PARAMETER ESTIMATION IN SWITCHING STOCHASTIC MODELS

A DISSERTATION SUBMITTED TO
THE DEPARTMENT OF INDUSTRIAL ENGINEERING
AND THE INSTITUTE OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
GÜLDAL GÜLERYÜZ
May, 2004

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Prof. Dr. Ülkü Gürler (Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Prof. Dr. Barbaros Tansel

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Prof. Dr. İsmihan Bayramoĝlu

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Assist. Prof. Dr. Murat Fadıloğlu

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of doctor of philosophy.

Assist. Prof. Dr. Emre Berk

Approved for the Institute of Engineering and Science:

Prof. Dr. Mehmet B. Baray
Director of the Institute

ABSTRACT

PARAMETER ESTIMATION IN SWITCHING STOCHASTIC MODELS

GÜLDAL GÜLERYÜZ
Ph.D. in Industrial Engineering
Supervisor: Prof. Dr. Ülkü Gürler
May, 2004

In this thesis, we suggest an approach to statistical parameter estimation when an estimator is constructed by the trajectory observations of a stochastic system and apply the approach to reliability models. We analyze the asymptotic properties of the estimators constructed by the trajectory observations using moments method, maximum likelihood method and least squares method. Using limit theorems for Switching Processes and the results for parameter estimation by trajectory observations, we study the behavior of moments method estimators which are constructed by the observations of a trajectory of a switching process and prove the consistency and asymptotic normality of such estimators. We consider four different reliability models with large number of devices. For each of the models, we represent the system process as a Switching Process and prove that the system process converges to the solution of a differential equation. We also prove the consistency of the moments method estimators for each model. Simulation results are also provided to support asymptotic results and to indicate the applicability of the approach to finite sample case for reliability models.

Keywords: Parameter estimation, Switching Processes, Reliability models.

ÖZET

DEĞİŞEN STOKASTİK MODELLERDE PARAMETRE TAHMİNLEMESİ

GÜLDAL GÜLERYÜZ
Endüstri Mühendisliği, Doktora
Tez Yöneticisi: Prof. Dr. Ülkü Gürler
Mayıs, 2004

Bu çalışmada stokastik sistemlerin örnek yollarını gözlemleyerek oluşturulan tahminleyicilerin bulunmasında kullanılan bir istatistiksel parametre tahmini yaklaşımı önerilmekte ve bu yaklaşım güvenilirlik modellerine uygulanmaktadır. Moment metodu, maksimum benzerlik metodu ve en küçük kareler toplamı metodu kullanılarak, örnek yolların gözlemleriyle oluşturulan tahminleyicinin asimtotik özellikleri incelenmiştir. Değişen süreçlerde örnek yol gözlemleriyle oluşturulan moment metodu tahminleyicisinin davranışı araştırılmış, tutarlılık ve asimtotik normalliği, Değişen süreçlerde limit teoremleri ve örnek yol gözlemleriyle yapılan parametre tahminleme sonuçları kullanılarak ispatlanmıştır. Çok sayıda parçadan oluşan dört farklı güvenilirlik modeli incelenmiştir. Her modelde, sistem süreci Değişen süreç olarak ifade edilmiş ve sistem sürecinin bir diferansiyel denklemin çözümüne yakınsaması ispatlanmıştır. Asimtotik sonuçları desteklemek ve yaklaşımın sonlu örnek durumlarında da güvenilirlik modellerinde kullanılabilirliğini belirtmek amacıyla simülasyon sonuçları verilmektedir.

Anahtar sözcükler: Parametre tahminleme, Değişen Süreçler, Güvenilirlik modelleri.

to my husband and son

Acknowledgement

First and foremost, I can not fully express my gratitude to my advisor, Prof. Dr. Ülkü Gürler whose guidance was most welcome and precious during the realization of this doctoral study. No less in importance is my utter appreciation that should immediately be extended to Assist. Prof. Dr. Murat Fadıloğlu, for his equally significant support that he was kind enough to supply in ample amounts.

I extend my deepest gratitude to Prof. Dr. Barbaros Tansel, a pure perfectionist, for long hours of listening and reading my thesis and letting me know that the less travelled road is usually the one that leads to deeper knowledge. I am also very indebted to Prof. Dr. İsmihan Bayramoğlu and Assist. Prof. Dr. Emre Berk for accepting to read the thesis and I give my humble appreciation in return for their kind assistances in shaping the final form of this study.

My gratitude would not be fully stated without mentioning Prof. Dr. Vladimir V. Anisimov. It has been a deep honor to receive his gracious help at all steps that finally led to a finished study.

I would like to express my sincere thanks to my dear friends Eren Miski Aydın and Banu Yüksel Özkaya for their continued encouragement and support.

I like to extend my gratitude to Güler and Erdal Tezcan whom I owe so much for their unconditional love and support which I could never appreciate more. I also would like to thank to Aynur and Vildan Güleryüz for being with me whenever I need, without asking why.

I extend my meek thanks to Onur Güleryüz, whose addition to our family brought so much happiness to my life, for all his love and understanding. Finally and by no means lastly, I like to express my deepest gratitude to Kemal Güleryüz, for his generous support, love and kindness and I find no words to articulate my appreciation, other than a simple but most momentous ”thank you” that comes sincerely from the bottom of my heart.

Contents

Contents viii

List of Figures x

List of Tables xii

1 Introduction 1

2 Literature Review and Preliminary Work 6

2.1 Literature review 6

2.2 Preliminary Work 11

2.2.1 Analysis of Solutions of Stochastic Equations 13

2.2.2 Asymptotic Behavior of Extreme Sets of Random Functions 15

2.2.3 Switching Processes 17

2.2.4 Averaging Principle and Diffusion Approximation for Switching Processes 21

3 Estimation by Trajectory Observations 26

3.1 Asymptotic Properties of Estimators Constructed by Trajectory Observations 26

3.2 Moments Method - Transient Case 27

3.3 Maximum Likelihood Method 32

3.4 Analysis of Least Squares Method Equation 38

3.5 Parameter Estimation in Switching Models 42

3.5.1 Preliminary Work 42

3.5.2 Moments Method for Switching Processes 46

4 Parameter Estimation in Reliability Models 49

4.1 Model 1: A Reliability Model without Replacement 49

4.2 Model 2: Estimation in a Reliability Model with Replacement . . 55

4.3 Model 3: Reliability Model with N Repairmen 62

4.4 Model 4: A Reliability Model with Probabilistic Chance of Repair 66

4.4.1 When Both Parameters are Unknown 80

4.5 Simulation Results 83

5 Conclusions 92

APPENDIX 101

A Figures 101

B Tables 109

List of Figures

2.1 Switching Processes: An illustration 24

2.2 RPSM: An illustration 25

4.1 Model 1: Illustration 50

4.2 Model 1: A trajectory for the initial transitions 51

4.3 Model 2: Illustration 55

4.4 Model 2: A trajectory for the initial transitions 57

4.5 Model 2: Behavior between the inspections 58

4.6 Model 3: Illustration 63

4.7 Model 4: Illustration 67

4.8 Model 4: Transactions 68

A.1 Model 2: Simulation of trajectory of failed devices for reliability model with replacement ($\lambda_0 = 0.5,\ \beta_0 = 1, T = 5, m = 10$) 101

A.2 Model2: Simulation of trajectory of failed devices for reliability model with replacement ($\lambda_0 = 0.8,\ \beta_0 = 0.6, a = 2,\ T = 2$) 102

A.3 Model3: Simulation of trajectory of failed devices for reliability model with N repairmen ($m = 5, \lambda_0 = 1.5, T = 2, N = 5,\ \mu_0 = 0.1$) 102

A.4 Model 3: Simulation of trajectory of failed devices for reliability model with N repairmen ($m = 10, \lambda_0 = 0.8,\ T = 5, N = 5, \mu_0 = 0.1$) 103

A.5 Model 4: Numerical solution and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$. 103

A.6 Model 4: Numerical solution and analytical approximation of numeric solution when $\lambda_0 = 2$, $\beta_0 = 0.8$ and $a = 2$. 104

A.7 Model 4: Observations y_{nk} for $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ 104

A.8 Model 4: Simulated values for $y_k/(1-r_k)$ versus time when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ 105

A.9 Model 4: Simulated values for $y_k/(1-r_k)$ and numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ 105

A.10 Model 4: Simulated values for y_k and numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ 106

A.11 Model 4: Simulated values for r_k and numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ 106

A.12 Model 4: Simulated values for $y_k/(1-r_k)$ and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ . . 107

A.13 Model 4: Simulated values of y_k and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ 107

A.14 Model 4: Simulated values of r_k and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$ 108

List of Tables

4.1 Model 1: Estimated values for λ_0 83

4.2 Model 1: Effect of h_n to the estimator 84

B.1 Model 2: Estimated values , bias, relative error and mean square error for $\hat{\lambda}$, T=10 (5 runs). 109

B.2 Model 2: Estimated values, bias, relative error and mean square error for $\hat{\lambda}$, T=4 (5 runs). 110

B.3 Model 2: Estimated values, bias, relative error and mean square error for $\hat{\lambda}$, T=1 (5 runs) 110

B.4 Model 2: Estimated values, bias and relative error for $\hat{\beta}$, T=10 . . 111

B.5 Model 2: Estimated values, bias and relative error for $\hat{\beta}$, T=4 . . 111

B.6 Model 2: Estimated values, bias and relative error for $\hat{\beta}$, T=1 . . 112

B.7 Model 3: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ (5 runs). 112

B.8 Model 4: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ using the numeric solution (5 runs). 113

B.9 Model 4: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ using the numeric solution (10 runs) 113

B.10 Model 4: Estimated values, bias, relative error, MSE for $\hat{\lambda}$ using the analytic approach (5 runs). 114

B.11 Model 4: Estimated values, bias, relative error and MSE for λ using the analytic approach (10 runs) 114

B.12 Model 4: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ when r(T) is known (5 runs). 115

B.13 Model 4: Estimated values, bias, relative error, MSE for $\hat{\lambda}$ when $r(T)$ is known (10 runs) 115

B.14 Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ using the numeric solutions (5 runs) 116

B.15 Model 4: Estimated values, bias, relative error, MSE for $\hat{\beta}$ using the analytic approximation (5 runs). 116

B.16 Model 4: Estimated values, bias, relative error, MSE for $\hat{\beta}$ when $r(T)$ is known (5 runs). 117

B.17 Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ and $\hat{\lambda}$ using numeric solutions (5 runs). 117

B.18 Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ and $\hat{\lambda}$ using analytic approximation (5 runs). 118

B.19 Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ and $\hat{\lambda}$ when r(T) is known (5 runs). 118

Chapter 1

Introduction

The data constructed by the observations on the stochastic systems, such as computer and communication systems, queueing and reliability models, are mostly dependent and non homogenous in time. Since the classical parameter estimation methods are mostly oriented to homogeneous and independent data, for dependent observations, for instance the trajectory observations under transient conditions, they are not appropriate for statistical estimation and can not be used to study the asymptotic behavior of estimators.

The main purpose of this study is to investigate the asymptotic behavior of estimators constructed by trajectory observations of Switching Processes and indicate the applicability of the method to statistical estimation problems in reliability models.

We suggest an approach to statistical parameter estimation from observations of trajectories of stochastic systems (trajectory of a stochastic system is a sample path or one particular realization of the system). According to this approach, using statistical estimation methods, we represent the estimators by the solutions of stochastic equations or extreme points of random functions which are integral type functions defined by the observations of the trajectories of stochastic systems.

For moments method and least squares method, we represent the estimator as the solution of a stochastic equation in the form $f(\theta) = 0$ where the function $f(\theta)$ is an additive function constructed by the trajectory observations of the stochastic system. For the maximum likelihood method, we represent the estimator as the extreme point of a random function variable, $F(\theta)$ where in this case $F(\theta)$ is the function constructed by the trajectory observations.

Using averaging type results for additive functions, along with the results about the behavior of solutions of stochastic equations, we study the asymptotic properties of the estimators.

To illustrate this approach consider the following example. Let $\{X(t),t \geq 0\}$ be a continuous time ergodic Markov process with following properties: Assume that $x_k = X(t_k)$ is the imbedded ergodic Markov chain which is homogenous and irreducible with finite state space, where $t_k, k = 1, 2, ..., n$, are the times of jumps. Also assume that the stationary probabilities of the imbedded process exists and defined by π_i, $i = \overline{1, m}$. Let us denote v_j as the exit rate of the process from state j and v_{ij} as the rate of transition from state i to state j so that $v_j = \sum_{i=1, i \neq j}^{m} v_{ji}$.

Suppose that we have an independent family of random variables $\{\gamma_k(i), i \in 1, 2, ..m\}, k = 1, 2, ...$ with distributions not depending on k. Also suppose that the first moments of random variables $\{\gamma_k(i), i \in 1, 2, ..m\}$ exist and belong to parametric family of functions $\{g(\theta, i)\ \theta \in \Theta,\ i \in \mathcal{R}^r\}$ where, $E\gamma_1(i) = g(\theta_0, i) = g(i)$. We observe the variables $x_k = X(t_k)$ and $y_k = \gamma_k(x_k)$ at the times of jumps t_k on the interval $[0, T]$ for $k \leq v(T)$, where $v(T)$ is the number of observations. Then the moments method estimator for the unknown parameter θ is the solution of the following equation:

$$\frac{1}{T} \sum_{k=1}^{v(T)} g(\theta, x_k) - \frac{1}{T} \sum_{k=1}^{v(T)} \gamma_k(x_k) = 0.$$

Let us denote

$$f_T(\theta) = \frac{1}{T} \sum_{k=1}^{v(T)} g(\theta, x_k) - \frac{1}{T} \sum_{k=1}^{v(T)} \gamma_k(x_k)$$

and the solution of equation $f_T(\theta) = 0$ by $\hat{\theta}$. Since x_k is an ergodic process, using the Law of large Numbers for Markov Processes, it is known that [44],

$$\frac{1}{T}v(T) \xrightarrow{\mathrm{P}} \frac{1}{\sum_{i=1}^{m} \pi_i / v_i}.$$

Multiplying and dividing f_T by $v(T)$ we have the following expression for f_T :

$$\frac{v(T)}{T}\left[\frac{1}{v(T)}\sum_{k=1}^{v(T)} g(\theta, x_k) - \frac{1}{v(T)}\sum_{k=1}^{v(T)} \gamma_k(x_k)\right].$$

Then the function $f_T(\theta)$ converges in probability to the function $f_0(\theta)$ where,

$$f_0(\theta) = \frac{1}{\sum_{i=1}^{m} \pi_i / v_i}\left[\sum_{i=1}^{m} \pi_i g(\theta, i) - \sum_{i=1}^{m} \pi_i g(\theta_0, i)\right].$$

It is obvious that, θ_0 is the solution of the equation $f_0(\theta) = 0$. The question of interest here is, under what conditions and in what sense $\hat{\theta}$ converges to θ_0.

For homogenous and ergodic Markov processes such convergence results are expected. However for more general classes of processes it may be difficult. Among those processes we can also consider the Switching Processes.

Switching Processes have the property that the character of the process switches in epochs of time which may be a random functional of the previous trajectory. At times t_k the switches occur and the behavior of the process depends only on X_k which is the discrete switching component and S_k which is the value of the previous trajectory at time t_k.

Switching Processes are very suitable in analyzing and asymptotically investigating stochastic systems with 'rare' and 'fast' switchings [6]. In particular, sums of random variables, processes with independent increments, random evolutions, dynamic systems with random perturbations, queuing systems, some stochastic networks and branching processes can be analyzed by using the properties of Switching Processes.

Let us illustrate a possible application of Switching Processes on the following example.

Consider a general queuing system $GI/M/1/\infty$. Assume that the incoming process is a recurrent process and the service rate is $\mu(Q)$ given that $Q(t) = Q$. We observe the value of queue $Q_k = Q(t_k)$ at times of arrivals t_k such that $t_1 < t_2 < t_3$.... The process between the arrival times t_k and t_{k+1} is a birth and death process with pure death property. In this case the process $Q(t)$ is not a Markov Process but it can be described as Switching Process with the switching times t_k, $k \geq 0$ where between switching times it behaves as a birth and death process.

We mainly consider parameter estimation when the trajectory of the stochastic system under investigation can be represented as a Switching Process. To illustrate the results, we apply our asymptotic results to several reliability models with large but finite number of devices. In applications, we represent the trajectory of the Reliability systems as Switching Processes, and using the result about the behavior of stochastic equations and extreme points of random functions along with the limit theorems for Recurrent Process of semi-Markov type (a special class of Switching Processes) we study the asymptotic behavior of moments method estimators. We prove the consistency and asymptotic normality of such estimators.

Even for stationary and homogenous systems, explicit characteristics and analytic representation may not be possible to find so that some results of the estimation may be difficult to achieve. But especially for nonstationary cases, this may cause bigger problems which are not easy or sometimes impossible to solve analytically. Simulation methods may help the asymptotic investigation in several ways. For example, in the case when the stationary distribution is unknown it may be possible to find it with simulation and use the results in the corresponding analytic relations for asymptotic properties. While finding a limiting point, we need to define some limiting function. If we don't have the underlying distribution for nonstationary cases we can approximate it by simulating corresponding random variables and functions on the trajectory of the stochastic system.

Simulation is also an important part of our study. Our theoretical calculations for reliability models are illustrated with the simulations. We simulate and observe the trajectories of these reliability systems. Using our theoretical calculations, we estimate the unknown parameters and verify our asymptotic results for finite samples also.

The thesis is organized as follows:

In the second Chapter, we give a literature review on parameter estimation approaches which are investigated in the literature. We also give the necessary definitions and theorems for further studies in second chapter. We present the Switching Processes and also give the limit theorems for Recurrent Process of semi-Markov type.

In the third Chapter, we consider the asymptotic properties of Moments method type, maximum likelihood and least squares method estimators, constructed by trajectory observations of stochastic processes. We also present moments method estimators which are constructed by the trajectory observations of Switching Processes.

The main part of the thesis is presented in Chapter four. The applications to Reliability models are considered, on four different but related models. Simulation results of the estimation procedure are also given to support our asymptotic results.

We finally give the conclusions in Chapter five.

Chapter 2

Literature Review and Preliminary Work

2.1 Literature review

In literature, parameter estimation studies for stochastic processes are usually devoted to diffusion processes and there are various types of estimation techniques, mostly related to martingale estimation.

Bibby and Sørensen [20] consider different martingale estimating functions of a diffusion process. They show that the estimators obtained are asymptotically normal and consistent and discuss the results of simulation studies of some specific examples. Kutoyants [35] considers the parameter estimation for the Gaussian, diffusion and non-homogenous Poisson processes. Barndorff-Nielsen and Sørensen [19] review the asymptotic likelihood theory for stochastic processes and particularly investigate the martingale properties. They also give some examples such as Birth-and-Death processes, Gaussian autoregressive processes and stochastic differential equations to show that the likelihood function for many situations are martingales and give the asymptotic results for maximum likelihood estimator.

We will briefly consider different parameter estimation approaches considered in the literature for different models and processes.

Kutoyants [35] considers a non-homogenous Poisson Process. Let $x_T = \{x(t), 0 \leq t \leq T\}$ be a Poisson process with intensity $S_T(\theta) = \{S(\theta,t), 0 \leq t \leq T\}$ and the unknown parameter $\theta \in (\alpha, \beta)$. Under their conditions, they form the likelihood function and find the estimator for the unknown parameter β for a particular case when $S_T(\theta, t) = \theta f(t)$.

Anisimov [3] and Anisimov, Orazklychev [12] consider asymptotic properties of parameter estimators for Poisson type processes switched by some ergodic sequence and asymptotic properties of maximum likelihood estimators constructed by observations on trajectories of recurrent processes of semi-Markov type.

Saldanha, et. at. [45] consider the estimation of rate of occurrence of failures (ROCOF) of a non-homogenous Poisson process when the rate of occurrence of failure depends on time. If we denote by $v(t)$ the rate of occurrence of failures, then $v(t)$ is defined as the time derivative of the number failures in the assigned time interval [18]. For two different forms of $v(t)$, ($v(t) = exp(\beta_0 + \beta_1 t)$ and $v(t) = \gamma\delta t^{-\delta-1}$), they consider the maximum likelihood estimation of parameters of $v(t)$ from observations at the times of failure of the system, for different values of stopping time (i.e. stop at a fixed time T, stop after n'th transition, stop at d'th departure, stop at m'th arrival).

Keiding [30] considers the maximum likelihood method for parameter estimation in Birth-and-Death processes. Let the population size at time t be X_t. With birth rate λ and death rate μ, they form the likelihood function in terms of λ, μ and X_t and estimate the unknown parameters λ and μ.

Keiding [30] also considers the case of discrete observations. Denote by $X_{n\tau}, n = 1, 2, 3, ..., k$ the observations at times $\tau, 2\tau, 3\tau, ..., k\tau$. The process studied, has particles which may or may not have offsprings, and the number of particles among the $X_{(n-1)\tau}$ that have 0 offspring is known and denoted by C_n. The likelihood function is represented in term of λ, μ, τ and C_n.

Applications of parameter estimation in Birth-and-Death Processes also vary according to environment. Phelan [39], [40], [41] considers the case of Birth-and-Death on a flow. Birth-and-Death on a flow refers to a particle system on a Brownian Motion [40]. Generally it is a Birth-and-Death process on a Brownian environment. Phelan [39] develops likelihood methods for parametric estimation of system parameters from a particle process which is observed over a fixed period of time. The follow up study of Phelan [41] considers the asymptotic properties of the estimators as the process is observed over a long period of time.

A different approach of estimation in Birth-and-Death processes is also considered by Zeifmann [54], [55]. Zeifmann [55] estimates the bounds for state probabilities for some nonhomogenous Birth-and-Death processes with known intensity functions and gives some examples of application.

Watson and Yip [52] extends the work of Chao and Severo [22] for parameter estimation for pure birth process. They consider a simple stochastic epidemic model with population size N and infection rate β. If we denote the number of infective at time t by $I(t)$, at times t_k the number of infective, I_k are observed. Note that the sequential observation times, t_k are nonrandom. Using martingale techniques they estimate the unknown parameter β.

Volokh [50] studies the parameter estimation on a function of random variables which have exponential type distributions.

Wolff [53] discusses the maximum likelihood estimating and likelihood ratio tests for a class of ergodic queueing models. Basawa and Prabhu (1998) proves the consistency and asymptotic normality of MLE for single server queues.

Acharya [1] also studies MLE estimators and rate of convergence of the distribution of MLE of the arrival and services rates in a GI/G/1 queueing system. As a special case, consider an M/M/1 queueing system. Interarrival times $u_k, k \geq 1$ and the service times $v_k, k \geq 1$ are independent and identically distributed random variables with densities $f(u,\theta) = \theta exp(-\theta u)$ and $g(v,\phi) = \phi exp(-\phi v)$ respectively. The system is observed in the time interval $(0,T]$, where T is a suitable stopping time. Let $A(T)$ be the number of arrivals and $D(T)$ be the number of departures in the time interval $(0,T]$. They form the log-likelihood function and estimate the unknown parameters θ and ϕ.

Maintenance related studies generally consider the cost optimization and finding the optimal maintenance policy. A survey of maintenance models for multi-component systems is given by Cho and Parlar [21].

An interesting study by Heidergott [26] considers a multicomponent maintenance system controlled by an age replacement policy. The main idea of the study is to estimate the threshold age θ of the components to minimize the total cost of operation. They consider a system with n components. The lifetimes of components are independent and identically distributed with distribution function F and F is assumed to be continuous. When a component fails, it is immediately replaced at a cost r and all components with age older then θ are preventively replaced at a cost p. The long range average cost per time unit for θ is denoted by $C(\theta)$. They obtain an estimator θ to minimize the long-run costs per time unit so that

$$C(\theta^*) = \min_{\theta \in \Theta} C(\theta) \tag{2.1}$$

where Θ is closed bounded region.

Without finding the explicit representation for $C(\theta)$, they use the stochastic approximation to solve (2.1).

Most of the studies in the parameter estimation literature consider the case of independent observations, such as Ibragimov, K'hasminski [28], Kutoyants [35] and Prakaso Rao [43]. Another main study direction necessarily uses the martingale techniques as in the works of Barndorf-Nielsen and Sørensen [19], Bibby and Sørensen [20] and Lipster and Shiryaev [36]. Some problems in the theory of statistical investigation are studied by Dupocava and Wets [23], Pflug [37] and Shapiro [47]. Kutoyants [35] considers some nonclassical problems using direct probabilistic methods. Kaniovski and Pflug [29] and Pflug [38] consider the stationary conditions for parameter estimation.

Using both simulation in finite samples and asymptotic theory for infinite samples, moments method estimators are derived and compared to maximum likelihood estimators for finite samples by Shi [48].

Some results on the statistical parameter estimation by trajectory observations are given by Anisimov [8], Bibby and Sørensen [20], Kutoyants [35].

Several results devoted to analysis of solutions of stochastic equations which are constructed for parameter estimation are considered by Anisimov and Kaibah [14], Anisimov [8] , Anisimov and Pflug [16] and Korolyuk and Swishchuk [33]. Asymptotic behavior of maximum likelihood estimators as function of the length of interval are considered in the papers of Anisimov and Orazklichev [12] and Anisimov [9].

Weak convergence and convergence in probability of sets of extreme points of random fields to the extreme point of some limiting field and basic applications to parameter estimation are studied by Anisimov and Seilhamer [17] . Their results are very closely connected with the results about the convergence of stochastic infima given by Dupacova and Wets [23] and Salinetti and Wets [46].

Parameter estimation for switching stochastic systems are not widely considered in the literature.

Switching Processes are described in the paper of Anisimov [7] as the generalization of Markov processes homogenous in the second component [24], processes with independent increments and semi-Markov switches [2], Markov processes with semi-Markov interference of chance and Markov and semi-Markov random evolutions [27], [31], [42].

Subclasses of Switching Processes are considered for different applications by Anisimov [2], [7] and Anisimov and Aliev [11]. For processes with independent increments and Markov and semi-Markov switches, law of large numbers and central limit theorem were proved in the literature [33], [34], [51]. Based on the asymptotic properties of Recurrent Processes of semi-Markov type (RPSM), a special type of Switching Processes, and theorems about the convergence of recurrent sequences to solutions of stochastic differential equations [2], [7], it is proved by Anisimov [7] that for the additive type functionals on the RPSM trajectories, the normed trajectory of the functional converges in probability to some non stochastic differential equation. Using another approach Averaging principle type results for stochastic differential equations are also given by Giego and Hersh [25], Hersh [27], Khas'minskii [32] and Skorokhod [49].

2.2 Preliminary Work

In different models that appear in statistical parameter estimation from observations of trajectories of stochastic systems, estimators can be represented by the solutions of stochastic equations or extreme points of random functions which are integral type functions defined by the observations on the trajectories of stochastic systems.

We consider a stochastic model in which different classes of problems appear during estimation process. Let $S(t)$ be the trajectory of a stochastic system observed on the interval [0,T], $T \geq 0$. Let $t_k,\ k = 1, 2, ...$ be times of observations. Assume that, we observe the variables $s_k = S(t_k)$ and $y_k = \gamma(s_k)$ where $\gamma(\alpha)$ is an independent family of random variables. Assume also that there is an unknown system parameter θ which we want to estimate.

Let the total number of observations on the interval [0,T] be n. Under different additional assumptions and situations, we can represent moments type, maximum likelihood and least squares method estimators of the unknown parameter θ in terms of solutions of equations which are in the form $f(\theta) = 0$ or extreme points of a function $F(\theta)$ where $\theta \in \Theta$ and Θ is a closed bounded set in $\mathcal{R}^r$. For each of these cases $f_n(\theta)$ and $F_n(\theta)$ are constructed by the trajectory observations. Note that, when we study the asymptotic behavior of the estimator we usually consider the case when T or n (or some other parameter) goes to infinity.

Assume that, the solution of the equation $f(\theta) = 0$ exists and is defined as $\{\theta\}$. Additionally, suppose that $f(\theta)$ converges (in some sense) to a limiting function $f_0(\theta)$ where θ_0 is the solution of equation $f_0(\theta) = 0$. The problem here is that under what conditions and in what sense the set of solutions of $f(\theta) = 0$ converges to θ_0 as $T \to \infty$.

Another problem can be described as finding the conditions of the convergence of sets of extreme points of random functions to some limiting point. Let us denote the set of points of global minimum for the function $F(\theta)$ by $\{\theta\} = \arg\min_{\theta\in\Theta} F(\theta)$. We can study the convergence of $\{\theta\}$ to θ_0 when the function $F(\theta)$ converges in some sense to a limiting function $F_0(\theta)$ as $n \to \infty$.

Another important but different kind of problem is to find the conditions of convergence $f(\theta) \to f_0(\theta)$ and $F(\theta) \to F_0(\theta)$ themselves. Usually, $f(\theta)$ and $F(\theta)$ are constructed as additive functions on the trajectories of the systems. In this case, to study the conditions of convergence of $f(\theta) \to f_0(\theta)$ and $F(\theta) \to F_0(\theta)$ on the trajectory of stochastic systems, we need to study the behavior of additive functional which can be found for wide classes of stochastic systems such as Markov processes.

We can also examine the behavior of the estimator which is constructed as a solution of some stochastic equation or as an extreme point of some random function on the trajectory of some stochastic system, in terms of the length of the interval of observations. Let $F(\theta, t)$, $\theta \in \Theta, t \in [0, T]$ be a random function and $\{\theta(t)\} = \arg\min_{\theta\in\Theta} F(\theta, t)$ be a set valued process. Consider the case where $F(\theta, t)$ converges in the region $\theta \in \Theta, t \in [0, T]$ to some limiting function $F_0(\theta, t)$. The problem in this case, is to find under which conditions and in what sense the sequence of set valued process $\{\theta(t)\}$ converges to $\theta_0(t) = \arg\min_{\theta\in\theta} F_0(\theta, t)$ on the interval [0,T].

In such cases, we need to study the asymptotic properties of solutions of stochastic equation and extreme sets of random functions in order to be able to analyze the problems of statistical parameter estimators.

Using these results and limiting theorems for Switching Proceesses along with statistical estimation methods we can study the asymptotic behavior of the statistical estimators for stochastic processes which can be described in terms of Switching Processes.

In this part we give the necessary definitions and theorems from the literature which are necessary for the further chapters of the thesis.

2.2.1 Analysis of Solutions of Stochastic Equations

This section mainly follows from the results of Anisimov and Pflug, [16] which are related to the asymptotic behavior of solutions of stochastic equations.

We now give necessary definitions in reference to Anisimov, Guleryuz [13].

Definition 2.2.1 *(Condition of Separateness): We say that the r dimensional function $g(\theta), \theta \in \Theta$ where Θ is a bounded region in $\mathcal{R}^r$, satisfies the condition of separateness* **S** *if there exists such $\delta > 0$ that for any $y \in \mathcal{R}^r, |y| < \delta$ the equation*

$$g(\theta) = y$$

has a unique solution and the solution θ_0 of the equation $g(\theta_0) = 0$ is the inner point of the region Θ.

Note that if the function $g(\theta)$ is random, and satisfies the condition **S** it means that the condition of separateness is satisfied with probability one.

We also like to mention that, if a function $f(\theta)$ is a random function then

1. for each θ, $f(\theta)$ is a random variable,
2. if $\theta \in [0, \infty)$, then $f(\theta)$ is a random process,
3. if $\theta \in \mathcal{R}^r$, then $f(\theta)$ is a random field.

Let $f_n(\theta), t \geq 0, \theta \in \Theta, n > 0$ be a sequence of continuous random functions with values in $\mathcal{R}^r$, where Θ is some bounded region in $\mathcal{R}^r$. Consider a stochastic equation in vector form

$$f_n(\theta) = 0, \tag{2.2}$$

and denote the set of all possible solutions by $\{\theta_n\}$. Hence, the random set $\{\theta_n\}$ is constructed as the solution set of the equation $f_n(\theta) = 0$.

