 1$. Now we can express the individual probabilities as

$$P_0 = q_0 P_0 \quad P_1 = q_0 q_1 P_0 \quad P_2 = q_0 q_1 q_2 P_0 \quad P_3 = q_0 q_1 q_2 q_3 P_0 \quad P_4 = q_0 q_1 q_2 q_3 q_4 P_0$$

and so on. The sum of all the probabilities is 1, so we have

$$P_0 = \frac{1}{q_0 + q_0 q_1 + q_0 q_1 q_2 + q_0 q_1 q_2 q_3 + q_0 q_1 q_2 q_3 q_4 + \cdots}.$$

The system failure rate is

$$\begin{aligned} \lambda &= \lambda(0) P_0 + \lambda(1\Delta t) P_1 + \lambda(2\Delta t) P_2 + \lambda(3\Delta t) P_3 + \lambda(4\Delta t) P_4 + \cdots \\ &= \left[\lambda(0) q_0 + \lambda(1\Delta t) q_0 q_1 + \lambda(2\Delta t) q_0 q_1 q_2 + \lambda(3\Delta t) q_0 q_1 q_2 q_3 + \cdots\right] P_0. \end{aligned}$$

Combining these expressions, gives the system failure rate entirely in terms of the individual component failure rates for all the "ages" from $t = 0$ to infinity, that is

$$\lambda = \frac{\lambda(0) q_0 + \lambda(1\Delta t) q_0 q_1 + \lambda(2\Delta t) q_0 q_1 q_2 + \lambda(3\Delta t) q_0 q_1 q_2 q_3 + \cdots}{q_0 + q_0 q_1 + q_0 q_1 q_2 + q_0 q_1 q_2 q_3 + q_0 q_1 q_2 q_3 q_4 + \cdots}.$$

Equivalently,

$$\begin{aligned} \lambda &= \frac{\sum\limits_{j=0}^{\infty} \lambda(j\Delta t) e^{\sum\limits_{i=0}^{j} \ln(q_i)}}{\sum\limits_{j=0}^{\infty} e^{\sum\limits_{i=0}^{j} \ln(q_i)}} \\ &= \frac{\sum\limits_{j=0}^{\infty} \lambda(j\Delta t) e^{-\sum\limits_{i=0}^{j} \ln(1+\Delta t \lambda(i\Delta t))}}{\sum\limits_{j=0}^{\infty} e^{-\sum\limits_{i=0}^{j} \ln(1+\Delta t \lambda(i\Delta t))}}. \end{aligned}$$

In view of the series expansion $\ln(1 + x) = x - x^2/2 + x^3/3 - \cdots$ we see that, in the limit as Δt goes to the infinitesimal dt, the natural logarithms go to $dt\lambda(idt)$. To avoid confusion we will let τ denote the time parameter appearing in the inner summation, and t will denote the time parameter in the outer summation. Then we have $\tau = id\tau$ and $t = jdt$, so if we multiply the numerator and denominator by dt, the above equation can

Andreas Makrides

be written in terms of continuous integrations as,

$$\lambda = \frac{\int_0^\infty \lambda(t) \left(e^{-\int_0^t \lambda(\tau)d\tau} \right) dt}{\int_0^\infty \left(e^{-\int_0^t \lambda(\tau)d\tau} \right) dt}.$$

In order for a function $\lambda(\tau)$ to be a rate function, its integral from 0 to t must increase monotonically to infinity as t increases. In that case, the numerator of the above expression (which we recognize as simply the integral of the density function corresponding to the rate λ) equals 1, so we have

$$\lambda = \frac{1}{\int_0^\infty \left(e^{-\int_0^t \lambda(\tau)d\tau} \right) dt}. \tag{2.4}$$

The last expression implies that the time independent mean recurrence (failure) rate is equal to the reciprocal of the average of sojourn times and is given by

$$\begin{aligned} \lambda_{ij} &= \frac{1}{E(T_{ij})} \\ &= \frac{1}{\int_0^\infty \left(e^{-\int_0^t \lambda_{ij}(x)dx} \right) dt} \end{aligned} \tag{2.5}$$

where

$$\begin{aligned} E(T_{ij}) &= \int_0^\infty R_{ij}(x)dx \\ &= \int_0^\infty e^{-\int_0^t \lambda_{ij}(x)dx} dt. \end{aligned} \tag{2.6}$$

Although the above equation is mathematically valid for the case $i = j$, for the estimation of λ_{ii} we use the Markov property that $\sum_{j=1}^{N} \lambda_{ij} = 0, \ \forall i$ so that $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$.

It is noted that all the above functions are assumed to depend on a parameter a_{ij} so that the sojourn times are independent but not necessarily identically distributed (*inid*) random variables.

Andreas Makrides

2.5 The problem setting - Characteristics

Recall that T_{ij} denotes the time spent in state i before moving (directly) to state j. We denote by $F_{ij}(t;\theta_{ij})$ its cumulative distribution function (cdf), where θ_{ij} is the m-dimensional parameter involved in the underlying distribution. We assume that the distribution of T_{ij} is absolutely continuous with respect to the Lebesgue measure; an associated density is denoted by $f_{ij}(t;\theta_{ij})$.

Note that in this section we are under the semi-Markov setting. The dynamic of the system under investigation is as follows: the next state to be visited after state i is the one for which T_{il} is the minimum. This is the way the next state to be visited, say j, is "chosen", namely

$$j = \underset{l \in E}{\operatorname{argmin}}(T_{il}).$$

From the theoretical point of view we state that $T_{il}(\omega)$ is minimum almost surely for any $\omega \in \Omega$.

Assume that the MSS is initially in state i at time instant $t = 0$. Then the probability $Q_{ik}(t)$ that *the first transit* from state i to a fixed (specific) state k before time t is called the transition probability from i to k and is given by

$$Q_{ik}(t) = \Pr\left(T_{ik} = \min\left\{T_{i1}, \ldots, T_{iN}\right\} \leq t\right). \tag{2.7}$$

In other words, $Q_{ik}(t)$ represents the probability that *the first transit* from state i will be in state k before time t conditioning on the fact that the unit is in state i at initial time instant $t = 0$. One can obtain the one-step probability $Q_{ik}(t)$, with k fixed, as the probability that under the condition $T_{ik} \leq t$ the random variable T_{ik} is the minimum of the random variables T_{ij}, $j = 1, \ldots, N$. In other words, if $T_{i(m)}$ represents the m^{th} order statistic, $m = 1, \ldots, N$ of the N random variables $T_{i1}, \ldots, T_{iN}$, then

$$T_{i(1)} = T_{ik}.$$

Andreas Makrides

Thus, for our semi-Markov system, the semi-Markov kernel becomes

$$\begin{aligned} Q_{ij}(t) &= \mathbb{P}(\min_l T_{il} \leq t \text{ \& the min occurs for } j | J_{n-1} = i) \\ &= \mathbb{P}(\min_l T_{il} \leq t, T_{ij} \leq T_{il}, \forall l | J_{n-1} = i) \\ &= \mathbb{P}(\min_l T_{il} \leq t | J_{n-1} = i, J_n = j) \times \mathbb{P}(T_{ij} \leq T_{il}, \forall l | J_{n-1} = i) \\ &= p_{ij} W_i(t), \end{aligned}$$

where

$$\begin{aligned} p_{ij} &= \mathbb{P}(J_n = j | J_{n-1} = i) \\ &= \mathbb{P}(T_{ij} \leq T_{il}, \forall l | J_{n-1} = i) \end{aligned}$$

and

$$\begin{aligned} W_{ij}(t) &= \mathbb{P}(S_n - S_{n-1} \leq t | J_{n-1} = i, J_n = j) \\ &= \mathbb{P}(\min_l T_{il} \leq t | J_{n-1} = i, J_n = j) \\ &= \mathbb{P}(\min_l T_{il} \leq t | J_{n-1} = i) =: W_i(t), \text{ independent of } j, \end{aligned}$$

which represents the cdf of the sojourn time in state i (unconditional to the next state to be visited).

Hence, for each $i = 1, \ldots, N$, using the cdf $F_{ij}(\cdot)$ we have

$$\begin{aligned} Q_{ik}(t) &= \mathbb{P}\left\{T_{ik} \leq t, T_{ik} = T_{i(1)}\right\} \\ &= \mathbb{P}\left\{T_{ik} \leq t, T_{ik} \leq T_{il}, \forall l \neq k\right\} \\ &= \int_0^t \left[1 - F_{i1}(u)\right] \ldots \left[1 - F_{i,k-1}(u)\right] \times \\ &\quad \times \left[1 - F_{i,k+1}(u)\right] ... \left[1 - F_{iN}(u)\right] dF_{ik}(u). \end{aligned} \tag{2.8}$$

Note that

$$\sum_j Q_{ij}(t) = W_i(t) \tag{2.9}$$

Andreas Makrides

which represents the cdf of the time till the first transit from i. In other words, while Q_{ik} represents the probability that the minimum of the random variables T_{ij}, $j = 1, \ldots, i, \ldots, N$ is the specific random variable T_{ik}, W_i represents the distribution function of the time till the first transit from state i irrespectively of the state to which the process arrives (including transits to state i itself).

Let us assume that $W_i(t)$ is absolutely continuous w.r.t. the Lebesgue measure and has a density denoted by $f_i(t)$.

As we will be dealing in the sequel with parametric inference, whenever a quantity of interest will depend on a parameter $\theta \in \Theta \subset \mathbb{R}^m$, we may set this parameter as an argument. For instance, if $Q_{ij}(t)$ depends on some parameter θ we could denote it by $Q_{ij}(t; \theta)$.

Our intention is to provide estimators of p_{ij}, $W_i(t)$, and $Q_{ij}(t)$ under a general class of distributions. This class of distributions is presented and discussed in the next chapter where we first provide estimators of the parameters involved in the class of distributions under investigation. The cases of four specific distributions, all of which are members of the general class, will be thoroughly discussed. These distributions are Exponential, Weibull, Raleigh and Pareto which are frequently encountered in Reliability theory.

Andreas Makrides

Chapter 3

The class of generalized Gamma distributions

The types of distributions considered in this work for the sojourn times are presented in this chapter. More specifically we concentrate on the generalized or three-parameter Gamma distribution. The generalized Gamma distribution is a continuous probability distribution with three parameters and is considered to be a generalization of the typical two-parameter Gamma distribution. The distribution is also known as *Stacy* distribution (Stacy, 1962) and if an extra fourth parameter (a location parameter δ) is introduced the distribution becomes the *Amoroso* distribution introduced by the Italian economist L. Amoroso (see Amoroso, 1925 and Crooks, 2010):

$$Amoroso(x|\delta, a_{ij}, \alpha, c) = \frac{1}{\Gamma(\alpha)}\left|\frac{c}{a_{ij}}\right|\left(\frac{x-\delta}{a_{ij}}\right)^{\alpha c-1}\exp\left\{-\left(\frac{x-\delta}{a_{ij}}\right)^{c}\right\}\ x, \delta, a_{ij}, \alpha, c \in \mathbb{R},\ \alpha > 0,\ a_ij \neq 0 \tag{3.1}$$

with support $x \geq \delta$ if $a_{ij} > 0$ and $x \leq \delta$ if $a_{ij} < 0$. The notation $\Gamma(\cdot)$ denotes the Gamma function given by

$$\Gamma(s) = \int_0^{\infty} z^{s-1}e^{-z}dz.$$

Table 3.1 provides the Amoroso family of distributions.

Many distributions commonly used in reliability theory like the Weibull distribution, the log-normal distribution, the Gamma distribution, the Rayleigh distribution and the Exponential distribution are special cases of the generalized Gamma.

Andreas Makrides

As a flexible skewed distribution, the generalized Gamma is frequently used for life-time analysis and reliability testing. In addition, it models fading phenomena in wireless communication, has been applied in automatic image retrieval and analysis (Choi and Tong, 2010; de Ves et al, 2010; Schutz et. al, 2013), was used to evaluate dimensionality reduction techniques (Li et. al, 2006), and also appears to be connected to diffusion processes in (social) networks (Lienhard and Meyer, 1967; Bauckhage et. al, 2013a; Bauckhage et. al 2013b). Accordingly, methods for measuring (dis)similarities between generalized Gamma distributions are of practical interest in data science because they facilitate model selection and statistical inference.

The cumulative distribution function (cdf) of the generalized Gamma distribution is given by

$$F_{ij}(t) = \frac{\gamma\left(d/c, (t/a_{ij})^c\right)}{\Gamma(d/c)}, \quad c,\ d,\ a_{ij} > 0$$

where $\gamma(\cdot)$ is the lower incomplete Gamma function given by $\gamma(s,x) = \int_0^x z^{s-1} e^{-z} dz$, a_{ij} the shape parameter and c and d the scale parameters of the distribution. The probability density function (pdf) of the generalized Gamma distribution is given by

$$f_{ij}(t) = \frac{c/a_{ij}^d}{\Gamma(d/c)} x^{d-1} e^{-(x/a_{ij})^c}, \quad x > 0,\ c,\ d,\ a_{ij} > 0.$$

Observe that the generalized Gamma is a special case of Amoroso distribution with $\delta = 0$ and $\alpha c = d$. The most popular special cases of the generalized Gamma distribution are:

For $d = c$ the generalized Gamma reduces to the Weibull distribution

For $d = c = 1$ the Exponential distribution is obtained

For $d = c = 2$ and $\sqrt{2}a_{ij}$ instead of a_{ij} the Rayleigh distribution is obtained

For $c = 1$ the typical 2-parameter Gamma distribution is obtained and

For $c \to 0$ the distribution reduces to the lognormal distribution.

For c negative the generalized Gamma distribution is not defined but the Amoroso distribution becomes the generalized inverse Gamma which is the parent of various inverse distributions, including the inverse Gamma, inverse Exponential etc.

Andreas Makrides

Amoroso	δ	a_{ij}	α	c	
Stacy	0	.	.	.	
gen. Fisher-Tippett	.	.	n^*	.	
Fisher-Tippett	.	.	1	.	
Fréchet	.	.	1	< 0	
generalized Fréchet	.	.	n	< 0	
scaled inverse chi	0	.	$\frac{1}{2}k^*$	-2	
inverse chi	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}k$	-2	
inverse Rayleigh	0	.	1	-2	
Pearson type V	.	.	.	-1	
inverse Gamma	0	.	.	-1	
scaled inverse chi-square	0	.	$\frac{1}{2}k$	-1	
inverse chi-square	0	$\frac{1}{2}$	$\frac{1}{2}k$	-1	
Lévy	.	.	$\frac{1}{2}$	-1	
inverse Exponential	0	.	1	-1	
Pearson type III	.	.	.	1	
Gamma	0	.	.	1	
Erlang	0	> 0	n	1	
standard Gamma	0	1	.	1	
scaled chi-square	0	.	$\frac{1}{2}k$	1	
chi-square	0	2	$\frac{1}{2}k$	1	
shifted Exponential	.	.	1	1	
Exponential	0	.	1	1	
standard Exponential	0	1	1	1	
Wien	0	.	4	1	
Nakagami	.	.	.	2	
scaled chi	0	.	$\frac{1}{2}k$	2	
chi	0	$\sqrt{2}$	$\frac{1}{2}k$	2	
half-normal	0	.	$\frac{1}{2}$	2	
Rayleigh	0	.	1	2	
Maxwell	0	.	$\frac{3}{2}$	2	
Wilson-Hilferty	0	.	.	3	
generalized Weibull	.	.	n	> 0	
Weibull	.	.	1	> 0	
pseudo-Weibull	.	.	$1 + \frac{1}{c}$	> 0	
stretched Exponential	0	.	1	> 0	
<u>Limits</u>					
log-Gamma	.	.	.	.	$\lim_{c\to\infty}$
power law	.	.	$\frac{1-p}{c}$	.	$\lim_{c\to 0}$
log-normal	.	.	$\frac{1}{(c\sigma)^2}$	.	$\lim_{c\to 0}$
normal	.	.	.	1	$\lim_{\alpha\to\infty}$
* where $k, n > 0$					

Table 3.1: The Amoroso Family of distributions

Andreas Makrides

3.1 INID random variables

We now focus on a special class of distributions within the generalized Gamma distribution of the previous section paying attention to the shape parameter of the distribution. In particular we view the distributional quantities of interest as functions of the shape parameter and assume that the cdf of sojourn times T_{ij} are of the same functional form for all i's and j's but with different shape parameters. Within the generalized Gamma family of distributions we concentrate on a (sub)class of distribution functions with the following property:

$$F(t; a_{ij}) = 1 - (1 - F(t; 1))^{a_{ij}} , \tag{3.2}$$

where $F(t; a_{ij})$ is absolutely continuous w.r.t. the Lebesgue measure and the pdf is denoted by $f(t; a_{ij})$.

In the previous section we found that the order statistics and their distribution are playing a key role in the proposed methodology. In general, order statistics of independent but not necessarily identically distributed (inid) random variables are not easy to deal with. In order to overcome practical and computational difficulties associated with the order statistics, we focus on random variables that belong to the above class which is closed under minima. This property of the above class of distributions results in producing recurrence relations that reduces the computational burden associated with functions of order statistics. The Theorem below presents this property by showing that the minimum order statistic from an inid random sample from the above class has a distribution belonging to the same class. In other words, the above class of distributions is closed under minima.

Theorem 3.1.1 *Let $X_1, \ldots, X_N$ be inid random variables such that $X_i \sim F(x; a_i)$ where $F(x; a_i)$ belong to the class (3.2). Then the distribution function $F^{(1)}$ of the minimum order statistic $X_{(1)}$ belongs also to the class (3.2).*

Proof. The distribution function of the minimum order statistic $X_{(1)}$ can be written

Andreas Makrides

as

$$\begin{aligned}
&F^{(1)}(t; a_i; i = 1, \ldots, N) \\
&= 1 - \prod_{i=1}^{N} (1 - F(t; a_i)) \\
&= 1 - \prod_{i=1}^{N} (1 - F(t; 1))^{a_i} \\
&= 1 - (1 - F(t; 1))^{\sum\limits_{i=1}^{N} a_i},
\end{aligned}$$

which belongs to class (3.2) with parameter $\sum\limits_{i=1}^{N} a_i$. □

Note that the following result (Balasubramanian et al. (1991) and Balakrishnan (2007)) holds for the distribution function $F^{(r)}$ of the r^{th} order statistic in terms of the distribution function $F^{(1)}$ of the minimum order statistic

$$\begin{aligned}
&F^{(r)}(t; a_i; i = 1, \ldots, N) = \\
&= \sum_{i=0}^{r-1} (-1)^{r-i-1} \binom{N-i-1}{N-r} \binom{N}{N-i} F^{(1)}(t; a_i; i = 1, \ldots, N)
\end{aligned}$$

which by Theorem 3.1.1 takes the final form

$$\begin{aligned}
&F^{(r)}(t; a_i; i = 1, \ldots, N) = \\
&= \sum_{i=0}^{r-1} (-1)^{r-i-1} \binom{N-i-1}{N-r} \binom{N}{N-i} \left[1 - (1 - F(t; 1))^{\sum\limits_{i=1}^{N} a_i}\right]
\end{aligned}$$

The above formula can be then used for the calculation of the expectation of functions of the r^{th} order statistic. Indeed, let $h(X_{(r)})$ be a function of $X_{(r)}$. Then, if the continuous case is assumed, we have

$$E\left[h(X_{(r)})\right] = \sum_{i=0}^{r-1} (-1)^{r-i-1} \binom{N-i-1}{N-r} \binom{N}{N-i} \int h(t) dF\left(t; \sum_{i=1}^{N} a_i\right). \tag{3.3}$$

For the special case of the minimum order statistic the above formula reduces to

$$E\left[h(X_{(1)})\right] = \int h(t)dF\left(t; \sum_{i=1}^{N} a_i\right). \tag{3.4}$$

Using $h(u) = u^k$ and $h(u) = \exp(uz)$, for some z one can obtain the k^{th} order moment and the moment generating function of the r^{th} order statistic $X_{(r)}$.

The simplest discrete distribution that belongs to the class defined by (3.2) is the Geometric distribution. Continuous distributions include the Pareto distribution, the Weibull distribution and its special cases like the Exponential, the Rayleigh and the Erlang truncated Exponential. As a general example for demonstrative purposes, we consider first the Weibull distribution. Finally, we also investigate the case of the Pareto distribution.

Example 3.1.1

Let $X_1, \ldots, X_N$ be *inid* random variables such that $X_i \sim F(x; a_i)$ which is a Weibull distribution with scale parameter a_i and common shape parameter c, namely,

$$F(t; a_i) = 1 - e^{-(t/a_i)^c}, \qquad t \geq 0, \quad c > 0, \quad a_i > 0. \tag{3.5}$$

Let now, $X_{(1)} < X_{(2)} < \cdots < X_{(N)}$, then

$$\begin{aligned} F^{(1)}\left(t; a_i; i = 1, \ldots, N\right) &= P\left(X_{(1)} \leq t\right) \\ &= 1 - e^{-t^c / \left(\left(\sum_{i=1}^{N} \frac{1}{a_i^c}\right)^{-1/c}\right)^c}, \end{aligned} \tag{3.6}$$

which is a Weibull distribution function with scale parameter $\gamma = \left(\sum_{i=1}^{N} \frac{1}{a_i^c}\right)^{-1/c}$.

Then, the moment generating function is

$$E\left[e^{tX_{(r)}}\right] = \sum_{i=0}^{r-1} (-1)^{r-i-1} \binom{N-i-1}{N-r} \binom{N}{N-i} \sum_{n=0}^{\infty} \frac{t^n \gamma^n}{n!} \Gamma\left(1 + \frac{n}{c}\right).$$

For the calculation of the k^{th} order moment of the r^{th} order statistic $X_{(r)}$ we consider in

Andreas Makrides

(3.3) $h(t)=t^k$, $k=1,2,\ldots$ and we obtain

$$E\left[X_{(r)}^k\right]=\sum_{i=0}^{r-1}(-1)^{r-i-1}\binom{N-i-1}{N-r}\binom{N}{N-i}\gamma^k\Gamma\left(\frac{k}{c}+1\right). \tag{3.7}$$

From (3.4) the mean of the minimum r.v. $X_{(1)}$ is

$$E\left[X_{(1)}\right]=\gamma\Gamma\left(\frac{1}{c}+1\right). \tag{3.8}$$

Example 3.1.2

Let $X_1,\ldots,X_N$ be *inid* random variables such that $X_i\sim F(x;a_i)$ which is a Pareto distribution with shape parameter a_i and scale parameter λ and given by,

$$F(t;a_i)=1-\left(\frac{\lambda}{t}\right)^{a_i}, \qquad t\geq\lambda, \quad a_i>0. \tag{3.9}$$

For the calculation of the k^{th} order moment of the r^{th} order statistic $X_{(r)}$ we consider in (3.3) $h(t)=t^k$, $k=1,2,\ldots$. Then

$$\begin{aligned}\int h(t)dF\left(t;\sum_{i=1}^N a_i\right) &= \int_\lambda^\infty t^k d\left(1-\left(\frac{\lambda}{t}\right)^{\sum\limits_{i=1}^N a_i}\right)=\sum_{i=1}^N a_i\int_\lambda^\infty t^k\lambda^{\sum\limits_{i=1}^N a_i}t^{-(\sum\limits_{i=1}^N a_i+1)}dt\\ &= \frac{\sum\limits_{i=1}^N a_i\lambda^k}{\sum\limits_{i=1}^N a_i-k}, \quad \sum_{i=1}^N a_i>k. \end{aligned}\tag{3.10}$$

Therefore,

$$E\left[X_{(r)}^k\right]=\sum_{i=0}^{r-1}(-1)^{r-i-1}\binom{N-i-1}{N-r}\binom{N}{N-i}\frac{\lambda^k\sum\limits_{i=1}^N a_i}{\sum\limits_{i=1}^N a_i-k}, \quad k<\sum_{i=1}^N a_i. \tag{3.11}$$

Andreas Makrides

From (3.4) the mean of the minimum r.v. $X_{(1)}$ is

$$E\left[X_{(1)}\right] = \frac{\lambda \sum_{i=1}^{N} a_i}{\sum_{i=1}^{N} a_i - 1}, \quad \sum_{i=1}^{N} a_i > 1. \tag{3.12}$$

3.2 General Case

Under the class of distributions given in (3.2), the following result concerning the main semi-Markov characteristics can be proved. Recall that T_{ij} represents the sojourn time with cdf F_{ij} and pdf f_{ij}. For notational convenience, we set $F(t) := F(t;1)$, $f(t) := f(t;1)$ and $Q_{ij}\left(t; a_{ik}; k = 1, \ldots, N\right) := Q_{ij}(t)$.

Proposition 3.2.1 *Under the setup of this section, the following results hold:*

$$Q_{ij}(t) = \frac{a_{ij}}{\sum_{k \in E} a_{ik}} \left[1 - (1 - F(t))^{\sum_{k \in E} a_{ik}}\right], \tag{3.13}$$

$$p_{ij} = \frac{a_{ij}}{\sum_{k \in E} a_{ik}}, \tag{3.14}$$

$$W_i(t) = 1 - [1 - F(t)]^{\sum_{j=1}^{N} a_{ij}} \tag{3.15}$$

and

$$f_i(t) = \sum_{j=1}^{N} a_{ij} \left(1 - F(t)\right)^{\sum_{j=1}^{N} a_{ij}} \frac{f(t)}{1 - F(t)}. \tag{3.16}$$

Consequently, the expectation of the minimum is given by

Andreas Makrides

$$\mathbb{E}\left(X_{(1)}\right) = \left(\sum_{j=1}^{N} a_{ij}\right) \int_0^{+\infty} t\left(1 - F(t)\right)^{\sum\limits_{j=1}^{N} a_{ij} - 1} f(t)dt.$$

Proof. Note that

$$\begin{aligned}
Q_{ij}(t) &= \int_0^t \left[1 - F(u)\right]^{a_{i1}} \dots \left[1 - F(u)\right]^{a_{i,j-1}} \times \\
&\quad \times \left[1 - F(u)\right]^{a_{i,j+1}} \dots \left[1 - F(u)\right]^{a_{iN}} d\left[1 - (1 - F(u))^{a_{ij}}\right] \\
&= \int_0^t a_{ij} \left[1 - F(u)\right]^{\sum\limits_{j\in E} a_{ij} - 1} dF(u) \\
&= -\frac{a_{ij}}{\sum\limits_{j\in E} a_{ij}} \left[(1 - F(t))^{\sum\limits_{j\in E} a_{ij}} - 1\right] \\
&= \frac{a_{ij}}{\sum\limits_{j\in E} a_{ij}} \left[1 - (1 - F(t))^{\sum\limits_{j\in E} a_{ij}}\right]. \qquad (3.17)
\end{aligned}$$

By taking the limit,

$$\lim_{t\to\infty} Q_{ij}(t) = p_{ij} = \frac{a_{ij}}{\sum\limits_{j\in E} a_{ij}}. \qquad (3.18)$$

We also need the distribution of the minimum which is given by

$$\begin{aligned}
W_i(t) &= F_{min}(t) = \mathbb{P}\left[T_{min} \le t\right] \\
&= 1 - \mathbb{P}\left[T_{min} > t\right] = 1 - \mathbb{P}\left[T_{ij} > t, \forall j\right] \\
&= 1 - \prod_{j=1}^{N} \left[1 - F_{ij}(t)\right] = 1 - \prod_{j=1}^{N} \left[1 - F(t)\right]^{a_{ij}} \\
&= 1 - \left[1 - F(t)\right]^{\sum\limits_{j=1}^{N} a_{ij}}. \qquad (3.19)
\end{aligned}$$

Andreas Makrides

So, the associated pdf is given by

$$
\begin{aligned}
f_i(t) &= \sum_{j=1}^{N} a_{ij} \left(1 - F(t)\right)^{\sum\limits_{j=1}^{N} a_{ij} - 1} \cdot f(t) \\
&= \sum_{j=1}^{N} a_{ij} \left(1 - F(t)\right)^{\sum\limits_{j=1}^{N} a_{ij}} \cdot \frac{f(t)}{1 - F(t)}. \qquad (3.20)
\end{aligned}
$$

Hence the expectation of the minimum follows, i.e.

$$
\begin{aligned}
E\left[X_{(1)}\right] &= \int t\, dF\left(t; \sum_{i=1}^{N} a_{ij}\right) \\
&= \int t \cdot \sum_{j=1}^{N} a_{ij} \left(1 - F(t)\right)^{\sum\limits_{j=1}^{N} a_{ij}} \cdot \frac{f(t)}{1 - F(t)}\, dt. \qquad (3.21)
\end{aligned}
$$

$\square$

Andreas Makrides

Chapter 4

The Moment Estimation Case

Andreas Makrides

In Chapter 2 we have provided the general formulas for the semi-Markov quantities of interest, namely the semi-Markov kernel $Q_{ij}(\cdot)$, the transition probabilities of the embedded Markov chain $p_{ij}(\cdot)$, and the cdf of the sojourn time $W_i(\cdot)$. In Chapter 3 these quantities have been expressed in terms of the unknown parameters involved in the general class of distributions given in (3.2). In this chapter the case of continuous time Markov process will be considered and the estimators of the unknown parameters will be given for a number of specific distributions belonging to the class (3.2). The estimates are provided for the Weibull, Exponential, Rayleigh and Pareto distributions. Then, the estimators of the Markov process quantities of interest will follow immediately.

INTENSITY RATE ESTIMATORS

Recall that p_{ij} defined in (2.2) can be viewed as the fraction of time the system stays in state i before moving to j. Ergodicity property holds, due to the assumptions of irreducibility and positive recurrence (see page 10), hence the above quantity can be consistently estimated by the method of moments as a ratio of the corresponding number of transitions, namely

$$\hat{p}_{ij}(M) = \frac{N_{ij}(M)}{N_i(M)}, \tag{4.1}$$

where

M is the total observation time,

$N_i(M) = \sum_{j=1}^{N} N_{ij}(M)$ is the accumulated number of system transits from state i to any state during observation time M and

$N_{ij}(M)$ is the accumulated number of system transitions from state i to state j during observation time M.

Combining expressions (2.2), (2.9) and (4.1) and solving the resulting system of equations we obtain the estimates of the parameters a_{ik} with $i, k = 1, \ldots, N$ of the distribution involved. Having available the estimates $\hat{a}_{ik}$, the estimate of the time independent mean intensity rate follows immediately by (2.5):

$$\hat{\lambda}_{ik} = \frac{1}{\int_0^\infty \left(e^{-\int_0^t \hat{\lambda}_{ik}(x)dx} \right) dt} \tag{4.2}$$

where by expression (2.3)

$$\begin{aligned} \hat{\lambda}_{ik}(t) &= \frac{\hat{f}_{ik}(t)}{\hat{R}_{ik}(t)} \\ &= \frac{d\ F(t; \hat{a}_{ik})/dt}{1 - F(t; \hat{a}_{ik})} \end{aligned} \tag{4.3}$$

is the variable intensity rate and $F(t; \hat{a}_{ij})$ the distribution function $F(t)$ with $\hat{a}_{ij}$ in place of a_{ij}. Note that by taking into consideration the class of distributions (3.2) the above expression takes the form

$$\begin{aligned} \hat{\lambda}_{ik}(t) &= \frac{\hat{a}_{ik} f(t;1) \left(1 - F(t;1)\right)^{\hat{a}_{ik}-1}}{\left(1 - F(t;1)\right)^{\hat{a}_{ik}}} \\ &= \frac{\hat{a}_{ik} f(t;1)}{1 - F(t;1)}. \end{aligned}$$

Furthermore, note that

$$\hat{\lambda}_{ii} = -\sum_{j \neq i} \hat{\lambda}_{ij}.$$

Before proceeding with the parameter estimators we provide below the expressions for obtaining the transition probabilities from state i to state j for any i and j. Let

$$\Lambda = (\lambda_{ij})_{i,j=1,..,N}$$

Andreas Makrides

be the intensity matrix of the multi-state system with eigenvalues denoted by γ_i and eigenvectors by u_i, $i = 1, \ldots, N$. The determination of the required probabilities is based on the estimator of the above intensity matrix Λ, given by

$$\hat{\Lambda} = \left(\hat{\lambda}_{ij}\right)_{i,j=1,..,N}, \tag{4.4}$$

with $\hat{\lambda}_{ij}$ as in (4.2).

More specifically, the estimators of the eigenvalues $\hat{\gamma}_i$, $i = 1, \ldots, N$ of $\hat{\Lambda}$ are obtained and then, the corresponding estimators of the eigenvectors $\hat{u}_1, \ldots, \hat{u}_N$ are evaluated. If $\hat{L}$ is the $N \times N$ matrix of the estimators of the eigenvectors and $\hat{D}(t)$ a $N \times N$ diagonal matrix with the (i,i) element equal to

$$\exp(-\hat{\gamma}_i \cdot t)$$

then the estimates of the transition probabilities are given by the matrix

$$\begin{aligned} \hat{P}(t) &= \left(\hat{p}_{ij}(t)\right)_{i,j=1,\ldots,N} \\ &= \hat{L} \cdot \hat{D}(t) \cdot \hat{L}^{-1}. \end{aligned} \tag{4.5}$$

Recall that for the case of the Exponential distribution for which the intensity rate is time independent, the origin $t = 0$ can be located at *any time instant* and therefore the above equation remains valid for *any* time t. On the other hand, due to time dependence of non-Exponential distributions, homogeneity is contrived to avoid any confusion with the time origin.

4.1 Weibull Model

We consider the Weibull model for sojourn times T_{ij} and obtain the estimators for the parameters of interest. Recall that both the Exponential and Rayleigh distributions are special cases of the Weibull class. The corresponding cumulative distribution function of the Weibull distribution is given by

$$F_{ij}\left(t\right) = 1 - e^{-\left(t/a_{ij}\right)^{c}}, \tag{4.6}$$

Andreas Makrides

where the scale parameter is denoted by a_{ij} and the shape parameter is represented by c and it is the same irrespectively of the state i or j. The intensity rate of the sojourn time for the case of the Weibull distribution associated with (4.6) is given by

$$\lambda_{ij}(t) = \frac{c}{a_{ij}} \left(\frac{t}{a_{ij}} \right)^{c-1}. \tag{4.7}$$

Note that for the case of the Exponential distribution where $c = 1$, the intensity rate is constant and equal to $1/a_{ij}$, while in the case of the Rayleigh distribution the intensity rate is given by (4.7) with $c = 2$ and $\sqrt{2}a_{ij}$ instead of a_{ij}. The Rayleigh distribution is presented here not only for illustrative purposes but also since it is the most commonly used distribution in reliability and life testing (Lawless, 2003).

4.1.1 The transition from state i to state i is allowed

Proposition 3.2.1 will be provided for the Weibull distribution under the assumption that the transition from state i to state i is allowed.

Using (2.8) we have that

$$\begin{aligned} Q_{ik}(t) &= \int_0^t \exp\left\{ -\frac{x^c}{a_{i1}^c} - \frac{x^c}{a_{i2}^c} - \cdots - \frac{x^c}{a_{ik-1}^c} - \frac{x^c}{a_{ik+1}^c} - \cdots - \frac{x^c}{a_{iN}^c} \right\} \times \\ &\times \frac{c}{a_{ik}^c} x^{c-1} \exp\left\{ -\frac{x^c}{a_{ik}^c} \right\} dx \\ &= \int_0^t \frac{c}{c a_{ik}^c} \exp\left\{ -x^c \sum_{j=1}^{N} \frac{1}{a_{ij}^c} \right\} dx^c \\ &= \frac{1}{a_{ik}^c \sum_{j=1}^{N} \frac{1}{a_{ij}^c}} \left[1 - e^{-t^c \sum_{j=1}^{N} \frac{1}{a_{ij}^c}} \right]. \end{aligned} \tag{4.8}$$

Andreas Makrides

For the cdf $W_i(t)$ we apply (4.8). Then, the expression (2.9) takes the form

$$\begin{aligned} W_i(t) &= \sum_{k=1}^{N} Q_{ik}(t) \\ &= \frac{\sum_{k=1}^{N} \frac{1}{a_{ik}^c}}{\sum_{j=1}^{N} \frac{1}{a_{ij}^c}} \left[1 - e^{-t^c \sum_{j=1}^{N} \frac{1}{a_{ij}^c}} \right] \\ &= 1 - e^{-t^c / \left(\left(\sum_{j=1}^{N} \frac{1}{a_{ij}^c} \right)^{-1/c} \right)^c}, \end{aligned} \tag{4.9}$$

which represents a Weibull distribution function with parameters c and $\left(\sum_{j=1}^{N} \frac{1}{a_{ij}^c} \right)^{-1/c}$ and with mean given by

$$\mu_i = \left(\sum_{j=1}^{N} \frac{1}{a_{ij}^c} \right)^{-1/c} \cdot \Gamma\left(1 + \frac{1}{c}\right)$$

where $\Gamma(\cdot)$ denotes the standard Gamma function. For the estimation of the μ_i's we take the standard moment estimator given by

$$\hat{\mu}_i = \frac{\sum_{j=1}^{N_i(M)} T_i^{(j)}}{N_i(M)}. \tag{4.10}$$

Thus,

$$\frac{\sum_{j=1}^{N_i(M)} T_i^{(j)}}{N_i(M)} = \left(\sum_{j=1}^{N} \frac{1}{a_{ij}^c} \right)^{-1/c} \cdot \Gamma\left(1 + \frac{1}{c}\right).$$

Equivalently,

$$\left(\sum_{j=1}^{N} \frac{1}{a_{ij}^c} \right)^{1/c} = \frac{\Gamma\left(1 + \frac{1}{c}\right) N_i(M)}{\sum_{j=1}^{N_i(M)} T_i^{(j)}}$$

Andreas Makrides

or,

$$\sum_{j=1}^{N} \frac{1}{a_{ij}^{c}} = \left[\frac{\Gamma\left(1+\frac{1}{c}\right) N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}} \right]^{c}. \tag{4.11}$$

Using expression (4.11) one can estimate the summation of the scale parameters for all transitions that exit from any state i. To estimate individual scale parameters a_{ij} for each $\{i, j\}$, we consider the probabilities:

$$\begin{aligned} p_{ik} &= \lim_{t\to\infty} Q_{ik}\left(t\right) \\ &= \lim_{t\to\infty} \frac{1}{a_{ik}^{c} \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^{c}}} \left[1 - e^{-t^{c} \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^{c}}} \right] \\ &= \frac{1}{a_{ik}^{c} \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^{c}}}. \end{aligned} \tag{4.12}$$

Equivalently,

$$a_{ik}^{c} = \frac{1}{p_{ik} \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^{c}}}, \tag{4.13}$$

which, with the use of (4.1) , gives

$$\begin{aligned} \hat{a}_{ik}^{c} &= \frac{1}{\frac{N_{ik}(M)}{N_i(M)} \widehat{\sum\limits_{j=1}^{N} \frac{1}{a_{ij}^{c}}}} \\ &= \frac{1}{\frac{N_{ik}(M)}{N_i(M)} \left[\frac{\Gamma\left(1+\frac{1}{c}\right) N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}} \right]^{c}}. \end{aligned} \tag{4.14}$$

The above results are summarized below:

Andreas Makrides

Theorem 4.1.1 *Consider the Weibull model given in (4.6) for the sojourn times in a MS system. Then, the estimators of the scale parameters a_{ij}, $i,j = 1,2,\ldots,N$ are given by:*

$$\hat{a}_{ij} = \frac{1}{\left(\frac{N_{ij}(M)}{N_i(M)}\right)^{1/c} \left[\frac{\Gamma\left(1+\frac{1}{c}\right)N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}\right]}, \quad i,j = 1,\ldots,N. \tag{4.15}$$

For an estimator of c see Section 4.1.3.