Definition 2.2.2 *(Modulus of Continuity): For any function $f(\theta), \theta \in \Theta$ modulus of continuity in the vicinity of c is defined as,*

$$\Delta_U(c, f(\cdot)) = \sup_{|\theta_1-\theta_2|<c, \theta_1\in\Theta, \theta_2\in\Theta} |f(\theta_1) - f(\theta_2)|.$$

Definition 2.2.3 *(Uniform Convergence): We say that the sequence of functions $f_n(\theta)$ uniformly converges (U-converges) to the function $f_0(\theta)$ on the set Θ if:*

1. *For any $k = 1, 2, ...$ and for any $\theta_1, \theta_2, ...\theta_k \in \Theta$ the multidimensional distribution function of vector $(f_n(\theta_i), i = \overline{1, k})$ weakly converges to the distribution function of vector $(f_0(\theta_i), i = \overline{1, k})$;*

2. *For any $\varepsilon > 0$*

$$\lim_{c\to+0} \limsup_{n\to\infty} P\{\Delta_U(c, f_n(\cdot)) > \varepsilon\} = 0.$$

We like to mention that, the function $f_0(\theta)$ can be random or deterministic.

The following theorem related to the solutions of stochastic equations follows from Anisimov, Kaibah [14] and Anisimov, Pflug [16].

Theorem 2.2.1 *1). Suppose that the sequence of functions $f_n(\theta)$ U-converges in each set $K \subset \Theta$ to the function $f_0(\theta)$ which can be random or deterministic. Suppose also that $f_0(\theta)$ satisfies the condition of separateness* **S**, *and the point θ_0 is the solution of a limiting equation:*

$$f_0(\theta_0) = 0. \tag{2.3}$$

Then with probability which tends to one the solution of the equation (2.2) exists and the sequence of sets $\{\theta_n\}$ converges in probability to θ_0. That is

$$\{\theta_n\} \xrightarrow{\text{P}} \theta_0. \tag{2.4}$$

2). Suppose further that θ_0 is a non-random point and there exists $\beta > 0$ and a non-random sequence $v_n \to \infty$ such that for any $L > 0$ the sequence of random functions $v_n^\beta f_n(\theta_0 + v_n^{-1} u)$ U-converges in the region $\{|u| \leq L\}$ to some (random) function $\eta_0(u)$, which satisfies the condition **S** *and the point κ_0 is the solution of the limiting equation*

$$\eta_0(\kappa_0) = 0. \tag{2.5}$$

Then there exists a solution $\widehat{\theta_n}$ of the equation (2.2) such that the sequence

$$v_n(\widehat{\theta_n} - \theta_0) \overset{\mathrm{w}}{\Rightarrow} \kappa_0. \tag{2.6}$$

We will use Theorem 2.2.1 to prove the consistency of the estimators when the estimators are represented as the solution of stochastic equation $f_n(\theta) = 0$.

2.2.2 Asymptotic Behavior of Extreme Sets of Random Functions

This section follows from the results of Anisimov, Seilhamer [17].

First, we give some necessary definitions in reference to Anisimov [8].

Definition 2.2.4 *Let G_n be a sequence of random sets in Θ. We say that the sequence G_n converges in probability to some point g_0 which can be random or non-random, if $\rho(g_0, G_n) \xrightarrow{\mathrm{P}} 0$, where $\rho(g, G) = \sup_{z \in G} ||z - g||$.*

We denote this convergence as $G_n \xrightarrow{\mathrm{P}} g_0$.

Definition 2.2.5 *Let G_n be a sequence of random sets in Θ. We say that the sequence G_n weakly converges to some random variable γ_0, if g_n weakly converges to γ_0 for any subsequence g_n such that $P\{g_n \in G_n\} = 1$.*

We denote this convergence as $G_n \overset{\mathrm{w}}{\Rightarrow} \gamma_0$.

Let at each $n \geq 0$, $F_n(\theta), \theta \in \Theta \subset R^r$ be a random function with values in R, Θ is a bounded closed set, n is the parameter of series.

Consider the function $\underline{F}(\theta) = \liminf_{\theta' \to \theta} F(\theta')$. If the function $F(\theta)$ is random then this limit is determined for any realization of $F(\theta)$. Let

$$\{\theta_n\} = \arg\min_{\theta \in \Theta} \underline{F}_n(\theta).$$

Here $\{\theta_n\}$ is the set of points of global minimum for the function $F_n(\theta)$. Hence, the random set $\{\theta_n\}$ is constructed as the points of global minimum for the function $F_n(\theta)$.

Definition 2.2.6 *(Condition of Separateness* ***S2****): The condition of separateness* ***S2*** *is satisfied if : with probability one $F_0(\theta_0) < F_0(\theta')$ for any random variable θ' given on the same probabilistic space and such that $\theta' \neq \theta_0$ with probability one, where*

$$\theta_0 = \arg\min_{\theta \in \Theta} F_0(\theta).$$

Now according to Anisimov, Seilhamer [17] we give two theorems, concerning the convergence of the sequence of sets $\{\theta_n\}$.

Theorem 2.2.2 *Let $F_n(\theta)$ be the sequence of random functions and following conditions are true:*

1) There exists a continuous random function $F_0(\theta)$ such that $F_n(\theta)$ U-converges to $F_0(\theta)$;

2) Condition of ***S2*** *is satisfied.*

Then

$$\{\theta_n\} \stackrel{w}{\Rightarrow} \theta_0. \tag{2.7}$$

Note that if the function $F_0(\theta)$ is non-random, then under the same conditions we have that

$$\{\theta_n\} \xrightarrow{\mathrm{P}} \theta_0. \tag{2.8}$$

The proof is given by Anisimov and Seilhamer [17].

Consider now the behavior of the normed deviation for $\{\theta_n\}$. Let us consider the random function

$$A_n(z) = \nu_n^{\beta}(F_n(\theta_0 + \frac{1}{\nu_n}z) - F_n(\theta_0))$$

as a function of a new argument $z \in R^r$.

Theorem 2.2.3 *Let the conditions of Theorem 2.2.2 hold and a nonrandom sequence $v_n \to \infty$ and a value $\alpha > 0$ exist such that for any $L > 0$ the sequence of functions $A_n(z)$ U-converges to some random function $A_0(z)$ in the region $|z| \leq L$. Suppose also that the point $\kappa_0 = \arg\min_z A_0(z)$ is a proper random variable (that is $P\{|\kappa_0| < \infty\} = 1$) and with probability one satisfies the condition **S2** of separateness.*

Then there exists a subsequence of points of local minimum $\tilde{\theta}_n$ for the function $F_n(\theta)$ such that

$$\nu_n(\tilde{\theta}_n - \theta_0) \overset{\mathrm{W}}{\Rightarrow} \kappa_0. \tag{2.9}$$

The proof is also given by Anisimov and Seilhamer [17].

2.2.3 Switching Processes

In this part we consider the description of Switching Processes (SP) and a subclass of Switching Processes, Recurrent Process of semi-Markov type ($RPSM$). We also give the limit theorems for Recurrent Process of semi-Markov type.

Switching Processes are described as two-component processes $(x(t), \zeta(t)),\ t \geq 0$, with the property that there exist a sequence of epochs $t_1 < t_2 < \cdots$ such that on each interval $[t_k, t_{k+1}),\ x(t) = x(t_k)$ and the behavior of the process $\zeta(t)$ depends on the value $(x(t_k), \zeta(t_k))$ only. The epochs t_k are switching times and $x(t)$ is the discrete switching component [6].

Note that switching times may be determined by external factors and also by inner and interconnected factors. In general switching times may be some random functions of the previous trajectory of the system [7].

2.2.3.1 Switching processes

Now we give a general construction of a Switching Process (SP). Let

$$\mathcal{F}_k = \{(\zeta_k(t, x, \alpha), \tau_k(x, \alpha), \beta_k(x, \alpha)), t \geq 0, x \in X, \alpha \in \mathcal{R}^r\}, k \geq 0$$

be jointly independent parametric families. At each fixed k, x, α, also let $\zeta_k(t, x, \alpha)$ be a random process in Skorokhod space $\mathcal{D}^r_\infty$. Note that Skorokhod space consists of the functions with discontinuities of type I. Such functions may have finite jumps and are right continuous at the time of jumps. The representation $\mathcal{D}^r_\infty$ indicates that the function is r dimensional and is defined on the interval $[0, \infty)$.

Let also at each fixed $k, x, \alpha,\ \ \tau_k(x, \alpha),\ \beta_k(x, \alpha)$ be random variables which are possibly dependent on $\zeta_k(\cdot, x, \alpha)$ and $\tau_k(\cdot) > 0,\ \beta_k(\cdot) \in X$. Let also (x_0, S_0) be an initial value, independent of $\mathcal{F}_k, k \geq 0$. We put

$$t_0 = 0,\ t_{k+1} = t_k + \tau_k(x_k, S_k),\ S_{k+1} = S_k + \xi_k(x_k, S_k),$$

$$x_{k+1} = \beta_k(x_k, S_k),\ k \geq 0, \tag{2.10}$$

where $\xi_k(x, \alpha) = \zeta_k(\tau_k(x, \alpha), x, \alpha)$, and set

$$\zeta(t) = S_k + \zeta_k(t - t_k, x_k, S_k), \tag{2.11}$$

$$x(t) = x_k, \quad \text{as} \quad t_k \leq t < t_{k+1},\ t \geq 0. \tag{2.12}$$

Then a two-component process $(x(t), \zeta(t)),\ t \geq 0$ is called a SP [6], [7]. In concrete applications the component $x(\cdot)$ usually means some random environment, and $S(\cdot)$ means the trajectory of the system. We should also mention that the general construction of a SP allows the dependence (feedback) between both components $x(\cdot)$ and $S(\cdot)$. Figure (2.1) illustrates a behavior of components $S(t), \zeta(t)$ and $x(t)$.

2.2.3.2 Recurrent Processes of semi-Markov Type

Let $\mathcal{F}_k = \{(\xi_k(\alpha), \tau_k(\alpha)), \alpha \in R^r\}, k \geq 0$, be jointly independent families of random variables with values in $R^r \times [0, \infty)$. Let also S_0 be a random variable which is independent of $\mathcal{F}_k, k \geq 0$ and with values in $\mathcal{R}^r$. We assume the measurability in α of variables introduced concerning σ-algebra $\mathcal{B}_{R^r}$. Denote

$$t_0 = 0,\ t_{k+1} = t_k + \tau_k(S_k),\ S_{k+1} = S_k + \xi_k(S_k),\ k \geq 0 \tag{2.13}$$

and

$$S(t) = S_k \text{ as } t_k \leq t < t_{k+1},\ t \geq 0. \tag{2.14}$$

Then the process $S(t)$ forms a Recurrent Process of a Semi-Markov type ($RPSM$) (Anisimov and Aliev [11]). Figure (2.2) shows and illustration of RPSM.

We mention that the representation may depend on scaling factors according to the construction of the process. We like to give another representation in reference to Anisimov and Guleryuz [13], which we will use in Chapter 4, for RPSM .

Consider the case when $\zeta_n(t, \theta)$ is a trajectory of a Switching Process. We fix θ and for simplicity omit it. Let for each n=1,2..., $\mathcal{F}_{nk} = \Big\{(\xi_{nk}(\alpha),\ \tau_{nk}(\alpha)), \alpha \in \mathcal{R}^r\Big\}, k \geq 0,$ be jointly independent families of random vectors with values in $\mathcal{R}^r \times [0, \infty)$ and distributions not depending on index k, and s_{no} be an initial value in $\mathcal{R}^r$ independent of $F_{nk},\ k \geq 0$. Let δ_n be some scaling factor, $\delta_n \to 0$ as $n \to \infty$. We construct the following recurrent sequences:

$$
\begin{aligned}
t_{no} &= 0, \\
t_{nk+1} &= t_{nk} + \tau_{nk}(s_{nk})\delta_n, \\
s_{nk+1} &= s_{nk} + \xi_{nk}(s_{nk})\delta_n
\end{aligned}
\tag{2.15}
$$

and denote $\zeta_n(t) = s_{nk}$, as $t_{nk} \leq t < t_{nk+1}$, $t \geq 0$. Then $\zeta_n(t), t \geq 0$, is a Recurrent Process of a semi-Markov type.

Consider following models of Switching Processes as example according to Anisimov [7].

Let $\{f(x, \alpha), \alpha \in \mathcal{R}^r\}$, $x \in X$ be a family of deterministic functions with values in $\mathcal{R}^r$, $\Gamma_k = \{\gamma_k(x, \alpha), x \in X, \alpha \in \mathcal{R}^r\}, k \geq 0$, be jointly independent families of random variables with values in $\mathcal{R}^r$ and $x(t), t \geq 0$ be a SMP in X independent of introduced families Γ_k. Put $x_k = x(t_k)$ and denote by $0 = t_0 < t_1 < ...$ sequential times of jumps for the process $x(t)$. We introduce the process $\zeta(t)$ as follows: $\zeta(0) = \zeta_0$ and

$$d\zeta(t) = f(x_k, \zeta(t))dt, t_k \leq t < t_{k+1},$$

$$\zeta(t_{k+1} + 0) = \zeta(t_{k+1} - 0) + \gamma_k(x_k, \zeta(t_{k+1} - 0)), k \geq 0.$$

Then the process $\zeta(t)$ forms a dynamical system with semi-Markov switches.

A class of SP's also gives possibility to describe various classes of stochastic queueing models such as some state-dependent queueing systems and networks.

For these models switching times are usually times of any changes in the system (Markov models), times of jumps of the environment (in case of external semi-Markov environment), times of exit from some regions for the process generated by queue, waiting times, etc. Several examples of switching queueing systems are given by Anisimov [10].

2.2.4 Averaging Principle and Diffusion Approximation for Switching Processes

This section exposes the results of Anisimov [7] for limit theorems of Recurrent process of semi-Markov type. We will consider the process on the interval $[0, nT]$, $n \to \infty$ and characteristics of the process depend on the parameter n in such a way that the number of switches on each interval $[na, nb]$, $0 < a < b < T$ tends, by probability, to infinity.

2.2.4.1 Averaging Principle (AP) for RPSM

Let us first consider Averaging Principle for simple $RPSM$. Note that Averaging Principle type theorems for Switching Processes are studied by Anisimov [5], [7], Anisimov and Aliev [11]. Below we give the construction and related theorem according to Anisimov [7].

Let for each n=1,2..., $\mathcal{F}_{nk} = \{(\xi_{nk}(\alpha)), \tau_{nk}(\alpha), \alpha \in \mathcal{R}^r\}$, $k \geq 0$ be jointly independent families of random variables taking values in $\mathcal{R}^r \times [0, \infty)$, with distributions do not depend on index k, and let S_{n0} be independent of $\mathcal{F}_{nk}, k \geq 0$ initial value in $\mathcal{R}^r$. Put

$$t_{n0} = 0, t_{nk+1} = t_{nk} + \tau_{nk}(S_{nk}), S_{nk+1} = S_{nk} + \xi_{nk}(S_{nk}),\ k \geq 0,$$

$$S_n(t) = S_{nk} \ \text{ as } \ t_{nk} \leq t < t_{nk+1},\ t \geq 0. \tag{2.16}$$

Assume that there exist functions $m_n(\alpha) = E\tau_{n1}(n\alpha)$, $b_n(\alpha) = E\xi_{n1}(n\alpha)$.

Theorem 2.2.4 *Averaging Principle*

Suppose that for any $N > 0$

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sup_{|\alpha|<N} \Big\{ E\tau_{n1}(n\alpha)\chi(\tau_{n1}(n\alpha) > L) +$$

$$+ E|\xi_{n1}(n\alpha)|\chi(|\xi_{n1}(n\alpha)| > L) \Big\} = 0, \tag{2.17}$$

as $\max(|\alpha_1|, |\alpha_2||) < N$,

$$|m_n(\alpha_1) - m_n(\alpha_2)| + |b_n(\alpha_1) - b_n(\alpha_2)| < C_N|\alpha_1 - \alpha_2| + \alpha_n(N), \tag{2.18}$$

where C_N *are some bounded constants,* $\alpha_n(N) \to 0$ *uniformly in* $|\alpha_1| < N, |\alpha_2| < N$*, and there exist functions* $m(a) > 0$*,* $b(a)$ *and a proper random variable* s_0 *such that as* $n \to \infty$ $n^{-1}S_{n0} \xrightarrow{\mathrm{P}} s_0$*, and for any* $\alpha \in \mathcal{R}^r$

$$m_n(\alpha) \to m(\alpha) > 0,\ b_n(\alpha) \to b(\alpha). \tag{2.19}$$

Then

$$\sup_{0 \le t \le T} |n^{-1}S_n(nt) - s(t)| \xrightarrow{\mathrm{P}} 0, \tag{2.20}$$

where

$$s(0) = s_0, ds(t) = m(s(t))^{-1}b(s(t))dt, \tag{2.21}$$

and T *is any positive number such that* $y(+\infty) > T$ *with probability one, where*

$$y(t) = \int_0^t m(\eta(u))du, \tag{2.22}$$

$$\eta(0) = s_0, d\eta(u) = b(\eta(u))du \tag{2.23}$$

(it is supposed that a solution of equation (2.23) exists on each interval and is unique).

We like to mention that the condition (2.18) is a modification of Lipschitz condition, and we use the form that, as $\max(|\alpha_1|, |\alpha_2|) < N$, $N > 0$, C_N are some bounded constants and $\alpha_n(N) \to 0$ uniformly in $|\alpha_1| < N, |\alpha_2| < N$, the following condition for a function $f(x)$ is satisfied:

$$|f(x, \alpha_1) - f(x, \alpha_2)| < C_N|\alpha_1 - \alpha_2| + \alpha_n(N).$$

2.2.4.2 Diffusion Approximation for RPSM

Now we consider a convergence of the process $\gamma_n(t) = n^{-1/2}(S_n(nt) - ns(t)), t \in [0, T]$ to some diffusion process according to Anisimov [7]. Denote

$$\tilde{b}_n(\alpha) = m_n(\alpha)^{-1}b_n(\alpha),\ \tilde{b}(\alpha) = m(\alpha)^{-1}b(\alpha),$$

$$\rho_n(\alpha) = \xi_{n1}(n\alpha) - b_n(\alpha) - \tilde{b}(\alpha)(\tau_{n1}(n\alpha) - m_n(\alpha)),$$
$$q_n(\alpha, z) = \sqrt{n}\Big(\tilde{b}_n(\alpha + \frac{1}{\sqrt{n}}z) - \tilde{b}(\alpha)\Big),\ D_n^2(\alpha) = E\rho_n(\alpha)\rho_n(\alpha)^* y$$

(We denote the conjugate vector by the symbol *).

Theorem 2.2.5 *DA (Diffusion approximation) Let conditions (2.2.4)-(2.20) be satisfied where in (2.18)* $\sqrt{n}\alpha_n(N) \to 0$ *, there exist continuous vector-valued function* $q(\alpha, z)$ *and matrix-valued function* $D^2(\alpha)$ *such that in any domain* $|\alpha| < N \quad |q(\alpha, z)| < C_N(1 + |z|)$*, and uniformly in* $|\alpha| < N$ *at each fixed* z

$$\sqrt{n}\left(\tilde{b}_n(\alpha + n^{-1/2}z) - \tilde{b}(\alpha)\right) \to q(\alpha, z), \tag{2.24}$$

$$D_n^2(\alpha) \to D^2(\alpha), \tag{2.25}$$

$\gamma_n(0) \overset{W}{\Rightarrow} \gamma_0,$ *and for any* $N > 0$

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sup_{|\alpha|<Nn} \Big\{E\tau_{n1}^2(\alpha)\chi(\tau_{n1}(\alpha) > L)$$
$$+E|\xi_{n1}(\alpha)|^2\chi(|\xi_{n1}(a)| > L)\Big\} = 0. \tag{2.26}$$

Then the sequence of the processes $\gamma_n(t)$ *J-converges on any interval [0,T] such that* $y(+\infty) > T$ *to the diffusion process* $\gamma(t)$ *which satisfies the following stochastic differential equation solution of which exists and is unique:* $\gamma(0) = \gamma_0$,

$$d\gamma(t) = q(s(t), \gamma(t))dt + D(s(t))m(s(t))^{-1/2}dw(t), \tag{2.27}$$

where $s(\cdot)$ *satisfies equation (2.21) (J-convergence denotes a weak convergence of measures in Skorokhod space* D_T*.)*

The detailed proofs of the Theorems 2.2.4 and 2.2.5 can be found in Anisimov [7].

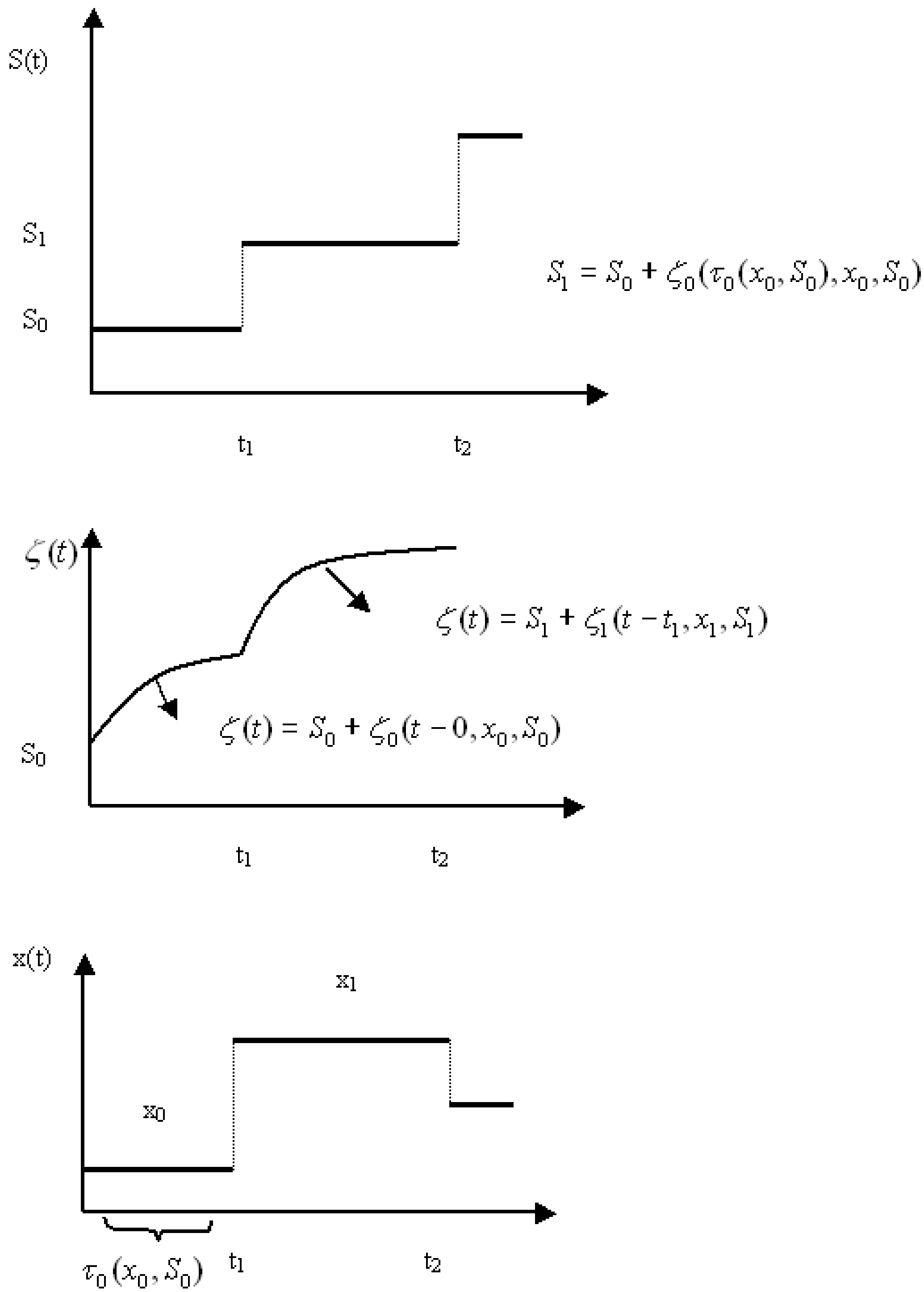

Figure 2.1: Switching Processes: An illustration

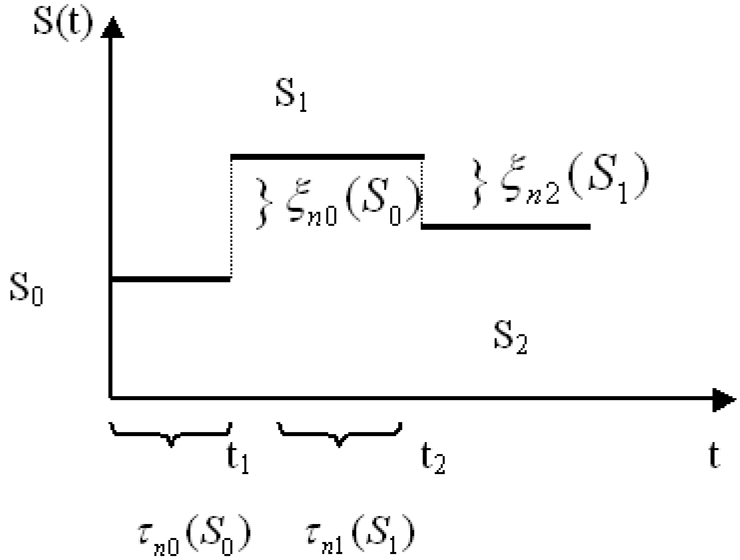

Figure 2.2: RPSM: An illustration

Chapter 3

Estimation by Trajectory Observations

3.1 Asymptotic Properties of Estimators Constructed by Trajectory Observations

In this chapter using the results of section 2.2.1 and 2.2.2, the analysis of stochastic equations and asymptotic properties of extreme sets of random functions, we consider a technique to solve the problems of statistical parameter estimation by observations of the trajectory of stochastic systems.

Our general construction explained below follows from Anisimov, [8].

Let $\{\gamma_k(\alpha), \alpha \in \mathcal{R}, k \geq 0\}$ be parametric families of random variables with values in $\mathcal{R}^r$. Let also $\{x_{nk}, k \geq 1\}$ be a trajectory of a (random or non-random) system with values in some space $S \subset \mathcal{R}^r$. Assume that, $\{\gamma_k(\alpha), \alpha \in \mathcal{R}, k \geq 0\}$ are jointly independent and independent of $\{x_{nk}, k \geq 1\}$.

Suppose that we have a complete scheme of observations. That is we observe variables x_{nk} and $y_k = \gamma_k(x_{nk}), k = 1, 2, ..., n,$ where n is the number of observations.

For simplicity we assume that distributions of random variables $\gamma_k(\alpha)$ do not depend on index k.

Let us consider the illustration how this general technique can be applied to statistical parameter estimation for several estimating methods: the method of moments, maximum likelihood method and least squares method in the non-classical situation when the observations are constructed on the trajectory of some random sequence.

3.2 Moments Method - Transient Case

We consider a one-dimensional case ($r = 1$) to illustrate the method. Suppose that first moment's of random variables $\{\gamma_k(\alpha), \alpha \in \mathcal{R}\}$ exist and belong to the parametric family of functions $\{g(\theta, \alpha),\ \theta \in \Theta \subset \mathcal{R},\ \alpha \in \mathcal{R}\}$. Also let $E\gamma_1(\alpha) = g(\theta_0, \alpha) = g(\alpha)$ where θ_0 is an inner point of the region Θ.

Then we can represent the moments method estimator as a solution of the equation

$$\frac{1}{n}\sum_{k=1}^{n} g(\theta, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} y_k = 0. \tag{3.1}$$

In this case, since the estimator is represented as a solution of a stochastic equation, we will use the results of Theorem 2.2.1.

Denote the set of possible solutions of the equation (3.1) by $\{\theta_n\}$. We study the asymptotic behavior $\{\theta_n\}$ as $n \to \infty$.

Let us give a necessary definition for an averaging condition which will be useful in the further studies.

Definition 3.2.1 *If there exists a continuous function $x(u)$ such that for any continuous bounded function $f(x),\ x \in X$*

$$\frac{1}{n}\sum_{k=0}^{n} f(x_{nk}) \xrightarrow{\mathrm{P}} \int_0^1 f(x(u))du, \tag{3.2}$$

is satisfied then we say that an averaging condition ***A*** *is satisfied.*

Note that the condition (3.2) is mostly oriented on non-stationary (transient) conditions. An average principle for rather general stochastic recurrent sequences in transient conditions is given by Anisimov, [4] and [7].

The following theorem is similar to Theorem 6.1 of Anisimov and Pflug [16] with the modification of condition 6.2. (more strong condition (6.2) is changed to a weaker condition of averaging type).

Theorem 3.2.1 *Suppose that the sequence* x_{nk} *satisfies following averaging condition* **A** *and variables* $\gamma_k(\alpha)$ *satisfy the following condition: for any* $L > 0$

$$\lim_{N\to\infty} \sup_{|\alpha|\le L} E\left(|\gamma_1(\alpha)|\chi\{|\gamma_1(\alpha)| > N\}\right) = 0, \tag{3.3}$$

the function $g(\theta, \alpha)$ *is continuous in both arguments* (θ, α) *and there exists* $\delta > 0$ *such that the equation*

$$\int_0^1 g(\theta, x(u))du - \int_0^1 g(\theta_0, x(u))du = v \tag{3.4}$$

has a unique solution for any $|v| < \delta$.

Then with probability which tends to one a solution of the equation (3.1) exists and $\{\theta_n\} \xrightarrow{\text{P}} \theta_0$.

Proof. We prove the convergence of second term in the left part of equation(3.1) under the conditions (3.2), (3.3).

Since $g(\theta, \alpha)$ is continuous and from the conditions (3.2) and (3.3), we have

$$\frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk}) \xrightarrow{\text{P}} \int_0^1 g(\theta_0, x(u))du. \tag{3.5}$$

We now can see that

$$\frac{1}{n}\sum_{k=1}^{n} y_k \xrightarrow{\text{P}} \int_0^1 g(\theta_0, x(u))du. \tag{3.6}$$

The first term of the left hand side of (3.1) for any $L > 0$ uniformly in $|\theta| \leq L$ converges to the function

$$\int_0^1 g(\theta, x(u))du.$$

And finally, since the equation

$$\int_0^1 g(\theta, x(u))du - \int_0^1 g(\theta_0, x(u))du = 0$$

has the unique solution (from (3.4)), it follows from the result of the Theorem (2.2.1) that $\theta_n \xrightarrow{\text{P}} \theta_0$, and this proves the Theorem (3.2.1).

Consider now the behavior of the normalized deviations $\sqrt{n}(\theta_n - \theta_0)$. The following theorem is similar to Theorem 3.3 of Anisimov [8] where he considers an estimator which depends on time also on the observation interval $[t_0, T]$.

Theorem 3.2.2 *Suppose that conditions of Theorem 3.2.1 hold and there exists a continuous in both arguments derivative*

$$R(\theta, \alpha) = \frac{\partial}{\partial \theta} g(\theta, \alpha)$$

and a continuous variance

$$\sigma^2(\alpha) = E(\gamma_1(\alpha) - g(\alpha))^2.$$

Denote

$$\widehat{R}(\theta_0) = \int_0^1 R(\theta_0, x(v))dv, \ \widehat{\sigma}^2 = \int_0^1 \sigma^2(x(v))dv. \tag{3.7}$$

Suppose that $\widehat{R}(\theta_0) > 0$ *and variables* $\gamma_k(\alpha)$ *satisfy Lindeberg condition: for any* $L > 0$

$$\lim_{N \to \infty} \sup_{|\alpha| \leq L} E\gamma_1(\alpha)^2 \chi\{|\gamma_1(\alpha)| > N\} = 0. \tag{3.8}$$

Then there exists a solution $\widehat{\theta}_n$ *of the equation (3.1) such that the sequence* $\sqrt{n}(\widehat{\theta}_n - \theta_0)$ *weakly converges to a normal random variable with mean 0 and variance* $\widehat{R}^{-2}\widehat{\sigma}^2$.