4.1.2 The transition from state i to state i is not allowed

In this section we provide again the expressions of Proposition 3.2.1 in the case that the transitions from state i to state i are not allowed. Thus, we assume that the process moves from i to j with $j \neq i$, i.e, $Q_{ii}(t) = 0$. Then, $T_{ik} \leq t$ is the minimum of the $N-1$ random variables T_{ij}, $j \neq i$, $j = 1,\ldots,N$, namely

$$T_{ik} = \min\{T_{i1},\ldots,T_{i,i-1},T_{i,i+1},\ldots,T_{iN}\}.$$

In other words, if $T_{i(m)}$ represents the m^{th} order statistic, $m = 1,\ldots,N-1$ of the $N-1$ random variables $T_{i1},\ldots,T_{i,i-1},T_{i,i+1},\ldots,T_{iN}$, then

$$T_{i(1)} = T_{ik}.$$

In this case and for the Weibull distribution, the quantity Q_{ik} takes the form:

$$Q_{ik}(t) = \frac{1}{a_{ik}^c \sum\limits_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{a_{ij}^c}} \left[1 - \exp\left\{-t^c \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{a_{ij}^c}\right\}\right]. \tag{4.16}$$

Andreas Makrides

Hence,

$$
\begin{aligned}
W_i(t) &= \sum_{\substack{k=1 \\ k\neq i}}^{N} Q_{ik}(t) \\
&= 1 - e^{-t^c / \left(\left(\sum\limits_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{a_{ij}^c} \right)^{-1/c} \right)^c} \qquad (4.17)
\end{aligned}
$$

which is a Weibull distribution function with parameters c and $\left(\sum\limits_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{a_{ij}^c} \right)^{-1/c}$ with mean

$$
\mu_i = \left(\sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{a_{ij}^c} \right)^{-1/c} \cdot \Gamma\left(1 + \frac{1}{c}\right). \qquad (4.18)
$$

As in the previous section, we have

$$
\sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{a_{ij}^c} = \left[\frac{\Gamma\left(1+\frac{1}{c}\right) N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}} \right]^c. \qquad (4.19)
$$

Similarly, as before,

$$
\begin{aligned}
p_{ik} &= \lim_{t\to\infty} Q_{ik}(t) \\
&= \frac{1}{a_{ik}^c \sum\limits_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{a_{ij}^c}}, \qquad (4.20)
\end{aligned}
$$

and finally, the following theorem holds true:

Theorem 4.1.2 *Consider the Weibull model given in (4.6) for the sojourn times in an MS system. Then, the estimators of the scale parameters* a_{ij}, $i, j = 1, 2, \ldots, N$ *are given*

Andreas Markides

by:

$$\hat{a}_{ij} = \frac{1}{\left(\frac{N_{ij}(M)}{N_i(M)}\right)^{1/c} \left[\frac{\Gamma\left(1+\frac{1}{c}\right)N_i(M)}{\sum\limits_{m=1}^{N_i(M)} T_i^{(m)}}\right]}, \quad i,j = 1,\ldots,N, \quad i \neq j. \tag{4.21}$$

4.1.3 Estimation of the shape parameter c

For the estimation of the shape parameter c we rely on the following regression setting. Let,

$$F(t) = 1 - e^{-\left(\frac{t}{a_{ij}}\right)^c}$$

which is equivalent to,

$$\begin{aligned} Y &= \ln\ln\frac{1}{1-F} \\ &= -c\ln a_{ij} + c\ln t. \end{aligned}$$

If we let $X = \ln t$, the following regression model could be considered,

$$Y_i = -c\ln a_{ij} + cX_i + \varepsilon_i, \qquad i = 1,2,\ldots,n.$$

where ε_i follow an appropriately chosen distribution. In the case of *inid* random variables T_{ij}, let us assume that

$$T_{ij} \sim \text{Weibull}(a_{ij}, c) \qquad i,j = 1,2,\ldots,N.$$

Then, the regression model takes the form,

$$Y_{i1_l} = -c\ln a_{i1} + c\ln T_{i1_l} + \varepsilon_{i1_l},$$

$$Y_{i2_l} = -c\ln a_{i2} + c\ln T_{i2_l} + \varepsilon_{i2_l},$$

$$\vdots$$

Andreas Makrides

$$Y_{iN_l} = -c \ln a_{im} + c \ln T_{im_l} + \varepsilon_{iN_l},$$

where $i = 1, 2, \ldots, N, \quad l = 1, 2, \ldots, k_{ij}$.

Following the standard least squares estimating method and substituting $\ln T_{ij} = X_{ij}$, the estimators become:

$$\begin{aligned}\ln \hat{a}_1 &= \frac{\sum_{i=1}^{k_1} X_{i1}}{k_1} - \frac{1}{c} \cdot \frac{\sum_{i=1}^{k_1} Y_{i1}}{k_1} \\ &= \overline{X}_1 - \frac{\overline{Y}_1}{c},\end{aligned}$$

$$\ln \hat{a}_2 = \overline{X}_2 - \frac{\overline{Y}_2}{c},$$

$$\vdots$$

$$\ln \hat{a}_m = \overline{X}_m - \frac{\overline{Y}_m}{c},$$

and

$$\hat{c} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{k_j} X_{ij} Y_{ij} - \sum_{j=1}^{m} k_j \overline{X}_j \overline{Y}_j}{\sum_{j=1}^{m} \sum_{i=1}^{k_j} X_{ij}^2 - \sum_{j=1}^{m} k_j \overline{X}_j^2}. \tag{4.22}$$

As an alternative the MLE estimator for the shape parameter c of the Weibull distribution can be obtained as follows (see for example Tableman and Kim, 2004, p. 105):
If T is considered to follow the Weibull distribution with parameters a_{ij} and c then $\ln(T)$ follows the standard extreme value distribution so that

$$\ln(T) = -ln(a_{ij}) + \frac{1}{c} \cdot \epsilon \equiv \mu + \sigma \cdot \epsilon,$$

where ϵ follows the standard extreme value distribution.
The MLE estimator of the shape parameter is then given by $\hat{c} = \frac{1}{\hat{\sigma}}$, where $\hat{\sigma}$ is the MLE of σ. This estimator can be obtained, for instance, via the *survReg* procedure of SAS.

Andreas Makrides

4.1.4 Estimation of transition intensities

The estimators for the intensity rate between states i and k is given in the following Corollary:

Corollary 4.1.1 *Consider the Weibull model given in (4.6) for the sojourn times in an MS system. Then, the estimate of the intensity rate between* i *and* j *is given by:*

$$\hat{\lambda}_{ij}(t) = \frac{\hat{c}}{\hat{a}_{ij}} \left(\frac{t}{\hat{a}_{ij}} \right)^{\hat{c}-1}. \tag{4.23}$$

Note that since $\sum_{k=1}^{N} \lambda_{ik}(t) = 0$, $\forall i$, then $\lambda_{ii}(t) = -\sum_{k \neq i} \lambda_{ik}(t)$ and threfore,

$$\hat{\lambda}_{ii}(t) = -\sum_{k \neq i} \hat{\lambda}_{ik}(t).$$

4.1.5 Homogeneous transition intensities

One can contrive homogeneous Markov models with variable transition rates as previously discussed. The integral of the failure rate is,

$$\begin{aligned} \int_0^t \lambda_{ik}(\tau) d\tau &= \int_0^t \frac{c}{a_{ik}} \left(\frac{\tau}{a_{ik}} \right)^{c-1} d\tau \\ &= \left(\frac{t}{a_{ik}} \right)^c. \end{aligned}$$

So the system failure rate using (2.4) takes the form,

$$\lambda_{ik} = \frac{1}{\int_0^\infty \left(e^{-(t/a_{ik})^c} \right) dt}.$$

Note that in the case of Exponential distribution (where the parameter $c = 1$), the intensity rate is constant and equal to $1/a_{ik}$.

In general, for arbitrary values of c, using the transformation $u = (t/a_{ik})^c$ the intensity

Andreas Makrides

rate is evaluated as,

$$\begin{aligned}
\lambda_{ik} &= \frac{1}{\int\limits_{u=0}^{\infty} \frac{e^{-u} a_{ik}^{c}}{c a_{ik}^{c-1} u^{1-1/c}} du} \\
&= \frac{1}{\frac{a_{ik}}{c} \int\limits_{u=0}^{\infty} e^{-u} u^{1/c-1} du} \\
&= \frac{1}{a_{ik}} \frac{c}{\Gamma(1/c)}.
\end{aligned}$$

Hence, using (4.15) and (4.22) the estimator of the homogeneous transition intensity between states i and k ($i \neq k$) is,

$$\begin{aligned}
\hat{\lambda}_{ik} &= \frac{1}{\hat{a}_{ik}} \frac{\hat{c}}{\Gamma(1/\hat{c})} \\
&= \left(\frac{N_{ik}(M)}{N_i(M)} \right)^{1/\hat{c}} \left[\frac{N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}} \right].
\end{aligned} \tag{4.24}$$

Although the above equation is mathematically valid for the case $i = k$, for the estimation of $\hat{\lambda}_{ii}$ we use the Markov property that $\sum\limits_{k=1}^{N} \lambda_{ik} = 0,\ \forall i$ so that $\lambda_{ii} = -\sum\limits_{k \neq i} \lambda_{ik}$ and therefore

$$\hat{\lambda}_{ii} = -\sum_{k \neq i} \hat{\lambda}_{ik}.$$

4.2 Exponential distribution

If the sojourn (residing) time T_{ij} is assumed to follow an Exponential distribution then the cdf is given by

$$F_{ij}(t) = 1 - e^{-t/a_{ij}}, \tag{4.25}$$

where $1/a_{ij}$ is the transition intensity from state i to state j.

Following the methodology of Section 3.2 for the case of the Exponential distribution, we

Andreas Makrides

have that

$$\begin{aligned}
Q_{ik}(t) &= \int_0^t \exp\left\{-\frac{x}{a_{i1}} - \frac{x}{a_{i2}} - \cdots - \frac{x}{a_{ik-1}} - \frac{x}{a_{ik+1}} - \cdots - \frac{x}{a_{iN}}\right\} \frac{1}{a_{ik}} \exp\left\{-\frac{x}{a_{ik}}\right\} dx \\
&= \int_0^t \frac{1}{a_{ik}} \exp\left\{-x\left(\frac{1}{a_{i1}} + \frac{1}{a_{i2}} + \cdots + \frac{1}{a_{iN}}\right)\right\} dx \\
&= \int_0^t \frac{1}{a_{ik}} \exp\left\{-x\sum_{j=1}^{N} \frac{1}{a_{ij}}\right\} dx \\
&= \frac{1}{a_{ik}\sum_{j=1}^{N}\frac{1}{a_{ij}}}\left[1 - e^{-t\sum_{j=1}^{N}\frac{1}{a_{ij}}}\right]. \qquad (4.26)
\end{aligned}$$

Hence,

$$\begin{aligned}
W_i(t) &= \sum_{k=1}^{N} Q_{ik}(t) \\
&= \frac{\sum_{k=1}^{N}\frac{1}{a_{ik}}}{\sum_{j=1}^{N}\frac{1}{a_{ij}}}\left[1 - e^{-t\sum_{j=1}^{N}\frac{1}{a_{ij}}}\right] \\
&= 1 - e^{-t\sum_{j=1}^{N}\frac{1}{a_{ij}}}, \qquad (4.27)
\end{aligned}$$

which is an Exponential distribution function with rate $\sum_{j=1}^{N}\frac{1}{a_{ij}}$ and mean

$$\mu_i = \frac{1}{\sum_{j=1}^{N}\frac{1}{a_{ij}}}. \qquad (4.28)$$

For an estimation of the mean we use the standard moment or maximum likelihood estimator, given by

$$\hat{\mu}_i = \frac{\sum_{j=1}^{N_i(M)} T_i^{(j)}}{N_i(M)}. \qquad (4.29)$$

So,

$$\widehat{\sum_{j=1}^{N}\frac{1}{a_{ij}}} = \frac{1}{\hat{\mu}_i} = \frac{N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}. \tag{4.30}$$

Using expression (4.30) one can estimate the summation of the scale parameters for all transitions that exit from any state i. To estimate individual scale parameters a_{ij} for each $\{i, j\}$, an additional expression can be obtained in the following way,

$$\begin{aligned} p_{ik} &= \lim_{t\to\infty} Q_{ik}(t) \\ &= \lim_{t\to\infty} \frac{1}{a_{ik}\sum\limits_{j=1}^{N}\frac{1}{a_{ij}}}\left[1 - e^{-t\sum\limits_{j=1}^{N}\frac{1}{a_{ij}}}\right] \\ &= \frac{1}{a_{ik}\sum\limits_{j=1}^{N}\frac{1}{a_{ij}}}. \end{aligned} \tag{4.31}$$

Hence,

$$a_{ik} = \frac{1}{p_{ik}\sum\limits_{j=1}^{N}\frac{1}{a_{ij}}}. \tag{4.32}$$

Substituting estimators (4.1) and (4.30) into expression (4.32) the following estimators will be obtained for the transition intensity:

$$\begin{aligned} \hat{a}_{ik} &= \frac{1}{\frac{N_{ik}(M)}{N_i(M)}\widehat{\sum\limits_{j=1}^{N}\frac{1}{a_{ij}}}} \\ &= \frac{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}{N_{ik}(M)}, \quad i, k = 1, \ldots, N. \end{aligned} \tag{4.33}$$

Andreas Makrides

The reciprocal of the scale parameter coincides with the transition rate, namely

$$\begin{aligned}
\hat{\lambda}_{ik}^{EXP} &= \frac{1}{\hat{a}_{ik}} \\
&= \left(\frac{\sum_{j=1}^{N_i(M)} T_i^{(j)}}{N_{ik}(M)} \right)^{-1} \\
&= \frac{N_{ik}(M)}{\sum_{j=1}^{N_i(M)} T_i^{(j)}}. \qquad (4.34)
\end{aligned}$$

For a Markov MSS with N states the sum $\sum_{j=1}^{N} \lambda_{ij} = 0$, therefore

$$\hat{\lambda}_{ii} = -\sum_{j \neq i} \hat{\lambda}_{ij}. \qquad (4.35)$$

In conclusion, the following theorem holds:

Theorem 4.2.1 *Consider the Exponential model given in (4.25) for the sojourn times in an MS system. Then, the estimators of the scale parameters a_{ik}, $i, k = 1, 2, \ldots, N$ are given by:*

$$\hat{a}_{ik} = \frac{\sum_{j=1}^{N_i(M)} T_i^{(j)}}{N_{ik}(M)}, \quad i, k = 1, \ldots, N. \qquad (4.36)$$

while the transition rate is given by:

$$\hat{\lambda}_{ik}^{EXP} = \frac{N_{ik}(M)}{\sum_{j=1}^{N_i(M)} T_i^{(j)}}. \qquad (4.37)$$

Andreas Makrides

4.3 Rayleigh Model

In this section we assume that $F_{ij}(t)$ is the cumulative distribution function of Rayleigh distribution

$$F_{ij}(t) = 1 - e^{-\left(t^2/2a_{ij}^2\right)}, \tag{4.38}$$

with intensity rate given by

$$\lambda_{ij}(t) = \frac{2}{\sqrt{2}a_{ij}}\left(\frac{t}{\sqrt{2}a_{ij}}\right). \tag{4.39}$$

According to the previous section,

$$\begin{aligned}
Q_{ik}(t) &= \int_0^t \exp\left\{-\frac{x^2}{2a_{i1}^2} - \cdots - \frac{x^2}{2a_{ik-1}^2} - \frac{x^2}{2a_{ik+1}^2} - \cdots - \frac{x^2}{2a_{iN}^2}\right\} \times \\
&\times \frac{2x}{2a_{ik}^2}\exp\left\{-\frac{x^2}{2a_{ik}^2}\right\} dx \\
&= \int_0^t \frac{1}{a_{ik}^2}\exp\left\{-\frac{x^2}{2}\sum_{j=1}^{N}\frac{1}{a_{ij}^2}\right\} d\left(\frac{x^2}{2}\right) \\
&= \frac{1}{a_{ik}^2\sum\limits_{j=1}^{N}\frac{1}{a_{ij}^2}}\left[1 - e^{-\frac{t^2}{2}\sum\limits_{j=1}^{N}\frac{1}{a_{ij}^2}}\right].
\end{aligned} \tag{4.40}$$

Hence,

$$\begin{aligned}
W_i(t) &= \sum_{k=1}^{N} Q_{ik}(t) \\
&= \frac{\sum\limits_{k=1}^{N}\frac{1}{a_{ik}^2}}{\sum\limits_{j=1}^{N}\frac{1}{a_{ij}^2}}\left[1 - e^{-\frac{t^2}{2}\sum\limits_{j=1}^{N}\frac{1}{a_{ij}^2}}\right] \\
&= 1 - e^{-\frac{t^2}{2}\sum\limits_{j=1}^{N}\frac{1}{a_{ij}^2}} \\
&= 1 - e^{-t^2/2\left(\left(\sum\limits_{j=1}^{N}\frac{1}{a_{ij}^2}\right)^{-1/2}\right)^2}
\end{aligned} \tag{4.41}$$

Andreas Makrides

which is a Rayleigh distribution function with parameter

$$\left(\sum_{j=1}^{N} \frac{1}{a_{ij}^2}\right)^{-1/2}$$

and mean equal to

$$\mu_i = \left(\sum_{j=1}^{N} \frac{1}{2a_{ij}^2}\right)^{-1/2} \cdot \Gamma\left(1 + \frac{1}{2}\right). \tag{4.42}$$

By using the estimation of the mean (4.42) one can obtain

$$\frac{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}{N_i(M)} = \left(\sum_{j=1}^{N} \frac{1}{2a_{ij}^2}\right)^{-1/2} \cdot \Gamma\left(1 + \frac{1}{2}\right).$$

Equivalently,

$$\left(\sum_{j=1}^{N} \frac{1}{2a_{ij}^2}\right)^{-1/2} = \frac{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}{\Gamma\left(1 + \frac{1}{2}\right) N_i(M)}.$$

Hence,

$$\sum_{j=1}^{N} \frac{1}{2a_{ij}^2} = \left[\frac{\Gamma\left(1 + \frac{1}{2}\right) N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}\right]^2. \tag{4.43}$$

Andreas Makrides

To estimate individual transition intensities we have,

$$\begin{aligned}
p_{ik} &= \lim_{t\to\infty} Q_{ik}(t) \\
&= \lim_{t\to\infty} \frac{1}{a_{ik}^2 \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^2}} \left[1 - e^{-\frac{t^2}{2} \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^2}} \right] \\
&= \frac{1}{a_{ik}^2 \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^2}}. \qquad (4.44)
\end{aligned}$$

Thus,

$$a_{ik}^2 = \frac{1}{p_{ik} \sum\limits_{j=1}^{N} \frac{1}{a_{ij}^2}}. \qquad (4.45)$$

Substituting equation (4.43) we obtain,

$$\begin{aligned}
\hat{a}_{ik}^2 &= \frac{1}{\frac{N_{ik}(M)}{N_i(M)} \widehat{\sum\limits_{j=1}^{N} \frac{1}{a_{ij}^2}}} \\
&= \frac{1}{\frac{2N_{ik}(M)}{N_i(M)} \widehat{\sum\limits_{j=1}^{N} \frac{1}{2a_{ij}^2}}} \\
&= \frac{1}{\frac{2N_{ik}(M)}{N_i(M)} \left[\frac{\Gamma\left(1+\frac{1}{2}\right) N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}} \right]^2}. \qquad (4.46)
\end{aligned}$$

Based on the above results, the following theorem holds:

Theorem 4.3.1 *Consider the Rayleigh model given in (4.38) for the sojourn times in an MS system. Then, the estimators of the scale parameters* a_{ik}*,* $i, k = 1, 2, \ldots, N$ *are given*

Andreas Makrides

by:

$$\hat{a}_{ik} = \frac{1}{\left(\frac{2N_{ik}(M)}{N_i(M)}\right)^{1/2}\left[\frac{\Gamma\left(1+\frac{1}{2}\right)N_i(M)}{\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}\right]}, \quad i,k = 1,\ldots,N, \quad i \neq k. \tag{4.47}$$

4.3.1 Estimation of the transition intensities

In this case with $c = 2$ and $\sqrt{2}a_{ik}$ in place of a_{ik} in (4.7), we have for $i,k = 1,\ldots,N$ with $i \neq j$, that the estimator of the intensity rate (4.39) is obtained analogously as in (4.23), namely

$$\hat{\lambda}_{ik}(t) = \frac{t}{\hat{a}_{ik}^2}. \tag{4.48}$$

and

$$\hat{\lambda}_{ii}(t) = -\sum_{k \neq i} \hat{\lambda}_{ik}(t).$$

According to the previous section,

$$\hat{\lambda}_{ik} = \frac{1}{\sqrt{2}\hat{a}_{ik}}\frac{2}{\Gamma(1/2)} = \sqrt{\frac{2}{\pi}}\frac{1}{\hat{a}_{ik}}$$

which by substituting (4.47) takes the form

$$\hat{\lambda}_{ik}^{RAL} = \left(\frac{2N_{ik}(M)}{N_i(M)}\right)^{1/2} \frac{\Gamma\left(1+\frac{1}{2}\right)\sqrt{2}N_i(M)}{\Gamma\left(\frac{1}{2}\right)\sum\limits_{j=1}^{N_i(M)} T_i^{(j)}}.$$

4.4 The Pareto Model

The Pareto distribution is also frequently encountered in reliability and engineering applications. Thus, in this subsection we assume that $F_{ij}(t)$ is the cumulative distribution

Andreas Makrides

function of the Pareto distribution, given by

$$F_{ij}\left(t\right)=1-\left(\frac{\lambda}{t}\right)^{a_{ij}},t\geq\lambda \tag{4.49}$$

where λ is the scale parameter and a_{ij} is the shape parameter. Then, the intensity rate (hazard rate) of the sojourn time is given by

$$\lambda_{ij}(t)=\frac{\frac{a_{ij}\lambda^{a_{ij}}}{t^{a_{ij}+1}}}{\frac{\lambda^{a_{ij}}}{t^{a_{ij}}}}=\frac{a_{ij}}{t} \tag{4.50}$$

where the pdf is given by

$$f_{ij}(t)=\frac{a_{ij}\lambda^{a_{ij}}}{t^{a_{ij}+1}}.$$

Following the same procedure as in all previous cases, we have

$$\begin{aligned}
Q_{ik}\left(t\right) &= \int_{\lambda}^{t}\frac{\lambda^{a_{i1}}}{x^{a_{i1}}}\cdots\frac{\lambda^{a_{i,k-1}}}{x^{a_{i,k-1}}}\cdot\frac{\lambda^{a_{i,k+1}}}{x^{a_{i,k+1}}}\cdots\frac{\lambda^{a_{iN}}}{x^{a_{iN}}}\cdot a_{ik}\frac{\lambda^{a_{ik}}}{x^{a_{ik}+1}}dx \\
&= a_{ik}\cdot\lambda^{\sum_{j=1}^{N}a_{ij}}\int_{\lambda}^{t}\frac{1}{x^{\sum_{j=1}^{N}a_{ij}+1}}dx \\
&= \frac{a_{ik}}{\sum_{j=1}^{N}a_{ij}}-\frac{a_{ik}}{\sum_{j=1}^{N}a_{ij}}\cdot\frac{\lambda^{\sum_{j=1}^{N}a_{ij}}}{t^{\sum_{j=1}^{N}a_{ij}}}.
\end{aligned} \tag{4.51}$$

Furthermore, we also have

$$\begin{aligned}
W_i(t) &= \sum_{k=1}^{N}Q_{ik}\left(t\right) \\
&= \frac{\sum_{k=1}^{N}a_{ik}}{\sum_{j=1}^{N}a_{ij}}-\frac{\sum_{k=1}^{N}a_{ik}}{\sum_{j=1}^{N}a_{ij}}\cdot\frac{\lambda^{\sum_{j=1}^{N}a_{ij}}}{t^{\sum_{j=1}^{N}a_{ij}}} \\
&= =1-\left(\frac{\lambda}{t}\right)^{\sum_{j=1}^{N}a_{ij}}
\end{aligned}$$

which is a Pareto distribution function with scale parameter λ and shape parameter $\sum_{j=1}^{N}a_{ij}$

Andreas Makrides

with mean

$$\mu_i = \frac{\lambda \sum_{j=1}^{N} a_{ij}}{\sum_{j=1}^{N} a_{ij} - 1}, \quad \text{for} \quad \sum_{j=1}^{N} a_{ij} > 1.$$

Observe that the distribution of the time till the first transit from state i is a Pareto distribution when the sojourn time T_{ij} from state i to j follows a Pareto distribution. Then, the standard moment estimator for the mean, gives

$$\frac{\sum_{j=1}^{k_i} T_i^{(j)}}{N_i(M)} = \frac{\lambda \sum_{j=1}^{N} a_{ij}}{\sum_{j=1}^{N} a_{ij} - 1}.$$

Hence,

$$\sum_{j=1}^{N} a_{ij} = \frac{\sum_{j=1}^{N_i(M)} T_i^{(j)}}{\sum_{j=1}^{N_i(M)} T_i^{(j)} - \lambda N_i(M)}. \tag{4.52}$$

For the individual shape parameters, consider (2.2) which in this case is given by

$$\begin{aligned} p_{ik} &= \lim_{t\to\infty} Q_{ik}(t) \\ &= \lim_{t\to\infty} \left(\frac{a_{ik}}{\sum_{j=1}^{N} a_{ij}} - \frac{a_{ik}}{\sum_{j=1}^{N} a_{ij}} \cdot \frac{\lambda^{\sum_{j=1}^{N} a_{ij}}}{t^{\sum_{j=1}^{N} a_{ij}}} \right) \\ &= \frac{a_{ik}}{\sum_{j=1}^{N} a_{ij}}, \end{aligned} \tag{4.53}$$

so that combining (4.52) and (4.53) we obtain the result below:

Theorem 4.4.1 *Consider the Pareto model given in (4.49) for the sojourn times in a MS* *system. Then, the estimators of the shape parameters* a_{ik}*,* $i, k = 1, 2, \ldots, N$ *are given by:*

$$\hat{a}_{ik} = \frac{N_{ik}(M)}{N_i(M)} \cdot \frac{\sum_{j=1}^{N_i(M)} T_i^{(j)}}{\sum_{j=1}^{N_i(M)} T_i^{(j)} - \lambda N_i(M)}.$$

If the scale parameter, λ, is not given then one estimator is given by,

$$\hat{\lambda} = X_{(1)} \tag{4.54}$$

Using the above formula for $\hat{a}_{ik}$ together with (4.50) we obtain the transition rate from i to k ($i \neq k$) for the Pareto distribution:

$$\hat{\lambda}_{ik}^{PAR}(t) = \frac{\hat{a}_{ik}}{t}$$

with

$$\hat{\lambda}_{ii}(t) = -\sum_{k \neq i} \hat{\lambda}_{ik}(t).$$

4.4.1 Homogeneous transition rates

In the case of the Pareto distribution the integral of the hazard function is,

$$\begin{aligned} \int_{\lambda}^{t} \frac{a_{ik}}{\tau}\, d\tau &= a_{ik}\,(\ln t - \ln \lambda) \\ &= \ln t^{a_{ik}} - \ln \lambda^{a_{ik}} = \ln \left(\frac{t}{\lambda}\right)^{a_{ik}}. \end{aligned}$$

Hence,

$$\begin{aligned} \lambda_{ik} &= \frac{1}{\int_{\lambda}^{\infty} e^{-(\ln t^{a_{ik}} - \ln \lambda^{a_{ik}})} dt} \\ &= \frac{1}{\int_{\lambda}^{\infty} \frac{\lambda^{a_{ik}}}{t^{a_{ik}}} dt} \\ &= \frac{a_{ik} - 1}{\lambda}, \quad \text{if} \quad a_{ik} > 1, \end{aligned}$$

Andreas Makrides

or

$$\lambda_{ik} = \infty, \quad \text{if} \quad a_{ik} \leq 1,$$

which is unusable.

So, if we are in the case where $\hat{a}_{ik} > 1$, the homogeneous transition intensity estimator is

$$\hat{\lambda}_{ik} = \frac{\hat{a}_{ik} - 1}{\hat{\lambda}}. \tag{4.55}$$

Remarks 4.4.1

1. In the present thesis we work with models of order 1. This is due to the fact that the Markov property which is associated with a first order model is very tractable mathematically and quite easy to present and explain. Of course it is equally interesting to explore models of a higher order. Formally, a suitable testing procedure could first be applied in order to determine whether the order of the model is equal to or greater than 1. It should be stated that in addition to classical tests based on the chi-squared distribution, model selection criteria like Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) could also be used. Both techniques use the likelihood ratio statistics and modify it by a penalty term. The model with the lowest value is to be preferred. Ideally we would get suggestions for the same order using both methods which is the case for a significant amount of datasets. In reality though, the disagreement between the two criteria is common with some instances providing ambiguous and non-interpretable results.

   In the case of an order greater than 1, the researcher may choose one of two directions for the statistical analysis:

   (a) Transform the original model to a model of order 1 by for instance, redefining the state space and work as we have done in this work or

   (b) Perform further analysis to explore and capture the effect of a high order. Indeed, although in such a case the exponential distribution is not appropriate, a formal investigation is required to determine, among other issues, the appro-

Andreas Makrides

priateness of the Weibull distribution.

2. The case of infinite number of performance rates may be considered although some problems can arise when working with infinite matrices; for instance, the product of two infinite matrices with arbitrary real entries is not always well defined, the matrix multiplication is not always associative, the inverse of an invertible infinite matrix is not uniquely defined. For work with infinite matrices for semi-Markov chains one may see Barbu and Limnios (2010). For works on infinite matrices for Markov chains see Kemeny et al. (1976), Seneta (1981) and Baldi et al. (2002) and for general reference for infinite matrices see Cooke (1955).

Andreas Makrides

Chapter 5

Maximum Likelihood Estimation

In this chapter we consider two statistical settings: in the first one, presented in Subsection 5.1, we dispose of one sample path of the system; in the second one, described in Subsection 5.2, several sample paths are available. On each situation we take into account two different cases: in the first case, we observe all the sojourn times; in the second one, the sojourn time in the last visited state can be right censored (lost to follow-up, for instance). For parameter estimation in this chapter we apply the maximum likelihood approach. Asymptotic results about the estimators are provided in Subsection 5.6.

5.1 Maximum Likelihood Estimation for one Trajectory

First, for any states $i, j \in E$ and $t \in \mathbb{R}_+$, let us introduce the counting processes $(N_i(t))_{t\geq 0}$, where $\forall t \geq 0$, $N_i(t)$ is the number of visits to state i of $(J_n)_{n\in\mathbb{N}}$ up to time t and $(N_{ij}(t))_{t\geq 0}$, where $\forall t \geq 0$, $N_{ij}(t)$ is the number of jumps of $(J_n)_{n\in\mathbb{N}}$ from state i to state j up to time t :

$$\begin{aligned} N_i(t) &:= \sum_{n=0}^{N(t)-1} \mathbb{1}_{\{J_n=i\}} = \sum_{n=0}^{\infty} \mathbb{1}_{\{J_n=i,\, S_{n+1}\leq t\}}, \\ N_{ij}(t) &:= \sum_{n=1}^{N(t)} \mathbb{1}_{\{J_{n-1}=i,\, J_n=j\}} = \sum_{n=1}^{\infty} \mathbb{1}_{\{J_{n-1}=i,\, J_n=j,\, S_n\leq t\}}, \end{aligned}$$

Andreas Makrides

where $N(t)$ was defined in Equation (2.1).

Recall that M is the total observation time. We consider first a sample path of a semi-Markov process $\left\{j_0, x_1, j_1, x_2, \ldots, j_{N(M)}\right\}$; the associated likelihood for 1 trajectory is

$$\begin{aligned}
\mathcal{L} &= \alpha_{j_0} p_{j_0 j_1} f_{j_0}(x_1) \ldots p_{j_{N(M)-1}, j_{N(M)}} f_{j_{N(M)-1}}(x_{N(M)}) \\
&= \alpha_{j_0} \left(\prod_{i,j \in E} p_{ij}^{N_{ij}(M)} \right) \left(\prod_{i \in E} \prod_{k=1}^{N_i(M)} f_i(x_i^{(1,k)}) \right),
\end{aligned} \tag{5.1}$$

where $x_i^{(1,k)}$ is the sojourn time in state i during the k^{th} visit, $k = 1, \ldots, N_i(M)$.

We wish now to calculate the likelihood in the case of the class of distributions given in (3.2) and provide the maximum likelihood estimators of the associated parameters. Then, the corresponding estimates of Q, p, W, and f will be easily obtained via expressions (3.13) – (3.16).

Using the previous results, the likelihood for this uncensored trajectory given in (5.1) takes the form

$$\mathcal{L} = \alpha_{j_0} \left(\prod_{i,j \in E} a_{ij}^{N_{ij}(M)} \right) \prod_{i,k} \left\{ \left(1 - F\left(x_i^{(1,k)}\right)\right)^{\sum\limits_{j \in E} a_{ij}} \frac{f\left(x_i^{(1,k)}\right)}{1 - F\left(x_i^{(1,k)}\right)} \right\}.$$

Therefore,

$$\begin{aligned}
\log \mathcal{L} &= \log \alpha_{j_0} + \sum_{i,j \in E} N_{ij}(M) \log a_{ij} \\
&+ \sum_{i \in E} \left(\sum_{j \in E} a_{ij} \right) \sum_{k=1}^{N_i(M)} \log \left(1 - F\left(x_i^{(1,k)}\right)\right) \\
&+ \log \left(\prod_{i \in E} \prod_{k=1}^{N_i(M)} \frac{f\left(x_i^{(1,k)}\right)}{1 - F\left(x_i^{(1,k)}\right)} \right).
\end{aligned} \tag{5.2}$$

Andreas Makrides

By taking the derivatives of the log-likelihood w.r.t. $a_{ij}, i, j \in E$, we get

$$\frac{\partial \log \mathcal{L}}{\partial a_{ij}} = \frac{N_{ij}(M)}{a_{ij}} + \sum_{k=1}^{N_i(M)} \log\left(1 - F\left(x_i^{(1,k)}\right)\right) \equiv 0 \tag{5.3}$$

and

$$\frac{\partial^2 \log \mathcal{L}}{\partial a_{ij}^2} = -\frac{N_{ij}(M)}{a_{ij}^2} < 0. \tag{5.4}$$

Finally, the estimator of a_{ij} is given by

$$\widehat{a}_{ij}(M) = -\frac{N_{ij}(M)}{\sum\limits_{k=1}^{N_i(M)} \log\left(1 - F\left(X_i^{(1,k)}\right)\right)}. \tag{5.5}$$

Let us now provide the general form of the likelihood in the case of censoring at time M. We consider $\left\{j_0, x_1, j_1, x_2, \ldots, j_{N(M)}, u_M\right\}$ a censored sample path, where

$$u_M := M - S_{N(M)}$$

is the last sojourn time that is considered to be censored. Then, the likelihood is given by

$$\begin{aligned}\mathcal{L} &= \alpha_{j_0} p_{j_0 j_1} f_{j_0}(x_1) \ldots f_{j_{N(M)-1}}(x_{N(M)}) \times \mathbb{P}\left(X_{j_{N(M)}} > u_M\right) \\ &= \alpha_{j_0} \left(\prod_{i,j \in E} p_{ij}^{N_{ij}(M)}\right) \left(\prod_{i \in E} \prod_{k=1}^{N_i(M)} f_i(x_i^{(1,k)})\right) \times \\ &\times \left(1 - W_{j_{N(M)}}(u_M)\right). \end{aligned} \tag{5.6}$$

For the class of distributions considered in (3.2), this likelihood takes the form

$$\begin{aligned}\mathcal{L} &= \alpha_{j_0} \left(\prod_{i,j \in E} a_{ij}^{N_{ij}(M)}\right) \prod_{i,k} \left\{\left(1 - F\left(x_i^{(1,k)}\right)\right)^{\sum\limits_{j \in E} a_{ij}} \frac{f\left(x_i^{(1,k)}\right)}{1 - F\left(x_i^{(1,k)}\right)}\right\} \\ &\times \left(\left(1 - F\left(u_M\right)\right)^{\sum\limits_{j \in E} a_{ij}}\right). \end{aligned}$$

Andreas Makrides

Hence,

$$
\begin{aligned}
\log \mathcal{L} &= \log \alpha_{j_0} + \sum_{i,j \in E} N_{ij}(M) \log a_{ij} \\
&+ \sum_{i \in E} \left(\sum_{j \in E} a_{ij} \right) \sum_{k=1}^{N_i(M)} \log \left(1 - F \left(x_i^{(1,k)} \right) \right) \\
&+ \log \left(\prod_{i \in E} \prod_{k=1}^{N_i(M)} \frac{f \left(x_i^{(1,k)} \right)}{1 - F \left(x_i^{(1,k)} \right)} \right) + \sum_{i,j \in E} a_{ij} \log \left(1 - F \left(u_M \right) \right). \qquad (5.7)
\end{aligned}
$$

By taking the derivative of the log-likelihood w.r.t. $a_{ij}, i, j \in E$, we get

$$
\begin{aligned}
\frac{\partial \log \mathcal{L}}{\partial a_{ij}} &= \frac{N_{ij}(M)}{a_{ij}} + \sum_{k=1}^{N_i(M)} \log \left(1 - F \left(x_i^{(1,k)} \right) \right) \\
&+ \log \left(1 - F \left(u_M \right) \right) \equiv 0. \qquad (5.8)
\end{aligned}
$$

Since

$$
\frac{\partial^2 \log \mathcal{L}}{\partial a_{ij}^2} = -\frac{N_{ij}(M)}{a_{ij}^2} < 0, \qquad (5.9)
$$

the maximum likelihood estimator of a_{ij} for the censored case is given by

$$
\widehat{a}_{ij}(M) = -\frac{N_{ij}(M)}{B_i(M) + C(M)}, \qquad (5.10)
$$

where we introduced

$$
B_i(M) := \sum_{k=1}^{N_i(M)} \log \left(1 - F \left(X_i^{(1,k)} \right) \right)
$$

and

$$
C(M) := \log \left(1 - F \left(U_M \right) \right).
$$

Andreas Makrides

5.2 Maximum Likelihood Estimation for Several Trajectories

In a similar way we calculate the general form of the likelihood formula for L trajectories.

First, we consider the no censoring case. Given L sample paths of a semi-Markov process, $\left\{j_0^{(l)}, x_1^{(l)}, j_1^{(l)}, x_2^{(l)}, \ldots, j_{N^l(M)}^{(l)}\right\}$, $l = 1, \ldots, L$, the associated likelihood is

$$
\begin{aligned}
\mathcal{L} &= \prod_{l=1}^{L} \alpha_{j_0}^{(l)} p_{j_0^{(l)} j_1^{(l)}} f_{j_0^{(l)}}(x_1^{(l)}) \ldots f_{j^{(l)}_{N^l(M)-1}}(x^{(l)}_{N^l(M)}) \qquad (5.11)\\
&= \left(\prod_{i \in E} \alpha_i^{N_{i,0}^{(L)}}\right) \left(\prod_{i,j \in E} p_{ij}^{\sum_{l=1}^{L} N_{ij}^{(l)}(M)}\right) \times \left(\prod_{l=1}^{L} \prod_{i \in E} \prod_{k=1}^{N_i^{(l)}(M)} f_i(x_i^{(l,k)})\right),
\end{aligned}
$$

where

- $N_{i,0}^{(L)} := \sum_{l=1}^{L} \mathbb{1}_{\{J_0^{(l)} = i\}}$,
- $N_i^{(l)}(M)$: the number of visits to state i up to time M of the l^{th} trajectory, $l = 1, \ldots, L$,
- $N_{ij}^{(l)}(M)$: the number of transitions from state i to state j up to time M during the l^{th} trajectory, $l = 1, \ldots, L$,
- $N_{ij}(L, M) := \sum_{l=1}^{L} N_{ij}^{(l)}(M)$,
- $x_i^{(l,k)}$: the sojourn time in state i during the k^{th} visit, $k = 1, \ldots, N_i^{(l)}(M)$ of the l^{th} trajectory, $l = 1, \ldots, L$.