Proof. We will use the second part of Theorem 2.2.1..

Let us denote

$$f_n(\theta) = \frac{1}{n}\sum_{k=1}^{n} g(\theta, \alpha) - \frac{1}{n}\sum_{k=1}^{n} y_k. \tag{3.9}$$

We then have

$$\nu_n^{\beta} f_n(\theta_0 + \frac{\nu}{\nu_n}) = \nu_n^{\beta}(\frac{1}{n}\sum_{k=1}^{n} g(\theta_0 + \frac{\nu}{\nu_n}, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} y_k). \tag{3.10}$$

Let us put $\nu_n = \sqrt{n},\ \beta = 1$. By adding and subtracting some terms we can write the right hand side of equation (3.10) as follows,

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} g(\theta_0 + \frac{\nu}{\sqrt{n}}, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} y_k \right.$$

$$\left. + \frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk})\right). \tag{3.11}$$

Rearranging the terms of (3.11) we have the right hand side of equation (3.10) equal to

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} g(\theta_0 + \frac{\nu}{\sqrt{n}}, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk})\right) - \sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} y_k - \frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk}).\right) \tag{3.12}$$

Consider the first part of (3.12). Using Taylor's formula we have;

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} g(\theta_0 + \frac{\nu}{\sqrt{n}}, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk})\right)$$

$$= \sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk}) + \frac{1}{n}\sum_{k=1}^{n} \frac{\partial g(\theta_0, x_{nk})}{\partial \theta}\frac{\nu}{\sqrt{n}} - \frac{1}{n}\sum_{k=1}^{n} g(\theta_0, x_{nk})\right) + o(.),$$

which is equal to

$$\frac{1}{n}\sum_{k=1}^{n} R(\theta_0, x_{nk})\nu + o(.) \tag{3.13}$$

Notice that, according to condition **A**, (3.13) uniformly converges in any bounded region $|\nu| \leq L$ to the value,

$$\int_0^1 R(\theta_0, x(u))\nu du = \hat{R}(\theta_0)\nu. \tag{3.14}$$

The second part of equation (3.12), due to the Lindeberg condition, weakly converges to a normal random variable $N(0, \hat{\sigma}^2)$, where

$$\hat{\sigma}^2 = \lim_{n\to\infty} \left(\frac{1}{n}\sum_{k=1}^{n} E(\gamma_k(x_{nk}) - g(\theta, x_{nk}))^2\right)$$

and from the conditions of theorem,

$$\hat{\sigma}^2 = \int_0^1 \sigma^2(x(u))du.$$

Then the limiting equation can be written as

$$\hat{R}(\theta_0)\nu + N(0, \hat{\sigma}^2) = 0$$

and

$$\nu = \frac{1}{\hat{R}(\theta_0)} N(0, \hat{\sigma}^2).$$

From Theorem 2.2.1 it follows that

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \overset{w}{\Rightarrow} \mathcal{N}\left(0, \frac{\hat{\sigma}^2}{\hat{R}(\theta_0)^2}\right).$$

This means that, $\hat{\theta}$ is the asymptotically normal estimator of θ_0 with coefficient $\frac{\hat{\sigma}}{\hat{R}(\theta_0)}$

3.3 Maximum Likelihood Method

Consider now the behavior of maximum likelihood method estimators. At the investigation we use the results of Theorem 2.2.2 about the behavior of extreme points.

Suppose that we have the same scheme of observations x_{nk} and y_{nk},for $k = 1, 2, ..., n$ as was described in the introduction of model. For simplicity we assume that distributions of random variables $\gamma_k(\alpha)$ do not depend on index k.

Let densities of random variables $\{\gamma_k(\alpha), \alpha \in R^r\}$ exist and belong to the parametric family of densities $\{p(z, \theta, \alpha), z \in R^d, \theta \in \Theta, \alpha \in R^r\}$ where Θ is some bounded closed region in R^d. Suppose that $p(z, \theta_0, \alpha)$ is the density of the variable $\gamma_k(\alpha)$ and θ_0 is the inner point of the region Θ. Note that, same scheme of observations and assumptions are given by Anisimov [8], and results are provided for RPSM.

We can write logarithmic maximum likelihood function $L_n(\theta)$ in the form:

$$L_n(\theta) = \frac{1}{n} \sum_{k=1}^{n} \ln p(y_{nk}, \theta, x_{nk}). \tag{3.15}$$

Let us denote $\{\theta_n\}$ as the set of points of maximum in the argument θ for $L_n(\theta)$ and let $f(\theta, \alpha) = E \ln p(\gamma_1(\alpha), \theta, \alpha)$.

The following Theorem about the behavior of the estimator is similar to Theorem 3.1 of Anisimov [8]. We have the relaxation that the estimator itself does not depend on time.

Theorem 3.3.1 *Suppose that the averaging condition* $\boldsymbol{A}$ *(see section 3.2) holds and the following conditions are true:*

1. $\sup_{\theta,\alpha} E|\ln p(\gamma_1(\alpha), \theta, \alpha)|^2 < \infty$;

2. *for any* $L > 0$

$$\lim_{c\to+0}\sup_{\alpha} E \sup_{|\theta_1-\theta_2|<c} |\ln p(\gamma_1(\alpha),\theta_1,\alpha)-$$
$$\ln p(\gamma_1(\alpha),\theta_2,\alpha)| = 0;$$

3. *the point* θ_0 *is the unique point of maximum for the function*

$$L_0(\theta) = \int_0^1 f(\theta, x(u))du. \tag{3.16}$$

Then

$$\{\theta_n\} \xrightarrow{\mathrm{P}} \theta_0.$$

Proof.

Let $f(\theta,\alpha) = E\ln((\gamma_1(\alpha),\theta,\alpha) = E\ln(z,\theta,\alpha)$. Consider the difference;

$$f(\theta,\alpha) - f(\theta_0,\alpha) = E\ln p(z,\theta,\alpha) - E\ln p(z,\theta_0,\alpha) = E(\ln\frac{p(z,\theta,\alpha)}{p(z,\theta_0,\alpha)})$$

since $\ln x \le x - 1$,

$$E\left(\ln\frac{p(z,\theta,\alpha)}{p(z,\theta_0,\alpha)}\right) \le \int\left(\frac{p(z,\theta,\alpha)}{p(z,\theta_0,\alpha)} - 1\right)p(z,\theta,\alpha)dz = 0. \tag{3.17}$$

The equation (3.17) indicates that $f(\theta,\alpha) - f(\theta_0,\alpha) \le 0$. This shows that, θ_0 is the point of maximum for $f(\theta,\alpha)$ and correspondingly point of maximum for $L_0(\theta)$.

From the condition **A**, it follows that at each fixed θ the sequence of functions $L_n(\theta)$ converges in probability to the function

$$L_0(\theta) = \int_0^1 f(\theta, x(u))du.$$

In order to prove the uniform convergence we need to check the modulus of continuity (see Definition(2.2.2))

$$P\{\Delta_u(c, F_n(\cdot)) > \varepsilon\} = P(\sup_{|\theta_1-\theta_2|<c} |L_n(\theta_1) - L_n(\theta_2)|) > \varepsilon$$

$$= P(\sup_{|\theta_1-\theta_2|<c} |\frac{1}{n}\sum_{k=1}^{n} \ln p(y_{nk}, \theta_1, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} \ln p(y_{nk}, \theta_2, x_{nk})|) > \varepsilon.$$

$$P(\sup_{|\theta_1-\theta_2|<c} \frac{1}{n}\sum_{k=1}^{n} |\ln p(y_{nk}, \theta_1, x_{nk}) - \ln p(y_{nk}, \theta_2, x_{nk})|) > \varepsilon. \tag{3.18}$$

Since we now have nonnegative jointly independent random variables, we can use the Chebychev Inequality in the form,

$$P\left\{\frac{1}{n}\sum_{k=1}^{n} x_k > \varepsilon\right\} \leq \frac{1}{\varepsilon}E(x_1).$$

to estimate the right hand side of (3.18).

Then an upper bound of the probability of (3.18) can be estimated by,

$$\frac{1}{\varepsilon}E \sup_{|\theta_1-\theta_2|<c} |\ln p(\gamma_1(\alpha), \theta_1, \alpha) - \ln p(\gamma_1(\alpha), \theta_2, \alpha)|.$$

According to condition 2 of the Theorem, we have

$$\lim_{c\to+0} \sup_{|\alpha|<L} \frac{1}{\varepsilon}E \sup_{|\theta_1-\theta_2|<c} |\ln p(\gamma_1(\alpha), \theta_1, \alpha) - \ln p(\gamma_1(\alpha), \theta_2, \alpha)| = 0.$$

Which means that

$$\lim_{c\to+0} \limsup_{n\to\infty} P\{\Delta_u(c, F_n(\cdot)) > \varepsilon\} = 0$$

hence, the modulus of continuity is equal to zero and $L_n(\theta)$ uniformly converges to $L_0(\theta)$. According to Theorem 2.2.2 this implies $\theta_n \xrightarrow{\text{P}} \theta_0$.

Based on Theorem 2.2.3 the convergence of deviations also can be studied. The following Theorem about the behavior of deviations is similar to Theorem 2.1 of Anisimov [9]. He considers the behavior of the process in time interval $[0, T]$, have an additional convergence assumption and their estimator depends on time.

Let a vector of first derivatives $\nabla_\theta \varphi(y,\theta,\alpha)$ and matrix of second derivatives

$$G(y,\theta,\alpha) = ||\frac{\partial^2}{\partial\theta_i \partial\theta_j}\varphi(y,\theta,\alpha)||, ij.$$

exist.

Theorem 3.3.2 *Assume that conditions of Theorem 3.3.1 hold and for any $L > 0$:*

1.
$$\lim_{N\to\infty} \sup_{|\alpha|\le L} E\left(|\nabla_\theta\varphi(\gamma_1(\alpha),\theta_0,\alpha)|^2 \chi|\nabla_\theta\varphi(\gamma_1(\alpha),\theta_0,\alpha)| > N\right) = 0;$$

2.
$$\lim_{c\to\ +0} \sup_{|\alpha|\le L} E\left(\sup_{|\theta-\theta_0|<c} |G(\gamma_1(\alpha),\theta,\alpha) - G(\gamma_1(\alpha),\theta_0,\alpha)|\right) = 0; \qquad (3.19)$$

3.
$$\lim_{N\to\infty} \sup_{|\alpha|\le L} E\left(|G(\gamma_1(\alpha),\theta_0,\alpha)|\chi|G(\gamma_1(\alpha),\theta_0,\alpha)| > N\right) = 0;$$

4. *functions*
$$B(\theta_0,\alpha)^2 = E\left(\nabla_\theta\varphi(\gamma_1(\alpha),\theta_0,\alpha)\nabla_\theta\varphi(\gamma_1(\alpha),\theta_0,\alpha)^*\right)$$
and $C(\theta_0,\alpha) = EG(\gamma_1(\alpha),\theta_0,\alpha)$ satisfy local Lipschitz condition in argument α.

Then there exist a sequence of random variables $\tilde{\theta}_n$ such that $\tilde{\theta}_n$ is a point of a local maximum of the function $L_n(\theta)$ and the sequence $\kappa_n = \sqrt{n}(\tilde{\theta}_n - \theta_0)$ weakly converges to κ_0, where

$$\kappa_0 = (\int_0^1 C(\theta_0, s(u))du)^{-1} \int_0^1 B(\theta_0, s(u))dw(u), \qquad (3.20)$$

and $w(t)$ is a standard Wiener process in R^r.

Proof. We use Theorem 2.2.3 about the behavior of extreme points. Consider the function

$$A_n(\nu) = \nu_n^{\beta}(L_n(\theta_0 + \frac{\nu}{\nu_n}) - L_n(\theta_0)).$$

Let $\nu = \sqrt{n}$ and $\beta = 2$. Then we have,

$$A_n(\nu) = n(\frac{1}{n}\sum_{k=1}^{n} \ln p(y_{nk}, \theta_0 + \frac{\nu}{\nu_n}, x_{nk}) - \frac{1}{n}\sum_{k=1}^{n} \ln p(y_{nk}, \theta_0, x_{nk})).$$

Using the Taylor expansion up to the second order we have $A_n(\nu)$ equal to

$$n(\frac{1}{n}\sum_{k=1}^{n} \ln p(y_{nk}, \theta_0, x_{nk}) + \frac{1}{n}\sum_{k=1}^{n} \nabla_\theta \ln p(y_{nk}, \theta_0, x_{nk})\frac{\nu}{\sqrt{n}})$$

$$+n(\frac{1}{2n}\sum_{k=1}^{n} G((y_{nk}, \theta_0, x_{nk})\nu, \nu)\frac{1}{n} + R_{\theta_0}(\cdot)) - \sum_{k=1}^{n} \ln p(y_{nk}, \theta_0, x_{nk})$$

Then

$$A_n(\nu) = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} \nabla_\theta \ln p(y_{nk}, \theta_0, x_{nk})\nu + \frac{1}{2n}\sum_{k=1}^{n} G((y_{nk}, \theta_0, x_{nk})\nu, \nu) + R_{\theta_0}(\cdot), \tag{3.21}$$

where $R_{\theta_0}(\cdot)$ is the remainder part of Taylor's formula up to second order and converges uniformly to 0 as $n \to \infty$.

Let us consider the expectation of the first part of equation (3.21) at point $\theta = \theta_0$ for $z = y_{nk}, \alpha = x_{nk}$.

$$E_{\theta=\theta_0}(\nabla_\theta \ln p(z, \theta, \alpha)) = E_{\theta=\theta_0}\frac{\nabla_\theta(p, z, \alpha)}{p(z, \theta, \alpha)}$$
$$= \int \nabla_\theta p(z, \theta_0, \alpha)\frac{1}{p(z, \theta_0, \alpha)}p(z, \theta_0, \alpha)dz$$
$$= \nabla_\theta \int p(z, \theta_0, \alpha)dz = 0.$$

Furthermore also at point $\theta = \theta_0$ we have,

$$E(\nabla_\theta \ln p(z, \theta_0, \alpha)\nabla_\theta \ln(p(z, \theta_0, \alpha))^*) = B(\theta_0, \alpha)^2 du.$$

Note that, the sum $\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\nabla_\theta \ln p(y_{nk},\theta_0,x_{nk})\nu$ forms a process with independent increments where increments have, in the limit, Normal distribution with expectation 0 and covariance matrix $B(\theta_0,x_{nk})^2$.

In reference to Anisimov [9] we write that,

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\nabla_\theta \ln p(y_{nk},\theta_0,x_{nk})\nu \to \int_0^1 B(\theta_0,x(u))\nu dw(u) \tag{3.22}$$

uniformly in ν.

The second term in the right hand side of equation (3.21) has the expectation $\frac{1}{2n}\sum_{k=1}^{n}C(\theta_0,x_{nk})$ and the variance tends to 0 as $n\to\infty$. Additionally, from conditions 3 and 4, we have

$$\frac{1}{2n}\sum_{k=1}^{n}C(\theta_0,x_{nk}) \xrightarrow{\mathrm{P}} \frac{1}{2}\int_0^1 C(\theta_0,x(u))du.$$

That is, the second term in the right-hand side of (3.21) converges in probability to the deterministic value $(\int_0^1 C(\theta_0,x(u))du\nu,\nu)$.

Then the limiting function $A_0(\nu)$ can be written as

$$A_0(\nu) = \int_0^1 B(\theta_0,x(u))dw(u)\nu + \frac{1}{2}\int_0^1 C((\theta_0,x(u))\nu du,\nu).$$

Matrix $C(\theta_0,\alpha)$ is negatively defined and self-conjugated. We can now find the solution κ_0 of the equation $A_0(\nu)$ as

$$\kappa_0 = (\int_0^1 C(\theta_0,x(u))du)^{-1}\int_0^1 B(\theta_0,x(u))dw(u).$$

Finally, following from Theorem (2.2.3) we have

$$\sqrt{n}(\hat{\theta_n}-\theta_0) \overset{\mathrm{w}}{\Rightarrow} \kappa_0.$$

3.4 Analysis of Least Squares Method Equation

This section exposes the results of Anisimov and Kaibah [15] for the analysis of least squares method in non-homogenous case.

For the same scheme of observations $\{x_{nk}, k \geq 0\}$ with values in the space X, which was given in the original construction, suppose that the parametric family of functions $g(\theta, x),\ \theta \in \Theta \subset \mathcal{R}^r, x \in X$ with values in $\mathcal{R}^r$ are given. Also let the jointly independent family of random vectors $\{\xi_k(x), k \geq 0\}$ with values in $\mathcal{R}^r$ with the same distributions be given.

For $k = 1, 2, ..., n$ we observe the following:

$$y_{nk} = g(\theta_0, x_{nk}) + \xi_k(x_{nk}), k = 0, 1, ..., n. \tag{3.23}$$

If the partial derivatives of $g(\theta, x)$ exist, so that,

$$\nabla_\theta g(\theta, x) \ = \Big|\Big|\frac{\partial}{\partial \theta_j} g_i(\theta, x)\Big|\Big|,\ i = \overline{1, m}, j = \overline{1, r},$$

then the least squares method estimator is a solution of the equation,

$$\frac{1}{n}\sum_{k=0}^{n} \nabla_\theta g(\theta, x_{nk})^* (y_{nk} - g(\theta, x_{nk})) = 0. \tag{3.24}$$

Let us denote

$$f_n(\theta) = \frac{1}{n}\sum_{k=0}^{n} \nabla_\theta g(\theta, x_{nk})^* (y_{nk} - g(\theta, x_{nk}))$$

and denote the set of all solutions of equation (3.24) by$\{\theta_n\}$. Let also $f_0(\theta)$ be the limiting function of $f_n(\theta)$.

Suppose that the sequence x_{nk} satisfies the averaging condition **A** of section 3.1 and the function

$$\int_0^1 \nabla_\theta g(\theta, x(u))^* \Big(g(\theta, x(u)) - g(\theta_0, x(u))\Big) du.$$

satisfies the condition of **S**.

The following two theorems follows from the theorems of Anisimov and Kaibah [15] and we give the extended proof of the theorems.

Theorem 3.4.1 *Let the function* $\nabla_\theta g(\theta, x)$ *be uniformly continuous in* $\Theta \times X$, *the function* $f_0(\theta)$ *satisfies the condition of separateness* $\boldsymbol{S}$, *and for any* $x \in X$

$$E\xi_1(x) \equiv 0, \ \ E\xi_1(x)\xi_1(x)^* = R(x)^2, \tag{3.25}$$

the condition $\boldsymbol{A}$ *holds and*

$$\sup_{x \in X} ||R(x)^2|| \leq C < \infty. \tag{3.26}$$

Then

$$\{\theta_n\} \xrightarrow{\mathrm{P}} \theta_0. \tag{3.27}$$

Proof. Under the conditions of theorem (3.4.1), it can be seen that $f_n(\theta)$ uniformly converges to $f_0(\theta)$ where,

$$f_0(\theta) = \int_0^1 \nabla_\theta g(\theta, x(u))^* \Big(g(\theta, x(u)) - g(\theta_0, x(u))\Big) du. \tag{3.28}$$

It follows from Theorem(2.2.1) that $\{\theta_n\} \xrightarrow{\mathrm{P}} \theta_0$.

Now, consider the behavior of deviations.

Theorem 3.4.2 *Let the conditions of Theorem (3.4.1) hold, and Lindeberg condition be satisfied in the following form:*

$$\lim_{L \to \infty} \sup_{x \in X} E||\xi_1(x)||^2 \chi(||\xi_1(x)|| > L) = 0. \tag{3.29}$$

Then there exists the sequence $\tilde{\theta}_n$ of the points of solutions of equation $f_n(\theta)$ such that

$$\sqrt{n}(\tilde{\theta}_n - \theta_0) \overset{w}{\Rightarrow} Q^{-2}B\mathcal{N}(0,1). \tag{3.30}$$

where

$$Q^2 = \int_0^1 g'_\theta(\theta_0, x(u))^* g'_\theta(\theta_0, x(u))du,$$

$$B^2 = \int_0^1 g'_\theta(\theta_0, x(u))^* R(x(u))^2 g'_\theta(\theta_0, x(u))du \tag{3.31}$$

(here for simplicity we denote $g'_\theta(\theta, x) = \nabla_\theta g(\theta, x)$).

Proof. Using the second part of Theorem 2.2.1 let us consider a random function

$$f_n(v) = v_n^\beta f_n(\theta_0 + \frac{v}{v_n})$$

$$= v_n^\beta(\frac{1}{n}\sum_{k=1}^{n} \nabla_\theta g(\theta_0 + \frac{v}{v_n}, x_{nk})^*(y_{nk} - g(\theta_0 + \frac{v}{v_n}, x_{nk})))$$

so that

$$f_n(v)$$

$$= v_n^\beta(\frac{1}{n}\sum_{k=1}^{n} \nabla_\theta g(\theta_0 + \frac{v}{v_n}, x_{nk})^*(g(\theta_0, x_{nk}) + \xi_k(x_{nk}) - g(\theta_0 + \frac{v}{v_n}, x_{nk}))). \tag{3.32}$$

Furthermore,

$$f_n(v) = v_n^\beta(\frac{1}{n}\sum_{k=1}^{n} \nabla_\theta g(\theta_0 + \frac{v}{v_n}, x_{nk})^*(g(\theta_0, x_{nk}) - g(\theta_0 + \frac{v}{v_n}, x_{nk})))$$

$$v_n^\beta(\frac{1}{n}\sum_{k=1}^{n} \nabla_\theta g(\theta_0 + \frac{v}{v_n}, x_{nk})^*\xi_k(x_{nk})). \tag{3.33}$$

Note that using Taylor expansion we can write,

$$g(\theta_0, x_{nk}) - g(\theta_0 + \frac{v}{v_n}, x_{nk})$$

$$= g(\theta_0, x_{nk}) - g(\theta_0, x_{nk}) - \nabla_\theta g(\theta_0, x_{nk}) \, \frac{v}{v_n} - o(\cdot)$$
$$= -\frac{v}{v_n} \nabla_\theta g(\theta_0, x_{nk}) - o(\cdot).$$

Let $v_n = \sqrt{n}$, $\beta = 1$. Using the uniform continuity of the gradient the first term of (3.33) can be written as

$$-\frac{1}{n} \sum_{k=1}^{n} \nabla_\theta g(\theta_0, x_{nk})^* \nabla_\theta g(\theta_0, x_{nk}) v - o(\cdot),$$

which by the conditions of the theorem converges to

$$-\int_0^1 \nabla_\theta g(\theta_0, x(u))^* \nabla_\theta g(\theta_0, x(u)) du = Q^2.$$

According to Lindeberg condition, (3.29),the second part of (3.33) has expectation 0 and the covariance matrix

$$\frac{1}{n} \sum_{k=1}^{n} E((\nabla_\theta g(\theta_0, x_{nk})^* \xi_k(x_{nk}))^2) = \frac{1}{n} \sum_{k=1}^{n} \nabla_\theta g(\theta_0, x_{nk})^* R(x_{nk})^2 \nabla_\theta g(\theta_0, x_{nk})$$

which converges to

$$\int_0^1 \nabla_\theta g(\theta_0, x(u))^* R(x(u))^2 \nabla_\theta g(\theta_0, x(u)) du = B^2.$$

Then the second part of (3.33) converges to Normal random variable with expectation 0 and covariance matrix B^2.

Following from those facts, the sequence of random functions $f_n(\theta)$ uniformly converges to $-Q^2 v + N(0, B^2)$. Therefore, according to the second part of Theorem (2.2.1) $\sqrt{n}(\hat{\theta}_n - \theta_0) \overset{\mathrm{w}}{\Rightarrow} \kappa_0$, where κ_0 is the solution of equation $-Q^2 v + N(0, B^2) = 0$ and in this case it is equal to $\kappa_0 = Q^{-2} N(0, B^2)$. This completes the proof.

As a special case we can consider the behavior of the least squares method estimator constructed by observations in a random external environment. For this case we can construct the estimator as an extreme point of a random function .

3.5 Parameter Estimation in Switching Models

3.5.1 Preliminary Work

In this section we consider the asymptotic behavior of moments method estimators constructed by observations on the trajectory of a switching stochastic system. We will prove the consistency and asymptotic normality of the moments method estimators.

We first consider the results of Anisimov and Guleryuz [13].

For each $n = 1, 2...$, $\theta \in \Theta$, let $\{\zeta_n(t, \theta), t \geq 0\}$ be a sequence of random processes in $\mathcal{D}^r_\infty$, where Θ is a bounded closed region in R^d, and $\{\gamma_{nk}(\theta, \alpha), \theta \in \Theta, \alpha \in \mathcal{R}^r\}$, $k \geq 1$, be an independent of $\zeta_n(\cdot)$ sequence of random variables with values in $\mathcal{R}^d$ and distributions not depending on k. Suppose that on the interval $[0, T]$ we observe variables $y_{nk} = \gamma_{nk}(\theta_0, \zeta_n(kah_n, \theta_0))$, $k = 1, 2, .., [T/(ah_n)]$, where $a > 0, h_n \to 0$ as $n \to \infty$.

Suppose that $\zeta_n(\cdot)$ satisfies the following property: there exists a deterministic function $s(t, \theta)$ in $\mathcal{D}^r_T$ such that at any $\theta \in \Theta$ and at some $T > 0$

$$\sup_{0 \leq t \leq T} |\zeta_n(t, \theta) - s(t, \theta)| \xrightarrow{\text{P}} 0. \tag{3.34}$$

Suppose also that there exist functions $g_n(\theta, \alpha) = \mathbf{E}\gamma_{nk}(\theta, \alpha)$, $\theta \in \Theta, \alpha \in R^r$. Then an analog of moments method equation can be written in the form:

$$h_n \sum_{k=1}^{[T/(ah_n)]} y_{nk} - \frac{1}{a} \int_0^T g_n(\theta, s(t, \theta))\mathrm{d}t = 0.$$

Let us denote

$$f_n(\theta) = h_n \sum_{k=1}^{[T/(ah_n)]} y_{nk} - \frac{1}{a} \int_0^T g_n(\theta, s(t, \theta))dt$$

Consider the equation

$$f_n(\theta) = 0. \tag{3.35}$$

Theorem 3.5.1 *Suppose that the condition (3.34) holds,*

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sup_{|\alpha|\le N} \mathbf{E}\left(|\gamma_{n1}(\theta_0,\alpha)|\chi(|\gamma_{n1}(\theta_0,\alpha)| > L)\right) = 0, \tag{3.36}$$

and there exists a continuous function $g_0(\theta,\alpha)$ *such that* $g_n(\theta,\alpha) \to g_0(\theta,\alpha)$ *as* $n \to \infty$ *uniformly in* θ, α *in each bounded region.*

Suppose also that the function

$$f_0(\theta) = \int_0^T \left(g_0(s(t,\theta_0),\theta_0) - g_0(s(t,\theta),\theta)\right) dt \tag{3.37}$$

satisfies condition $\boldsymbol{S}$*. Then* $\{\boldsymbol{\theta_n}\} \xrightarrow{\mathrm{P}} \theta_0$*, where* $\{\boldsymbol{\theta_n}\}$ *is the set of solutions of (3.35).*

Proof.

First we prove that

$$\Delta_{1n} = h_n \sum_{k=1}^{[T/h_n]} y_{nk} - h_n \sum_{k=1}^{[T/h_n]} g_n(\theta_0,\zeta_{nk}) \xrightarrow{\mathrm{P}} 0, \tag{3.38}$$

where $\zeta_{nk} = \zeta_n(kh_n,\theta_0)$. Condition (3.36) implies that $E\exp\{i\phi\gamma_{nk}(\theta_0,\alpha)h_n\} = 1 + i\phi g_n(\theta_0,\alpha)h_n + o(h_n)$, where $h_n^{-1}o(h_n) \to 0$ uniformly in $|\alpha| \le N$. Now, using formula of conditional expectations and a method of characteristic functions, we get

$$E\exp\{i\phi\Delta_{1n}\} = E\left(E\left[\exp\{i\phi\Delta_{1n}\}\ |\zeta_{nk}\right]\right)$$

$$= E\left(\prod_{k=1}^{[T/h_n]} (1 + i\phi g_n(\theta_0,\zeta_{nk})h_n + o(h_n)) \exp\left\{-i\phi h_n \sum_{k=1}^{[T/h_n]} g_n(\theta_0,\zeta_{nk})\right\}\right)$$

$$\approx E\exp\left\{\sum_{k=1}^{[T/h_n]} (i\phi g_n(\theta_0,\zeta_{nk})h_n + o(h_n)) - i\phi h_n \sum_{k=1}^{[T/h_n]} g_n(\theta_0,\zeta_{nk})\right\} \to 1,$$

which proves (3.38). According to condition (3.34), $\sup_{k\le T/h_n}|\zeta_{nk} - s(kh_n,\theta_0)| \xrightarrow{\mathrm{P}} 0$. Then, as $g_n(\theta_0,\alpha)\to g_0(\theta_0,\alpha)$ uniformly in $\alpha\le N$ and $g_0(\theta_0,\alpha)$ is continuous, we can easy prove that

$$h_n\sum_{k=1}^{[T/h_n]} g_n(\theta_0,\zeta_{nk}) - h_n\sum_{k=1}^{[T/h_n]} g_0(\theta_0,s(kh_n,\theta_0)) \xrightarrow{\mathrm{P}} 0.$$

Further,

$$h_n\sum_{k=1}^{[T/h_n]} g_0(\theta_0,s(kh_n,\theta_0)) \to \int_0^T g_0(\theta_0,s(t,\theta_0))dt,$$

and also $\int_0^T g_n(\theta,s(t,\theta))dt \to \int_0^T g_0(\theta,s(t,\theta))dt$ uniformly in θ as $n\to\infty$.

Therefore, $f_n(\theta)$ in (3.35) U-converges to the function $f_0(\theta)$ in (3.37), and our result follows from Theorem (2.2.1).

Consider now the behavior of deviations. Let

$$R_n(\theta,\alpha)^2 = \mathbf{E}\left((\gamma_{n1}(\theta_0,\alpha) - g_n(\alpha,\theta_0,\alpha))(\gamma_{n1}(\theta_0,\alpha) - g_n(\alpha,\theta_0,\alpha))^*\right)$$

.