Note that for $L = 1$ the likelihood (5.11) reduces to the likelihood of the 1 trajectory case given in (5.1).

We give now the likelihood in the case of the class of distributions given in (3.2) and provide the maximum likelihood estimators of the associated parameters.

Andreas Makrides

The likelihood for L uncensored trajectories given in (5.11) takes the form

$$
\begin{aligned}
\mathcal{L} &= \prod_{i\in E} \alpha_i^{N_{i,0}^{(L)}} \left(\prod_{l=1}^{L} \prod_{i,j\in E} a_{ij}^{N_{ij}^{(l)}(M)} \right) \\
&\times \prod_{l,i,k} \left\{ \left(1 - F\left(x_i^{(l,k)}\right)\right)^{\sum\limits_{j\in E} a_{ij}} \frac{f\left(x_i^{(l,k)}\right)}{1 - F\left(x_i^{(l,k)}\right)} \right\}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\log \mathcal{L} &= \sum_{i\in E} N_{i,0}^{(L)} \log \alpha_i + \sum_{l=1}^{L} \sum_{i,j\in E} N_{ij}^{(l)}(M) \log a_{ij} \\
&+ \sum_{l=1}^{L} \sum_{i\in E} \left(\sum_{j\in E} a_{ij} \right) \sum_{k=1}^{N_i^{(l)}(M)} \log\left(1 - F\left(x_i^{(l,k)}\right)\right) \\
&+ \log \left(\prod_{l=1}^{L} \prod_{i\in E} \prod_{k=1}^{N_i^{(l)}(M)} \frac{f\left(x_i^{(l,k)}\right)}{1 - F\left(x_i^{(l,k)}\right)} \right).
\end{aligned} \tag{5.12}
$$

By taking the derivative w.r.t. a_{ij}, we get

$$
\frac{\partial \log \mathcal{L}}{\partial a_{ij}} = \sum_{l=1}^{L} \frac{N_{ij}^{(l)}(M)}{a_{ij}} + \sum_{l=1}^{L} \sum_{k=1}^{N_i^{(l)}(M)} \log\left(1 - F\left(x_i^{(l,k)}\right)\right) \equiv 0. \tag{5.13}
$$

Since

$$
\frac{\partial^2 \log \mathcal{L}}{\partial a_{ij}^2} = -\frac{N_{ij}(L,M)}{a_{ij}^2} < 0, \tag{5.14}
$$

we obtain the estimator of a_{ij} given by

$$
\widehat{a}_{ij}(L,M) = -\frac{N_{ij}(L,M)}{\sum\limits_{l=1}^{L} \sum\limits_{k=1}^{N_i^{(l)}(M)} \log\left(1 - F\left(X_i^{(l,k)}\right)\right)}. \tag{5.15}
$$

Taking the derivative w.r.t. α_i, and taking into account the fact that $\sum\limits_{i\in E} \alpha_i = 1$, we obtain the MLE of the initial distribution given by

$$
\widehat{\alpha}_i(L,M) = \frac{N_{i,0}^{(L)}}{L}. \tag{5.16}
$$

Andreas Makrides

Let us now consider the case of L trajectories with censoring at time M. Given L censored sample paths of a semi-Markov process,
$\left\{j_0^{(l)}, x_1^{(l)}, j_1^{(l)}, x_2^{(l)}, \ldots, j_{N^l(M)}^{(l)}, u_M^{(l)}\right\}$, $l = 1, \ldots, L$, the associated likelihood is

$$\begin{aligned}
\mathcal{L} &= \left(\prod_{i \in E} \alpha_i^{N_{i,0}(L)}\right) \left(\prod_{i,j \in E} p_{ij}^{\sum\limits_{l=1}^{L} N_{ij}^{(l)}(M)}\right) \times \\
&\times \left(\prod_{l=1}^{L} \prod_{i \in E} \prod_{k=1}^{N_i^{(l)}(M)} f_i(x_i^{(l,k)})\right) \prod_{i \in E} \prod_{k=1}^{N_{i,M}(L)} \left(1 - W_i(u_i^{(k)})\right), \qquad (5.17)
\end{aligned}$$

where

- $u_M^{(l)} := M - S_{N^l(M)}$ is the observed censored time of the l^{th} trajectory;

- $N_{i,M}(L) = \sum\limits_{l=1}^{L} \mathbb{1}_{\{J_{N^l(M)}^{(l)} = i\}}$ is the number of visits of state i, as last visited state, over the L trajectories; note that $\sum\limits_{i \in E} N_{i,M}(L) = L$;

- $u_i^{(k)}$ is the observed censored sojourn time in state i during the k^{th} visit, $k = 1, \ldots, N_{i,M}(L)$.

Note that, if the censoring time M in a certain trajectory l is a jump time, then for the corresponding observed censored time we have that $u_M^{(l)} = 0$. Consequently, the contribution to the likelihood of the associated term will be equal to 1. For this reason, if no censoring is involved, the uncensored likelihood given in (5.11) is just a particular case of (5.17). Note also that for $L = 1$ the above likelihood expression given in (5.17) becomes the likelihood of the 1 trajectory case with censoring, given in (5.6).

For the class of distributions given in (3.2), the likelihood for the case of L trajectories

Andreas Makrides

with censoring given in (5.17) takes the form

$$
\begin{aligned}
\mathcal{L} &= \left(\prod_{i\in E}\alpha_i^{N_{i,0}^{(L)}}\right)\left(\prod_{l=1}^{L}\prod_{i,j\in E}a_{ij}^{N_{ij}^{(l)}(M)}\right)\times \\
&\quad\times\prod_{l,i,k}\left[\left(1-F\left(x_i^{(l,k)}\right)\right)^{\sum\limits_{j\in E}a_{ij}}\left(\frac{f\left(x_i^{(l,k)}\right)}{1-F\left(x_i^{(l,k)}\right)}\right)\right]\times \\
&\quad\times\left(\prod_{i\in E}\prod_{k=1}^{N_{i,M}(L)}\left(1-F\left(u_i^{(k)}\right)\right)^{\sum\limits_{j\in E}a_{ij}}\right). \qquad (5.18)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\log\mathcal{L} &= \sum_{i\in E}N_{i,0}^{(L)}\log\alpha_i+\sum_{l=1}^{L}\sum_{i,j\in E}N_{ij}^{(l)}(M)\log a_{ij} \\
&+ \sum_{l=1}^{L}\sum_{i\in E}\left(\sum_{j\in E}a_{ij}\right)\sum_{k=1}^{N_i^{(l)}(M)}\log\left(1-F\left(x_i^{(l,k)}\right)\right) \\
&+ \log\left(\prod_{l=1}^{L}\prod_{i\in E}\prod_{k=1}^{N_i^{(l)}(M)}\frac{f\left(x_i^{(l,k)}\right)}{1-F\left(x_i^{(l,k)}\right)}\right) \\
&+ \sum_{i,j\in E}a_{ij}\sum_{k=1}^{N_{i,M}(L)}\log\left(1-F\left(u_i^{(k)}\right)\right). \qquad (5.19)
\end{aligned}
$$

By taking the derivatives w.r.t. a_{ij}, we get

$$
\begin{aligned}
\frac{\partial\log\mathcal{L}}{\partial a_{ij}} &= \sum_{l=1}^{L}\frac{N_{ij}^{(l)}(M)}{a_{ij}}+\sum_{k=1}^{N_{i,M}(L)}\log\left(1-F\left(u_i^{(k)}\right)\right) \\
&+ \sum_{l=1}^{L}\sum_{k=1}^{N_i^{(l)}(M)}\log\left(1-F\left(x_i^{(l,k)}\right)\right)\equiv 0. \qquad (5.20)
\end{aligned}
$$

and

$$
\frac{\partial^2\log\mathcal{L}}{\partial a_{ij}^2}=-\frac{N_{ij}(L,M)}{a_{ij}^2}<0, \qquad (5.21)
$$

Andreas Makrides

Consequently, we obtain that the estimator of a_{ij} in this case is given by

$$\widehat{a}_{ij}(L,M) = -\frac{N_{ij}(L,M)}{B_i(L,M)+C(L,M)}, \tag{5.22}$$

where

$$B_i(L,M) := \sum_{l=1}^{L} \sum_{k=1}^{N_i^{(l)}(M)} \log\left(1-F\left(X_i^{(l,k)}\right)\right)$$

and

$$C(L,M) := \sum_{k=1}^{N_{i,M}(L)} \log\left(1-F\left(U_i^{(k)}\right)\right).$$

As for the initial law, the estimator as before is given by

$$\widehat{\alpha}_i(L,M) = \frac{N_{i,0}^{(L)}}{L}. \tag{5.23}$$

Among the estimators of a_{ij} given in (5.5), (5.10), (5.15) and (5.22), one should choose, according to the case under investigation, the appropriate expression for the determination of the estimators of p_{ij}, W_i and Q_{ij}. The following result is immediate:

Theorem 5.2.1 *For all $i,j \in E$, the estimators $\widehat{p}_{ij}(M)$ of p_{ij}, $\widehat{Q}_{ij}(t,M)$ of Q_{ij} and $\widehat{W}_i(t,M)$ are given by*

$$\widehat{p}_{ij}(M) = \frac{\widehat{a}_{ij}(L,M)}{\sum\limits_{l\in E}\widehat{a}_{il}(L,M)} = \frac{N_{ij}(M)}{N_i(M)}, \tag{5.24}$$

$$\widehat{W}_i(t,M) = \left[1-(1-F(t))^{\sum\limits_{j\in E}\widehat{a}_{ij}(L,M)}\right] \tag{5.25}$$

and

$$\widehat{Q}_{ij}(t,M) = \frac{\widehat{a}_{ij}(L,M)}{\sum\limits_{k\in E}\widehat{a}_{ik}(L,M)}\left[1-(1-F(t))^{\sum\limits_{k\in E}\widehat{a}_{ik}(L,M)}\right]. \tag{5.26}$$

Andreas Makrides

5.3 Parameter Estimation for Special Distributions

In the case of the *Weibull distribution* we have

$$F(x;a_{ij}) = 1 - e^{-a_{ij}x^c}, \quad F(x;1) = 1 - e^{-x^c}, \quad f(x;1) = cx^{c-1}e^{-x^c},$$

where c is the shape parameter. The estimator of the scale parameter in the uncensored case takes the form

$$\widehat{a}_{ij}(L,M) = \frac{N_{ij}(L,M)}{\sum_{l=1}^{L}\sum_{k=1}^{N_i^{(l)}(M)}\left(X_i^{(l,k)}\right)^c} \tag{5.27}$$

and in the censored case becomes

$$\widehat{a}_{ij}(L,M) = \frac{N_{ij}(L,M)}{\sum_{l=1}^{L}\sum_{k=1}^{N_i^{(l)}(M)}\left(X_i^{(l,k)}\right)^c + \sum_{k=1}^{N_{i,M}(L)}\left(U_i^{(k)}\right)^c}. \tag{5.28}$$

For the maximum likelihood estimator of the shape parameter c, one should solve the equation bellow, following the form of likelihood (5.11)

$$\frac{\partial \log \mathcal{L}}{\partial c} = \sum_{l,i,j}\left(\frac{N_{ij}(L,M)}{\sum_{l,k}\left(x_i^{(l,k)}\right)^c}\right)\sum_{k=1}^{N_i^{(l)}(M)}\left(x_i^{(l,k)}\right)^c \log x_i^{(l,k)} + \frac{\sum_{l,i,k}\left(1 + c\log x_i^{(l,k)}\right)}{c} = 0.$$

In the case of the *Exponential distribution* we have

$$F(x;a_{ij}) = 1 - e^{-a_{ij}x}, \quad F(x;1) = 1 - e^{-x}, \quad f(x;1) = e^{-x}.$$

The estimator of the scale parameter in the uncensored case becomes

$$\widehat{a}_{ij}(L,M) = \frac{N_{ij}(L,M)}{\sum_{l=1}^{L}\sum_{k=1}^{N_i^{(l)}(M)}\left(X_i^{(l,k)}\right)} \tag{5.29}$$

Andreas Makrides

and in the censored case

$$\widehat{a}_{ij}(L,M) = \frac{N_{ij}(L,M)}{\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{N_i^{(l)}(M)}\left(X_i^{(l,k)}\right) + \sum\limits_{k=1}^{N_{i,M}(L)}\left(U_i^{(k)}\right)}. \tag{5.30}$$

In the *Rayleigh distribution* case, i.e.,

$$F(x;a_{ij}) = 1 - e^{-\sqrt{2}a_{ij}x^2}, \quad F(x;1) = 1 - e^{-\sqrt{2}x^2}, \quad f(x;1) = 2\sqrt{2}xe^{-\sqrt{2}x^2}.$$

The estimator of the scale parameter in the uncensored case can be written as

$$\widehat{a}_{ij}(L,M) = \frac{N_{ij}(L,M)}{\sqrt{2}\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{N_i^{(l)}(M)}\left(X_i^{(l,k)}\right)^2}, \tag{5.31}$$

while in the censored case it takes the form

$$\widehat{a}_{ij}(L,M) = \frac{N_{ij}(L,M)}{\sqrt{2}\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{N_i^{(l)}(M)}\left(X_i^{(l,k)}\right)^2 + \sqrt{2}\sum\limits_{k=1}^{N_{i,M}(L)}\left(U_i^{(k)}\right)^2}. \tag{5.32}$$

In the case of the *Pareto distribution* we have

$$F(x;a_{ij}) = 1 - \left(\frac{c}{x}\right)^{a_{ij}}, \quad F(x;1) = 1 - \frac{c}{x}, \quad f(x;1) = \frac{c}{x^2},$$

where c is the shape parameter. The MLE of the scale parameter in the case of no censoring takes the form

$$\widehat{a}_{ij}(L,M) = \frac{-N_{ij}(L,M)}{\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{N_i^{(l)}(M)}\log\left(\frac{c}{X_i^{(l,k)}}\right)}, \tag{5.33}$$

while the corresponding estimator in the case of censoring is given by

$$\widehat{a}_{ij}(L,M) = \frac{-N_{ij}(L,M)}{\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{N_i^{(l)}(M)}\log\left(\frac{c}{X_i^{(l,k)}}\right) + \sum\limits_{k=1}^{N_{i,M}(L)}\log\left(\frac{c}{U_i^{(k)}}\right)}. \tag{5.34}$$

Andreas Makrides

Note that the maximum likelihood estimator of the shape parameter c for the case of the Pareto distribution is obtained through a graphical representation and is given by

$$\widehat{c} = \min_{i,l,k} X_i^{(l,k)}.$$

5.4 Markov Renewal Function and semi-Markov Transition Matrix

In this Section, we are interested in two quantities that are important in the study of the behaviour of a semi-Markov process, namely the Markov renewal function, defined in (5.35), and the semi-Markov transition function, defined in (5.37). These two quantities are important either by themselves (see the meaning/interpretation of each one), or by their role in more complex quantities (cf. Section 5.5 for different reliability indicators, where these quantities appear). For details see Limnios and Ouhbi (2003), Ouhbi and Limnios (1999) or Pyke and Schaufele (1964).

An important role in Markov renewal theory is played by the *Markov renewal function* $\Psi_{ij}(t)$, $i, j \in E$, $t \geq 0$, which is defined as the expected value of the number of visits $N_j(t)$ to state j observed up to some time t knowing that the process started at state i at initial time $t = 0$. Hence the Markov renewal function is given by (see Limnios and Oprisan, 2001 or Limnios and Ouhbi, 2006)

$$\begin{aligned} \Psi_{ij}(t) &:= \mathbb{E}_i[N_j(t)] \\ &= \sum_{n=1}^{\infty} Q_{ij}^{(n)}(t) \\ &= \sum_{n=1}^{\infty} \sum_{k \in E} \int_0^t Q_{ik}(s) Q_{kj}^{(n-1)}(t-s) ds \end{aligned} \tag{5.35}$$

Andreas Makrides

which by (3.13) takes the form

$$\Psi_{ij}(t) = \sum_{n=1}^{\infty}\sum_{k\in E}\int_0^t \frac{a_{ik}}{\sum\limits_{k\in E} a_{ik}} \times \times\left[1-(1-F(s))^{\sum\limits_{k\in E} a_{ik}}\right]\times Q_{kj}^{(n-1)}(t-s)ds, \tag{5.36}$$

where the n^{th} convolution $Q_{ij}^{(n)}$ of Q by itself is given for any $i,j\in E$ by

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

The *semi-Markov transition matrix (function)* is defined as

$$P_{ij}(t) := \mathbb{P}(Z_t = j|Z_0 = i), i,j\in E. \tag{5.37}$$

It is shown in Limnios and Oprisan (2001) that the semi-Markov transition function satisfies, in matrix notation, the Markov renewal equation (MRE)

$$P(t) = I_N - diag\left(\sum_j Q_{ij}(t); i\in E\right) + (Q\star P)(t),$$

where $P(t) = (P_{ij}(t))_{i,j\in E}$, $Q(t) = (Q_{ij}(t))_{i,j\in E}$, I_N is the $N\times N$ identity matrix and $(Q\star P)(t)$ is the convolution product of Q and P defined by $(Q\star P)(t) = ((Q\star P)_{ij}(t))_{i,j\in E}$, with

$$(Q\star P)_{ij}(t) := \sum_{k\in E}\int_0^t Q_{ik}(s)P_{kj}(t-s)ds.$$

Let

$$W(t) := diag\left(W_i(t); i\in E\right) = diag\left(\sum_j Q_{ij}(t); i\in E\right) = diag\left(Q\cdot \mathbf{1}_N\right)(t)$$

be the diagonal matrix with the (i,i) element equal to $W_i(t) = \sum\limits_{j\in E} Q_{ij}(t)$, where $\mathbf{1}_N =$

Andreas Makrides

$\underbrace{(1,\ldots,1)}_{N}^{\top}$, $()^{\top}$ denoting the transposed of a vector.

Then, the unique solution of the above MRE is given by (Limnios and Oprisan, 2001)

$$\begin{aligned} P(t) &= \Big((I_N - Q)^{(-1)} \star (I_N - W)\Big)(t) \\ &= (\Psi \star (I_N - W))(t), \end{aligned} \tag{5.38}$$

where $\Psi(t) = (\Psi_{ij}(t))_{i,j\in E}$ and it is shown that

$$(I_N - Q)^{(-1)}(t) = \Psi(t).$$

Consequently, for any $i, j \in E$ and $t > 0$, the estimators of $P_{ij}(t)$ and $\Psi_{ij}(t)$ are given by:

$$\widehat{P}_{ij}(t, M) = \left(\widehat{\Psi}_{ij} \star \left(1 - \sum_{k\in E} \widehat{Q}_{jk}\right)\right)(t, M) \tag{5.39}$$

and

$$\begin{aligned} \widehat{\Psi}_{ij}(t, M) &= \sum_{n=1}^{\infty} \widehat{Q}_{ij}^{(n)}(t, M) \\ &= \sum_{n=1}^{\infty}\sum_{k\in E}\int_0^t \widehat{Q}_{ik}(s, M)\widehat{Q}_{kj}^{(n-1)}(t-s, M)ds, \end{aligned} \tag{5.40}$$

where among the estimators of a_{ij} given in (5.5), (5.10), (5.15) and (5.22), the appropriate one is chosen for obtaining $\widehat{Q}_{ij}(t, M)$.