Theorem 3.5.2 *Suppose that conditions of Theorem 3.5.1 are satisfied,*

$$h_n^{-1/2}\sup_{0\le t\le T}|\zeta_n(t,\theta_0) - s(t,\theta_0)| \xrightarrow{\mathrm{P}} 0, \tag{3.39}$$

there exist continuous functions $B(\theta,\alpha), R(\theta,\alpha)$ such that

$$\begin{aligned} g_n(\theta_0+z\sqrt{h_n}, s(t,\theta_0+z\sqrt{h_n})) &= g_n(\theta_0,s(t,\theta_0)) \\ &+\sqrt{h_n}B(\theta_0,s(t,\theta_0))z + \sqrt{h_n}o_n(1), \end{aligned} \tag{3.40}$$

where $o_n(1)\to 0$ uniformly in each bounded region $|z|\le L, |\alpha|\le N$, for any $N>0$ $R_n(\theta_0,\alpha)\to R(\theta_0,\alpha)$ uniformly in $|\alpha|\le N$, and

$$\lim_{L\to\infty}\limsup_{n\to\infty}\sup_{|\alpha|\le N}\mathbf{E}\left(|\gamma_{n1}(\theta_0,\alpha)|^2\chi(|\gamma_{n1}(\theta_0,\alpha)|>L\right) = 0.$$

Then there exists a solution $\widetilde{\theta}_n$ of equation (3.35) such that $h_n^{-1/2}(\widetilde{\theta}_n-\theta_0)$ weakly converges to the gaussian vector with mean 0 and covariance matrix $aB^{-1}R^2(B^*)^{-1}$, where

$$R^2 = \int_0^T R(\theta_0,s(t,\theta_0))^2\mathrm{d}t,\ \ B = \int_0^T B(\theta_0,s(t,\theta_0))\mathrm{d}t.$$

Proof. We will apply the results of Theorem 2.2.1. Consider the function,

$$v_n^\beta f_n(\theta_0 + \frac{z}{v_n}).$$

For $v_n = h_n^{-1/2}$ and $\beta = 1$, we have

$$h_n^{-1/2} f_n(\theta_0 + \frac{z}{h_n^{-1/2}})$$

$$= \frac{1}{\sqrt{h_n}} \left(h_n \sum_{k=1}^{[T/ah_n]} \gamma_{nk}(\theta_0, \zeta_n(kah_n, \theta_0)) - \frac{1}{a} \int_0^T g_n \left((\theta_0 + z\sqrt{h_n}), s(t, \theta_0 + z\sqrt{h_n}) \right) dt \right). \tag{3.41}$$

Note that for any bounded function $f(x)$ the integral $\int_0^T f(x)dx$ can be approximated by an integral sum up to the order of h_n so that

$$\int_0^T f(x)dx = \sum_{k=1}^{T/h_n} f(kh_n)h_n + o(\cdot).$$

Using this fact, the integral

$$\int_0^T g_n(\theta_0, s(t, \theta_0))dt$$

can be represented in the summation form as

$$ah_n \sum_{k=1}^{[T/ah_n]} g_n(\theta_0, \zeta_n(kah_n, \theta_0)) + o(1).$$

From this representation and the condition (3.40) we can write the right hand side of (3.41) as follows:

$$\left[\sqrt{h_n} \sum_{k=1}^{[T/ah_n]} \gamma_{nk}(\theta_0, \zeta_n(kah_n, \theta_0)) - \frac{1}{a\sqrt{h_n}} (ah_n \sum_{k=1}^{[T/ah_n]} g_n(\theta_0, \zeta_n(kah_n, \theta_0))) \right]$$

$$- \frac{1}{a} \int_0^T B(\theta_0, s(t, \theta_0)) z dt. \tag{3.42}$$

According to conditions of the theorem, the first two terms of (3.42) uniformly converges to a Normal random variable with mean zero and covariance matrix $\frac{1}{a}R^2$ such that,

$$\sqrt{h_n}\sum_{k=1}^{[T/ah_n]} R_n(\theta_0, \zeta_n(kah_n, \theta_0))^2 \to \frac{1}{a}\int_0^T R(\theta_0, s(t, \theta_0))^2 dt = \frac{1}{a}R^2.$$

The last term of (3.42) is equal to the deterministic function $\frac{1}{a}Bz$. Finally, $h_n^{-1/2} f_n(\theta_0 + \frac{z}{h_n^{-1/2}})$ uniformly converges in any bounded region $|z| \leq L$ to $N(0, \frac{1}{a}R^2) - \frac{1}{a}Bz$ and $aB^{-1}R^2(B^*)^{-1}$ is the solution of equation

$$N(0, \frac{1}{a}R^2) - \frac{1}{a}Bz = 0$$

According to Theorem 2.2.1, $h_n^{-1/2}(\tilde{\theta}_n - \theta_0)$ weakly converges to the Gaussian vector with mean 0 and variance $aB^{-1}R^2(B^*)^{-1}$.

3.5.2 Moments Method for Switching Processes

In applications we need to check the conditions (3.34) and (3.39). Consider now the case when we observe data on the trajectory of a Switching Process. In this case (3.34) and (3.39) can be verified in terms of individual characteristics of the process such as switching intervals and increments of the process on the switching intervals.

Let $\zeta_n(t, \theta)$ be a trajectory of a Switching Process . Let at each $\theta \in \Theta$ for each n=1,2... $\mathcal{F}_{nk} = \left\{(\xi_{nk}(\theta, \alpha),\ \tau_{nk}(\theta, \alpha)),\ \alpha \in \mathcal{R}^r\right\}, k \geq 0$, be jointly independent families of random vectors with values in $\mathcal{R}^r \times [0, \infty)$ and distributions not depending on index k. Also let $S_{no}(\theta)$ be an initial vector in $\mathcal{R}^r$ independent of $\{\mathcal{F}_{nk},\ k \geq 0\}$.

Let

$$t_{no}(\theta) = 0,$$

$$t_{nk+1}(\theta) = t_{nk}(\theta) + \frac{1}{n}\tau_{nk}(\theta, S_{nk}(\theta)),$$

$$S_{nk+1}(\theta) = S_{nk}(\theta) + n^{-1}\xi_{nk}(\theta, S_{nk}(\theta)),$$

and denote

$$\zeta_n(t, \theta) = S_{nk}(\theta),$$

as $t_{nk}(\theta) \le t < t_{nk+1}(\theta),\ t \ge 0$. Then $\zeta_n(t, \theta), t \ge 0$, is a recurrent process of a semi-Markov type (RPSM).

Assume that at each $\theta \in \Theta$ and for any $N > 0$

$$\lim_{L\to\infty} \limsup_{n\to\infty} \sup_{|\alpha|\le N}$$

$$\left\{\mathbf{E}\left(\tau_{n1}(\theta, \alpha)\chi(\tau_{n1}(\theta, \alpha) > L) + \mathbf{E}\left(|\xi_{n1}(\alpha)|\chi(\xi_{n1}(\theta, \alpha)| > L\right)\right\} = 0, \tag{3.43}$$

there exist functions

$$m_n(\theta, \alpha) = \mathbf{E}\tau_{n1}(\theta, \alpha), \quad b_n(\theta, \alpha) = \mathbf{E}\xi_{n1}(\theta, \alpha)$$

such that

$$|m_n(\theta, \alpha_1) - m_n(\theta, \alpha_2)| + |b_n(\theta, \alpha_1) - b_n(\theta, \alpha_2)| \le C_N|\alpha_1 - \alpha_2| + \alpha_n(N), \tag{3.44}$$

as $\max(|\alpha_1|, |\alpha_2|) \le N$, where C_N are some constants, $\alpha_n(N) \to 0$ uniformly in $\max(|\alpha_1|, |\alpha_2|) \le N$, and there exist functions $m(\theta, a) > 0$, $b(\theta, \alpha)$ such that for any $\alpha \in \mathcal{R}^r$ as $n \to \infty$

$$m_n(\theta, \alpha) \to m(\theta, \alpha),$$

$$b_n(\theta, \alpha) \to b(\theta, \alpha),$$

and

$$S_{no}(\theta) \xrightarrow{\mathrm{P}} s_0(\theta).$$

Then all the conditions of Averaging Principle (Theorem 2.2.4) for RPSM are satisfied hence, the relation (3.34) holds where $s(t, \theta)$ satisfies the differential equation:

$$ds(t, \theta) = m(\theta, s(t, \theta))^{-1}b(\theta, s(t, \theta))dt,$$

and

$$s(0, \theta) = s_0(\theta),$$

and T is chosen in such a way that $y(+\infty, \theta) > T$ a.s., where

$$y(t, \theta) = \int_0^t m(\theta, \eta(u, \theta))du, \ \eta(0, \theta) = s_0(\theta),$$

$$d\eta(u, \theta) = b(\theta, \eta(u))du$$

(it is supposed that a unique solution $\eta(u, \theta)$ exists on each interval).

Suppose that in addition, (3.43) holds for second moments, in (3.44) $\sqrt{n}\alpha_n(N) \to 0$, there exist continuous functions $D^2(\theta, \alpha), \sigma^2(\theta, \alpha), Q(\theta, \alpha), q(\theta, \alpha)$ such that uniformly in $|\alpha| \le N$,

$$\mathbf{E}\xi_{n1}(\theta, \alpha)\xi_n(\theta, \alpha)^* \to D^2(\theta, \alpha),$$

$$\mathbf{E}\tau_{n1}(\theta, \alpha)^2 \to \sigma^2(\theta, \alpha),$$

$$\sqrt{n}\Big(m_n(\theta, \alpha + z/\sqrt{n})^{-1}b_n(\theta, \alpha + z/\sqrt{n}) - m(\theta, \alpha)^{-1}b(\theta, \alpha)\Big) \to Q(\theta, \alpha)z + q(\theta, \alpha),$$

and

$$\sqrt{n}(S_{n0}(\theta) - s_0(\theta)) \overset{\mathrm{w}}{\Rightarrow} \gamma_0.$$

Then all the conditions of Diffusion Approximation (section 2.2.4) are satisfied and (3.39) holds at $nh_n \to \infty$.

In the following chapter, we will study the moments method estimators for different reliability models.

Chapter 4

Parameter Estimation in Reliability Models

In this chapter, we will use the results of section 3.5.2 to estimate the unknown parameters for several reliability models and study the asymptotic properties of the estimators. We show that the trajectories of reliability models we consider can be represented as Switching Processes. Our first model a reliability model without replacement and was included for illustration of our approach for parameter estimation. Second and third models are constructed with several changes of the first model. The final model, is the most extended model and includes the case when it is not possible to find an exact representation of the solutions of differential equations describing the system. We also give estimation results of two unknown parameters for that final model.

4.1 Model 1: A Reliability Model without Replacement

Suppose that a system consists of n devices subject to random failures. Any working device is considered as 'good' and any failed device is considered as 'bad'.

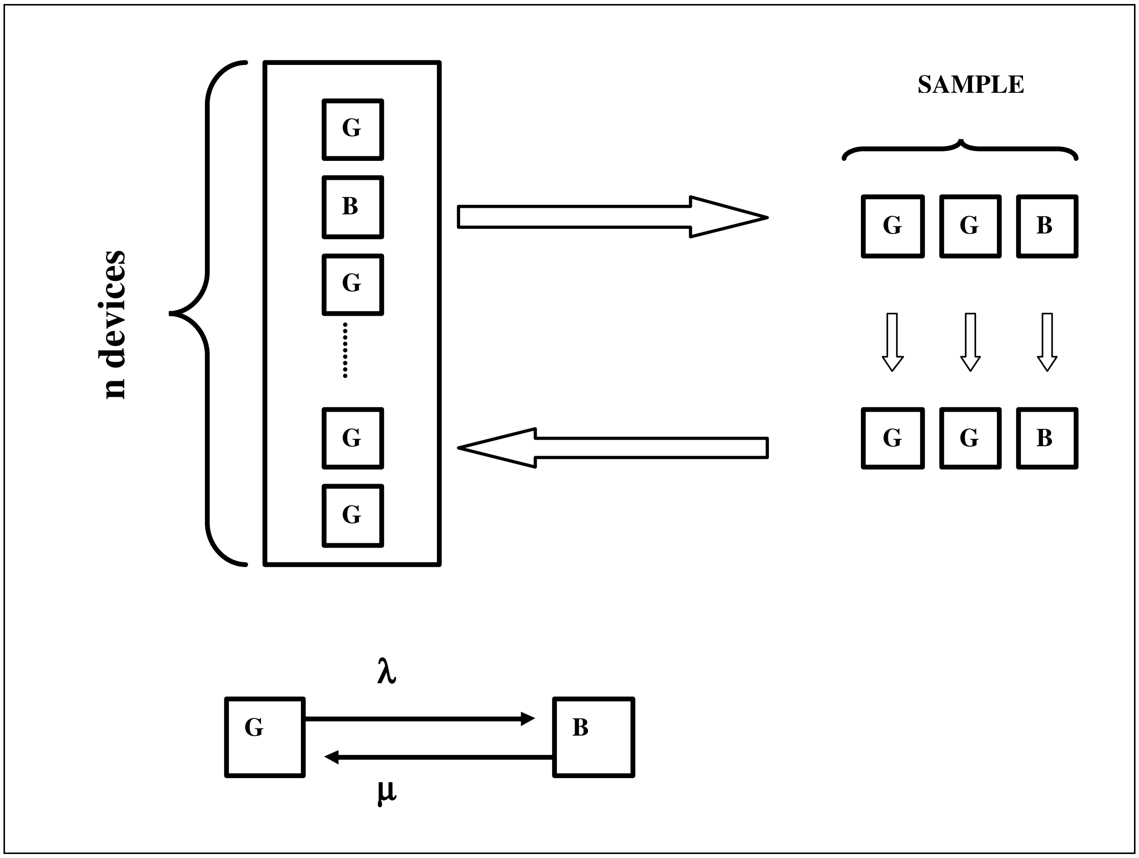

Figure 4.1: Model 1: Illustration

Let $S_n(t)$ be the number of failed devices at time t. If at time t, $\frac{1}{n}S_n(t) = s$ then each 'good' device has a local failure rate $\lambda(\theta_0, s)$, and each 'bad' device has a local repair rate $\mu(\theta_0, s)$ where the nonnegative functions $\lambda(\theta_0, \alpha), \mu(\theta_0, \alpha)$, $\Theta \in R^d$ are given. Here and further we assume both the time until the next failure and the time until the next repair are exponentially distributed.

Suppose that on the interval [0,T] we can provide a sample inspection at times kah_n, $k = 1, 2, ..., [T/ah_n]$ and a is a constant. At the time of inspection, we take a random sample of size m (m is fixed) and observe the number of failed devices in that sample without any repair. After the inspection, we return the sample back. Our goal is to estimate the parameter θ_0 from the observations of a trajectory of the system. Figure (4.1) illustrates the structure of our model.

Let y_{nk} be the number of failed devices in the observed sample taken at time

kah_n and also let $S_{ni} = S_n(t_{ni}), i = 1, 2, ...$ where t_{ni} are the times of sequential jumps of the process $S_n(t)$. Note that, since both the time until the failure and the time until a repair are exponentially distributed then the times between the jumps are also exponentially distributed. Let $s = \frac{1}{n}S_{ni}$. Then the process $S_n(t)$ is a Birth-and-Death Process. In state ns, birth and death rates are $(n - ns)\lambda(\theta, s)$ and $ns\mu(\theta, s)$.

Figure (4.2) illustrates the first transitions of a system trajectory.

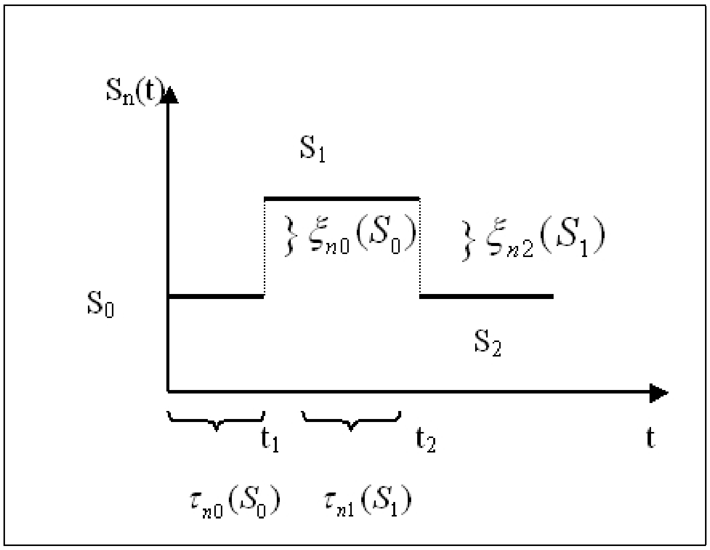

Figure 4.2: Model 1: A trajectory for the initial transitions

We can now write the following representation:

$$S_{ni+1} = S_{ni} + \xi_{ni}(\theta, s) \tag{4.1}$$

where

$$\xi_{ni}(\theta, s) = \begin{cases} 1 & \textit{with probability} \quad \frac{(n-ns)\lambda(\theta,s)}{(n-ns)\lambda(\theta,s)+ns\mu(\theta,s)} \\ -1 & \textit{with probability} \quad \frac{ns\mu(\theta,s)}{(n-ns)\lambda(\theta,s)+ns\mu(\theta,s)} \end{cases}$$

We can now represent the normalized process $s_n(t) = \frac{S_n(t)}{n}$ as follows

$$\begin{aligned} t_{ni+1} &= t_{ni} + \tau_{ni}(\theta, s)\frac{1}{n} \\ s_{ni+1} &= s_{ni} + \xi_{ni}(\theta, s)\frac{1}{n} \\ s_n(t) &= s_{ni}, t_{ni} \leq t < t_{ni+1} \end{aligned} \tag{4.2}$$

where $\tau_{ni}(\theta, s)$ is exponential with rate $(1-s)\lambda(\theta, s) + s\mu(\theta, s)$.

Then letting $\delta_n = \frac{1}{n}$ in the representation of RPSM (see section 2.2.3), the process $s_n(t)$ is a Recurrent Process of semi Markov type. We need to check the conditions of Averaging Principle for RPSM (Theorem 2.2.4) for the process $s_n(t)$. Note that the process $s_n(t)$ corresponds to the process $\frac{1}{n}S(n\alpha)$ in Theorem 2.2.4.

Following the notation of Theorem 2.2.4, we have the expectations

$$E(\tau_{ni}(\theta, s)) = m_n(\theta, s) = \frac{1}{(1-s)\lambda(\theta, s) + s\mu(\theta, s)} \tag{4.3}$$

and

$$E(\xi_{ni}(\theta, s)) = b_n(\theta, s) = \frac{(1-s)\lambda(\theta, s) - s\mu(\theta, s)}{(1-s)\lambda(\theta, s) + s\mu(\theta, s)} \tag{4.4}$$

hence

$$\begin{aligned} m(\theta, s) &= \frac{1}{(1-s)\lambda(\theta, s) + s\mu(\theta, s)}, \\ b(\theta, s) &= \frac{(1-s)\lambda(\theta, s) - s\mu(\theta, s)}{(1-s)\lambda(\theta, s) + s\mu(\theta, s)}. \end{aligned}$$

From continuous and finite expectation functions we see that the conditions of Averaging Principle are satisfied and

$$\sup_{t \in [0,T]} |s_n(t) - s(t, \theta)| \xrightarrow{\text{P}} 0, \tag{4.5}$$

where,

$$ds(t, \theta) = m(s(t, \theta))^{-1} b(s(t, \theta))dt$$

so that,

$$ds(t,\theta) = [\lambda(\theta, s(t,\theta)) - (\lambda(\theta, s(t,\theta)) + \mu(\theta, s(t,\theta)))s(t,\theta)]dt, \tag{4.6}$$

with some initial condition $s(0,\theta) = s_0$

Note that, the solution of differential equation (4.6), $s(t,\theta)$, represents the proportion of failed devices at time t.

Assume that $S_n(0) = 0$, hence, $s(0,\theta) = 0$. Let us take some function $\varphi(y)$ which has the same dimension d as the dimension of the vector of unknown parameters θ_0, where $\varphi(y) = [\varphi^{(1)}(y), \varphi^{(2)}(y), ..., \varphi^{(d)}(y)], y = 0, 1, 2, ..., m.$ and denote a binomial random variable with parameters (m,p) by $B(m,p)$. Let us denote the expectation as $g(p) = E(\varphi(B(m,p)), 0 \le p \le 1$. Consider the following equation constructed as analog of Moments Method equation:

$$h_n \sum_{k=1}^{[T/ah_n]} y_{nk} = \frac{1}{a}\int_0^T g(s(t,\theta))dt \tag{4.7}$$

where $s(t,\theta)$ satisfies the differential equation (4.6).

Let

$$f_n(\theta) = h_n \sum_{k=1}^{[T/ah_n]} y_{nk} - \frac{1}{a}\int_0^T g(s(t,\theta))dt. \tag{4.8}$$

Suppose that the functions $\lambda(\theta,s), \mu(\theta,s)$ satisfy the local Lipschitz condition (see section 2.2.4) uniformly in θ, $nh_n \to \infty$, and the function $f_0(\theta) = \int_0^T (g(s(t,\theta_0)) - g(s(t,\theta)))dt$ satisfies the condition of separateness **S**. Then according to Theorem 2.2.1, $\theta_n \xrightarrow{\mathrm{P}} \theta_0$, where θ_n is the set of solutions of (4.8).

Suppose that $\lambda(\theta,s) = \lambda$, $\mu(\theta,s) = \mu$. Consider the case when μ is known and we want to estimate the actual failure rate λ_0. Since the unknown parameter is one dimensional we can take $\varphi(y) = y$. The differential equation (4.6) is now reduced to

$$ds(t,\lambda) = [\lambda - (\lambda+\mu)s(t,\lambda)]dt \tag{4.9}$$

and

$$s(t,\lambda) = \frac{1}{\lambda+\mu}\lambda(1-e^{-(\lambda+\mu)}) \tag{4.10}$$

satisfies equation (4.9) with $s(0,\lambda) = 0$. In this case, $f_n(\theta)$ uniformly converges to the limiting function $f_0(\lambda)$, which satisfies the condition **S**, where

$$f_0(\lambda) = \int_0^T (ms(t,\lambda_0) - ms(t,\lambda))dt.$$

Note that λ_0 is the solution of equation $f_0(\lambda) = 0$. Thus the estimator $\hat{\lambda}$ is a consistent estimator of λ_0.

We can find the estimator $\hat{\lambda}$ by substituting $s(t,\lambda)$ in to the equation (4.7) and solving the resulting equation

$$\frac{m\lambda}{a(\lambda+\mu)}\left(T - \frac{e^{-(\lambda+\mu)T}}{\lambda+\mu} + \frac{1}{\lambda+\mu}\right) = h_n \sum_{k=1}^{[T/h_n]} y_{nk}$$

for λ.

In the steady state, as $t \to \infty$, total repair rate of 'bad' devices and the total failure rate of 'good' devices will be equal. Then the trajectory of the system will be stochastically trembling around a constant level. The function $s(t,\lambda)$ also indicates the same result since

$$s(\infty,\lambda) = \frac{\lambda}{\lambda+\mu}.$$

This means, the proportion of failed devices in steady state will be $\lambda/(\lambda+\mu)$.

Let us also prove the asymptotic normality. According to Theorem 3.5.2, we have to check the condition (3.40).

Let $z = \sqrt{n}(\hat{\lambda} - \lambda_0)$. From Taylor's formula we can write,

$$s(t,\lambda) = s(t,\lambda_0) + B'\ z\sqrt{h} + o(\cdot)\sqrt{h}$$

where $B' = (\frac{d}{d\lambda}s(t,\lambda))_{\lambda_0}$. As in our case $g(p) = m\,p$ then in (3.40) $B(\theta_0, s(t,\theta_0)) = mB'$

Let us define, $B = \int_0^T B'dt$ and $R^2 = \int_0^T ms(t,\lambda_0)(1-s(t,\lambda_0))dt$. Then according to Theorem 3.5.2, z weakly converges to a normal random variable with mean zero and variance aR^2/B^2.

4.2 Model 2: Estimation in a Reliability Model with Replacement

Consider n devices subject to independent random failures (any 'good' device has a failure rate λ_0). Assume that on the interval $[0,T]$ we have the possibility to provide a sample inspection at times $kh_n,\ k = 1,2,..,[T/h_n]$, that means we take at random a sample of a fixed size m and can observe the number of failed devices Q_{nk} in it. Each failed device in the sample with probability $\beta_0,\ 0 \leq \beta_0 \leq 1$, is immediately replaced by a new one. Otherwise, (if not replaced) the failed device remains failed in the sample. After the inspection we return the sample back.

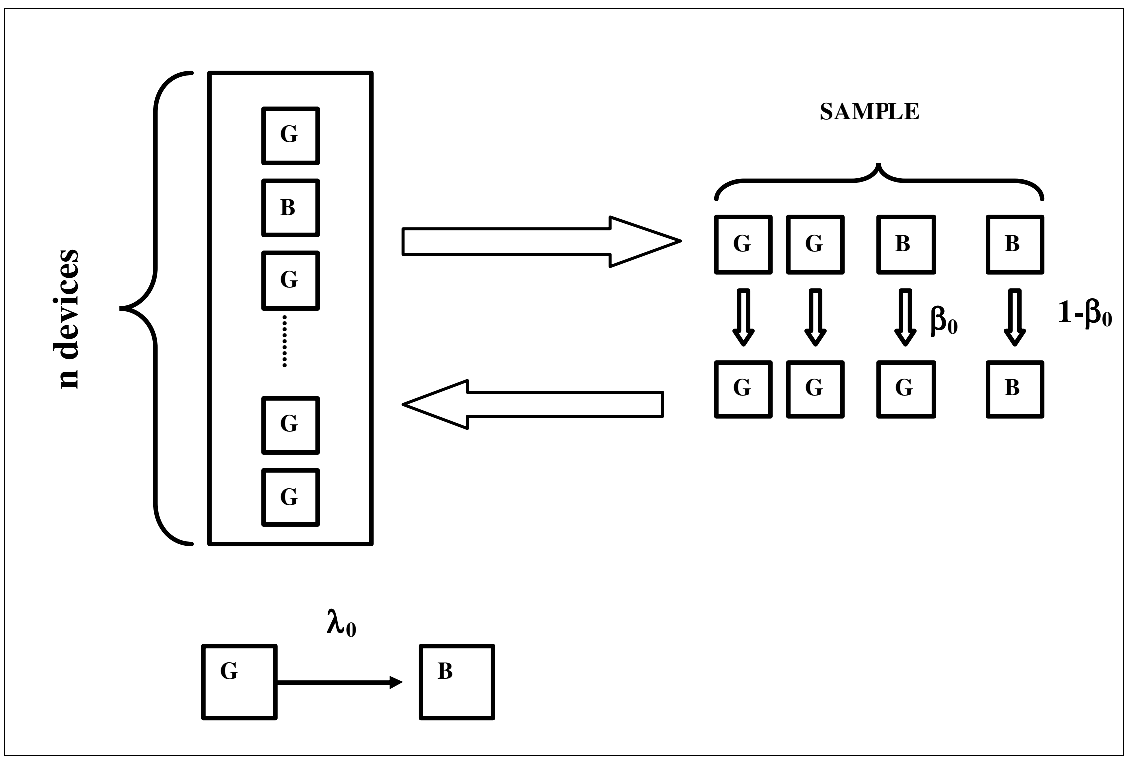

Figure 4.3: Model 2: Illustration

Figure (4.3) illustrates the structure of the model. Our goal is to estimate the failure rate λ_0 and the probability β_0.

Suppose that $\lambda_0 = \lambda(\theta_0)$, $\beta_0 = \beta(\theta_0)$, where the functions $\lambda(\theta), \beta(\theta)$, $\theta \in \Theta$, are some continuous functions, Θ is a bounded closed set in $\mathcal{R}^d$, and θ_0 is the unknown parameter.

Consider a d-dimensional function $\varphi(y) = (\varphi^{(1)}(y), ..., \varphi^{(d)}(y))$ of a discrete argument $y = 0, 1, .., m$. Denote by $B(m,p)$ a binomial random variable with parameters (m,p). For fixed m, $g(p) = \mathbf{E}\varphi(B(m,p))$ be a function of argument p, $0 \leq p \leq 1$. Denote also $y_{nk} = \varphi(Q_{nk})$. Assume that $h_n = a/n$, $a > 0$, and without loss of generality assume that initially all devices are good. Consider the moments method equation:

$$\int_0^T g(s(t,\theta))\mathrm{d}t = h_n \sum_{k=1}^{[T/h_n]} y_{nk}, \tag{4.11}$$

where

$$s(t,\theta) = \frac{\lambda(\theta)}{\lambda(\theta) + m\beta(\theta)/a}(1 - e^{-(\lambda(\theta)+m\beta(\theta)/a)t}). \tag{4.12}$$

Theorem 4.2.1 *Suppose that the function*

$$f_0(\theta) = \int_0^T \Big(g(s(t,\theta_0)) - g(s(t,\theta))\Big)\mathrm{d}t, \tag{4.13}$$

as a function of θ*, satisfies condition* $\boldsymbol{S}$ *. Then* $\{\boldsymbol{\theta_n}\} \xrightarrow{\mathrm{P}} \theta_0$*, where* $\{\boldsymbol{\theta_n}\}$ *is the set of solutions (4.11) in* θ*.*

Proof.

Denote by $S_n(t)$ the number of failed devices at time t. If we observe $S_n(t)$ at times kh_n, for the initial transactions $k = 1, 2$ we would observe a behavior similar to Figure (4.4).

First we study the asymptotic behavior of the trajectory of the normalized process $\zeta_n(t) = S_n(t)/n$. For this, we represent $\zeta_n(t)$ as a Switching Process at points kh_n.

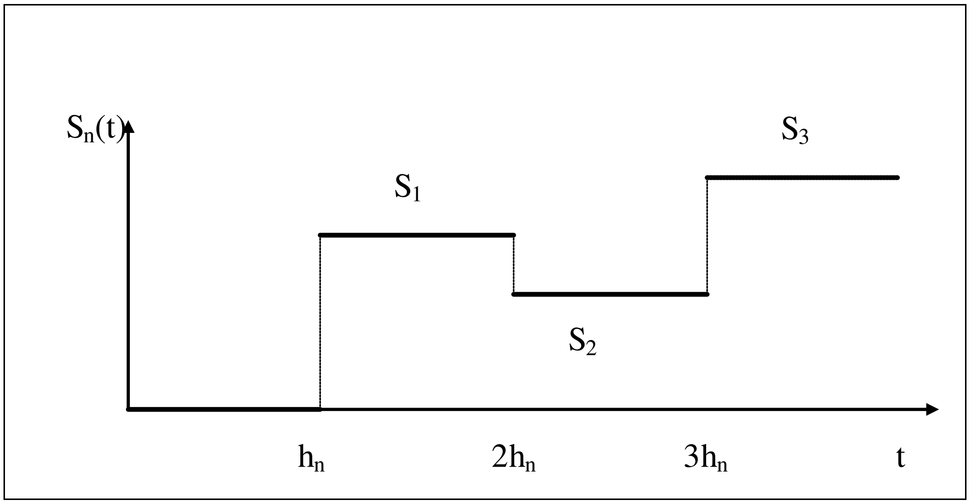

Figure 4.4: Model 2: A trajectory for the initial transitions

Denote

$$s_{nk}^{-} = \zeta_n(kh_n - 0),\ \ s_{nk}^{+} = \zeta_n(kh_n + 0),$$

so that s_{nk}^{-} is the number of failed devices just before k-th inspection normalized by n, and s_{nk}^{+} is the number of failed devices just after k-th inspection and replacement, normalized by n. Between the inspections, the system will show a behavior similar to Figure (4.5).

According to the figure, let us denote h_n^{-} as the time just before the first inspection and $2h_n^{-}$ as the time just before the second inspection. Between $2h_n^{-}$ and h_n^{-}, an inspection is applied just after h_n^{-}, if a failed device is found then it is replaced. After inspection some of the 'good' devices may fail until the next inspection. Hence, we have two actions to include in our representation of the trajectory.

Denote by $H(n, j, m)$ a hypergeometric random variable with parameters (n, j, m):

$$P(H(n, j, m) = i) = \frac{\binom{j}{i}\binom{n-j}{m-i}}{\binom{n}{m}},\ \ i = 0, 1, .., m. \tag{4.14}$$

Let λ be the failure rate and β be the probability of a correct replacement. Denote by $p_n(\lambda) = 1 - e^{-\lambda h_n}$ a probability that a good device fails during time

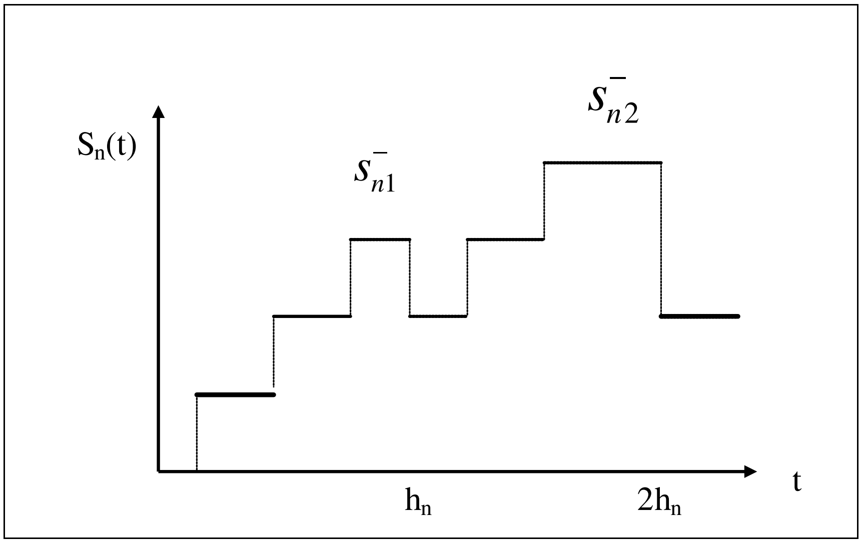

Figure 4.5: Model 2: Behavior between the inspections

h_n (between two successive inspections).