5.5 Reliability/Survival Analysis Indicators

We recall the division of the state space E into two subsets, U (containing the up-states) and D (containing the down-states), such that $E = U \cup D$ and $U \cap D = \varnothing$, where we assume that $U = \{1, \ldots, n\}$ and $D = \{n+1, \ldots, N\}$. So the matrices can be written

Andreas Makrides

as

$$p = \begin{array}{cc} \begin{array}{cc} \text{U} & \text{D} \end{array} & \\ \begin{bmatrix} p_{UU} & p_{UD} \\ p_{DU} & p_{DD} \end{bmatrix} & \begin{array}{c} \text{U} \\ \text{D} \end{array} \end{array}, \qquad Q(t) = \begin{array}{cc} \begin{array}{cc} \text{U} & \text{D} \end{array} & \\ \begin{bmatrix} Q_{UU}(t) & Q_{UD}(t) \\ Q_{DU}(t) & Q_{DD}(t) \end{bmatrix} & \begin{array}{c} \text{U} \\ \text{D} \end{array} \end{array},$$

$$W(t) = \begin{array}{cc} \begin{array}{cc} \text{U} & \text{D} \end{array} & \\ \begin{bmatrix} W_U(t) & 0 \\ 0 & W_D(t) \end{bmatrix} & \begin{array}{c} \text{U} \\ \text{D} \end{array} \end{array}, \qquad \alpha = \begin{array}{c} \begin{array}{cc} \text{U} & \text{D} \end{array} \\ \begin{bmatrix} \alpha_U & \alpha_D \end{bmatrix} \end{array},$$

where $W_U(t) = diag\,(W_i(t); i \in U)$, $W_D(t) = diag\,(W_i(t); i \in D)$, $\alpha_U = (\alpha_i; i \in U)$ and $\alpha_D = (\alpha_i; i \in D)$.

For the matrices $\Psi(t)$ and $P(t)$ we consider their partitions induced by the corresponding partitions of the semi-Markov kernel $Q(t)$. For example,

$$P_{UU}(t) = (\Psi_{UU} \star (I_n - W_U))\,(t), \tag{5.41}$$

$$\Psi_{UU}(t) = (I_n - Q_{UU})^{(-1)}(t) \tag{5.42}$$

and

$$Q_{UU}(t) = \left(\frac{a_{ij}}{\sum\limits_{k \in E} a_{ik}} \left[1 - (1 - F(t))^{\sum\limits_{k \in E} a_{ik}} \right] \right)_{ij \in U}. \tag{5.43}$$

In the rest of this section we are interested in providing estimators for various reliability indicators: reliability function, availability, maintainability, failure rate, mean time to failure. These results are stated in the case of a single sample path (cf. Subsection 5.1). For several sample paths (cf. Subsection 5.2), the estimators are given in the same way but the contributing estimators of a_{ij} and a_i are given by (5.22) and (5.23) respectively.

Andreas Makrides

5.5.1 Reliability

It is reminded that the *reliability* or *survival function* of the system at time t, $R(t)$, is defined as the probability of being in acceptable states for $s \leq t$, i.e.,

$$\begin{aligned} R(t) &= \mathbb{P}(T_D > t) \\ &= \mathbb{P}(Z_s \in U, s \leq t), \end{aligned}$$

where $T_D := \inf\{t \mid Z_t \in D\}$ is the lifetime of the system. The following result presents the estimator of the reliability of a semi-Markov system in terms of estimators of basic quantities of a semi-Markov process.

Proposition 5.5.1 *For a semi-Markov system, the estimator of the reliability at time $t > 0$ is given by*

$$\widehat{R}(t, M) = \widehat{\alpha}_U(M)\widehat{P}_{UU}(t, M)\mathbf{1_n}, \tag{5.44}$$

where $\widehat{\alpha}_U(M)$ is an estimator of α_U and $\widehat{P}_{UU}(t, M)$ is an estimator of $P_{UU}(t)$.

Proof. From Ouhbi and Limnios (1996) we have that

$$R(t) = \alpha_U P_{UU}(t)\mathbf{1_n}. \tag{5.45}$$

For the multi-state system under study in the present thesis, we have

$$(Q_{UU})_{ij}(t) = \frac{a_{ij}}{\sum\limits_{k \in E} a_{ik}} \left[1 - (1 - F(t))^{\sum_{k \in E} a_{ik}}\right], \quad i, j \in U.$$

Using the estimators of the parameters a_{ij} as well as those of the initial distribution α_i obtained earlier in this chapter, the result is immediate. According to the available data, one can use one of the estimators of a_{ij} obtained in (5.5), (5.10), (5.15) or (5.22). As for the initial distribution, one can take the estimators proposed in (5.16) or (5.23), if several sample paths are available. Otherwise, depending on the case under study, one can assume that the initial distribution is known/given, or that it is the stationary distribution of the

Andreas Makrides

semi-Markov chain. For an estimator of the stationary distribution of a semi-Markov process, one can see Limnios et al. (2005). □

Remark 5.5.1

Using the estimator of the reliability obtained in (5.44), we immediately have an estimator of the failure rate of the system, given by

$$\widehat{\lambda}(t, M) := -\frac{\widehat{R}'(t, M)}{\widehat{R}(t, M)}, t > 0.$$

5.5.2 Availability

As mentioned in Definition 2.1.3, the *pointwise* (or *instantaneous*) *availability* of the system, $A(t)$, is defined as the probability of being in acceptable state at time instant t, i.e.,

$$A(t) = \mathbb{P}(Z_t \in U). \tag{5.46}$$

Proposition 5.5.2 *For a semi-Markov system, the estimator of the availability at time $t > 0$ is given by*

$$\widehat{A}(t, M) = \widehat{\alpha}(M)\widehat{P}(t, M)\mathbf{1}_{N;n}, \tag{5.47}$$

where $\widehat{P}(t, M)$ is the estimator of $P(t)$, $\mathbf{1}_{N;n} = (\underbrace{1, \cdots, 1}_{n}, \underbrace{0, \cdots, 0}_{N-n})^\top$ and $\widehat{\alpha}(M)$ is the estimator of α.

Proof. The result is an immediate consequence of the estimators presented in Sections 5 and 5.4 and of the following expression of the availability given in Ouhbi and Limnios (1996):

$$A(t) = \alpha P(t)\mathbf{1}_{N;n}. \tag{5.48}$$

Andreas Makrides

For more details concerning the estimators of a_{ij} and α_i involved in these expressions, see the proof of Proposition 5.5.1. □

5.5.3 Maintainability

The *maintainability* of the system, $M(t)$, is the probability that the system is repaired up to time t, given that it has failed at time $t = 0$, i.e.,

$$\begin{aligned} M(t) &= \mathbb{P}(T_U \leq t) \\ &= 1 - \mathbb{P}(Z_s \in D, s \leq t), \end{aligned} \tag{5.49}$$

where $T_U := \inf\{t \mid Z_t \in U\}$ is the duration of repair.

The following result gives the estimator of the maintainability of a semi-Markov system, the proof of which is omitted. For more details concerning the estimators of a_{ij} and α_i involved in the expression of this estimator, see the proof of Proposition 5.5.1. For the expression of the maintainability of a semi-Markov system, one can see Limnios and Oprisan (2001).

Proposition 5.5.3 *For a semi-Markov system, the estimator of the maintainability at time $t > 0$ is given by*

$$\widehat{M}(t, M) = 1 - \widehat{\alpha}_D(M)\widehat{P}_{DD}(t, M)\mathbf{1}_{\mathbf{N-n}}, \tag{5.50}$$

where $\widehat{\alpha}_D(M)$ is an estimator of α_D and $\widehat{P}_{DD}(t, M)$ is an estimator of $P_{DD}(t)$.

5.5.4 The Mean Time to Failure

The *mean time to failure* (*MTTF*) is defined as the mean lifetime, i.e., the expectation of the hitting time to the down set D,

$$MTTF := \mathbb{E}\left(T_D\right).$$

Andreas Makrides

Before giving the estimator of the MTTF, let us introduce the vector of the mean sojourn times $m := (m_1, \ldots, m_N)^\top$, where the *mean sojourn time in state* i is

$$\begin{aligned} m_i : &= \mathbb{E}(S_1 \mid J_0 = i) \\ &= \int_0^\infty (1 - W_i(t))\, dt \\ &= \int_0^\infty [1 - F(t)]^{\sum_{j=1}^{N} a_{ij}}\, dt. \end{aligned}$$

Note that for a regular and positive recurrent MRP, we have $m_i < \infty, i \in E$ (cf., e.g., Limnios and Oprisan, 2001).

We can estimate the mean sojourn time in state i in two different ways, namely using a plug-in MLE (considering the estimators obtained in Section 5) or the empirical estimator. Consequently, we propose the following estimators:

$$\begin{aligned} \widehat{m}_i^{(1)}(M) : &= \int_0^\infty \left(1 - \widehat{W}_i(t, M)\right) dt \\ &= \int_0^\infty [1 - F(t)]^{\sum_{j=1}^{N} \widehat{a}_{ij}(M)}\, dt \end{aligned}$$

and

$$\widehat{m}_i^{(2)}(M) := \frac{\sum_{k=1}^{N_i(M)} X_i^{(k)}}{N_i(M)},$$

where the estimators of a_{ij} are obtained in (5.5), (5.10), (5.15) or (5.22). For the expression of the MTTF of a semi-Markov system, one can see Limnios and Oprisan (2001).

Proposition 5.5.4 *For a semi-Markov system, assuming that the matrix $(I_n - \widehat{p}_{UU}(M))$ is nonsingular, the MTTF can be estimated by one of the following two ways:*

$$\begin{aligned} \widehat{MTTF}^{(1)}(M) &= \widehat{\alpha}_U(M)(I_n - \widehat{p}_{UU}(M))^{-1}\widehat{m}_U^{(1)}(M), \\ \widehat{MTTF}^{(2)}(M) &= \widehat{\alpha}_U(M)(I_n - \widehat{p}_{UU}(M))^{-1}\widehat{m}_U^{(2)}(M), \end{aligned}$$

where, $\widehat{p}_{UU}(M)$ is the estimator of p_{UU}, $\widehat{\alpha}_U(M)$ is an estimator of α_U, $\widehat{m}_U^{(1)}(M)$ and $\widehat{m}_U^{(2)}(M)$ are the restrictions to set U of $\widehat{m}^{(1)}(M)$ and $\widehat{m}^{(2)}(M)$, respectively.

Andreas Makrides

Note that in a similar way one can estimate the mean time to repair (MTTR), defined as

$$MTTR := \mathbb{E}\left(T_U\right).$$

5.6 Consistency of the Proposed Estimators

In this section we investigate the consistency of the estimators of $a_{ij}, i, j \in E$, and of the initial distribution $\alpha_i, i \in E$. This implies the consistency of most of the estimators proposed in this chapter, a fact that is stated in Corollary 5.6.1.

Theorem 5.6.1 *For all $i, j \in E$, the estimators $\widehat{a}_{ij}(M)$ of a_{ij} given in (5.5), $\widehat{a}_{ij}(L, M)$ of a_{ij} given in (5.15) and $\widehat{\alpha}_i(L, M)$ of α_i given in (5.16), are strongly consistent, i.e.,*

$$\begin{aligned}
&\widehat{a}_{ij}(M) \xrightarrow[M\to\infty]{a.s.} a_{ij}, \\
&\widehat{a}_{ij}(L, M) \xrightarrow[L\to\infty]{a.s.} a_{ij}, \\
&\widehat{\alpha}_i(L, M) \xrightarrow[L\to\infty]{a.s.} \alpha_i.
\end{aligned}$$

Proof. We give here the proof of the first statement only. The second part is similar; the third one is straightforward.

The estimator $\widehat{a}_{ij}(M)$ given in (5.5) can be written as

$$\begin{aligned}
\widehat{a}_{ij}(M) &= -\frac{N_{ij}(M)}{\sum\limits_{k=1}^{N_i(M)} \log\left(1 - F(X_i^{(1,k)})\right)} \\
&= \frac{N_i(M)}{\sum\limits_{k=1}^{N_i(M)} -\log\left(1 - F(X_i^{(1,k)})\right)} \times \frac{N_{ij}(M)}{N_i(M)}.
\end{aligned} \tag{5.51}$$

For the second factor of (5.51) we have the typical result

$$\frac{N_{ij}(M)}{N_i(M)} \xrightarrow[M\to\infty]{a.s.} p_{ij} = \frac{a_{ij}}{\sum\limits_{j\in E} a_{ij}}. \tag{5.52}$$

Andreas Makrides

For the first factor of (5.51), let us set

$$Y_{i,k} := -\log\left(1 - F(X_i^{(1,k)})\right).$$

Since $(Y_{i,k})_k$ are iid and $N_i(M) \to +\infty$, as $M \to +\infty$ (since the MRP is regular, cf. Section 2), we get that

$$\frac{\sum_{k=1}^{N_i(M)} Y_{i,k}}{N_i(M)} \xrightarrow[M\to+\infty]{a.s} \mathbb{E}(Y_{i,1}).$$

Some computations yield

$$\mathbb{E}(Y_{i,1}) = \frac{1}{\sum_{j\in E} a_{ij}}$$

and we get the desired result. $\square$

Corollary 5.6.1 *For all $i, j \in E$, and $t > 0$, we have:*

1. *Under the setting of Section 5.1, for one sample path of length M of a semi-Markov process, the estimators of $Q_{ij}(t)$, p_{ij}, $W_i(t)$, $f_i(t)$, $\Psi_{ij}(t)$, $P_{ij}(t)$, $R(t)$, $\lambda(t)$, $A(t)$, $M(t)$, $MTTF$, obtained by considering the estimators $\widehat{a}_{ij}(M)$ of a_{ij} given in (5.5), are strongly consistent, as M goes to infinity.*

2. *Under the setting of Section 5.2, for L sample paths of length M of a semi-Markov process, the estimators of $Q_{ij}(t)$, p_{ij}, $W_i(t)$, $f_i(t)$, $\Psi_{ij}(t)$, $P_{ij}(t)$, $R(t)$, $\lambda(t)$, $A(t)$, $M(t)$, $MTTF$, obtained by considering the estimators $\widehat{a}_{ij}(L, M)$ of a_{ij} given in (5.15) and $\widehat{\alpha}_i(L, M)$ of α_i given in (5.16), are strongly consistent, as L goes to infinity.*

Note that, as we already discussed in the proof of Proposition 5.5.1, in the case of a single sample path, one can assume that the initial distribution is known or that it is the stationary distribution of the semi-Markov chain.

Andreas Makrides

Chapter 6

An Application to Geosciences - The case of South America region

Andreas Makrides

6.1 Introduction

In this section we are in a continuous time framework under the homogeneous Markov setting of Section 2.4 and we consider as an application of the multi state methodology a data set of great earthquakes with the purpose of making earthquake forecasts. The most common model for earthquake occurrence is the Poisson one which assumes spatial and temporal independence of all earthquakes including large events. For instance, the occurrence of one earthquake does not affect the likelihood of a similar earthquake at the same location in the next time unit. From comparison of various models, it is observed that while the Poisson model may be applied to seismic regions which show a moderate or high frequency of earthquakes (Anagnos and Kiremidijan, 1988), other stochastic methods such as Markov chains and Renewal processes (Pyke, 1961a & 1961b; Andersen et al., 1993) outline the sequence of seismic events more appropriately in regions with large earthquakes, like South America. Large earthquakes are usually infrequent and in such cases simulation methods (e.g. Monte Carlo) are the most common tools in order to overcome such inconsistencies. The effort made in an earthquake hazard study is to have long term predictions of earthquakes and commonly this is expressed as probabilities of exceedance of a specific magnitude of earthquake over a period of time (see for example

Natvig and Tvete (2007)). One of the important objectives in seismology is the estimation of the aforementioned quantities. Usually such methods are applied in active faults or very often the whole examined area is divided in seismic zones (under certain criteria). South America was examined (see Tsapanos, 2001) in the light of the Markov model, in order to define large earthquake recurrences. For this purpose Tsapanos considered as states the seismic zones, pre-defined by other authors (Kelleher, 1972; Beck and Ruff, 1989; Beck and Nishenko, 1990; Papadimitriou, 1993; Cernadas et al., 1998). It is easy then to inspect the earthquake occurrence throughout the states and to estimate their genesis in a statistical way by the transition probabilities.

In Tsapanos (2001, 2011) the number of transitions from one seismic zone to another was used as the basis for providing transition rates confirmed via Monte Carlo simulations. In the present study we describe seismic zoning data as data of an MSS system by incorporating into the procedure introduced in Tsapanos (2001) the effect, via the underlying distribution, of sojourn times between transitions. In other words, we assume that the seismic activity of a region is described as a discrete-state continuous-time Markov process with N possible seismic zones (states) and evaluate intensity rates and transition probabilities between zones.

In what follows we analyze a seismic data set from a region of South America bounded between latitudes $47^o S - 0^o$ and longitudes $85^o - 65^o W$. The data set covers the period 1899-2010, consists of six seismic zones and includes 113 earthquakes of shallow depth, with magnitude $M \geq 6.5$ and 112 sojourn times (see Figure 6.1). The number of transitions among the 6 seismic zones from state i to state j, for $i, j = 1, 2, \ldots, 6$ are given in Table 6.1. The data set used is restricted to main shocks only, considering as dependent events the fore- and aftershocks. These events are removed by the application of a method introduced by Musson et al. (2002). We choose South America for our application because it is one of the most seismically active areas of the world. The area often experienced destructive earthquakes with $M \geq 8.0$; The largest event, ever recorded in the world, occurred there during 1960 with magnitude $M = 9.5$. For instance, Tsapanos and Burton (1991) ranked Chile and Peru in the second and third place, respectively, among 50 countries of high seismicity.

Andreas Makrides

Table 6.1: Number of Transitions among the 6 states (seismic zones)

i/j	**1**	**2**	**3**	**4**	**5**	**6**	**Row total**
1	1	2	2	4	2	5	16
2	3	1	2	3	1	1	11
3	4	2	4	4	4	5	23
4	3	2	8	3	2	5	23
5	1	0	4	6	3	2	16
6	4	4	3	2	4	6	23

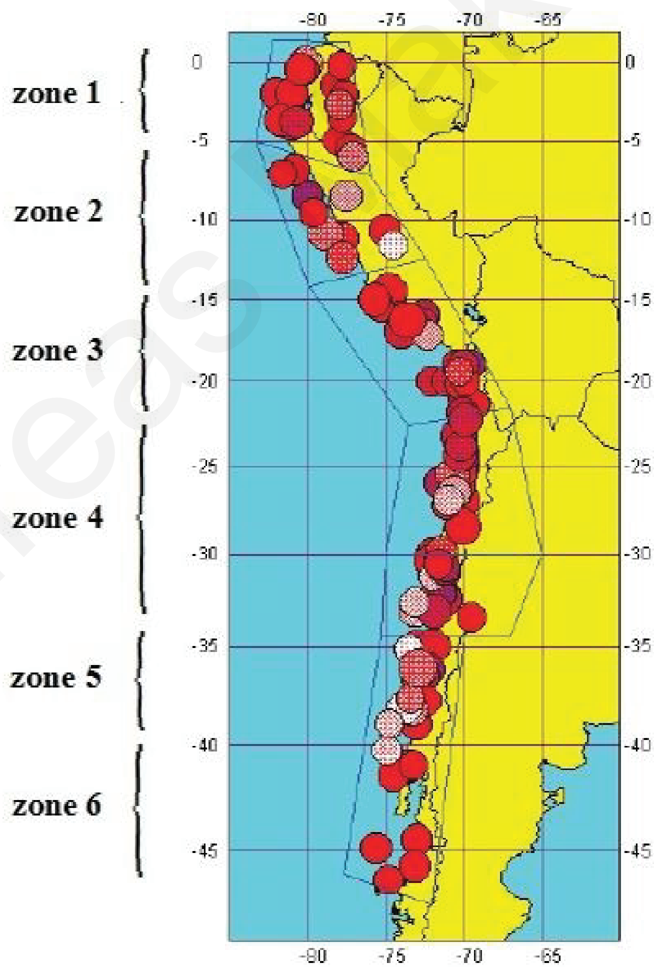

Figure 6.1: Seismic data (M $\geq$ 6.5) 1899-2010 South America - 6 zones.