According to the evolution of the process, we can write stochastic relations:

$$s_{nk}^+ = s_{nk}^- - \frac{1}{n} B\left(H(n, ns_{nk}^-, m), \beta\right), \tag{4.15}$$

$$s_{nk+1}^- = s_{nk}^+ + \frac{1}{n} B\left(n - ns_{nk}^+, p_n(\lambda)\right), k = 0, 1, ... \tag{4.16}$$

Assuming initially all devices are good, we have $s_{n0}^- = 0$.

Thus, we represented $\zeta_n(t)$ as an RPSM with switching points kh_n taking $\delta_n = h_n$, $\tau_{nk}(\cdot) = 1$ (see Section 2.2.3). As $h_n = a/n$, we get from (4.15) and (4.16) that

$$s_{nk+1}^- = s_{nk}^- + \xi_{nk}(s_{nk}^-)h_n, \tag{4.17}$$

$$t_{nk+1} = t_{nk} + \tau_{nk}(\cdot)h_n \tag{4.18}$$

where, given $s_{nk}^- = s$,

$$\xi_{nk}(s) = -B\left(H(n, ns, m), \beta\right) a^{-1} + B\left(n - ns^+(s), p_n(\lambda)\right) a^{-1},$$

and

$$s^{+}(s) = s - \frac{1}{n} B\left(H(n, ns, m), \beta\right).$$

We now need to check the conditions of Averaging Principle to prove the convergence of trajectory of $\zeta_n(t)$. Note that $\zeta_n(t)$ corresponds to $\frac{1}{n}S_n(nt)$ of Theorem 2.2.4.

As $p_n(\lambda) \approx \lambda h_n = \lambda a/n$,then

$$E\xi_{nk}(s) \to -sm\beta/a + (1-s)\lambda \tag{4.19}$$

uniformly in $s \in [0, 1]$.

Following from the notation of Theorem 2.2.4, we have

$$m(s) = 1$$

and

$$b(s) = -sm\beta/a + (1-s)\lambda$$

Now, using the averaging principle for RPSM, and recurrent relation (4.17), we get

$$\sup_{t\in[0,T]} |\zeta_n(t) - s(t)| \xrightarrow{\mathrm{P}} 0, \tag{4.20}$$

where the function $s(t)$ is a solution of a linear differential equation

$$ds(t) = m(s(t))^{-1}b(s(t))dt.$$

Using the functions $m(s)$ and $b(s)$ obtained, we have,

$$ds(t) = \Big(\lambda - s(t)(\lambda + m\beta/a)\Big)dt, \ s(0) = 0. \tag{4.21}$$

Note that, λ and β are the functions of θ, therefore $s(t) = s(t,\theta)$. Solving equation (4.21) with initial value $s_0 = 0$ we get,

$$s(t,\theta) = \frac{\lambda(\theta)}{\lambda(\theta) + m\beta(\theta)/a}(1 - e^{-(\lambda(\theta)+m\beta(\theta)/a)t}),$$

which is the same expression with (4.12) for the function $s(t,\theta)$.

Therefore, for any θ, the condition (3.34) is satisfied and we can use Theorem 4.2.1.

In our case, $y_{nk} = \varphi(Q_{nk})$. As y_{nk} are bounded, condition (3.36) is true. Now, given $s_{nk}^- = s$, the variable Q_{nk} (the number of failed devices in a sample) has hypergeometric distribution $H(n, ns, m)$. Therefore, in notation of Theorem 3.5.1, $g_n(\theta, s) = E\varphi(H(n, ns, m))$.

If $n \to \infty$ and $j = [ns]$, then we can check directly using (4.14) that for any $i = 0, .., m$ uniformly in $s \in [0,1]$,

$$P(H(n, ns, m) = i) \to \binom{m}{i} s^i (1-s)^{m-i}.$$

The right-hand side corresponds to Binomial distribution with parameters (m, s). This means, as $n \to \infty$, $H(n, ns, m)$ converges in distribution to $B(m, s)$ and $E\varphi(H(n, ns, m)) \to g(s)$ uniformly in s. Thus, all conditions of Theorem 3.5.1 are satisfied and our statement is true.

Note that if $nh_n \to \infty$, then the replacement is too slow and our system is asymptotically equivalent to the system without replacement which can be obtained by putting $\beta = 0$. In this case $s(t,\theta) = 1 - e^{-\lambda(\theta)t}$.

If initially $S_n(0) \approx ns_0$, then in the equation (4.21) $s(0) = s_0$ and

$$s(t,\theta) = \frac{\lambda(\theta)}{\lambda(\theta) + m\beta(\theta)/a}(1 - e^{-(\lambda(\theta)+m\beta(\theta)/a)t}) + s_0 e^{-(\lambda(\theta)+m\beta(\theta)/a)t}.$$

Consider some particular cases.

Assume that $\beta_0 = \beta$ and $\lambda_0 = \lambda$ are given, probability β_0 is known and we need to estimate only the failure rate λ_0. In this case the unknown parameter is

$\theta = \lambda$, and

$$s(t, \lambda) = \frac{\lambda}{\lambda + m\beta/a}(1 - e^{-(\lambda+m\beta/a)t}). \tag{4.22}$$

The function $s(t, \lambda)$ represents the proportion of failed devices at time t.

Let us take $\varphi(y) = y$ so that $y_{nk} = Q_{nk}$. Then $g(p) = mp$, and the moments method equation has the form:

$$m \int_0^T s(t, \lambda) \mathrm{d}t = h_n \sum_{k=1}^{[T/h_n]} y_{nk}, \tag{4.23}$$

and the function $f_0(\lambda)$ in (3.5.1) has the form:

$$f_0(\lambda) = m \int_0^T (s(t, \lambda_0) - s(t, \lambda))dt. \tag{4.24}$$

It is easy to check that both functions $\frac{\lambda}{\lambda+m\beta/a}$ and $1 - e^{-(\lambda+m\beta/a)t}$ strictly monotonically increase in λ, therefore, the function $s(t, \lambda)$ at each t strictly monotonically increases in λ and the function $f_0(\lambda)$ satisfies condition $\boldsymbol{S}$. Thus, the left-hand side of the equation (4.23) also strictly monotonically increases in λ. Integrating $s(t, \lambda)$ in t, we can write (4.23) in the form

$$\frac{m\lambda}{\lambda + m\beta/a} \left(T - \frac{1 - e^{-(\lambda+m\beta/a)T}}{\lambda + m\beta/a} \right) = h_n \sum_{k=1}^{[T/h_n]} y_{nk}. \tag{4.25}$$

The right-hand side of equation (4.25) is changing in the interval $(0, mT)$. When λ is changing from 0 till ∞, the left-hand side is monotonically changing from 0 till mT. Therefore, a unique solution $\widehat{\lambda}_n$ of equation (4.25) exists and the estimator $\widehat{\lambda}_n$ is consistent according to Theorem 3.5.1

Suppose now that the failure rate $\lambda_0 = \lambda$ is known but the probability $\beta_0 = \beta$ is unknown. In this case the parameter $\theta = \beta$. The moments method equation has the form (4.23) where $s(\cdot) = s(t, \beta)$ now depends on parameter β.

It is easy to check that $\frac{\partial s(t,\beta)}{\partial \beta} < 0$. Thus, the left-hand side in (4.23) with function $s(t, \beta)$ is strictly monotonically decreasing. Therefore, the function $f_0(\lambda)$ satisfies condition $\boldsymbol{S}$. In this case, the solution $\widehat{\beta}_n$ of the equation (4.25) exists at large n, is unique and the estimator $\widehat{\beta}_n$ is consistent according to Theorem 2.2.1.

If both λ_0 and β_0 are unknown, then $\theta = (\lambda, \beta)$ and we have to solve a system of two equations. Therefore, let us consider, for example, a two dimensional function $\varphi(y) = (y, y^2)$. Then

$$g^{(1)}(p) = EB(m,p) = mp,$$

$$g^{(2)}(p) = E(m,p)^2 = mp(1-p) + m^2p^2.$$

Thus, a system of moments method equations for two parameters (λ, β) has the form: the first equation is (4.23) and the 2nd equation has the form

$$\int_0^T \Big(ms(t,\theta)(1 - s(t,\theta)) + m^2 s(t,\theta)^2 \Big)\, dt = h_n \sum_{k=1}^{[T/h_n]} y_{nk}^2, \tag{4.26}$$

where $s(t,\theta)$ is given by the expression (4.23). The left-hand side of (4.26) can be calculated explicitly and we get a system of two nonlinear equations for two unknown parameters (λ_0, β_0).

4.3 Model 3: Reliability Model with N Repairmen

Consider n devices subject to independent random failures with rate λ_0. There are N repairmen each with large repair rate $n\mu_0$. Each failed device is immediately taken for repair if at least one repairman is available, otherwise the device is waiting its turn. After repair, the device is considered to be as good as new and immediately starts to work again.

Suppose that on the interval $[0, T]$ we provide a sample inspection at times $kh_n,\ k = 1, 2, .., [T/h_n]$. That is, we take at random a sample of a fixed size m and can observe (without repair) the number of failed devices Q_{nk} in it. After inspection we return the sample back. Assume that initially all devices are 'good'. We will consider the behavior of the system under heavy traffic, so that $\lambda_0 > N\mu_0$.

Suppose that $\lambda_0 = \lambda(\theta_0)$, $\mu_0 = \mu(\theta_0)$, where the functions $\lambda(\theta), \mu(\theta)$ $\theta \in \Theta$, are continuous functions, Θ is a bounded closed set in $\mathcal{R}^d$, and θ_0 is the unknown parameter.

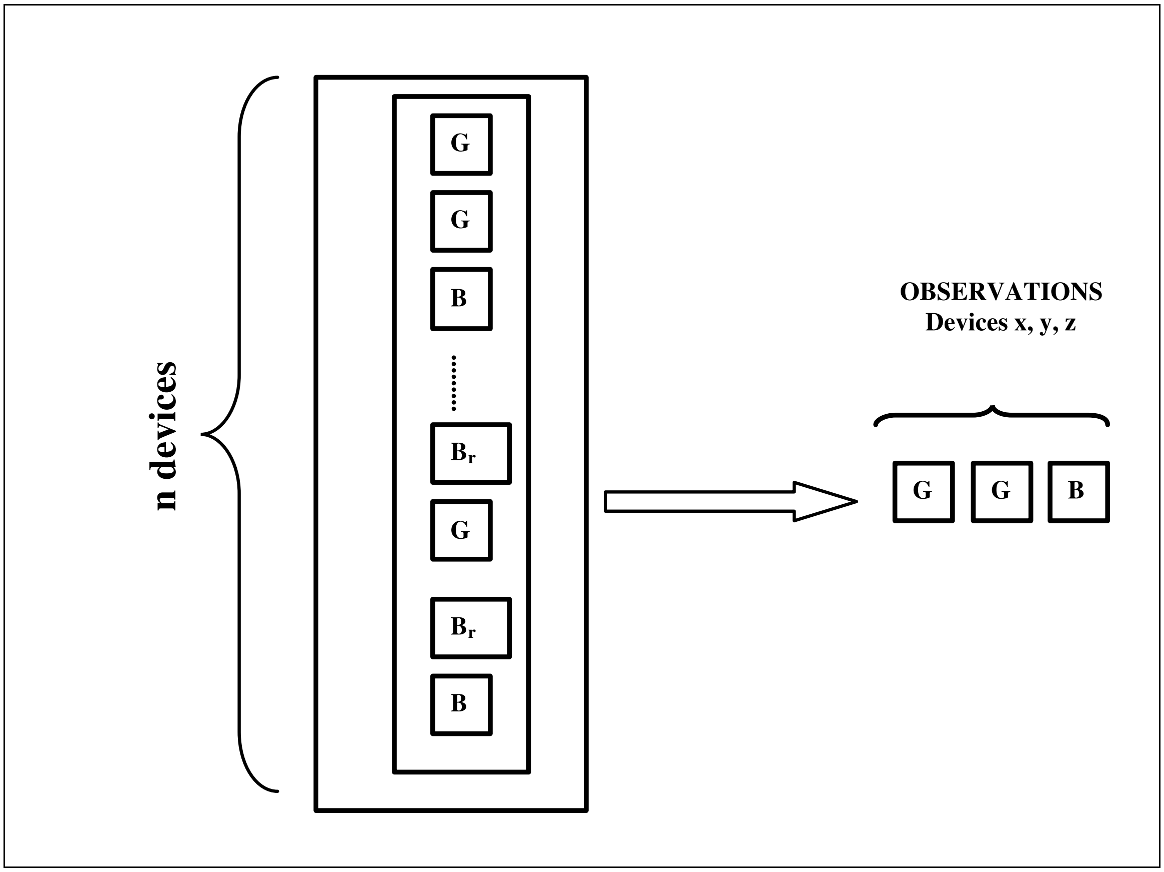

Figure 4.6: Model 3: Illustration

The Figure (4.6) presents the case when there are two repairmen and the devices which are in repair are indicated as B_r.

We consider a d-dimensional function $\varphi(y) = (\varphi^{(1)}(y), .., \varphi^{(d)}(y)), y = 0, 1, .., m$. Let $g(p) = \mathbf{E}\varphi(B(m,p)), 0 \leq p \leq 1$, where $B(m,p)$ is a binomial random variable with parameters m and p. Denote $y_{nk} = \varphi(Q_{nk})$ so that y_{nk} is a function of our observations . Assume that $\lambda_0 > N\mu_0$ and consider the equation:

$$\int_0^T g(s(t,\theta))dt = h_n \sum_{k=1}^{[T/h_n]} y_{nk}, \tag{4.27}$$

where $s(t,\theta) = (1 - N\mu(\theta)/\lambda(\theta))(1 - e^{-\lambda(\theta)t})$.

Theorem 4.3.1 *Suppose that $h_n \to 0$, and the function in the left-hand side of (4.27) satisfies condition $\boldsymbol{S}$.*

Then $\{\boldsymbol{\theta_n}\} \xrightarrow{\text{P}} \theta_0$, *where* $\{\boldsymbol{\theta_n}\}$ *is the set of solutions of (4.27).*

Proof. Let $S_n(t)$ be the number of failed devices at time t and $\zeta_n(t) = S_n(t)/n$. Also let the failure and the repair rates be λ and $n\mu$. First we need to check the condition (3.34). For this, we represent $\zeta_n(t)$ as a Switching process. Note that $S_n(t)$ is a Birth-and-Death process with birth and death rates in state j, $(n-j)\lambda$ and $\min(j, N)n\mu$, respectively. Denote by t_{nk} the sequential times of jumps of $S_n(\cdot)$. We construct a switching process by the times t_{nk}.

Denote $s_{nk} = \zeta_n(t_{nk}+0)$ and let $h_n = 1/n$. Note that for $j \geq N$, the time spent in state j is exponential with rate $(n-j)\lambda + Nn\mu$. Otherwise, the time spent in state j is exponential with rate $(n-j)\lambda + jn\mu$. But since we consider the system under heavy traffic, asymptotically $j \geq N$. Denote $j/n = s$. Then at $s > 0$ and large n such that $ns \geq N$, the time spent in state j can be represented as $\tau_{nk}(s)/n$, where $\tau_{nk}(s)$ has an exponential distribution with rate $(1-s)\lambda + N\mu$.

For $\delta_n = \frac{1}{n}$, we can write the following recurrent relations:

$$s_{nk+1} = s_{nk} + \xi_{nk}(s_{nk})\delta_n$$
$$t_{nk+1} = t_{nk} + \tau_{nk}(s_{nk})\delta_n,$$

where given $s_{nk} = s$,

$$\xi_{nk}(s) = \begin{cases} 1 & \text{with probability} \quad \frac{(1-s)\lambda}{(1-s)\lambda + N\mu}, \\ -1 & \text{with probability} \quad \frac{N\mu}{(1-s)\lambda + N\mu}. \end{cases}$$

Hence we represented the process $\zeta_n(t)$ as an RPSM (see section 2.2.3) with switching points t_{nk}.

Following the notation of Theorem 2.2.4, we now have

$$m(s) = \frac{1}{(1-s)\lambda + N\mu}$$

and

$$b(s) = \frac{(1-s)\lambda - N\mu}{(1-s)\lambda + N\mu}.$$

Note that, $\zeta_n(t)$ corresponds to $\frac{1}{n}S_n(n\alpha)$ in Theorem 2.2.4. Since all the conditions of Averaging principle are satisfied, for $\lambda > N\mu$, the condition

$$\sup_{0 \le t \le T} |\zeta_n(t) - s(t)| \xrightarrow{\mathrm{P}} 0.$$

is also satisfied for any $T > 0$, where $s(t)$ is the solution of differential equation

$$ds(t) = m(s(t))^{-1} b(s(t))dt$$

so that

$$ds(t) = ((1 - s(t))\lambda - N\mu)dt.$$

Assume that initially all devices are good. Since $\lambda > N\mu$, we can find the exact solution of the differential equation with an initial condition $s_0 = 0$ as

$$s(t) = (1 - N\mu/\lambda)(1 - e^{-\lambda t}).$$

Taking into account that λ and β are functions of θ, we get

$$s(t, \theta) = (1 - N\mu(\theta)/\lambda(\theta))(1 - e^{-\lambda(\theta)t}).$$

As the right-hand side of (4.27) has the same form as in (4.23), then using the same arguments as at the proof of Theorem (4.2.1) we get our statement.

Consider as an example the case when the functions $\mu(\theta) = \mu$ and $\lambda(\theta) = \lambda$ are given, $\mu = \mu_0$ is known, and λ_0 is unknown. Now $\theta = \lambda$ and we can take $\varphi(y) = y$ so that $y_{nk} = Q_{nk}$. Then $g(p) = mp$ and after integration the moments method equation can be represented as the equation

$$m(1 - N\mu/\lambda)\Big(\lambda T - 1 + e^{-\lambda T}\Big)/\lambda = h_n \sum_{k=1}^{[T/h_n]} y_{nk} \tag{4.28}$$

with respect to λ. Note that the function $s(t,\lambda) = (1 - N\mu/\lambda)(1 - e^{-\lambda t})$ at each $t > 0$ is strictly monotonically increasing in λ. Also, the left-hand side of (4.28) is changing from 0 till mT as λ is changing from $N\mu$ till ∞. Thus, condition **S** is satisfied and a unique solution $\widehat{\lambda}_n$ of (4.28) exists and is consistent according to Theorem 3.5.1. Note that the case $\mu_0 = 0$ (no repair) is equivalent to the case $\beta_0 = 0$ considered in the previous section.

4.4 Model 4: A Reliability Model with Probabilistic Chance of Repair

We have n devices subject to random failures. Any 'good' device has a failure rate λ_0. There are N repairmen in the system each working with a large repair rate of nμ. When a device fails, it goes to repair only with probability β_0. That is, with probability $1 - \beta_0$, a failed device may not be sent to repair and just stays failed. $1 - \beta_0$ can be interpreted as the error of the recognition of the failed devices and in this case that failed device can be found only during further inspections. A device which is sent to repair, starts to get repaired if at least one repairman is available, otherwise it waits its turn in queue.

For constant h_n, at times $kh_n, k = 1, 2, ..[T/h_n]$, we provide a sample inspection as follows: we take, at random, sample of size m from the machines which are not in repair and observe the number of failed devices, Q_{nk},in that sample. The observed failed devices are immediately sent to repair and the remaining 'good' devices are returned back to operation. We study the behavior of the system under heavy traffic, so that the number of devices in the repair is large and $\lambda_0\beta_0 > N\mu$.

Figure (4.7) illustrates the structure of the model. Our goal is to estimate the failure rate λ_0 and the probability β_0 by sample observations Q_{nk}.

We consider a d-dimensional function $\varphi(x) = (\varphi^{(1)}(x), .., \varphi^{(d)}(x)), x = 0, 1, .., m$, and let $g(p) = \mathbf{E}\varphi(B(m,p)), 0 \le p \le 1$. Denote $y_{nk} = \varphi(Q_{nk})$, where

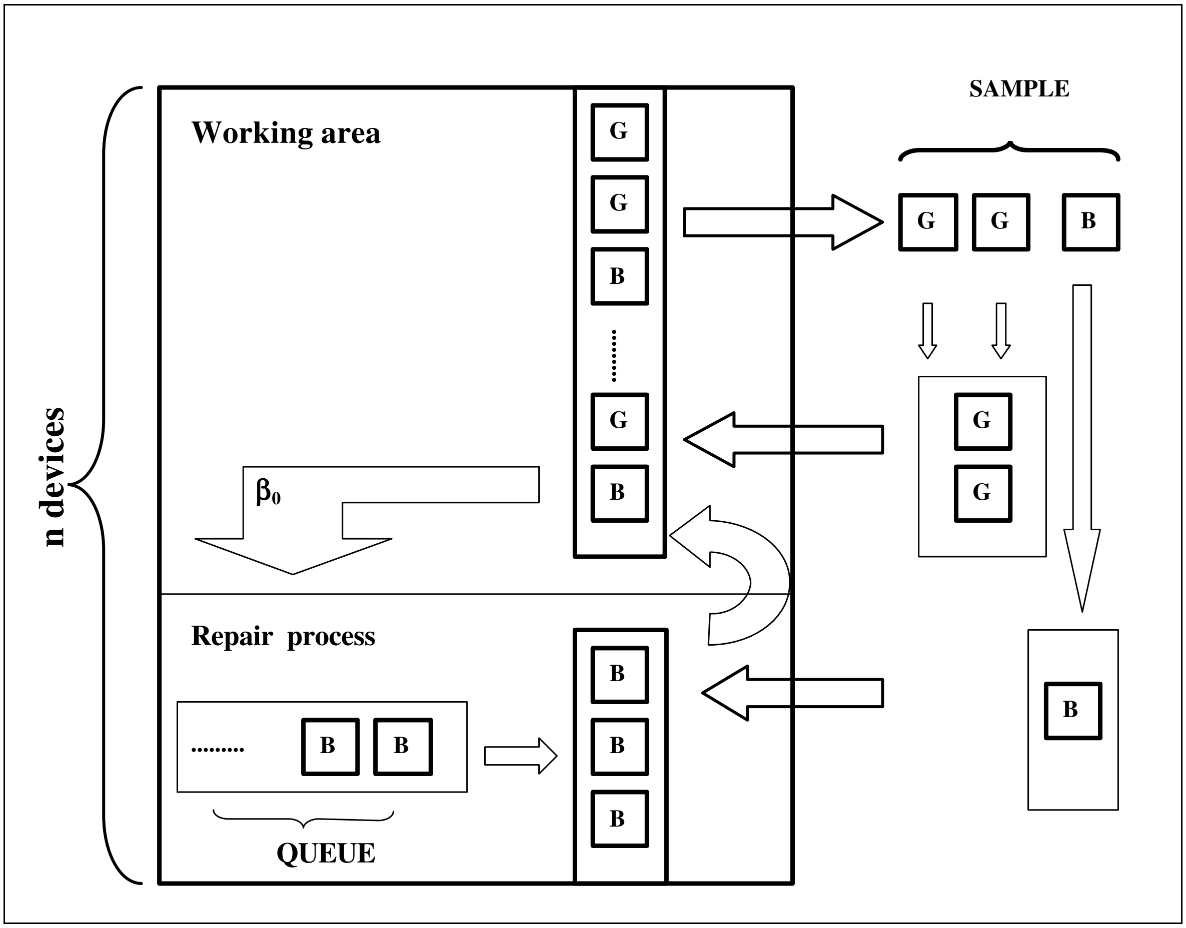

Figure 4.7: Model 4: Illustration

$B(m,p)$ is a binomial random variable with parameters m and p. Assume that $\lambda_0\beta > N\mu_0$ and consider the moments method equation:

$$\int_0^T g(p(t,\theta))dt = h_n \sum_{k=1}^{[T/h_n]} y_{nk}, \tag{4.29}$$

where $p(t,\theta)$ is a continuous deterministic function, satisfying the condition $h_n \sum_{k=1}^{[T/h_n]} y_{nk} \to \int_0^T g_0(p(t,\theta_0))dt$.

In the following parts we will give the analytical and computational approaches for how to calculate the function $p(t,\theta)$.

Theorem 4.4.1 *Suppose that $h_n \to 0$, and the function in the left-hand side of (4.29) satisfies condition $\boldsymbol{S}$.*

Then $\{\boldsymbol{\theta_n}\} \xrightarrow{\mathrm{P}} \theta_0$, *where* $\{\boldsymbol{\theta_n}\}$ *is the set of solutions of (4.29).*

Proof.

Denote $S_n(t)$ as the total number of failed devices at time t. Also denote $R_n(t)$ as the number of devices in repair or waiting in queue to be repaired (in repair process) and $Y_n(t)$ as the number of failed devices which are not in repair or waiting to be repaired (not in repair process) at time t. From the construction, we have $S_n(t) = Y_n(t) + R_n(t)$.

Also denote $Z_n(t) = (R_n(t), Y_n(t))$ as a two component process of number of failed devices. Given that $(R_n(t), Y_n(t)) = (R, Y)$ we have three possible transactions at the time of next jump as follows:

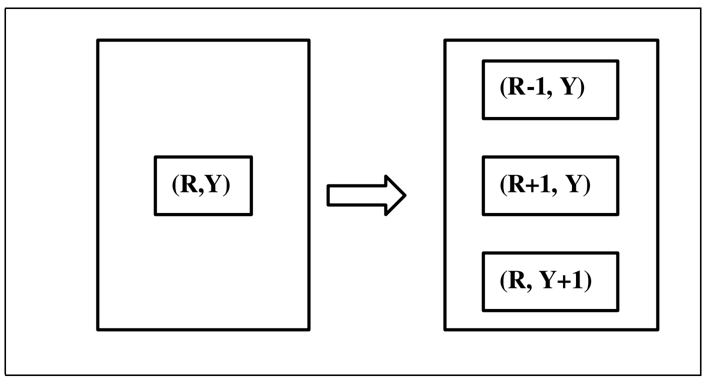

Figure 4.8: Model 4: Transactions

We will study the asymptotic behavior of the normalized process $z_n(t) = Z_n(t)/n = (R_n(t)/n, Y_n(t)/n) = (r_n(t), y_n(t))$.

Let $Z_k = Z_n(t_k)$ and $z_{nk} = z_n(t_k)$ so that $z_n(t_k) = (r_n(t_k), y_n(t_k))$. We first will consider the behavior of the process $z_n(t)$ on the time interval $[0, T]$ in case when no inspection is provided.

Let $r_n(t) = r, y_n(t) = y$ so that $z_n(t) = (r, y)$. While at state (r, y), there are three possible transitions at the time of next jump: to the states $(r - \frac{1}{n}, y), (r + \frac{1}{n}, y)$ and $(r, y + \frac{1}{n})$ with respective rates $nN\mu, n(1 - r - y)\lambda\beta$ and $n(1 - r - y)\lambda(1 - \beta)$.

Time spent in each state (r, y) is exponential with parameter $nN\mu + (n - nr - ny)\lambda$. Let $\tau_{nk}(r, y)$ be an exponential random variable with parameter $N\mu + (1 - r - y)\lambda$. Then the two component process $Z_n(t) = (R_n(t), Y_n(t))$ is a quasi Birth-and-Death process. Let $\frac{1}{n}\xi_{nk}(r, y)$ be the size of jump from state (r, y) at the time of exit from this state. Then taking $\delta_n = \frac{1}{n}$, we can write the following stochastic relations:

$$\begin{aligned} z_{nk+1} &= z_{nk} + \xi_{nk}(z_{nk})\delta_n \\ t_{nk+1} &= t_{nk} + \tau_{nk}(z_{nk})\delta_n, \end{aligned}$$

Note that given $r_{nk} = r$ and $y_{nk} = y$, $\xi_{nk}(r, y)$ is a random vector such that,

$$\xi_{nk}(r, y) = \begin{cases} (-1, 0) & with \;\; probability \;\; \frac{N\mu}{N\mu + (1-r-y)\lambda} \\ (1, 0) & with \;\; probability \;\; \frac{(1-r-y)\lambda\beta}{N\mu + (1-r-y)\lambda} \\ (0, 1) & with \;\; probability \;\; \frac{(1-r-y)\lambda(1-\beta)}{N\mu + (1-r-y)\lambda} \end{cases}$$

with vector expectation

$$E(\xi_{nk}(r, y)) = \left[\frac{(1 - r - y)\lambda\beta - N\mu}{N\mu + (1 - r - y)\lambda}, \frac{(1 - r - y)\lambda(1 - \beta)}{N\mu + (1 - r - y)\lambda} \right]. \tag{4.30}$$

Hence we represented the process $z_n(t)$ as RPSM with switching points t_{nk}. Following the notation from Theorem 2.2.4, and taking $\alpha = (r, y), m_n(\alpha) = E\tau_{nk}(r, y)$ and $b_n(\alpha) = E(\xi_{nk}(r, y))$ we have

$$m(r, y) = \frac{1}{N\mu + (1 - r - y)\lambda}$$

and

$$b(r, y) = E(\xi_{nk}(r, y)) = \left[\frac{(1 - r - y)\lambda\beta - N\mu}{N\mu + (1 - r - y)\lambda}, \frac{(1 - r - y)\lambda(1 - \beta)}{N\mu + (1 - r - y)\lambda} \right]$$

Note that, $z_n(t)$ corresponds to $\frac{1}{n}S(n\alpha)$ of Theorem 2.2.4. Then the conditions of Averaging Principle are satisfied so that the condition

$$\sup_{0 \le t \le T} |z_n(t) - z(t)| \xrightarrow{\text{P}} 0$$

is also satisfied, where, $z(t)$ is the solution of differential equation

$$dz(t) = m(z(t))^{-1}b(z(t)).$$

Since $z(t) = (r(t), y(t))$, we have,

$$dz(t) = [((1 - r(t) - y(t))\lambda\beta - N\mu)\,dt, ((1 - r(t) - y(t))\lambda(1 - \beta))]\,dt.$$

Note that, the individual components $r(t)$ and $y(t)$ satisfy

$$\begin{aligned} dr(t) &= [(1 - r(t) - y(t))\lambda\beta - N\mu]dt && (4.31) \\ dy(t) &= (1 - r(t) - y(t))\lambda(1 - \beta)dt && (4.32) \end{aligned}$$

with some initial condition $(r(0), y(0))$.

Let us denote $s(t) = r(t) + y(t)$. Adding equations (4.31) and (4.32)we obtain a differential equation for $s(t)$ such that

$$ds(t) = ((1 - s(t))\lambda - N\mu)dt$$

with the solution

$$s(t) = 1 + (s(0) - 1)e^{-\lambda t} - N\mu(1 - e^{-\lambda t})/\lambda.$$

Let $s(0) = 0$. Then substituting s(t) into the equations (4.31), (4.32)we can also find the exact solutions for $r(t), y(t)$ so that,

$$r(t) = -\beta e^{-\lambda t} + N\mu\beta t + \frac{N\mu}{\lambda}e^{-\lambda t} - N\mu t,$$

$$y(t) = -(1 - \beta)e^{-\lambda t} + N\mu(1 - \beta)t + \frac{N\mu(1 - \beta)e^{-\lambda t}}{\lambda}$$

Note that, these solutions are valid only in the initial interval when both $0 < r(t) < 1, 0 < y(t) < 1$. For large t, $y(t)$ becomes greater than 1, and $r(t)$ become negative thus having no relation to the interpretation of the model.

Now consider the behavior of the process on the interval $[t_{kh_n}, t_{(k+1)h_n})$, when inspection is provided at times kh_n where t_{kh_n} is the time of k'th inspection.

Let t_{nd} be the sequential times of jumps of $z_n(t)$ on the interval $[t_{kh_n}, t_{(k+1)h_n})$. The time between the sequential jumps $(t_{nd+1} - t_{nd})$, are exponentially distributed with rate $(n - nr_{nd} - ny_{nd})\lambda + n\mu N$ where $r_{nd} = r_n(t_{nd}), y_{nd} = y_n(t_{nd})$.

Note that from (4.30) and the exponential time between the sequential jumps, the process $z_n(t)$ is a Quasi-Birth-and-Death process on each interval $[t_{kh_n}, t_{(k+1)h_n})$. Now, taking $\delta_n = \frac{1}{n}$ we can write the following stochastic equations

$$\begin{aligned}(r_{nd+1}, y_{nd+1}) &= (r_{nd}, y_{nd}) + \xi_{nd}(r_{nd}, y_{nd})\delta_n \\ t_{nd+1} &= t_{nd} + \tau_{nd}(r_{nd}, y_{nd})\delta_n\end{aligned}$$

where given $r_{nd} = r$ and $y_{nd} = y$

$$\xi_{nd}(r, y) = \begin{cases} (-1, 0) & \textit{with probability} \ \frac{N\mu}{N\mu + (1-r-y)\lambda} \\ (1, 0) & \textit{with probability} \ \frac{(1-r-y)\lambda\beta}{N\mu + (1-r-y)\lambda} \\ (0, 1) & \textit{with probability} \ \frac{(1-r-y)\lambda(1-\beta)}{N\mu + (1-r-y)\lambda} \end{cases}$$

and $\tau_{nd}(r, y)$ has exponential distribution with rate $(1 - r - y)\lambda + N\mu$.

Then $z_n(t)$ forms an RPSM on the time interval $t_{kh_n} \leq t < t_{(k+1)h_n}$.

We can now check if the Averaging Principle holds. Let $\alpha = (r, y)$. Following from the notation of Theorem 2.2.4, we have $m_n(\alpha) = (1 - r - y)\lambda + N\mu$ and $b_n(\alpha)$ is the expectation of $\xi_{nd}(r, y)$ in the vector form so that

$$b_n(\alpha) = \left\{ \frac{(1-r-y)\lambda\beta - N\mu}{(1-r-y)\lambda + N\mu}, \frac{(1-r-y)\lambda(1-\beta)}{(1-r-y)\lambda + N\mu} \right\}.$$

Hence we have

$$m(r, y) = (1 - r - y)\lambda + N\mu$$

$$b(r, y) = \left\{ \frac{(1-r-y)\lambda\beta - N\mu}{(1-r-y)\lambda + N\mu}, \frac{(1-r-y)\lambda(1-\beta)}{(1-r-y)\lambda + N\mu} \right\}.$$

Note that, since y_{nk} are bounded the condition (3.36) holds. Then the conditions of Averaging principle holds hence the condition

$$\sup_{0 \le t \le T} |z_n(t) - z(t)| \xrightarrow{\text{P}} 0$$

is satisfied. The process $z_n(t)$ converges in probability to the process $z(t) = (r(t), y(t))$ which satisfy a system of differential equations such that

$$\frac{dz(t)}{dt} = m(z(t))^{-1} b(z(t)) \tag{4.33}$$

and

$$\frac{dz(t)}{dt} = ((1 - r(t) - y(t))\lambda\beta - N\mu, (1 - r(t) - y(t))\lambda(1-\beta)) \tag{4.34}$$

with some initial condition $(r_0, y_0) = z_0$.

At times kh_n, due to the inspection, there is an additional jump to the process $z_n(t)$. Let $z^+_{nk+1} = (r^+_{nk+1}, y^+_{nk+1})$ be the value of process $z(t)$ just after the k+1'st inspection and let $z^-_{nk+1} = (r^-_{nk+1}, y^-_{nk+1})$ be the value of process $z_n(t)$ just before the k+1'st inspection. Let also H be a hypergeometric random variable with parameters (c, j, m) such that

$$P(H(c, j, m) = i) = \frac{\begin{pmatrix} j \\ i \end{pmatrix} \begin{pmatrix} c - j \\ m - i \end{pmatrix}}{\begin{pmatrix} c \\ m \end{pmatrix}}.$$

We can now write the following stochastic equation:

$$z^+_{k+1} = (r^+_{k+1}, y^+_{k+1})$$

$$= (r^-_{k+1}, y^-_{k+1}) + \frac{1}{n}(H(n(1 - r^-_{k+1}), ny^-_{k+1}, m), -H(n(1 - r^-_{k+1}), ny^-_{k+1}, m)) \tag{4.35}$$

We take $h_n = a/n$. Note that for given $(r^-_{k+1}, y^-_{k+1}) = (r, y)$ as $n \to \infty$, using the result about asymptotic behavior of $H(n, ns, m)$ (section 4.2) we have:

$$E(H(n(1-r),ny,m),-H(n(1-r),ny,m)) \to (\frac{my}{1-r},-\frac{my}{1-r}).$$

Consider now the total increment of the process $z_n(t)$ on the interval $[t_{kh_n}, t_{(k+1)h_n}]$ taking into account the change on this interval without replacement and the change due to the replacement at the inspection point. Then, according to (4.34) we can represent the main parts of the increment for each component as follows:

$$r_n((k+1)h_n) \approx r_n(kh_n) + (1 - r_n(kh_n) - y_n(kh_n))\lambda\beta h_n - N\mu h_n$$

$$+\frac{1}{n}H\left(n(1-r_n(kh_n)), ny_n(kh_n), m\right)$$

and

$$y_n((k+1)h_n) \approx y_n(kh_n) + (1 - r_n(kh_n) - y_n(kh_n))\lambda(1-\beta)h_n$$

$$-\frac{1}{n}H\left(n(1-r_n(kh_n)), ny_n(kh_n), m\right).$$

As $h_n = a/n$, it can be seen from (4.35) that the individual components $r(t)$ and $y(t)$ satisfy the differential equations

$$\frac{dr(t)}{dt} = (1-r(t)-y(t))\lambda\beta - N\mu + \frac{a^{-1}my(t)}{1-r(t)} \tag{4.36}$$

and

$$\frac{dy(t)}{dt} = (1-r(t)-y(t))\lambda(1-\beta) - \frac{a^{-1}my(t)}{1-r(t)} \tag{4.37}$$

with an initial condition $z_0 = (r_0, y_0)$.

Let us denote $s(t) = r(t) + y(t)$. Note that we can now rewrite the equations (4.36) and (4.37) in the following form;

$$\frac{dr(t)}{dt} = (1-s(t))\lambda\beta - N\mu + \frac{a^{-1}my(t)}{1-s(t)+y(t)} \tag{4.38}$$

and

$$\frac{dy(t)}{dt} = (1 - s(t))\lambda(1 - \beta) - \frac{a^{-1}my(t)}{1 - s(t) + y(t)} \tag{4.39}$$

Then the sum of equations (4.36) and (4.37) reduces to

$$\frac{ds(t)}{dt} = (1 - s(t))\lambda - N\mu \tag{4.40}$$

with $s(0) = s_0 = r_0 + y_0$ and general solution

$$s(t) = 1 + (s(0) - 1)e^{-\lambda t} - \frac{N\mu(1 - e^{-\lambda t})}{\lambda}$$

Since by construction $0 \leq y(u) \leq 1, 0 \leq r(u) \leq 1$, and $0 \leq s(u) \leq 1$, the solution $s(t)$ in this form can be written on some interval $[0, t]$ only if for all $0 < u < t$, we have the following restrictions: $0 < y(u) < 1$, $0 < r(u) < 1$ and $0 < s(u) < 1$.

Without loss of generality, assume that initially all devices are 'good' so that $s(0) = 0$. Then, for $\lambda\beta > N\mu$ (due to the fact that (4.36) can not be negative at t=0) the exact solution of (4.40) is found as

$$s(t) = \frac{\lambda - N\mu}{\lambda}(1 - e^{-\lambda t}). \tag{4.41}$$

Since all the conditions are satisfied we can now use the results of Theorem 3.5.1.

To complete the proof of Theorem 4.4.1 we need to show that the left hand side of (4.29) satisfies the condition **S**. In our case $p(t, \theta) = \frac{y(t)}{1-r(t)}$. Unfortunately, we could not find the exact solutions of differential equations (4.36) and (4.37), hence could not determine $p(t, \theta)$, so we can not exactly determine the left side of (4.29). But, if these solutions can be found and they satisfy the condition of **S**, this will be sufficient to complete the proof.

The following three different methods are suggested to approximate the solutions for differential equations (4.36) and (4.37) and then to estimate the unknown parameters.

Numerical Calculations and Statistical Estimation:

Let y_{nk} be the number of failed devices in the sample at the time of the control ka/n. Assuming that β_0 is known, for one-dimensional parameter λ_0 consider the estimation of λ.

Consider the relation (4.29). Since the Averaging principle for $z_n(t)$ holds it follows that it also holds for for the sequences y_{nk} and r_{nk}. So the condition $h_n \sum_{k=1}^{[T/h_n]} y_{nk} \rightarrow \int_0^T g(p(t,\theta))dt$ is satisfied. We will now give the expressions for functions $g(p(t,\theta))$ and $p(t,\theta)$.

At the time of k'th inspection we take a sample from $ny_n(kh_n)$, the devices which are not in repair (not being repaired or waiting in queue to be repaired). Given $r_n(t) = r$ and $y_n(t) = y$ the number of failed devices in the sample we observe is a Hypergeometric random variable, with parameters $n(1-r)$, ny and m, that is $H(n(1-r), ny, m)$.

As asymptotically

$$E(H(n(1-r), ny, m)) \rightarrow \frac{my}{1-r}$$

in this case $p(t) = \frac{y(t)}{1-r(t)}$. Taking $g(p(t)) = mp(t)$, we can re-write the equation (4.29) as follows.

$$\frac{1}{a}\int_0^T \frac{my(t)}{1-r(t)}dt = \frac{1}{n}\sum_{k=1}^{[Tn/a]} y_{nk} \tag{4.42}$$

The right hand side of the equation (4.42) is calculated by observations. For each fixed λ, β we can find a numerical solution of the system together with the left hand side of (4.36) using the recurrent procedure below.

Consider the system of equations (4.36) and (4.37) and rewrite the system in the following form:

$$dr(t) = A(r(t), y(t), \lambda, \beta)dt$$
$$dy(t) = B(r(t), y(t), \lambda, \beta)dt.$$

Let us denote

$$C(t, \lambda, \beta) = \frac{1}{a}\int_0^t \frac{my(t)}{1-r(t)}dt$$

so that the equation (4.42) can be written as

$$C(T,\lambda,\beta) = \frac{1}{n}\sum_{k=0}^{[T/h_n]} y_{nk}. \tag{4.43}$$

Also let $y(0) = 0$, $r(0) = 0$ and $C_0(0) = 0$. For some value of δ we can now construct recurrent sequences:

$$\begin{aligned} r_{k+1} &= r_k + A(r_k, y_k, \lambda, \beta)\delta, & (4.44)\\ y_{k+1} &= y_k + B(r_k, y_k, \lambda, \beta)\delta, & (4.45)\\ C_{k+1} &= C_k + \frac{m y_k}{a(1-r_k)}\delta. & (4.46) \end{aligned}$$

The sequence of r_k at points $k\delta$ gives a numerical solution for $r(t)$ (and the sequence of y_k at points $k\delta$ gives a numerical solution for $y(t)$) and $C_{T/\delta}$ is the approximation of the numerical value of $C(T,\lambda,\beta)$.

If β_0 is known and λ_0 is unknown, initially a value of λ and a default lag l is chosen. Keeping β fixed to its known value, in order to make left and right hand sides of equation (4.43) equal, the value of λ is changed accordingly and the best λ value to make this equation closest is chosen as a solution. Denote this value of λ by $\hat{\lambda}$.

Note that since $C(T,\lambda,\beta)$ is a strictly monotonically increasing function, it satisfies the condition **S**.

Then the unknown parameter λ_0 can be numerically calculated using the equation (4.43) and $\hat{\lambda}$ is the consistent estimator of λ_0.

Analytical Approximation and Statistical Estimation:

The behavior of numerical solutions suggests , that both $r(t)$ and $y(t)$ as well as the division $p(t) = \frac{y(t)}{1-r(t)}$ shows a behavior in the form of $C_0+C_1e^{-\lambda t}+C_2e^{-mt/a}$. Even though we cannot find the explicit solution we can try to approximate the numerical solution with some error component. This would give the opportunity to visualize the underlying function and having an analytical representation on hand would make it easier to solve for unknown parameters.

Assume that

$$y(t) = c_0 + c_1 e^{-\lambda t} - c_2 e^{-mt/a} \tag{4.47}$$

where $c_0 + c_1 = c_2$ and let $u = (\lambda - N\mu)/\lambda$. Substituting $r(t) = s(t) - y(t)$ we then get

$$p(t) = \frac{c_0 + c_1 e^{-\lambda t} - c_2 e^{-mt/a}}{(1-u+c_0) + (c_1+u)e^{-\lambda t} - c_2 e^{-mt/a}}$$

$$\ln p(t) = \ln\left(\frac{c_0 + c_1 e^{-\lambda t} - c_2 e^{-mt/a}}{(1-u+c_0) + (c_1+u)e^{-\lambda t} - c_2 e^{-mt/a}}\right)$$

$$\ln p(t) \cong \ln\left(\frac{c_0}{1-u+c_0}\right) + \left(\frac{c_1}{c_0} - \frac{c_1+u}{1-u+c_0}\right)e^{-\lambda t} - \left(\frac{c_2}{c_0} - \frac{c_2}{1-u+c_0}\right)e^{-mt/a}$$

and

$$p(t) \cong \left(\tfrac{c_0}{1-u+c_0}\right) + \left(\tfrac{c_0}{1-u+c_0}\right)\left(\tfrac{c_1}{c_0} - \tfrac{c_1+u}{1-u+c_0}\right)e^{-\lambda t} - \left(\tfrac{c_0}{1-u+c_0}\right)\left(\tfrac{c_2}{c_0} - \tfrac{c_2}{1-u+c_0}\right)e^{-mt/a}. \tag{4.48}$$

For $y(0) = 0$ we have $p(0) = 0$. For an approximation of $p(t) = C_0 + C_1 e^{-\lambda t} - C_2 e^{-mt/a}$ we should have $C_0 + C_1 = C_2$ where $C_0 = \frac{c_0}{(1-c_0+u)}$.

We now consider the differential equation (4.39),using (4.47) with this approximation:

$$\frac{dy(t)}{dt} = -c_1 \lambda e^{-\lambda t} + c_2 \frac{m}{a} e^{-mt/a} \tag{4.49}$$

and

$$\frac{dy(t)}{dt} = (1 - (u - ue^{-\lambda t}))\lambda(1-\beta) - \frac{m}{a}\left(C_0 + C_1 e^{-\lambda t} - C_2 e^{-mt/a}\right) \tag{4.50}$$

Equating the related parts of equations (4.49) and (4.50) the explicit solution of differential equation (4.50) can be found if,

$$C_0 = \frac{(1-u)\lambda(1-\beta)}{m/a}$$

$$c_0 = \frac{C_0(1-u)}{1-C_0} = \frac{\lambda(1-\beta)(1-u)^2}{(m/a)-(1-u)\lambda(1-\beta)}$$
$$C_1 = \frac{(\lambda - N\mu)(1-\beta) + (C_0 - c_0)\lambda}{(m/a-\lambda)}$$
$$C_0 + C_1 = c_0 + c_1 = c_2 = C_2$$

and

$$C_0 = \frac{a}{m}N\mu(1-\beta), \tag{4.51}$$
$$c_0 = \frac{\frac{a}{m}(N\mu)^2(1-\beta)}{\lambda(1-\frac{a}{m}N\mu(1-\beta))} \tag{4.52}$$

$$C_1 = \frac{(\lambda - N\mu)(1-\beta) + aN\mu(1-\beta)\lambda[1-\frac{N\mu}{\lambda(1-aN\mu(1-\beta)/m)}]/m}{m/a-\lambda}$$
$$C_2 = \frac{a}{m}N\mu(1-\beta) + \frac{(\lambda - N\mu)(1-\beta) + aN\mu(1-\beta)\lambda[1-\frac{N\mu}{\lambda(1-aN\mu(1-\beta)/m)}]/m}{m/a-\lambda}$$

The coefficients of $p(t)$ found by division are not same with the coefficients found by this differential equation solution, except the terms c_0 and C_0 as expected. The error consists of terms not included both in $y(t)$ and $p(t)$ and some remainder which is not included in the explicit solution. But both show good approximations to the numerical solutions of $y(t)$ and $p(t)$.

We can check if the solution found by the approximation captures, in some sense, the behavior of original process by considering the balance equations when $t \to \infty$. Note that $s(t) \to (1 - N\mu/\lambda)$ when $t \to \infty$. Denote $r_* = \lim r(t), y_* = \lim y(t)$ as $t \to \infty$. From (4.39) we get the equation

$$N\mu(1-\beta) = \frac{my_*}{a(1-\frac{\lambda - N\mu}{\lambda} + y_*)}$$

which has a solution as

$$y_* = \frac{\frac{a}{m}(N\mu)^2(1-\beta)}{\lambda(1-\frac{a}{m}N\mu(1-\beta))} \tag{4.53}$$

As $t \to \infty$, $\lim y(t) = c_0$ where c_0 was found by equation (4.52) which is equal to the value of the limit in (4.53). We also like to mention that these calculations are valid for $\frac{a}{m}N\mu(1-\beta) < 1$ along with the condition that $\lambda\beta > N\mu$.

Assume that the probability $\beta_0 = \beta$ is known and we need to estimate only the failure rate λ_0.

The left hand side of equation (4.42) is now changed and we consider the moments method equation

$$m\int_0^T p(t,\lambda)dt = h_n \sum_{k=1}^{[T/h_n]} y_{nk}, \tag{4.54}$$

where $p(t,\lambda) = C_0 + C_1 e^{-\lambda t} - C_2 e^{-mt/a}$.

Integrating $p(t,\lambda)$ in t on the interval [0,T], we can write (4.54) in the form

$$C_0 T - \frac{C_1}{\lambda} e^{-\lambda T} + \frac{C_2}{m/a} e^{-mT/a} + \frac{C_1}{\lambda} - \frac{C_2}{m/a} = \frac{h_n}{m} \sum_{k=1}^{[T/h_n]} y_{nk}. \tag{4.55}$$

Since the left hand side of equation (4.55) now satisfies the condition **S** we can calculate the consistent estimator for the unknown parameter λ_0 by solving equation (4.55) for λ .

Note that, since the $p(t)$ is just an approximation to a function which is represented in terms of numerical solutions of differential equations (4.36) and (4.37), it may not always be possible to find the solution of equation (4.55). But, solving analytical equation (4.55) is simpler then solving the equation (4.43) and needs less computing time.

We also would like to mention that, this approximation is better when u, C_0, C_1 and C_2 are smaller compared to 1. The steady state values will be equal in any case, but as u gets closer to 1, the transient values will be departing from each other. For visualization of this fact, we give the Figures (A.5) and (A.6) which show the numerical solution and analytical approximation of numerical solution when $N = 5, \mu = 0.1, T = 10, n = 1000$ and $h_n = 1/1000$ for two different cases of $\lambda_0 = 1, \beta_0 = 0.8, a = 1$ and $\lambda_0 = 2, \beta_0 = 0.8, a = 2$.

For $\lambda_0 = 1, \beta_0 = 0.8, a = 1$, we have

$$p(t) = 0.01 + 0.11661e^{-t} - 0.0221661e^{-10t}$$

and for $\lambda_0 = 2, \beta_0 = 0.8, a = 2$ we have

$$p(t) = 0.02 + 0.1099e^{-2t} - 0.1299e^{-5t}.$$

Approximation when the Number of Machines in Repair is Known:

If the number of machines in repair at the end of the inspection period (t=T) is known,we can rewrite (4.36) using equation (4.29) as follows:

$$\int_0^T dr(t) = \int_0^T \{(1 - r(t) - y(t))\lambda\beta - N\mu + a^{-1}(\frac{my(t)}{1 - r(t)})\}dt. \tag{4.56}$$

Assume that $r(0) = 0$. Then (4.56), using (4.29) reduces to

$$r(T) + N\mu(1-\beta)t|_0^T + (1 - \frac{N\mu}{\lambda})\beta e^{-\lambda t}|_0^T = \int_0^T a^{-1}(\frac{my(t)}{1 - r(t)})dt \tag{4.57}$$

so that,

$$r(T) + N\mu(1-\beta)T + (1 - \frac{N\mu}{\lambda})\beta e^{-\lambda T} - (1 - \frac{N\mu}{\lambda})\beta = \int_0^T a^{-1}(\frac{my(t)}{1 - r(t)})dt \tag{4.58}$$

and using (4.42)

$$r(T) + N\mu(1-\beta)T + (1 - \frac{N\mu}{\lambda})\beta e^{-\lambda T} - (1 - \frac{N\mu}{\lambda})\beta = \frac{1}{n}\sum_{k=1}^{[T/h_n]} y_{nk}. \tag{4.59}$$

Since the left hand side of equation (4.59) satisfies the condition of S. The consistent estimator for the unknown parameter λ_0 can be calculated solving the equation (4.58) for λ.

4.4.1 When Both Parameters are Unknown

If both λ_0 and β_0 are unknown, then $\theta = (\lambda, \beta)$ and we have to solve a system of two equations. Therefore, let us consider, for example, a two dimensional function $\varphi(y) = (y, y^2)$. Then

$$g^{(1)}(p) = EB(m, p) = mp,$$

$$g^{(2)}(p) = EB(m,p)^2 = mp + (m^2 - m)p^2.$$

Thus, a system of equations for two parameters (λ, β) has the form:

$$m \int_0^T p(t,\theta)dt = h_n \sum_{k=1}^{[T/h_n]} y_{nk} \tag{4.60}$$

$$\int_0^T \left(mp(t,\theta)) + (m^2 - m)p(t,\theta)^2\right) dt = h_n \sum_{k=1}^{[T/h_n]} y_{nk}^2, \tag{4.61}$$

Then we get a system of two nonlinear equations for two unknown parameters (λ, β) which can be solved in a similar way using the analytic and numeric technique discussed above.

Numerical Calculations and Statistical Estimation:

Consider the equations (4.60) and (4.61). Using the iterative method which was used for the single unknown parameter case, and equations (4.44),(4.45) we can find a numerical solution for the system of equations (4.60), (4.61) using the following:

Let us denote

$$C^*(t,\lambda,\beta) = \int_0^t mp(t,\theta)dt + \int_0^t (m^2 - m)p^2(t,\theta)$$

and,

$$C^*_{k+1} = C^*_k + C_k + \frac{1}{a}(m^2 - m)(\frac{y_k}{1 - r_k})^2\delta.$$

Then $C^*_{T/\delta}(T,\lambda,\beta)$ gives the approximation of the numeric value of $C^*(T,\lambda,\beta)$.

We can now estimate the unknown parameters λ_0, β_0 using the system of equations,

$$C(T,\lambda,\beta) = \frac{1}{n} \sum_{k=1}^{[T/h_n]} y_{nk} \tag{4.62}$$

$$C^*(T,\lambda,\beta) = \frac{1}{n} \sum_{k=1}^{[T/h_n]} y_{nk}^2. \tag{4.63}$$

In this case, we choose an initial value both for λ and β and adjust the values until both of the equations (4.62) and (4.63) have left and right hand side closest to each other and find the estimators of unknown parameters.

Analytical Approximation and Estimation:

The system of equations (4.60) and (4.61) is represented in terms of $p(t)$, which was approximated analytically with $p(t) = C_0 + C_1 e^{-\lambda t} - C_2 e^{-mt/a}$. Using this approximation, we have the system of equations,

$$\frac{m}{a}\int_0^T (C_0 + C_1 e^{-\lambda t} - C_2 e^{-(m/a)t})dt = \frac{1}{n}\sum_{k=1}^{[T/h_n]} y_{nk} \tag{4.64}$$

$$\frac{m}{a}\int_0^T (C_0 + C_1 e^{-\lambda t} - C_2 e^{-(m/a)t})dt +$$
$$\frac{1}{a}\int_0^T (m^2 - m)(C_0^2 + C_1^2 e^{-2\lambda t} + C_2^2 e^{-2mt/a} + C_0 C_1 e^{-\lambda t} - C_0 C_2 e^{-mt/a} - C_1 C_2 e^{-(\lambda + m/a)t})dt =$$
$$\frac{1}{n}\sum_{k=1}^{[T/h_n]} y_{nk}^2. \tag{4.65}$$

Then the estimators for the unknown parameters (λ_0, β_0) can be calculated by simultaneously solving the equations (4.64) and (4.65).

Approximation when the Number of Machines in Repair is Known:

If the number of failed machines in the repair is known, we can rewrite the system of equations (4.60), (4.61) using equation(4.59) in the following way:

$$r(T) + N\mu(1-\beta)T + (1 - \frac{N\mu}{\lambda})\beta e^{-\lambda T} - (1 - \frac{N\mu}{\lambda})\beta = \frac{1}{n}\sum_{k=1}^{[T/h_n]} y_{nk}. \tag{4.66}$$

$$\frac{1}{n}\sum_{k=1}^{[T/h_n]} y_{nk} + \frac{1}{a}\int_0^T (m^2 - m)p^2(t)dt = \frac{1}{n}\sum_{k=1}^{[T/h_n]} y_{nk}^2. \tag{4.67}$$

And finally, we can find the estimators for λ_0 and β_0 by solving the system of equations 4.66) and (4.67) for two unknowns, λ and β. Note that, the equation (4.67) can be solved using either the numerical method or the analytical approximation.

4.5 Simulation Results

The theoretical results obtained in the sections 4.1 through 4.4 are devoted to the asymptotic analysis, when $n \to \infty$. In this section we consider the simulation of the system for finite samples and estimate the unknown parameters using trajectory observations. For Model 1: A reliability model without replacement, we also provide the comparison between theoretical results and simulation results. For the other models, we provide the estimators and their properties obtained by simulations.

Model 1: A Reliability Model without Replacement

The system is simulated for the failure rate $\lambda_0 = 0.1$ and the repair rate $\mu = 0.3$ and different values of n, m, h_n and total observation time T. Consider the case when μ is known and λ is unknown, and we want to estimate λ. We find the estimator solving the moments method equation obtained in Section 4.1, which is

$$\frac{m\lambda}{a(\lambda+\mu)}\left(T - \frac{e^{-(\lambda+\mu)T}}{\lambda+\mu} + \frac{1}{\lambda+\mu}\right) = h_n \sum_{k=1}^{[T/h_n]} y_{nk}$$

for λ.

The summation on the right hand side of (4.7) is obtained by observations, and the left hand side was calculated using the function $s(t)$, which is given by (4.10).

The estimated values of λ_0 and the bias $|\hat{\lambda} - \lambda_0|$ for $T = 10$ using 5 run results and different values of n, m, ah_n are given in Table (4.1).

n	m	ah_n	$\hat{\lambda}$	$\lvert\hat{\lambda} - \lambda_0\rvert$
10000	500	0.001	0.09926	0.00074
5000	500	0.001	0.09855	0.0145
1000	100	0.001	0.097775	0.002225
100	10	0.01	0.095375	0.04625

Table 4.1: Model 1: Estimated values for λ_0

We will now study the effect of values of the system parameters n, h_n and a

to the value of the estimator.

When $n \to \infty$, we have the relations (3.34) and (3.39) valid. For simulation, if n is not large enough, we can see from the results that the mean square error increases significantly. The mean square error for $n = 10000$ was calculated as $3.54 * 10^{-6}$ while for $n = 100$ it was calculated as 0.00803.

Note that theoretical calculations were based on the assumptions that, $n \to \infty$, $h_n \to 0$ and $nh_n \to \infty$. It is obvious from the weak convergence $(\hat{\lambda} - 0.1)/\sqrt{h_n} \overset{w}{\Rightarrow} N(0, aR^2/B^2)$, that $\hat{\lambda} - 0.1$ asymptotically has normal distribution with mean zero and variance $h_n\ aR^2/B^2$. The asymptotic variance of the difference $\hat{\lambda} - 0.1$ depends on the values of h_n, a, T and m.

According to this relation, as h_n decreases the variance decreases as well. For practical reasoning, while a is constant, if h_n is selected too small, then the time between the inspections will be small that it will be close to continuously observing the system. If ah_n is constant, then we can choose the value of h_n as small as possible, but this would force a to be very large so the resulting variance would not reduce. Also, for simulation choosing h_n too small forces nh_n be too small which violates the assumption of our theoretical calculations that nh_n is large.

For $n = 5000$, $T = 3.28$, $a = 1$ and $m = 100$ the variances of the limiting distribution of the estimators and estimated values of λ for different values of h_n are given below.

h_n	variance	$\hat{\lambda}$	$(\hat{\lambda} - 0.1)$
0.001	0.009934*10^{-5}	0.0995	0.0005
0.01	0.496718*10^{-5}	0.102	0.002
0.05	0.0.99344*10^{-5}	0.107	0.007

Table 4.2: Model 1: Effect of h_n to the estimator

Keeping other parameters constant, the result is that the smallest h_n choice gave the best estimator in terms of bias, which was expected.

The results of simulation are very much in agreement with our theoretical results. Even for $n = 100$, which is very small compared to other chosen n values, the bias is 0.0426. This result would improve if we would use more then five runs for estimation. For other cases, the quality of the estimators even for single runs are very good and this is reflected in the five run results.

Model 2: Reliability Model with Replacement:

The model was simulated for $n = 1000$, $h_n = 1/n$, and different values of $T, a, \lambda_0,\ m$ and β_0. In the Figure (A) for $\lambda_0 = 0.5, \beta_0 = 1, a = 1, T = 5, m = 10$ we give the trajectory observations and $s(t, \lambda)$ which was calculated as

$$s(t,\lambda) = \frac{\lambda}{\lambda + m\beta/a}(1 - e^{-(\lambda+m\beta/a)t}).$$

With the given values of parameters $s(t) = 0.0476(1 - e^{-10.5\ t})$. Note that, this is the case when replacement is perfect, i.e. all failed devices in the sample are replaced immediately.

In Figure (A) we give the trajectory observations and $s(t, \lambda)$ for $\lambda_0 = 0.8,\ \beta_0 = 0.6,\ a = 2$ and $T = 2$. In this case $s(t) = 0.347(1 - e^{-2.3\ t})$.

The function $s(t, \lambda)$ behaves very much like the trajectory of the process for both of the cases. They are also good examples of why we can study the transient conditions. The function captures the behavior of the trajectory even for small t.

Consider the case when β_0 is known and we want to estimate the unknown parameter λ_0. We now use the equation

$$\frac{m\lambda}{\lambda + m\beta/a}\left(T - \frac{1 - e^{-(\lambda+m\beta/a)T}}{\lambda + m\beta/a}\right) = h_n \sum_{k=1}^{[T/h_n]} y_{nk},$$

where y_{nk} are our observations, to find the estimator.

In Tables (B.1), (B.2) and (B.3) the estimated values of λ_0, bias $|\hat{\lambda} - \lambda_0|$, relative error $|\hat{\lambda} - \lambda_0|/\lambda_0$ and Mean Square Error (MSE) using 5 runs for the reliability model with replacement grouped according to different values of T are given.