6.2 Preliminary Statistical Analysis

The basic statistical characteristics of the magnitudes of the 113 earthquakes and the 112 sojourn times (in days) are presented in Table 6.2.

Table 6.2: Basic Statistical Characteristics

Variable	N	Mean	SE	Std	min	Q1	Q2	Q3	max
Magnitude	113	7.082	0.048	0.051	6.5	6.7	7.0	7.3	9.5
Time	112	381.5	37.6	398	3	102	266.5	575.8	2458

The associated boxplots for the two variables of interest are given in Figures 6.2 and 6.3.

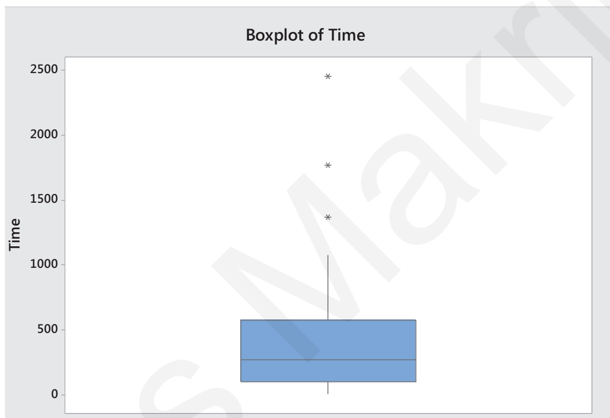

Figure 6.2: Boxplot of sojourn times 1899-2010 South America - 6 zones.

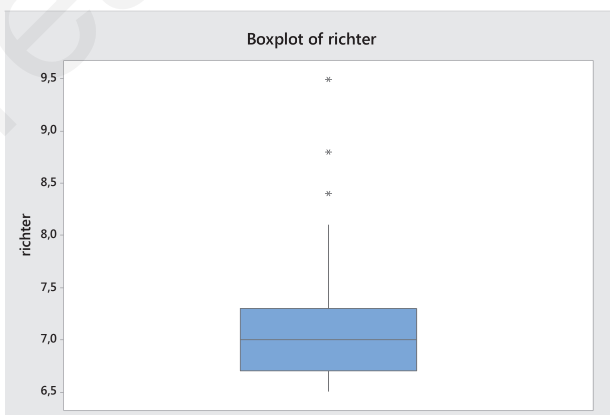

Figure 6.3: Boxplot of magnitudes of earthquakes 1899-2010 South America - 6 zones.

A reliability analysis has been performed to examine candidate parametric models for the sojourn times. Among various distributions the most probable ones have been used and depicted in Figure 6.4 (for details see Pardo, 2006).

Andreas Makrides

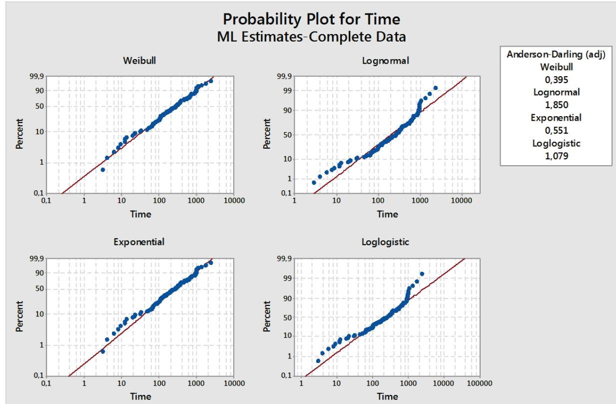

Figure 6.4: Distribution comparison for sojourn times.

It is clear that according to the Anderson-Darling goodness of fit (gof) tests (see D'Agostino and Stephens, 1986, Jaynes, 1957, Mattheou and Karagrigoriou, 2010, Read and Cressie, 1988 or Zografos et al., 1990) Exponential and Weibull are the 2 distributions most likely describing the underlying distribution of the sojourn times. Indeed, according to the Anderson-Darling(AD) test the best fit to the data is provided by the Weibull distribution (AD statistic=0.395) which is slightly better than the Exponential distribution (0.551). The other two models stay (relatively) close behind. The details for the best two fits are presented in Figures 6.5 and 6.6 with appropriate 95% confidence intervals for the probability plots.

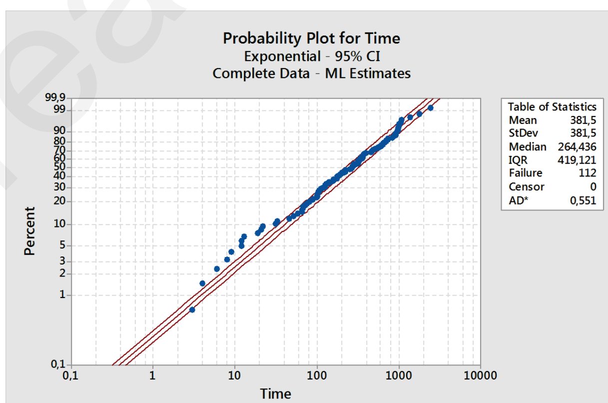

Figure 6.5: Fitting of Exponential distribution to sojourn times.

For a further confirmation of the gof test conclusion, we decided to apply to all candidate models the model selection (information) criteria AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) that take into consideration not only the log-

Andreas Makrides

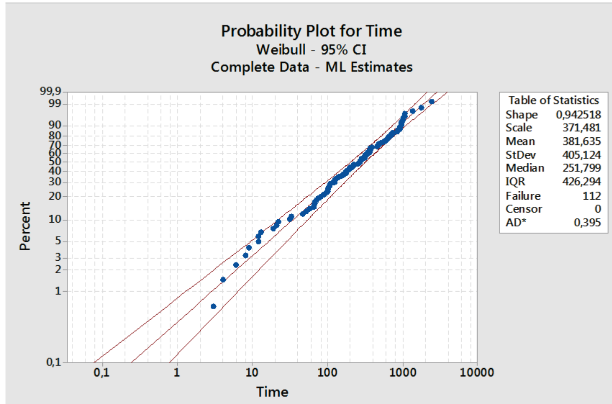

Figure 6.6: Fitting of Weibull distribution to sojourn times.

Table 6.3: Model Selection for Sojourn Times

Distribution	Parameters	Log-Likelihood	AIC	BIC
Weibull	scale= 371.48 - shape= 0.94	-777.413	1558.826	1564.263
Exponential	scale= 381.50	-777.74	1557.48	1560.198

likelihood but also the complexity of the model:

$$AIC(p) = -2\log(likelihood) + 2p$$

and

$$BIC(p) = -2\log(likelihood) + \log(N)p$$

The results for the two best fits are presented in Table 6.3.

Clearly the simplicity of the Exponential distribution makes it a better candidate for the sojourn times. Observe also that an estimate of the shape parameter for the Weibull distribution was found to be equal to 0.9425 (with a standard error equal to 0.07). Naturally, the Weibull distribution with parameter equal to 0.9425 is close to the Exponential distribution. Note that the corresponding 95% confidence interval for the shape parameter of the Weibull distribution is $(0.8148, 1.09)$ making the assumption of unity for the shape parameter an acceptable one (at least at the standard 5% level). Furthermore, appropriate exponentiality test procedures (see Csiszar, 1963, Forte, 1984, Forte and Hughes, 1988 or Henze and Meintanis, 2005) could verify that the difference observed is not statistically significant at the usual 5% level. Finally, it is noted that the mean time to failure (MTTF) is 381.5 (std. error=36.05) for the Exponential and 381.635 (std. error=38.28) for the Weibull distribution.

Recall that due to time dependence of the Weibull distribution, the time origin is taken to be the time origin of the data set so that any value of $t > 0$ describes the time since the instant of the occurrence of the first reported event.

6.3 Sojourn times analysis by the proposed methodology

In this Section we employ our methodology applying, for comparative purposes, the Exponential and Weibull distributions.

Tables 6.4 - 6.7 provide the probabilities of occurrence of an event in zone j provided that in time instant $t = 0$ an event took place in zone i for the Exponential distribution with $i = 1, \ldots, 6$ and $j = 1, \ldots, i \ldots, 6$. Tables 6.8 - 6.11 refer to the Weibull distribution.

Having available data up to 2010 we obtain the results for the two distributions for the years 2016, 2017, 2018 and 2020 using for the time t the values 2160 for 2016, 2520 for 2017, 2880 for 2018 and 3600 for 2020. Notice though that the transition probabilities hold for any 6-, 7-, 8- and 10- year periods starting from any time point $t = 0$. For the calculations we make use of equation (4.5) where t measures the time period in days from the day associated with the time instant $t = 0$ to the day of the event. The values in Tables 6.4 - 6.7 represent the transition probabilities from state i to state j for 2016, 2017, 2018 and 2020 respectively. For instance, the value 19.30% in the upper right corner of Table 6.4 represents the probability of occurrence of an event at time $t = 2160$ (year 2016) in zone 6 provided that the event at time instant $t = 0$ (year 2010: time of the last event in the data set) took place in zone 1. Note that due to time-homogeneity of the Exponential distribution, the time origin has been chosen to be the time instant of the last event in the data set. Thus, knowing the zone of the last event one could obtain via the tables, the zone of an event, for a specific t, where it will most likely occur together with the associated (transition) probability for this event.

In all the above cases the probabilities range from 19.98% (for the transition from zone 2

Andreas Makrides

to zone 1) to 21.56% (for the transition from zone 1 to zone 1). We also observe that as time progresses from 2016 to 2020 an event is more likely to occur in zone 1 irrespectively of the zone of the previous event with probabilities for the year 2020 being between 19.82% to 21.70%. Zones 4 and 6 follow closely with probabilities for the year 2020 ranging from 18.42% to 21.53%. The results suggest that it is unlikely for an event to take place in zone 2. Indeed, the probabilities of occurrence of an event in zone 2 are around 6.5%.

The results for the distributions examined are quite similar. The most significant difference for the case of the Weibull distribution, for all 6-, 7-, 8- and 10- year periods, is that lower probabilities are observed for almost all transitions from zone 4 as compared to those for the Exponential distribution irrespectively of the state visited by the system. Recall that the Exponential distribution is a special case of the Weibull distribution with $c = 1$.

Examining Table 6.8 (Weibull distribution) one observes that for the year 2016 an event is more likely to occur

(a) in zone 1 if we know that 6 years earlier (i.e. 2160 days earlier) occurred in zones 1 or 2

(b) in zone 4 if 6 years earlier occurred in zone 5 &

(c) in zone 6 if 6 years earlier occurred in zones 3, 4 or 6.

The above comments hold also for any 6-year period starting from any time point $t = 0$.

Similar results, for both distributions, can be obtained from Tables 6.5 and 6.9 which correspond to 7- year period and Tables 6.6 and 6.10 which correspond to 8- year period.

It should be emphasis that the Markov Processes are not able to describe the stochastic evolution of the system by taking into account the real time (due to the lack of memory of the Exponential distribution) and even if the differences between the two models seem to be slight, these results are dataset-dependent.

Andreas Makrides

Table 6.4: Exponential Distribution: Transition Probabilities for 2016

State	1	2	3	4	5	6
1	0.216	0.068	0.153	0.204	0.154	0.194
2	0.208	0.073	0.156	0.205	0.157	0.191
3	0.216	0.069	0.156	0.211	0.159	0.201
4	0.209	0.068	0.157	0.210	0.157	0.199
5	0.206	0.067	0.159	0.211	0.164	0.193
6	0.215	0.066	0.156	0.212	0.161	0.202

Table 6.5: Exponential Distribution: Transition Probabilities for 2017

State	1	2	3	4	5	6
1	0.212	0.068	0.155	0.205	0.156	0.196
2	0.208	0.077	0.153	0.202	0.154	0.194
3	0.214	0.065	0.161	0.211	0.161	0.199
4	0.209	0.068	0.158	0.209	0.158	0.196
5	0.206	0.067	0.158	0.210	0.160	0.194
6	0.214	0.066	0.159	0.212	0.162	0.198

Table 6.6: Exponential Distribution: Transition Probabilities for 2018

State	1	2	3	4	5	6
1	0.211	0.069	0.156	0.206	0.157	0.193
2	0.203	0.075	0.159	0.202	0.161	0.189
3	0.215	0.065	0.161	0.212	0.156	0.201
4	0.210	0.068	0.158	0.208	0.159	0.196
5	0.208	0.067	0.157	0.209	0.159	0.195
6	0.214	0.068	0.160	0.213	0.156	0.199

Table 6.7: Exponential Distribution: Transition Probabilities for 2020

State	1	2	3	4	5	6
1	0.209	0.068	0.156	0.207	0.156	0.194
2	0.201	0.078	0.158	0.199	0.162	0.189
3	0.214	0.061	0.161	0.210	0.163	0.201
4	0.210	0.067	0.157	0.208	0.158	0.196
5	0.209	0.067	0.159	0.208	0.158	0.195
6	0.212	0.066	0.161	0.211	0.161	0.199

Andreas Makrides

Table 6.8: Weibull Distribution: Transition Probabilities for 2016

State	1	2	3	4	5	6
1	0.217	0.068	0.156	0.196	0.152	0.203
2	0.202	0.071	0.164	0.194	0.161	0.199
3	0.208	0.064	0.161	0.202	0.159	0.212
4	0.196	0.072	0.163	0.195	0.164	0.201
5	0.196	0.064	0.165	0.212	0.162	0.208
6	0.215	0.060	0.159	0.204	0.154	0.219

Table 6.9: Weibull Distribution: Transition Probabilities for 2017

State	1	2	3	4	5	6
1	0.211	0.069	0.155	0.193	0.157	0.205
2	0.203	0.074	0.159	0.191	0.166	0.200
3	0.209	0.064	0.161	0.202	0.151	0.212
4	0.193	0.073	0.171	0.191	0.169	0.198
5	0.203	0.061	0.158	0.208	0.168	0.212
6	0.216	0.053	0.151	0.208	0.158	0.220

Table 6.10: Weibull Distribution: Transition Probabilities for 2018

State	1	2	3	4	5	6
1	0.206	0.072	0.156	0.193	0.161	0.204
2	0.199	0.074	0.163	0.190	0.168	0.198
3	0.209	0.064	0.162	0.202	0.157	0.212
4	0.193	0.081	0.159	0.189	0.171	0.197
5	0.207	0.056	0.166	0.207	0.159	0.213
6	0.218	0.053	0.158	0.210	0.151	0.221

Table 6.11: Weibull Distribution: Transition Probabilities for 2020

State	1	2	3	4	5	6
1	0.202	0.073	0.162	0.193	0.161	0.203
2	0.197	0.076	0.171	0.189	0.165	0.198
3	0.208	0.065	0.160	0.202	0.158	0.212
4	0.192	0.081	0.163	0.188	0.173	0.196
5	0.210	0.060	0.164	0.205	0.156	0.214
6	0.216	0.051	0.155	0.209	0.160	0.220

Andreas Makrides

Remarks 6.3.1

1. The number of available data plays a crucial role in such type of problems. Recall that for the South America case under study we focus on events of high magnitude ($M \geq 6.5$) and observed in total 113 such events for a period of more than 110 years. At the same time we considered 6 zones which in turn requires the evaluation of a 6 x 6 matrix with 36 unknown scale parameters. As expected, the estimations would have been even more accurate if more data were available. Furthermore, if sufficient data is available then data validation should be performed to increase overall data quality and workflow efficiency. Note that data validation is a key issue, mainly used in settings where the goal is prediction, and one wants to evaluate the accuracy of the predictive model. Note that due to limited data available for the earthquake application, a data validation analysis was not possible to be undertaken without losing significant modelling or testing capability.

2. An important property of the Weibull distribution is the power-law behavior of its hazard function. Values of the shape parameter equal to 1 indicate that the rate remains constant over time. At the same time values less than 1 indicate that the rate decreases over time. On the other hand if the shape parameter is bigger than 1 then the hazard rate increases as a power of the time moment of the last event. The hazard rate is expected to increase monotonically for a driven system. Note that major plate boundary faults can be considered to be driven systems due to the fact that motions of tectonic plates systematically increase the stress of such faults. As a result a value of the shape parameter larger than 1 as indicated by the 95% confidence interval presented earlier, is legitimate so that thinner tails than the exponential distribution are likely to occur for the holding times.

3. Observe that on the above matrices the elements on each column appear with similar values (different for each column). This happens due to the fact that not only the Exponential but also the Weibull distribution with shape parameter been estimated close to 1, have fast convergence rate so that the elements on each column tend quickly towards the same value and the steady-state distribution, namely the limiting transition matrix is quickly attained (for relatively small values of t). Having this

Andreas Makrides

Table 6.12: Exponential Distribution: Transition Probabilities for 2011

State	1	2	3	4	5	6
1	0.569	0.048	0.070	0.112	0.067	0.127
2	0.189	0.295	0.131	0.184	0.089	0.106
3	0.137	0.055	0.401	0.139	0.124	0.152
4	0.099	0.048	0.153	0.495	0.072	0.128
5	0.063	0.019	0.127	0.182	0.512	0.092
6	0.131	0.080	0.095	0.095	0.115	0.493

in mind and for the sake of completeness, we present in Table 6.12 the results for 1-year predictions in the case of the Exponential distribution where the elements on each column appear significantly different.

Andreas Makrides

Chapter 7

Conclusions

Markov processes represent typical tools for modelling multi-state systems. In this thesis we focus on multi state systems that we model by means of semi-Markov processes, which generalize typical Markov jump processes by allowing general distributions for sojourn times. For this reason, the semi-Markov processes are more adapted for reliability studies (and for applications in general).

The proposed methodology provides estimating methods for obtaining transition rates together with transition probabilities in multi-state systems with sojourn times following a number of special distributions belonging to a subclass of the generalized Gamma distribution.

An application of the proposed methodology was presented in Chapter 6. The application uses a data set of 113 great earthquakes (with magnitude M over 6.5) from the South America region covering the period 1899-2010 with the purpose of making earthquake forecasts. The seismology application shows that if the seismic zones are considered to be the states of a multi-state system then such a statistical approach can be found useful in determining the probability of occurrence of a physical phenomenon like the earthquake. The case of South America reveals that the Exponential assumption for the residing times is quite satisfactory since the implementation of the Weibull distribution shows that the shape parameter is sufficiently close to 1 (the case of the Exponential distribution). For more accurate conclusions one may have to depend on the (2-parameter) Weibull distri-

Andreas Makrides

bution which includes a flexible positive shape parameter. The analysis presented in this thesis appears to suggest that the shape parameter should not be expected to exceed the value of 1 which implies that distributions like the Rayleigh distribution (for which the shape parameter is equal to 2) may not be useful for the analysis of seismic data although they could be considered for different types of systems.

We also assume a general class of distributions for the sojourn times and provide the likelihood for the censored and uncensored case for one or several trajectories. The maximum likelihood estimators for the parameters of a general class of distributions together with the initial law estimators have been furnished. We further obtain estimators of other measures related to such systems such as the Markov renewal function and the semi-Markov transition matrix. Finally, reliability indicators like the reliability, availability, maintainability and mean time to failure (MTTF) are also estimated.

The expressions developed can be easily used for obtaining the corresponding expressions for particular distributions belonging to the general class of distributions like the Exponential, Weibull, Pareto and Rayleigh.