For all T values, the maximum bias obtained is 0.2639, maximum relative error is 0.0034133779 and maximum mean square error value obtained is 0.00264134, all of which was obtained when $T = 1$. When $T = 10$ the bias is less then 0.014, and when $T = 4$ the bias is less then 0.02. From the tables we see that, as T increases, the quality of the estimator increases also.

Since h_n was fixed to the value $1/n = 1/1000$, the choice of a did not affect the results of simulation significantly. Even though increasing a from 1 to 2 resulted in higher bias and relative error values for the same parameter sets, in several cases the situation is otherwise, i.e. bias and relative error decreases as it can be seen when $T = 1, \lambda_0 = 1, \beta_0 = 0.6$

In tables (B.4), (B.5) and (B.6) we give the simulation results when λ_0 is known and we want to estimate β_0.

All simulation results agree with our theoretical results.

Model 3: Reliability Model with N Repairmen.

The model was simulated for different values of T, λ, μ, N and m when $n = 1000$, $h_n = 1/1000$ and $a = 1$. For $\lambda_0 = 1.5$, $\mu_0 = 0.2$, $m = 5$ and $N = 5$, Figure (A) shows the trajectory observations and the function $s(t, \lambda)$. It was obtained in section 4.3 that

$$s(t, \lambda) = (1 - N\mu/\lambda)(1 - e^{-\lambda t}).$$

With the given values of parameters we now have $s(t) = \frac{1}{3}(1 - e^{-1.5t})$

Figure (A.4) shows $s(t, \lambda)$ and trajectory observations for $m = 10$, $\lambda = 0.8$, $T = 5, N = 5, \mu = 0.1$. In this case $s(t) = 0.375(1 - e^{-0.8t})$

The function $s(t)$ very much captures the behavior of the trajectory for both of the cases.

We consider the case when μ_0 is known and we want to estimate the unknown parameter λ_0.

We use the equation

$$m(1 - N\mu/\lambda)\Big(\lambda T - 1 + e^{-\lambda T}\Big)/\lambda = h_n \sum_{k=1}^{[T/h_n]} y_{nk}$$

to estimate the unknown parameter and y_{nk} are the observations.

Table (B.7) summarizes the results of estimation for different cases of parameters T, λ_0, μ_0, N.

The condition $\lambda > N\mu$ is satisfied by all parameter sets chosen for each realization. But the ratio $N\mu/\lambda$ is not the same for all cases. We see from our estimation function that as this ratio gets closer to 1, the estimation function gets closer to 0, independent of observations, making estimation procedure less effective. Due to this fact, the worst estimator obtained was with $\lambda_0 = 1.5, \mu_0 = 0.2, N = 5$ and $T = 1$ having the ratio $N\mu/\lambda = 2/3$ which is the biggest ratio in the simulation.

The simulation results supported our theoretical results indicating a bias for estimators less than 0.029, relative error less then 0.019 and mean square error less than 0.0045.

Model 4: Reliability Model with Probabilistic Chance of Repair

The simulation was performed for different values of λ_0, β_0 and a for $h_n = 1/n$, $n = 1000$, $\mu = 0,1$, $N = 5$, and $T = 10$. The Figure (A.8) shows the simulated values of $y_k/(1 - r_k)$ for $\lambda_0 = 1$, $\beta_0 = 0,8$ and $a = 1$ and Figure (A.7) shows the simulated values of observations y_{nk}. Due to great variation in the observations, total time of inspection is chosen as $T = 10$.

Single Unknown Parameter Case.

Consider the case when β_0 is known and we want to estimate the value of the parameter λ_0.

Numerical Solution: In the Figure (A.9), we give the simulated values $y_k/(1 - r_k)$ for $\lambda_0 = 1$, $\beta_0 = 0.8$ and the numerical approximation $C(t, \lambda, \beta)$. For illustration we also provide the figures for simulated values of y_k and the numerical solution of $y(t)$ and simulated values of r_k and numerical solution of $r(t)$ in the Figures (A.10) and (A.11).

It can be seen from Figures (A.9), (A.10) and (A.11) that the numerical solution captures the behavior of trajectories for all cases.

Consider the case when the parameter β_0 is known and we want to estimate the parameter λ_0. For estimation we use the equation

$$C(T, \lambda, \beta) = \frac{1}{n} \sum_{k=0}^{[T/h_n]} y_{nk}$$

where y_{nk} are observations.

Tables (B.8) and (B.9) give the results of estimation of λ_0 for 5 and 10 runs respectively.

Analytical Approximation: The Figure (A.12) shows the simulated values $y_k/(1 - r_k)$ for $\lambda = 1$, $\beta = 0.8$, $a = 1$ and the analytical approximation $p(t)$. $p(t)$ was obtained in the section 4.4 as

$$p(t) = C_0 + C_1 e^{-\lambda t} - C_2 e^{-mt/a}.$$

With the given values of parameters we have

$$p(t) = 0.01 + 0.011661 e^{-t} - 0.021661 e^{-10t}.$$

We also like to mention that with the same values of parameters we have

$$s(t) = 0.5(1 - e^{-t}),$$

$$y(t) = 0.00505 + 0.01661 e^{-t} - 0.02166 e^{-10t},$$

and

$$r(t) = 0.49495 - 0.51661 e^{-t} + 0.02166 e^{-10t}.$$

We also give the simulated values of the number of failed devices which are not in repair y_k and the number of failed devices which are in repair, r_k and their analytical approximation of numerical solution, $y(t)$ and $r(t)$ for visualization in Figures (A.13) and (A.14).

Consider the case when the unknown parameter is λ_0

In the estimation we use the equation

$$C_0T - \frac{C_1}{\lambda}e^{-\lambda T} + \frac{C_2}{m/a}e^{-mT/a} + \frac{C_1}{\lambda} - \frac{C_2}{m/a} = \frac{h_n}{m}\sum_{k=1}^{[T/h_n]} y_{nk}.$$

Tables (B.10) and (B.11) summarize the results of the estimation of λ_0 for 5 and 10 runs respectively.

Estimation when r(T) is Known:

In the case when r(T) is known we give the results of estimation of λ_0 in the Tables (B.12) and (B.13). We use the numerical solution to find the right hand side of equation

$$r(T) + N\mu(1-\beta)T + (1-\frac{N\mu}{\lambda})\beta e^{-\lambda T} - (1-\frac{N\mu}{\lambda})\beta = \frac{1}{n}\sum_{k=1}^{[T/h_n]} y_{nk}.$$

When λ_0 is known, the estimation results of the β_0 by simulation is given in the tables (B.14), (B.15) and (B.16) . Note that, h_n is chosen as $1/n$, $n = 1000$, $\mu = 0.1$, $N = 5$, and $T = 10$.

When both Parameters are Unknown:

For 5 runs, we give the simulation results in the Tables (B.17), (B.18) and (B.19), for $n = 1000$, $h_n = 1/1000$ and $T = 10$ and different values of λ_0, β_0 and using three approaches presented in section (4.4).

We see from Figures (A.9), (A.10) and (A.11) that, numerical solutions of functions $p(t)$, $y(t)$ and $r(t)$ give very good approximation to the simulated values of trajectories. Hence, our simulation results agree with our theoretical calculations.

For 10 runs, maximum relative error using numerical solution is less than 0.19, using analytical approximation it is less than 0.082 and when $r(T)$ is known it is less than 0.074. With two exceptions, which belong to numerical solution, relative error of the estimator for λ_0 is less then 0.082 for all methods. Even though, we were not able to represent the function $p(t)$ exactly, we still can use the approach for parameter estimation.

As was expected, analytical approximation to numerical solution did not provide a solution of equation (4.29), so the estimator could not be found for several cases using analytical approximation when both parameters are unknown. For a single unknown parameter case, there was only one set of observations yielding the same no solution situation when λ_0 is the unknown parameter.

For comparison, the results given in tables are provided so that we can find a solution of equation (4.55) and same set of observations are used to estimate the parameters with each method for respective values of λ_0, β_0 and a. We also like to mention that, when analytical approximation does not provide a solution, numerical solution provides an estimator which has very high bias and absolute error which would not be appropriate.

Due to the fact that, the analytical approximation undervalues the numerical solution for our parameter sets (in different degrees for different values of parameters for the chosen sets), the estimators obtained by analytical approximation are always smaller then the ones obtained by numeric solution. But, since in most of the cases numerical solution tends to overestimate the parameter, the analytical approximation gave better results then numerical solution case.

When only a single parameter is unknown, for several estimators of λ_0 the absolute error and mean square error were relatively high. This is due to the fact that observations have great variance and we were not able to solve the differential equations but used the approximations. These errors are considerably reduced when $r(T)$ is known.

When we analyze the equation (4.55) we see that the function on the left hand side is linear with respect to β and non linear with respect to λ. That is why it is not a surprise that when β_0 is unknown, the estimation results are better then when λ_0 is unknown. All methods results are quite satisfactory with very small bias, absolute errors and mean square errors for the estimators of β_0. The simulation results of estimation when λ_0 is known and β_0 is unknown for each of three methods explained are given in Tables (B.14), (B.15) and (B.16).

The value of the information of $r(T)$ even more visible when both parameters are unknown. Generally since we have two unknowns and two equations it was expected to have higher error values of the estimators. The fact that both λ_0 and β_0 are unknown did not effect the estimators for β_0 and the errors are very small.

Chapter 5

Conclusions

In this thesis, we consider an approach to statistical parameter estimation in stochastic systems. Depending on the nature of the problem, an estimator is represented as one of the following ways:

1. As a solution of stochastic equation $f_n(\theta) = 0$, with an additive type of function constructed on the trajectory of the observed system,

2. As the extreme point (set) of a random function $F_n(\theta)$ constructed on the trajectory of the observed system.

In order to be able to analyze the asymptotic behavior of the estimator constructed as a solution of stochastic equation, such as moments type estimators, we give the results about the solutions of stochastic equations. We present the result that, if the functions $f_n(\theta)$ uniformly converges to a limiting function $f_0(\theta)$ such that $f_0(\theta)$ satisfies the condition S, then the solution set of equation $f_n(\theta) = 0$ converges in probability to the solution of equation $f_0(\theta) = 0$. We also consider the asymptotic normality of such estimators and show that the normed deviation weakly converges to a random variable which is the solution of a limiting equation.

We also consider the asymptotic behavior of extreme points of random functions $F_n(\theta)$. We give the result that if $F_n(\theta)$ uniformly converges to some limiting function $F_0(\theta)$ where, $F_0(\theta)$ satisfies the condition $S2$, then the extreme points of $F_n(\theta)$ converges in probability to the extreme point of $F_0(\theta)$. We also give the result of weak convergence of normed deviation.

Definition and properties of Switching Processes and a subclass of Switching Processes are also given. We give the results of averaging principle and diffusion approximation for Recurrent Processes of semi-Markov (RPSM) type.

Using these results, we study the asymptotic properties of estimators constructed by trajectory observations of stochastic systems. For moments type, maximum likelihood and least squares method estimators, we show that the estimators are consistent and asymptotically normal.

Combining the results of solutions of stochastic equations and the limit theorems for Switching Processes, we further investigate the properties of the moments method type estimators when estimators are constructed on the trajectory of a Switching Process.

The approach of representing the estimator by trajectory observations is illustrated on the applications of four different but related reliability models. For each model, we represent the trajectory of the process as a switching process and prove that the system process converges to the solution of a differential equation.

Using our previous results, we estimate the unknown parameters. Commonly in all models, we consider a large number (n) of devices which are subject to independent random failures with failure rate of each device given as λ. All systems are inspected at the sequential times t_k so that $t_1 < t_2 < t_3$... on the time interval [0, T]. The inspection is provided instantaneously.

Model 1 is a reliability model without any disturbance to the system. In addition to the failures, each device has a repair rate μ. The inspection is performed as follows: At the time of inspection a sample is selected at random and the number of failed devices in that sample is observed. Without any replacement the sample is returned back. We estimate the unknown parameter λ and prove the consistency and asymptotic normality of the estimator.

In the second model we consider a reliability model with replacement. In this case no repair is available. At the inspection times, a sample is selected in random and the number of failed devices in that sample is observed. The observed (detected) failed devices are immediately replaced by new ones, and sample is returned back. However, the inspection is imperfect so that, a failed device in the sample is observed (or detected) only with probability β. We estimate the unknown parameters (for a single unknown case) λ and β and prove the consistency of the estimators.

The case where we have N repairmen each with large repair rate in a reliability system is considered in third model and we have no replacement. At the inspection time, we take a sample at random and observe the failed devices in that sample. After the inspection, we return the sample back immediately. In this case, we estimate the unknown parameter λ and prove the consistency of the estimator.

The final model has a more general and complicated structure than the other ones, including N repairmen each having large repair rate. When a device fails, there is a chance that it will not be sent to the repair immediately. In this case, when a device fails, it is sent to repair process only with probability β, so that a failed device will be considered as working and stay in the workplace with probability $1-\beta$. If at least one repairman is available, the device which is sent to repair starts being repaired and otherwise, it waits in queue (FCFS) for its turn. At the times of inspections, a sample is selected at random from the devices which are not in repair process and the number of failed devices are observed in that sample. The observed failed devices are sent to repair process immediately and the remaining of the sample is returned back to working. We estimate the unknown parameters λ and β for single unknown parameter and when both parameters are unknown.

For this model, we were unable to find a solution to system of differential equations which represent the behavior of the system. We estimated the unknown parameters with three different approaches: using numerical solution, using an analytical approximation for numerical solution and finally assuming that the number of devices which are in repair at the end of the inspection period (at time T) is known.

We also would like to mention that, this model could be also analyzed by assuming that the number of failed devices in repair process is always known. This assumption is reasonable since we assume that the failed devices are not being repaired in their working place, but sent somewhere else. This information is very much likely to increase the quality of the estimators.

We could extend our final model for future studies in several ways. Since the number of devices in repair process is very large, it is possible to replace the failed devices which are observed in the inspections. We can also consider the case when the devices have more than one type of failures, or when the time between the inspections is not constant, possibly random, depending on the previous value of the trajectory.

The simulation results are very much in agreement with the theoretical results. We successfully proved the convergence of the trajectories of systems to a limiting deterministic function. The calculated functions representing the system process behave very much like the trajectories of simulated systems.

Although we studied only the consistency of the moments method estimators, except for the first model, it is possible to study the asymptotic normality using our previous results. It is also possible to consider maximum likelihood and least squares method estimators for Switching Processes for future works.

Bibliography

[1] S.K. Acharya, "On normal approximation for maximum likelihood estimation from single server queues", *Queueing Systems*, **31**,207-216,1999.

[2] V.V Anisimov, "Limit theorems for Switching stochastic processes and their applications", *Cybernetics*, **14(6)**, 917-929, 1978.

[3] V.V Anisimov, "On estimation of the parameters of switched Poisson processes",*Theory Probability and Mathematical Statistics*,**31**, 1-11, 1985.

[4] V.V Anisimov, "Averaging principle for switching recurrent sequences", *Theory Probability and Mathematical Statistics*, **45**, 1-8, 1991.

[5] V.V Anisimov, "Averaging principle for the processes with fast switchings", *Random Operations and Stochastic Equations*, **1(2)**, 151-161, 1993.

[6] , V.V Anisimov, "Limit theorems for processes with semi-Markov switchings and their applications", *Random Operations and Stochastic Equations*, **2(4)**, 335-354, 1994.

[7] V.V Anisimov, "Switching processes: Averaging principle, diffusion approximation and applications", *Acta Applicandae Mathematicae*, **40**, 95-141, 1995.

[8] V.V Anisimov, "Extremal problems for random functions and applications in statistical parameter estimation", *Theory of Stochastic Processes*, **2(18)(1-2)**, 28-43, 1996.

[9] V.V Anisimov, "Asymptotic behavior of trajectories of extreme sets of random functions and stochastic estimation", *Theory Probability & Mathematical Statistics*, **55**, 1-12, 1997.

[10] V.V Anisimov, "Diffusion approximation in overloaded switching queueing models", *Queueing Systems*, **40**, 141-180, 2002.

[11] V.V Anisimov and A.O. Aliev, "Limit theorems for recurrent processes of semi-Markov type", *Theory Probability and Mathematical Statistics*, **41**, 7-13, 1990.

[12] V.V Anisimov and A. Orazklychev, "Asymptotic parameter estimation of recurrent processes of semi-Markov type", *Theory Probability and Mathematical Statistics*, **41**, 1-13, 1993.

[13] V.V Anisimov and G. Guleryuz, " Statistical estimation in switching systems in transient conditions and applications to reliability models ", *Cybernetics and System Analysis*, **40(3)**, (to appear), 2004.

[14] V.V. Anisimov and H. Kaibah, "Asymptotic behavior of solutions of stochastic equations and applications in statistics", *İstatistik, Journal of Turkish Statistical Association*, **1(3)**, 31-42, 1998.

[15] V.V. Anisimov and H. Kaibah, "The asymptotic properties of estimates of the parameters of non-linear time series ", *Cybernetics and System Analysis*, **36(2)**, 203-211, 2000.

[16] V.V. Anisimov and G. Pflug, "Z-theorems: Limits of stochastic equations", *Bernoulli Journal*, **6(5)**, 917-938, 2000.

[17] V.V. Anisimov and A.V. Seilhamer, "Asymptotic properties of extremal sets of random fields", *Theory Probability & Mathematical Statistics*, **51**, 1-9, 1994.

[18] H. Ascher and H. Feingold, Repairable system reliability, modelling, inference, misconceptions and their causes, Marcel Dekker, Washington, DC, 1994.

[19] O.E. Barndorff-Nielsen and M. Sørensen, "A review of some aspects of asymptotic likelihood theory for stochastic processes", *International Statistical Review*, **62(1)**, 133-165, 1994.

[20] Bo M. Bibby and M. Sørensen, "Martingale estimation functions for discretely observed diffusion processes", *Bernoulli Journal*, **1(2)**, 17-39, 1995.

[21] D.I Cho and M. Parlar, "A survey of maintenance models for multi-unit systems", *European Journal of Operations Research*, **51**, 1-23, 1991.

[22] Y.J Choi and N.C. Severo, "An approximation for the maximum likelihood estimator of the infection rate in the simple stochastic epidemic", *Biometrika*, **75**, 392-394, 1988.

[23] J. Dupacova and R. Wets, "Asymptotic behavior of statistical estimator and optimal solutions of stochastic optimization problems", *Annals of Statistics*, **16(4)**, 1517-1549, 1988.

[24] I.I. Ezov and A.V Skorokhod, "Markov processes which are homogeneous in the second component II", *Theory Probability & Applications*, **14**, 679-692, 1969.

[25] R.J. Giego and R. Hersh, "Random evolutions, Markov chains and system of partial differential equations", *Proceedings of National Academy of Science*, **62**, 305-308, 1969.

[26] B. Heidergott, "A weak derivative approach to optimization of threshold parameters in a multicomponent maintenance system", *Journal of Applied Probability*, **38**, 386-406, 2001.

[27] R. Hersh, "Random evolutions: survey of results and problems", *Rocky Mountain Journal of Mathematics*, 4(3), 443-475, 1974.

[28] I.A. Ibragimov and R.Z Khas'minskii, Statistical Estimation - Asymptotic Theory, Springer-Verlag, New-York, 1981.

[29] Y. M. Kaniovski and G.C. Pflug, "Limit theorems for stationary distributions of birth and death processes", *Communications in Statistics, Stochastic Models*, **15(1)**, 103-123, 1999.

[30] N. Keiding, "Maximum likelihood estimation in the birth-and-death processes", *The Annals of Statistics*, **3(2)**, 363-372, 1975.

[31] R.P. Kertz, "Random evolutions with underlying semi-Markov processes", *Publications Research Institute Mathematical Sciences*, **14(3)**, 589-614, 1978.

[32] R. Z. Khas'minskii, "About the averaging principle for ITO stochastic differential equations", *Kybernetika*, **4(3)**, 260-279, 1968.

[33] V.S. Korolyuk and A.V. Swishchuk, "Central limit theorem for semi-Markov random evolutions", *Ukranian Mathematics Journal*, **38**, 330-334, 1986.

[34] T. Kurtz, "A limit theorem for perturbed operator semigroups with applications to random evolutions", *Journal of Functional Analysis*, **12**, 55-67, 1973.

[35] Yu. A. Kutoyants, Parameter Estimation for Stochastic Processes, Heldermann Verlag, Berlin, 1984.

[36] R.L. Liptser and A.N. Shiryaev, Statistics of Random Processes, Springer, New-York, 1977.

[37] G. Pflug, "Asymptotic stochastic programs", *Technical Report, International Institute Applied System Analysis, Austria*, 1992.

[38] G.C. Pflug, "Stochastic programs and statistical data, stochastic programming. State of art", *Annals of Operations Research*, **85**, 59-78, 1998.

[39] M.J. Phelan, "A quasi likelihood for integral data on birth-and-death on flows", *Stochastic Processes and Their Applications*, **53**, 379-392, 1994.

[40] M.J. Phelan, "Transient tracers and birth-and-death on flows: parametric estimation of rates of drift, injection and decay", *Journal of Statistical Planning and Inference*, **51**, 19-40, 1996.

[41] M.J. Phelan, "Asymptotic likelihood estimation from birth and death on a flow", *The Annals of Statistics*, **24(3)**, 1161-1184, 1996.

[42] M. Pinsky, "Multiplicative operator functionals of a Markov process", *Bull. AMS*, **77**, 377-380, 1971.

[43] B.L.S Prakasa Rao, Asymptotic Theory of Stochastic Inference, John Wiley & Sons, 1987.

[44] S. Ross, Stochastic Processes, Wiley, New-York, 1996.

[45] P.L.C Saldanha, E.A. de Simone and P.F. Frutuoso e Melo, "An application of non-homogenous Poisson point processes to the reliability analysis of service water pumps", *Nuclear Engineering and Design*, **210**, 125-133, 2001.

[46] G. Salinetti and R. Wets, "Convergence of infima, especially stochastic infima", *Technical Report, University of Roma, La Sapienza*, 1986.

[47] A. Shapiro, "Asymptotic behavior of optimal solutions in stochastic programming", *Mathematics of Operations Research*, **18(4)**, 829-845, 1993.

[48] D. Shi, "Moment estimation for multivariate extreme value distribution in a nested logistic model", *Annals of Institute of Statistical Mathematics*, **51(2)**, 253-264, 1999.

[49] A.V. Skorokhod, "Limit theorems for random processes", *Theory Probability & Applications* , **1(3)**, 289-319, 1956.

[50] L.V. Volokh, "An estimate of parameters of random variables of exponential distributions", *Cybernetics and System Analysis*, 36(5), 786-790, 2000.

[51] J.C. Watkins, "A central limit problem in random evolutions", *Annals of Probability*, **12(2)**, 480-513, 1984.

[52] R. Watson and P. Yip, "A note on estimation of the infection rate", *Stochastic Processes and Their Applications*, **41**, 257-260, 1992.

[53] R.W. Wolff, "Problems of statistical inference for birth and death queueing models", *Operations Research*, **13(3)**, 343-357, 1965.

[54] A.I. Zeifman, "On the estimation of probabilities for birth and death processes", *Journal of Applied Probability*, **32**, 623-634, 1995.

[55] A.I. Zeifman, "Estimation of probabilities for some birth and death processes", *Journal of Mathematical Sciences*, **111(6)**, 3918-3921, 2002.

Appendix A

Figures

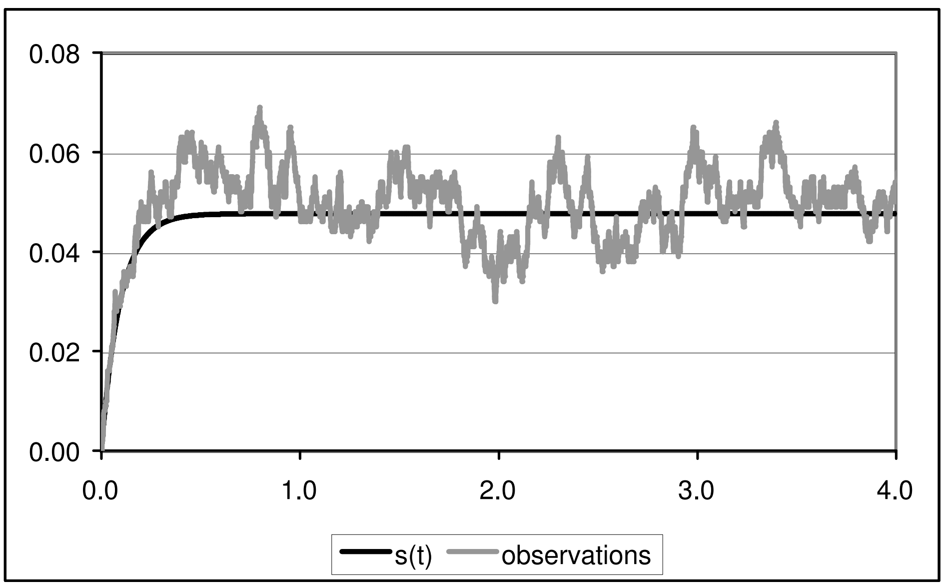

Figure A.1: Model 2: Simulation of trajectory of failed devices for reliability model with replacement ($\lambda_0 = 0.5,\ \beta_0 = 1, T = 5, m = 10$)

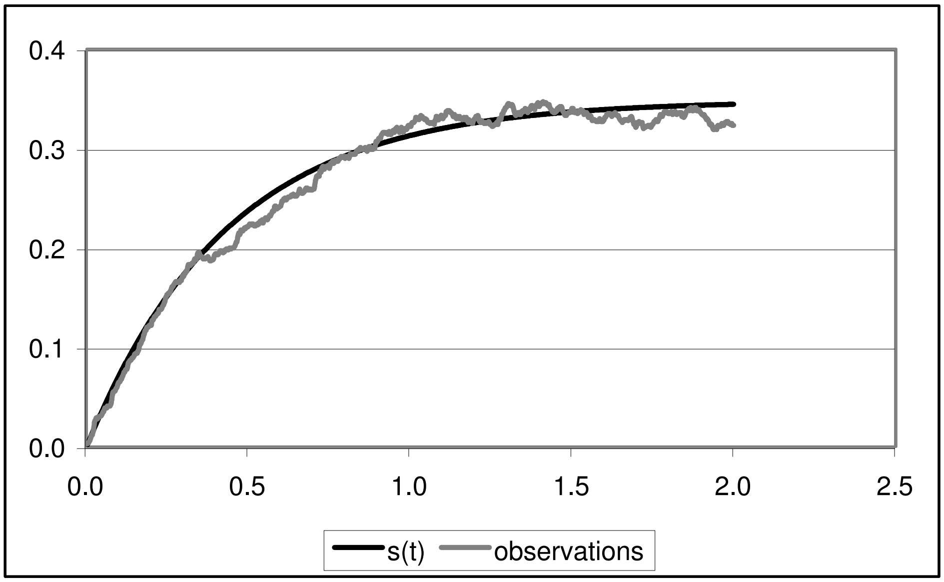

Figure A.2: Model2: Simulation of trajectory of failed devices for reliability model with replacement ($\lambda_0 = 0.8,\ \beta_0 = 0.6, a = 2,\ T = 2$)

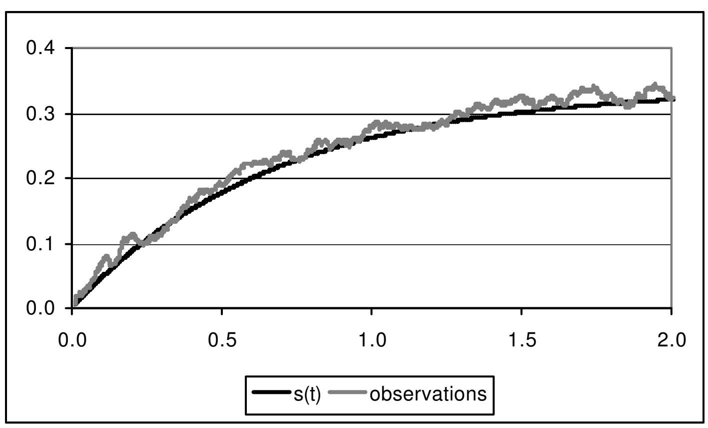

Figure A.3: Model3: Simulation of trajectory of failed devices for reliability model with N repairmen ($m = 5, \lambda_0 = 1.5, T = 2, N = 5,\ \mu_0 = 0.1$)

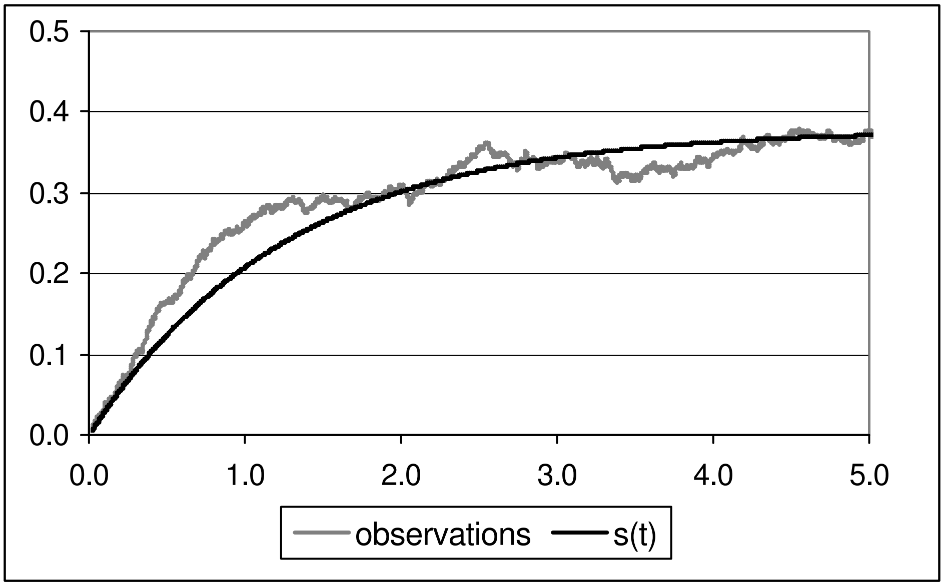

Figure A.4: Model 3: Simulation of trajectory of failed devices for reliability model with N repairmen ($m = 10, \lambda_0 = 0.8,\ T = 5, N = 5, \mu_0 = 0.1$)

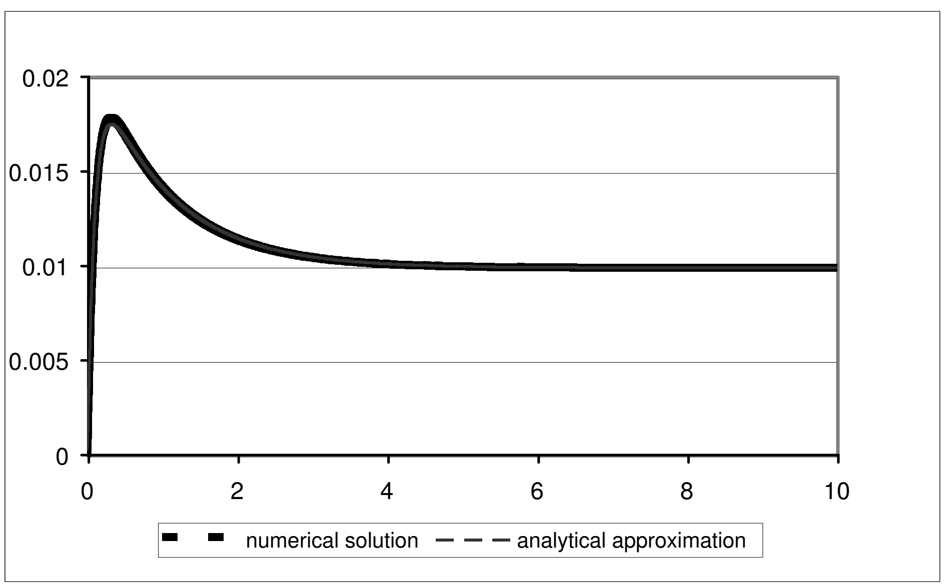

Figure A.5: Model 4: Numerical solution and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$.