Andreas Makrides

Chapter 8

Future work

The results of this thesis can be extended and generalized in different ways.

8.1 Asymptotic Theory

Regarding the estimators proposed in this work, one could investigate their asymptotic properties. More specifically, we intend to investigate the asymptotic normality of the estimators of $a_{ij}, i, j \in E$, and of the initial distribution $\alpha_i, i \in E$ that have been studied in Section 5. The following theorem remains to be established:

Theorem 8.1.1 *For all $i, j \in E$ and $t \in [0, M]$,*

$$\sqrt{M}\,(\hat{a}_{ij}(t, M) - a_{ij})$$

converges in law to a zero mean normal random variable with an (asymptotic) variance, say σ_{ij}^2.

Having the above result we will investigate the asymptotic normality of all the other quantities that have been introduced in Section 5, namely the Markov renewal function, the semi-Markov transition matrix and the reliability indicators.

Andreas Makrides

8.2 Class of distributions closed under maxima

Order statistics from non-identically but independently distributed random variables are not easy to deal with. Despite that if these variables belong to families of random variables closed under maximum or minimum elegant simplifications are possible. In this thesis we explored a class of distributions closed under minima and obtain statistical inference for multi-state systems. Similar analysis can be performed if we consider instead, a class of distributions closed under maxima. Indeed, we could examine the following class of distributions given by

$$F^{\lambda}(x) = [F(x)]^{\lambda}, \quad \lambda > 0. \tag{8.1}$$

Then exact and explicit expressions for expectations of functions of single order statistics can be obtained. These results will then be used for statistical inference in the same way as in the present thesis by considering continuous or even discrete distributions belonging to this class of distributions.

8.3 Time Varying Models

We also intend to deal with time varying models. The basic assumption of such models is that the parameters of the distribution follow a specific temporal pattern which can be captured by the time-varying parameters of the underlying distribution. The time-varying model accounts for parameter uncertainty by maximizing the time-varying likelihood function, which is to estimate time-trend parameters and the distributional parameters in one step.

The general model we plan to consider is a Weibull model with constant shape parameter and time-varying scale parameter, which can be a function of either the system age or the number of cumulative failures. A number of parametric models for the scale parameter will be proposed and thoroughly examined. The models are suitable for situations where the system's mean time between failures (MTBF) can be monotonic or bathtub-shaped, and bounded. The proposed model can be viewed as a Weibull process model. Potential

Andreas Makrides

applications of the models include modeling manufacturing defect occurrence processes and evaluating the effectiveness of maintenance.

Let $T_i, i = 1, 2, \dots,$ be the successive failure times of a system, $T_i < T_{i+1}$, and $T_0 := 0$. Let also

$$X_i := T_i - T_{i-1}, i = 1, 2, \dots,$$

be the times between two successive failures (TBF). We assume that $X_i, i = 1, 2, \dots,$ are inid random variables such that

$$X_i \sim F(x; a_i),$$

belonging to the class (3.2) with different scale parameters a_i and possible other common parameters (which are suppressed in the notation below), i.e.:

$$F(x; a_i) = 1 - (1 - F(x; 1))^{a_i}. \tag{8.2}$$

Let also $x_1, x_2, \dots; t_1, t_2, \dots$ denote a realization of $X_1, X_2, \dots; T_1, T_2, \dots,$ with

$$T_i = \sum_{j=1}^{i} X_j.$$

In order to examine a time-varying scale parameter a_i for $i = 1, 2, \dots$ for modeling failure processes we consider the model presented in the definition below:

Definition 8.3.1 (time varying models) *For $i = 1, 2, \dots$ we consider the following model:* Model 1 *defined by*

$$a_i = a_\infty \left(1 - e^{-t_i/e_1}\right) \tag{8.3}$$

where a_∞, e_1 and are nonnegative model parameters.

Two special cases could be considered when the time varying behaviour is associated to the time index. Thus, if in Definition 8.3.1 we replace:

- t_i by i, $i = 1, 2, \dots,$ we obtain the analogous model that we will call *Model 2*;

- t_i by $\sum_{j=1}^{i} j = i(i+1)/2$, $i = 1, 2, \ldots$ (that is, x_j is replaced by its index j in $t_i = \sum_{j=1}^{i} x_j$), we obtain the analogous model that we will call *Model 3* .

Thus, *Model 2* is defined by

$$a_i = a_\infty \left(1 - e^{-i/e_1}\right); \tag{8.4}$$

whereas *Model 3* is defined by

$$a_i = a_\infty \left(1 - e^{-i(i+1)/(2e_1)}\right). \tag{8.5}$$

Remarks 8.3.1

1. Note that in the above models, the limit of a_i, as i goes to ∞, is a_∞; this is the reason for using this notation.
2. Note that in the above models, the limit of a_i, as i goes to 0, is the constant a_0.
3. It is clear also that, in *Model 1*, a_i is a strictly increasing function of t_i, with the limit being equal to a_∞.
4. The same type of remarks apply to *Models 2* and *3*.

Consider a sample path $x_1, x_2, \ldots, x_n$ from the general class of distribution given in (8.2) under the assumption that the scale parameter follows Model 1 given in (8.3). Then, the likelihood function can be easily obtained:

$$\begin{aligned} \mathcal{L}(a_\infty, e_1) &= \prod_{i=1}^{n} f(x_i; a_\infty, e_1) \\ &= a_\infty^n \prod_{i=1}^{n} \left(1 - e^{-t_i/e_1}\right) \prod_{i=1}^{n} (1 - F(x_i; 1))^{a_\infty\left(1-e^{-t_i/e_1}\right)-1} \prod_{i=1}^{n} f(x_i; 1). \end{aligned} \tag{8.6}$$

Our intension is to provide the expressions for the maximum likelihood estimators for the parameters involved in the time-varying models. The asymptotic theory of the proposed estimators will also be investigated. Finally, model selection techniques could be used for choosing between the various models.

Bibliography

[1] Alvarez, E.E. (2005). Estimation in Stationary Markov Renewal Processes, with Application to Earthquake Forecasting in Turkey, *Methodol. Comput. Appl. Probab.*, **7**, 119-130.

[2] Amoroso, L. (1925). Ricerche intorno alla curve dei redditi. *Annali di Matematica Pura ed Applicata*, **2(1)**, 123-159.

[3] Anagnos, T. and Kiremidjian, A. S. (1988). A review of earthquake occurrence models for seismic hazard analysis. *Probab. Engin. Mechanics* **3 (1)**, 3-11.

[4] Andersen, P. K., Borgain, O., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer.

[5] Aven, T. and Jensen, U. (1999). *Stochastic Models in Reliability*. Springer, New York.

[6] Balakrishnan, N. (2007). Permanents, order statistics, outliers, and robustness, *Rev. Math. Comput.*, **20 (1)**, 7-107.

[7] Balasubramanian, K., Beg, M.I., and Bapat, R.B. (1991). On families of distributions closed under extrema, *Sankhya A*, **53**, 375-388.

[8] Baldi, P., Mazliak, L., and Priouret P. (2002). *Martingales and Markov chains: Solved exercises and elements of theory*. Chapman and Hall/CRC, London.

[9] Barbu, V., Boussemart, M. and Limnios, N. (2004). Discrete time semi-Markov model for reliability and survival analysis, *Comm. Statist. Theory Methods*, **33(11)**, 2833b•“2868.

Andreas Makrides

[10] Barbu, V. and Limnios, N. (2006). Empirical estimation for discrete time semi-Markov processes with applications in reliability, *J. Nonparametr. Statist*, **18(7-8)**, 483b•“498.

[11] Barbu, V. and Limnios, N. (2010). Some algebraic methods in semi-Markov processes, *Algebraic Methods in Statistics and Probability*, **2**, 19-35.

[12] Barlow, R.E. and Proschan, F. (1965). *Mathematical Theory in Reliability*. Wiley, New York.

[13] Barlow, R.E. and Wu, AS. (1978). Coherent systems with multi-state components. *Math. Oper. Res.* **3**, 275-281.

[14] Bauckhage, C., Kersting, K., Hadiji, F.(2013). Mathematical Models of Fads Explain the Temporal Dynamics of Internet Memes. *Proc. ICWSM, Association for the Advancement of Artificial Intelligence.* URL: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6022/6340

[15] Bauckhage, C., Kersting, K., Rastegarpanah, B. (2013). The Weibull as a Model of Shortest Path Distributions in Random Networks. *Proc.* 11^{th} *Workshop on Mining and Learning with Graphs, Chicago, Illinois, USA Copyright 2013, ACM 978-1-4503-2322-2.* URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.367.869&rep=rep1&type=pdf

[16] Baxter, V. (2002). Advances in supermarket refrigeration systems. *Proceedings of the 7th international Energy Agency heat pump conference*, Beijing, China.

[17] Beck, S. L. and Nishenko, S. (1990). Variations in the mode of great earthquake rupture along the central Peru subduction zone. *Geophys Res Lett* **17**, 1969-1972.

[18] Beck, S. L. and Ruff, L. (1989). Great earthquakes and subduction along the Peru trench. *Physics Earth Planet Inter* **57**, 199-224.

[19] Billinton, R. and Allan, R. (1996). *Reliability Evaluation of Power Systems*. Plenum, New York.

[20] Block, H. and Savits, T. (1982). A decomposition of multistate monotone system. *J. Appl. Probab.* **19**, 391-402.

Andreas Makrides

[21] Cernadas, D. , Osella, A. and Sabbione, N. (1998). Self-similarity in the seismicity of the South American subduction zone, *Pageoph*, **152**, 57-73.

[22] Choi, S., Tong, C. (2010). Statistical Wavelet Subband Characterization based on Generalized Gamma Density and Its Application to Texture Retrieval, *IEEE Trans. on Image Processing*, **19(2)**, 281-289.

[23] Chryssaphinou, O., Limnios, N. and Malefaki, S. (2010). Multistate reliability systems under discrete time semi-Markovian hypothesis, *IEEE Transactions on Reliability* , **60 (1)**, 80b•“87.

[24] Cooke, R.G.(1955). *Infinite matrices and sequence spaces.* Dover Publications, New York.

[25] Crooks, G. (2010). The Amoroso Distribution. arXiv:1005.3274 [math.ST]. https://arxiv.org/pdf/1005.3274v2.pdf

[26] Csiszar, I. (1963). Eine Informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten, *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, **8**, 84-108.

[27] D'Agostino, R. B. and Stephens, M. A. (1986). *Goodness-of-Fit Techniques.* Marcel Dekker, New York.

[28] de Ves, E., Benavent, X., Ruedin, A., Acevedo, D., Seijas, L. (2010). Wavelet-based Texture Retrieval Modeling the Magnitudes of Wavelet Detail Coefficients with a Generalized Gamma Distribution. *Proc. ICPR.*, 221-224.

[29] El-Neveihi, E. and Proschan, F. (1984). Degradable systems: a survey of multistate system theory. *Comm. Statist. Theory Methods*, **13**, 405-432.

[30] El-Neveihi, E., Proschan, F. and Sethuraman, J. (1978). Multistate coherent systems. *J. Appl. Probab.* **15**, 675-688.

[31] Forte, B. (1984). Entropies with and without probabilities. Applications to questionnaires, *Inform. Process. and Manag.*, **20**, 397-405.

[32] Forte, B. and Hughes, W. (1988). The maximum entropy principle: a tool to define new entropies, *Rep. Math. Phys.*, **26**, 227-235.

Andreas Makrides

[33] Fox, B. (1968). *Semi Markov Processes: A primer*, Springfield, Va. : Clearinghouse for Federal Scientific and Technical Information.

[34] Giard, N., Lichtenstein, P. and Yashin, A. (2002) A multi-state model for genetic analysis of the aging process. *Stat. Med.*, **21**, 2511-2526.

[35] Henze, N. and Meintanis, S.G. (2005). Recent and classical tests for exponentiality: a partial review with comparisons, *Metrika*, **61**, 29-45.

[36] Hudson, JC. and Kapur, KC. (1982). Reliability theory for multistate systems with multistate components. *Microelectron Reliab.*, **22**, 1-7.

[37] Jaynes, E. T. (1957). Information Theory and Statistical Mechanics, *Phys. Rev.*, **106**, 620-630.

[38] Kelleher, J. A. (1972). Rupture zones of large South American earthquakes and sime predictions. *J Geophys Res.*, **77**, 2087-2103.

[39] Kemeny, J.G., Snell, J.L. and Knapp, A.W. (1976). *Denumerable Markov chains*, 2nd ed. Series Graduate Texts in Mathematics, vol. 40, Springer, New York.

[40] Kovalenko, I., Kuznetsov, N. and Pegg, Ph. (1997). *Mathematical Theory of Reliability of Time Dependent Systems with Practical Applications*. Wiley, Chichester, UK.

[41] Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*, New York: Wiley.

[42] Levy, P. (1954). Process semi-markoviens. In: *Proc. of the International Congress on Mathematics*, Amsterdam, 416-426.

[43] Li, P., Hastie, T., Church, K. (2006). Improving Random Projections Using Marginal Information. *Proc. COLT. G. Lugosi and H.U. Simon, eds., Springer-Verlag*, 635-649.

[44] Lienhard, J. and Meyer, P. (1967). A Physical Basis for the Generalized Gamma Distribution. *Quarterly of Applied Mathematics*, **25(3)**, 330-334.

[45] Limnios, N. and Oprişan, G. (2001). *Semi-Markov Processes and Reliability*, Birkhauser, Boston.

Andreas Makrides

[46] Limnios, N. and Ouhbi, B. (2003). Empirical estimators of reliability and related functions for semi-Markov systems, In: Lindqvist, B. H., Doksum, K. A. (eds), *Mathematical and Statistical Methods in Reliability*, **7**, World Scientific, Singapore, 469-484.

[47] Limnios, N. and Ouhbi, B. (2006). Nonparametric estimation of some important indicators in reliability for semi-Markov processes, *Stat. Methodol.*, **3**, 341-350.

[48] Limnios, N., Ouhbi, B. and Sadek, A. (2005). Empirical estimator of stationary distribution for semi-Markov processes, *Comm. Statist. Theory Methods*, **34** (4), 987–995.

[49] Lisnianski, A., Frenkel, I. and Ding, Y. (2010). *Multi-state System Reliability Analysis and Optimization for Engineers and Industrial Managers,* Springer, London.

[50] Lisnianski, A. and Levitin, G. (2003). *Multi-state System Reliability: Assessment, Optimization and Applications*, World Scientific, Singapore.

[51] Marshall, G. and Jones, R. (2007) Multi-state models in diabetic retinopathy. *Stat. Med.*, **14(18)**, 1975- 1983.

[52] Masala, G. (2012). Earthquakes occurrences estimation through a parametric semi-Markov approach, *J. Appl. Stat.*, **39 (1)**, 81b•“96.

[53] Mattheou, K. and Karagrigoriou, A. (2010). A new family of divergence measures for tests of fit, *Aust. N. Z. J. Stat.*, **52**, 187-200.

[54] McClean, S.I., Papadopoulou, A.A. and Tsaklidis, G. (2004). Discrete time reward models for homogeneous semi-Markov systems. Semi-Markov processes: theory and applications, *Commun. Stat. Theor. M.*, **33 (3)**, 623-638.

[55] Moore, H. and Pyke, R. (1968). Estimation of the transition distribution of a Markov renewal process, *Annals of the Institute of Statistical Mathematics*, **20** , 411-424.

[56] Murchland, J. (1975). Fundamental concepts and relations for reliability analysis of Multistate systems. In: Barlow RE, Fussell JB and Singpurwalla N (eds) *Reliability and fault tree analysis: theoretical and applied aspects of system reliability*, SIAM, Philadelphia, 581-618.

Andreas Makrides

[57] Musson, R. M. W., Tsapanos, T. M. and Nakas, C. F. (2002). A power-law function for earthquake interarrival time and magnitude. *Bull. Seismol. Soc. Am.* **92(5)**, 1783-1794.

[58] Natvig, B. (1982). Two suggestions of how to define a multistate coherent system. *Adv. in Appl. Probab.*, **14**, 434-455.

[59] Natvig, B. (1985). Multi-state coherent systems. *Johnson N, Kotz S (eds) Encyclopedia of statistical sciences*, **5**, Wiley, New York, 732-735.

[60] Natvig, B. (2011). *Multistate Systems Reliability. Theory with Applications.* Wiley, New York.

[61] Natvig, B. and Tvete, I. (2007). Bayesian hierarchical space - time modeling of earthquake data. *Meth. Comput. Applied Probab.* **9**, 89-114.

[62] Ouhbi, B. and Limnios, N. (1996). Non-parametric estimation for semi-Markov kernels with application to reliability analysis, *Appl. Stoch. Models Data Anal.*, **12**, 209-220.

[63] Ouhbi, B. and Limnios, N. (1999). Non-parametric estimation for semi-Markov processes based on its hazard rate functions, *Statist. Infer. Stoch. Processes*, **2 (2)**, 151-173.

[64] Papadimitriou, E. E. (1993). Long-term earthquake prediction along the western coast of South and Central America based on a time predictable model. *Pageoph.* **140**, 301-316.

[65] Pardo, L. (2006). *Statistical Inference Based on Divergence Measures*, Chapman & Hall.

[66] Pyke, R. (1961a). Markov Renewal Processes: Definitions and Preliminary Processes, *Ann. Math. Stat.*, **32**, 1231-1242.

[67] Pyke, R. (1961b). Markov Renewal Processes with Finitely Many States, *Ann. Math. Stat.*, **32**, 1243-1259.

[68] Pyke, R. and Schaufele, R. (1964). Limit theorems for Markov renewal process, *Ann. Math. Stat.*, **35**, 1746-1764.

Andreas Makrides

[69] Read, R. C. and Cressie, N. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*. Springer, New York.

[70] Ross, SM. (1979). Multivalued state component systems. *Ann. Probab.* **7**, 379-383.

[71] Sahner R., Trivedi K. and Poliafito, A. (1996). *Performance and Reliability Analysis of Computer Systems: an Example-based Approach Using the SHARPE Software Package*. Kluwer, Boston.

[72] Schutz, A., Bombrum, L., Berthoumieu, Y. and Najim, M. (2013). Centroid-Based Texture Classification Using the Generalized Gamma Distribution. *Proc. EUSIPCO. Marrakech, Morocco*, 1280-1284.

[73] Seneta, E. (1981). *Non-negative matrices and Markov chains*. Springer, New York.

[74] Stacy, E. (1962). A Generalization of the Gamma Distribution. *The Annals of Mathematical Statistics* **33(3)**, 1187-1192.

[75] ableman, M.and Kim, J. S. (2004). *Survival Analysis using S*, Chapman & Hall.

[76] Tsapanos, T. M. (2001). The Markov model as a pattern for earthquakes recurrence in South America. *Bull. Geol. Soc. Gr. XXXIV*, **4**, 1611-1617.

[77] Tsapanos, T. M. (2011). The Markov-chains as a tool for very large earthquakes in South America. STATSEI7 - Intern Workshop on Statistical Seismology 25-27 May 2011 Thera Greece.

[78] Tsapanos, T. M. and Burton P. W. (1991). Seismic hazard evaluation for specific seismic regions of the world. *Tectonophysics* **194**, 153-169.

[79] Van den Hout, A. and Matthews, F. (2008). Multi-state analysis of cognitive ability data: a piecewise-constant model and a Weibull model, *Stat. Med.*, **27**, 5440-5455.

[80] Votsi, I., Limnios, N., Tsaklidis, G. and Papadimitriou, E. (2014). Hidden semi-Markov modeling for the estimation of earthquake occurrence rates, *Communications in Statistics: Theory and Methods*, **43**, 1484-1502.

Andreas Makrides

[81] Zografos, K., Ferentinos, K. and Papaioannou, T. (1990). ϕ-divergence statistics: sampling properties, multinomial goodness of fit and divergence tests. *Comm. Statist. Theory Methods*, **19**, 1785-1802.

Andreas Makrides