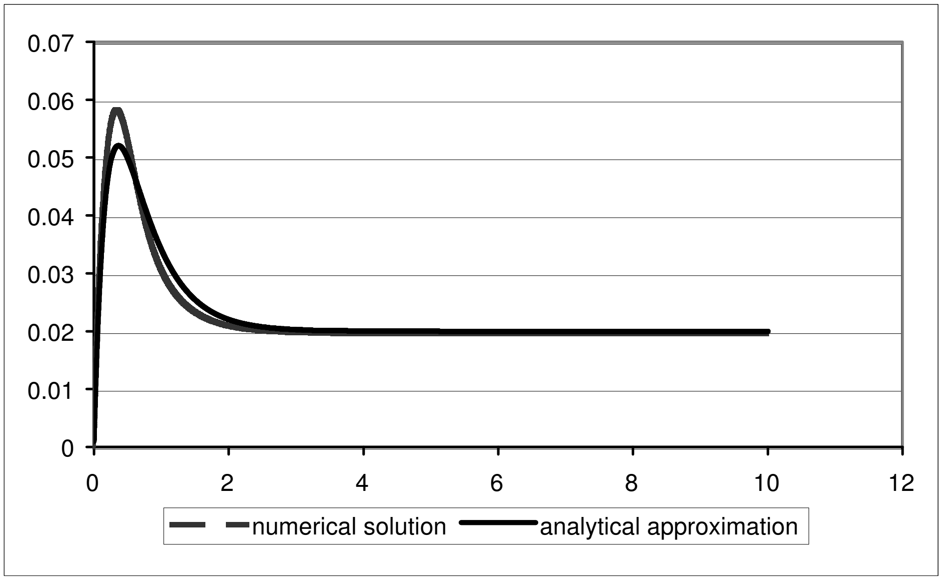

Figure A.6: Model 4: Numerical solution and analytical approximation of numeric solution when $\lambda_0 = 2$, $\beta_0 = 0.8$ and $a = 2$.

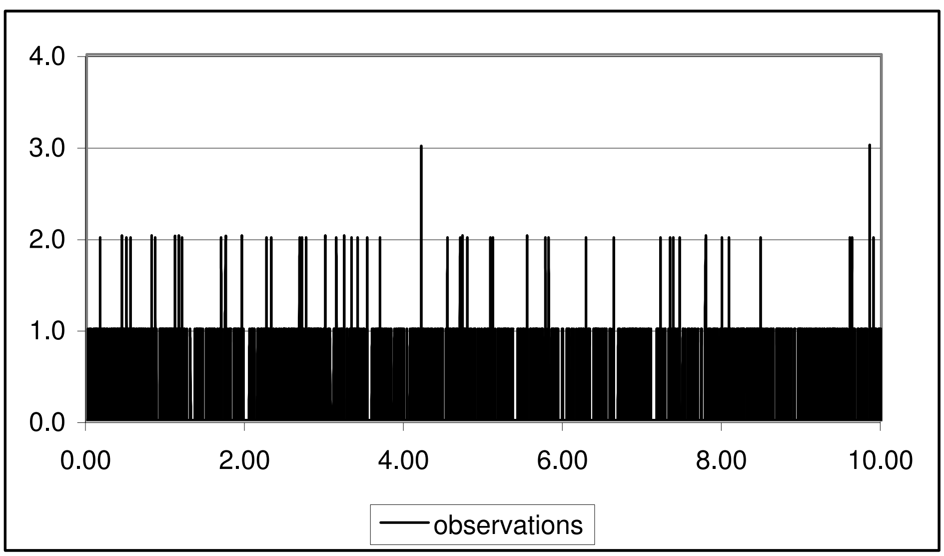

Figure A.7: Model 4: Observations y_{nk} for $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

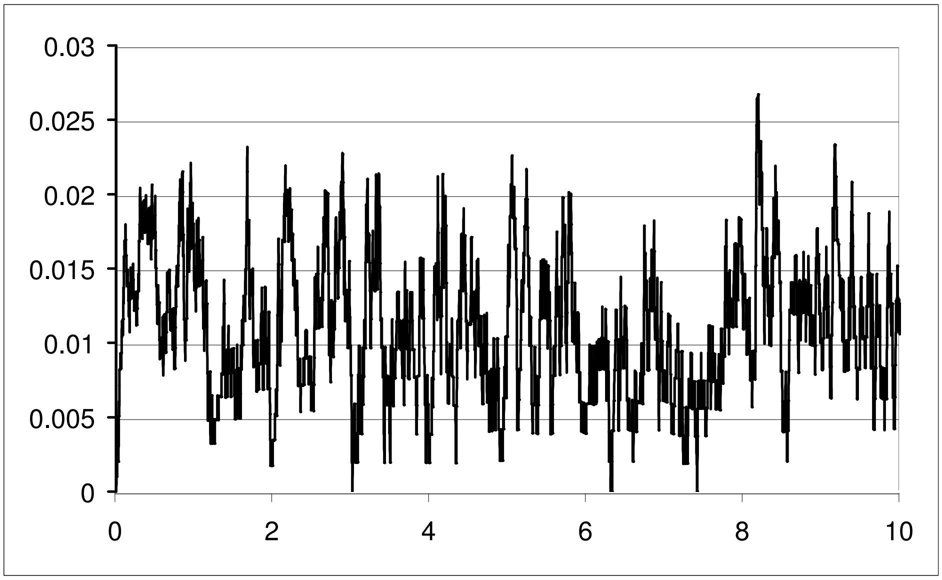

Figure A.8: Model 4: Simulated values for $y_k/(1-r_k)$ versus time when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

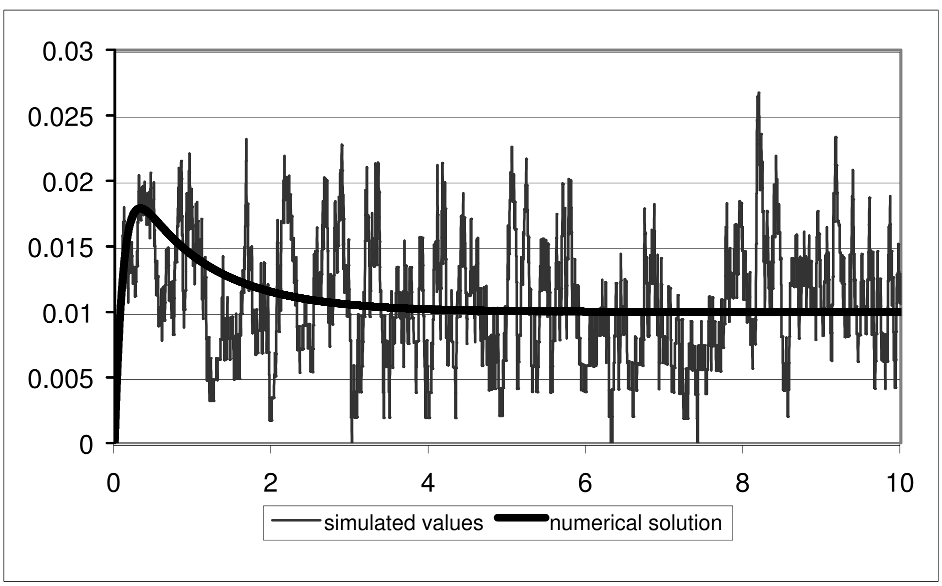

Figure A.9: Model 4: Simulated values for $y_k/(1-r_k)$ and numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

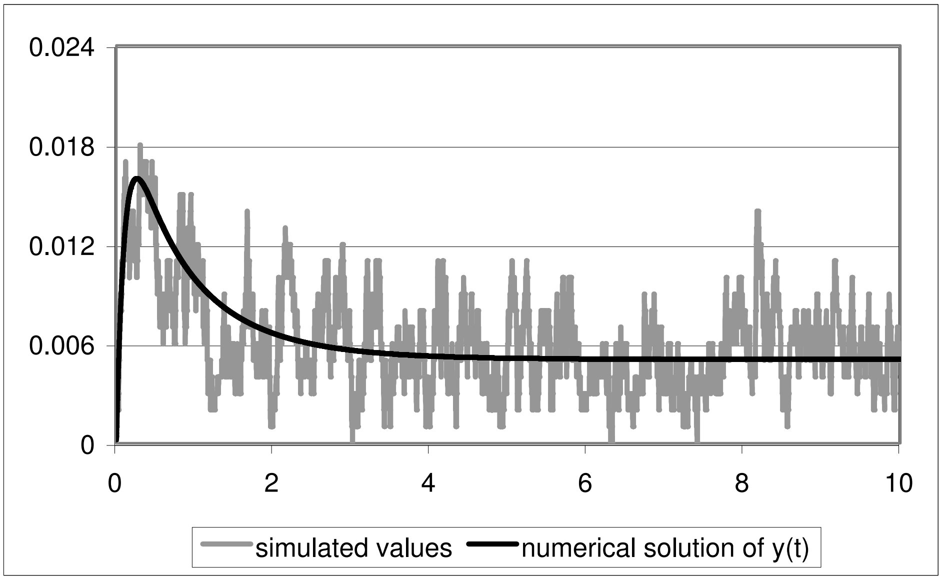

Figure A.10: Model 4: Simulated values for y_k and numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

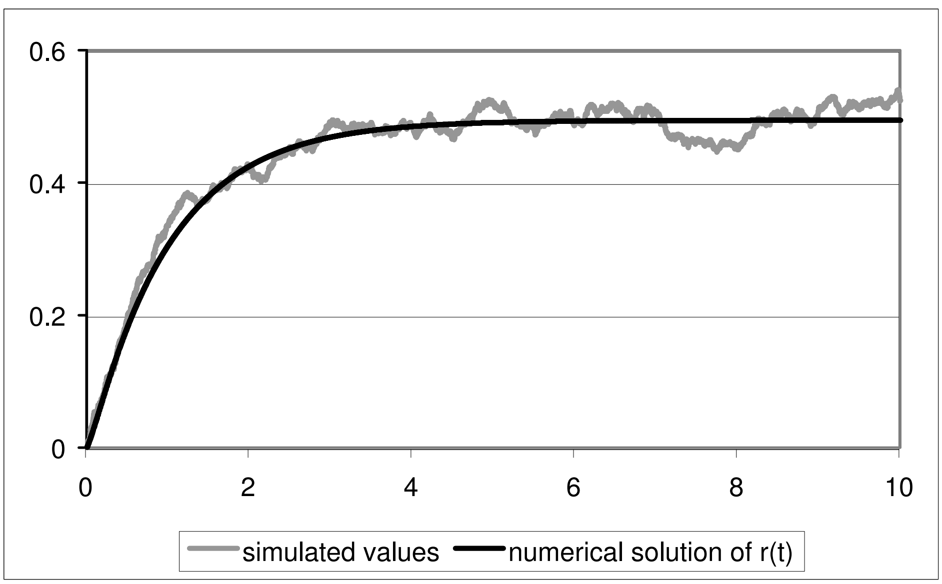

Figure A.11: Model 4: Simulated values for r_k and numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

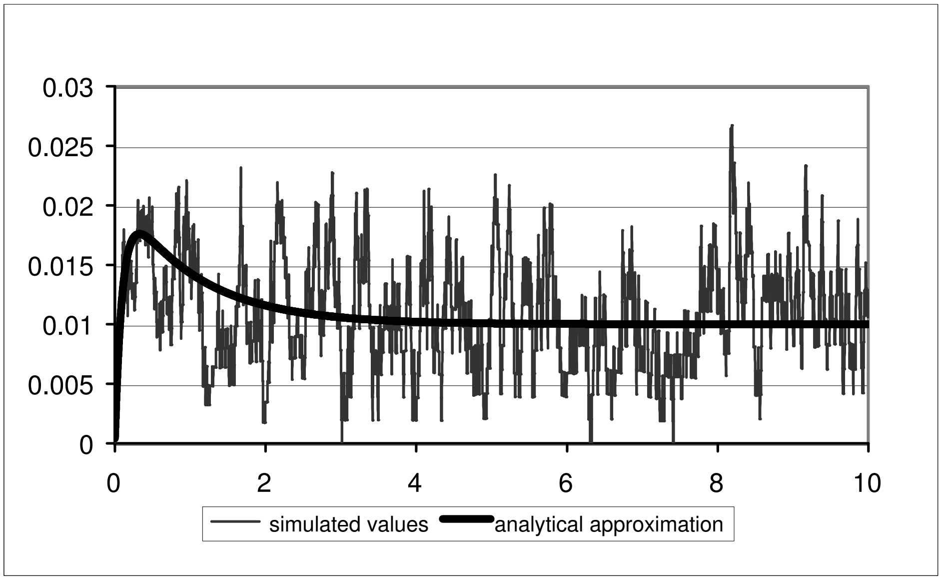

Figure A.12: Model 4: Simulated values for $y_k/(1-r_k)$ and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

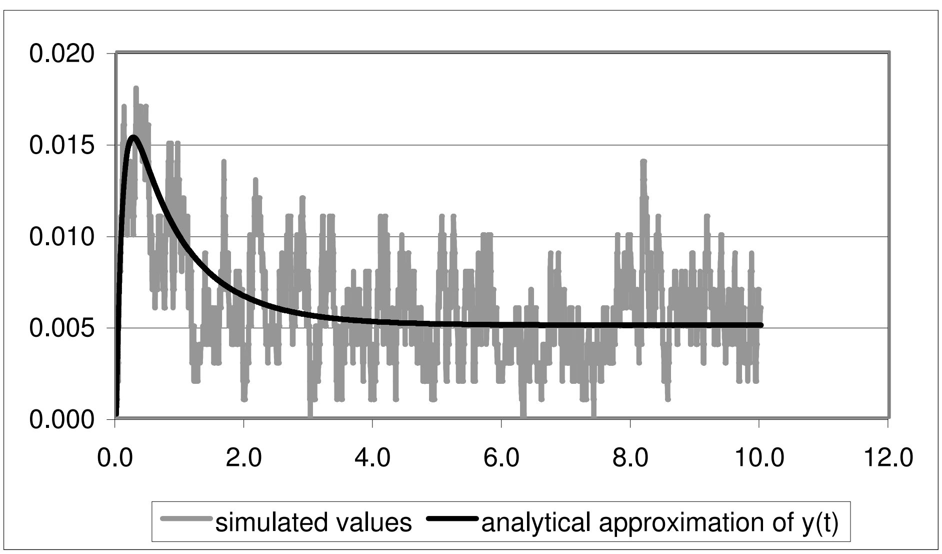

Figure A.13: Model 4: Simulated values of y_k and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

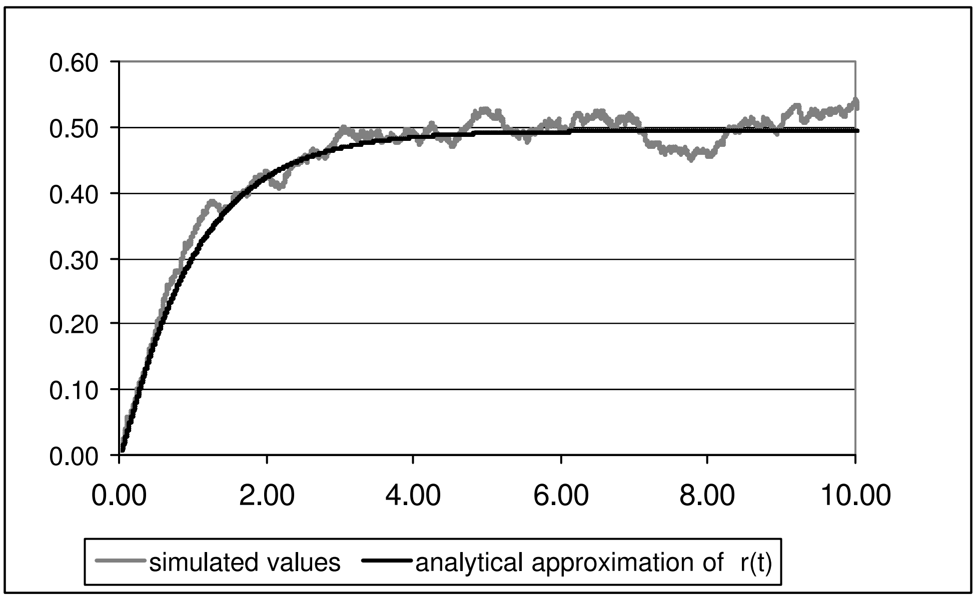

Figure A.14: Model 4: Simulated values of r_k and analytical approximation of numerical solution when $\lambda_0 = 1$, $\beta_0 = 0.8$ and $a = 1$

Appendix B

Tables

T	λ_0	β_0	a	$\hat{\lambda_0}$	$\lvert\hat{\lambda}-\lambda_0\rvert$	$\lvert\hat{\lambda}-\lambda_0\rvert/\lambda_0$	MSE
10	0.5	1	1	0.50096931	0.00096931	0.001938619	7.76242E-05
10	0.5	1	2	0.501820594	0.001820594	0.003641188	6.74354E-06
10	0.5	0.6	1	0.499942028	5.79724E-05	0.000115945	0.000166628
10	0.5	0.6	2	0.49841524	0.00158476	0.00316952	6.98167E-05
10	0.8	0.8	1	0.791839167	0.008160833	0.010201041	0.000138495
10	0.8	0.8	2	0.813604912	0.013604912	0.017006141	0.000260837
10	1	1	1	1.001518792	0.001518792	0.001518792	0.000154411
10	1	1	2	1.0002315	0.0002315	0.0002315	2.72806E-05
10	1	0.6	1	0.987725583	0.012274417	0.012274417	0.000306165
10	1	0.6	2	1.00150455	0.00150455	0.00150455	0.000518706

Table B.1: Model 2: Estimated values , bias, relative error and mean square error for $\hat{\lambda}$, T=10 (5 runs).

T	λ_0	β_0	a	$\hat{\lambda}$	$\hat{\lambda} - \lambda_0$	$\lvert\hat{\lambda} - \lambda_0\rvert/\lambda_0$	MSE
4	0.5	1	1	0.500133952	0.000133952	0.000267904	0.000259277
4	0.5	1	2	0.515968582	0.015968582	0.031937164	0.000303882
4	0.5	0.6	1	0.502947142	0.002947142	0.005894284	0.000162302
4	0.5	0.6	2	0.498818593	0.001181407	0.002362813	0.000180248
4	0.8	0.8	1	0.78648003	0.01351997	0.016899963	0.000490233
4	0.8	0.8	2	0.80742046	0.00742046	0.009275575	0.000327522
4	1	1	1	0.997878267	0.002121733	0.002121733	0.000168736
4	1	1	2	0.99441705	0.00558295	0.00558295	0.000124261
4	1	0.6	1	0.980128565	0.019871435	0.019871435	0.000565488
4	1	0.6	2	0.995710834	0.004289166	0.004289166	0.000101969

Table B.2: Model 2: Estimated values, bias, relative error and mean square error for $\hat{\lambda}$, T=4 (5 runs).

T	λ_0	β_0	a	$\hat{\lambda}$	$\hat{\lambda} - \lambda_0$	$\lvert\hat{\lambda} - \lambda_0\rvert/\lambda_0$	MSE
1	0.5	1	1	0.514292605	0.014292605	0.02858521	0.000442754
1	0.5	1	2	0.517066889	0.017066889	0.034133779	0.000394497
1	0.5	0.6	1	0.492029461	0.007970539	0.015941078	0.000913741
1	0.5	0.6	2	0.491485229	0.008514771	0.017029543	0.001126099
1	0.8	0.8	1	0.773609776	0.026390224	0.03298778	0.001367691
1	0.8	0.8	2	0.800458801	0.000458801	0.000573501	0.000725807
1	1	1	1	1.013997088	0.013997088	0.013997088	0.001113522
1	1	1	2	0.995500596	0.004499404	0.004499404	0.001001699
1	1	0.6	1	0.986711215	0.013288785	0.013288785	0.00264134
1	1	0.6	2	0.98983021	0.01016979	0.01016979	0.000772969

Table B.3: Model 2: Estimated values, bias, relative error and mean square error for $\hat{\lambda}$, T=1 (5 runs)

T	λ_0	β_0	a	$\hat{\beta}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$
10	0.5	1	1	0.998353	0.001647	0.001647058
10	0.5	1	2	0.996311	0.003689	0.003689408
10	0.5	0.8	1	0.799985	0.000015	1.88159E-05
10	0.5	0.8	2	0.790380	0.009620	0.012025549
10	0.5	0.6	1	0.600470	0.000470	0.000783956
10	0.5	0.6	2	0.602147	0.002147	0.003577824
10	0.8	1	1	1.002609	0.002609	0.002608958
10	0.8	1	2	0.996262	0.003738	0.003738204
10	0.8	0.8	1	0.808445	0.008445	0.010555632
10	0.8	0.8	2	0.786366	0.013634	0.017042914
10	0.8	0.6	1	0.599340	0.000660	0.001100468
10	0.8	0.6	2	0.600774	0.000774	0.001290252
10	1	1	1	0.998620	0.001380	0.001379511
10	1	1	2	0.999792	0.000208	0.000208398
10	1	0.8	1	0.807329	0.007329	0.009161841
10	1	0.8	2	0.800380	0.000380	0.000475528
10	1	0.6	1	0.607681	0.007681	0.012802373
10	1	0.6	2	0.599393	0.000607	0.001010954

Table B.4: Model 2: Estimated values, bias and relative error for $\hat{\beta}$, T=10

T	λ_0	β_0	a	$\hat{\beta}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$
4	0.5	1	1	1.000798	0.000798	0.000797504
4	0.5	1	2	0.967535	0.032465	0.032465117
4	0.5	0.8	1	0.854106	0.054106	0.067632947
4	0.5	0.8	2	0.788455	0.011545	0.014430989
4	0.5	0.6	1	0.596693	0.003307	0.005511412
4	0.5	0.6	2	0.602028	0.002028	0.003379871
4	0.8	1	1	1.003356	0.003356	0.003356284
4	0.8	1	2	1.003925	0.003925	0.003924998
4	0.8	0.8	1	0.814631	0.014631	0.018288654
4	0.8	0.8	2	0.792491	0.007509	0.009385756
4	0.8	0.6	1	0.598919	0.001081	0.001801463
4	0.8	0.6	2	0.606624	0.006624	0.011040659
4	1	1	1	1.002351	0.002351	0.002350862
4	1	1	2	1.006019	0.006019	0.006019373
4	1	0.8	1	0.805840	0.005840	0.007300236
4	1	0.8	2	0.795309	0.004691	0.005864239
4	1	0.6	1	0.612824	0.012824	0.021372717
4	1	0.6	2	0.602886	0.002886	0.004810483

Table B.5: Model 2: Estimated values, bias and relative error for $\hat{\beta}$, T=4

T	λ_0	β_0	a	$\hat{\beta}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$
1	0.5	1	1	0.969734	0.030266	0.030265772
1	0.5	1	2	0.957369	0.042631	0.042631287
1	0.5	0.8	1	0.810403	0.010403	0.013003315
1	0.5	0.8	2	0.758366	0.041634	0.052042868
1	0.5	0.6	1	0.614664	0.014664	0.02444035
1	0.5	0.6	2	0.620395	0.020395	0.033992328
1	0.8	1	1	1.023811	0.023811	0.023810752
1	0.8	1	2	1.014220	0.014220	0.014220079
1	0.8	0.8	1	0.832733	0.032733	0.040916107
1	0.8	0.8	2	0.800561	0.000561	0.000701582
1	0.8	0.6	1	0.607227	0.007227	0.012044887
1	0.8	0.6	2	0.578141	0.021859	0.036432477
1	1	1	1	0.985490	0.014510	0.014509584
1	1	1	2	1.007070	0.007070	0.007069815
1	1	0.8	1	0.813064	0.013064	0.016330311
1	1	0.8	2	0.778592	0.021408	0.026760006
1	1	0.6	1	0.611750	0.011750	0.019582669
1	1	0.6	2	0.610308	0.010308	0.017180276

Table B.6: Model 2: Estimated values, bias and relative error for $\hat{\beta}$, T=1

T	λ	μ	N	$\hat{\lambda}$	$\hat{\lambda}-\lambda$	$\lvert\hat{\lambda}-\lambda\rvert/\lambda$	MSE
1	1	0.1	5	0.987050727	0.012949273	0.012949273	0.000722053
2	1	0.1	5	1.002429756	0.002429756	0.002429756	0.000899369
1	2	0.1	10	1.997986067	0.002013933	0.001006966	0.002951017
2	2	0.1	10	1.994035697	0.005964303	0.002982152	0.002176102
1	2	0.2	5	1.990849951	0.009150049	0.004575024	0.002873985
2	2	0.2	5	1.99006516	0.00993484	0.00496742	0.000901891
1	1.5	0.2	5	1.528625077	0.028625077	0.019083385	0.004431895
2	1.5	0.2	5	1.505427764	0.005427764	0.003618509	0.00049552

Table B.7: Model 3: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ (5 runs).

λ_0	β_0	a	$\hat{\lambda}$	$\hat{\lambda}-\lambda_0$	$\lvert\hat{\lambda}-\lambda_0\rvert/\lambda_0$	MSE
1	0.8	1	1.039	0.039	0.039	0.0546926
1	0.8	1	0.93546	0.06454	0.06454	0.027104178
1	0.8	2	1.0478	0.0478	0.0478	0.0502278
1	0.8	2	1.1426	0.1426	0.1426	0.1290346
1	0.6	1	0.98	0.02	0.02	0.0200272
1	0.6	1	1.0158	0.0158	0.0158	0.036437
1	0.6	2	0.9518	0.0482	0.0482	0.0083754
1	0.6	2	1.2244	0.2244	0.2244	0.1426488
2	0.8	1	2.0934	0.0934	0.0467	0.2645554
2	0.8	1	2.1542	0.1542	0.0771	1.253969
2	0.8	2	2.5486	0.5486	0.2743	0.417305
2	0.8	2	2.0876	0.0876	0.0438	0.4976952
2	0.6	1	2.1138	0.1138	0.0569	0.054767
2	0.6	1	1.8026	0.1974	0.0987	0.3459062
2	0.6	2	2.2584	0.2584	0.1292	0.2911372
2	0.6	2	2.4914	0.4914	0.2457	1.0060054

Table B.8: Model 4: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ using the numeric solution (5 runs).

λ_0	β_0	a	$\hat{\lambda}$	$\hat{\lambda}-\lambda_0$	$\lvert\hat{\lambda}-\lambda_0\rvert/\lambda_0$	MSE
1	0.8	1	0.98723	0.01277	0.01277	0.040898389
1	0.8	2	1.0952	0.0952	0.0952	0.0896312
1	0.6	1	0.9979	0.0021	0.0021	0.0282321
1	0.6	2	1.0881	0.0881	0.0881	0.0755121
2	0.8	1	2.1238	0.1238	0.0619	0.7592622
2	0.8	2	2.3181	0.3181	0.15905	0.4575001
2	0.6	1	1.9582	0.0418	0.0209	0.2003366
2	0.6	2	2.3749	0.3749	0.18745	0.6485713

Table B.9: Model 4: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ using the numeric solution (10 runs)

λ_0	β_0	a	$\hat{\lambda}$	$\hat{\lambda} - \lambda_0$	$\lvert\hat{\lambda} - \lambda_0\rvert/\lambda_0$	MSE
1	0.8	1	1.0228	0.0228	0.0228	0.0493316
1	0.8	1	0.92476	0.07524	0.07524	0.027032248
1	0.8	2	1.0316	0.0316	0.0316	0.0441192
1	0.8	2	1.119	0.119	0.119	0.1086678
1	0.6	1	0.96906	0.03094	0.03094	0.018795098
1	0.6	1	1.0024	0.0024	0.0024	0.03241
1	0.6	2	0.94144	0.05856	0.05856	0.009086288
1	0.6	2	1.1978	0.1978	0.1978	0.118391
2	0.8	1	1.9584	0.0416	0.0208	0.1690748
2	0.8	1	1.934	0.066	0.033	0.7574628
2	0.8	2	2.3308	0.3308	0.1654	0.1812344
2	0.8	2	1.934	0.066	0.033	0.3040444
2	0.6	1	1.9892	0.0108	0.0054	0.0284808
2	0.6	1	1.706	0.294	0.147	0.3046364
2	0.6	2	2.0948	0.0948	0.0474	0.1558252
2	0.6	2	2.2326	0.2326	0.1163	0.4730562

Table B.10: Model 4: Estimated values, bias, relative error, MSE for $\hat{\lambda}$ using the analytic approach (5 runs).

λ	β	a	$\hat{\lambda}$	$\hat{\lambda} - \lambda$	$\lvert\hat{\lambda} - \lambda\rvert/\lambda$	MSE
1	0.8	1	0.97378	0.02622	0.02622	0.038181924
1	0.8	2	1.0753	0.0753	0.0753	0.0763935
1	0.6	1	0.98573	0.01427	0.01427	0.025602549
1	0.6	2	1.06962	0.06962	0.06962	0.063738644
2	0.8	1	1.9462	0.0538	0.0269	0.4632688
2	0.8	2	2.1324	0.1324	0.0662	0.2426394
2	0.6	1	1.8476	0.1524	0.0762	0.1665586
2	0.6	2	2.1637	0.1637	0.08185	0.3144407

Table B.11: Model 4: Estimated values, bias, relative error and MSE for λ using the analytic approach (10 runs)

λ_0	β_0	a	$\hat{\lambda}$	$\hat{\lambda} - \lambda_0$	$\|\hat{\lambda} - \lambda_0\|/\lambda_0$	MSE
1	0.8	1	1.01932	0.01932	0.01932	0.011467432
1	0.8	1	0.9978	0.0022	0.0022	0.0046754
1	0.8	2	1.0372	0.0372	0.0372	0.005452
1	0.8	2	1.0018	0.0018	0.0018	0.0131486
1	0.6	1	1.0072	0.0072	0.0072	0.017818
1	0.6	1	1.0824	0.0824	0.0824	0.0515
1	0.6	2	1.0806	0.0806	0.0806	0.0152486
1	0.6	2	0.931	0.069	0.069	0.0129558
2	0.8	1	2.1298	0.1298	0.0649	0.0346554
2	0.8	1	2.1624	0.1624	0.0812	0.1847108
2	0.8	2	2.0318	0.0318	0.0159	0.036183
2	0.8	2	1.9946	0.0054	0.0027	0.0703102
2	0.6	1	1.8916	0.1084	0.0542	0.0641188
2	0.6	1	2.4022	0.4022	0.2011	0.3184426
2	0.6	2	2.1824	0.1824	0.0912	0.28052
2	0.6	2	2.0622	0.0622	0.0311	0.1313546

Table B.12: Model 4: Estimated values, bias, relative error and MSE for $\hat{\lambda}$ when r(T) is known (5 runs).

λ_0	β_0	a	$\hat{\lambda}$	$\hat{\lambda} - \lambda_0$	$\|\hat{\lambda} - \lambda_0\|/\lambda_0$	MSE
1	0.8	1	1.00856	0.00856	0.00856	0.008071416
1	0.8	2	1.0195	0.0195	0.0195	0.0093003
1	0.6	1	1.0448	0.0448	0.0448	0.034659
1	0.6	2	1.0058	0.0058	0.0058	0.0141022
2	0.8	1	2.1461	0.1461	0.07305	0.1096831
2	0.8	2	2.0132	0.0132	0.0066	0.0532466
2	0.6	1	2.1469	0.1469	0.07345	0.1912807
2	0.6	2	2.1223	0.1223	0.06115	0.2059373

Table B.13: Model 4: Estimated values, bias, relative error, MSE for $\hat{\lambda}$ when $r(T)$ is known (10 runs)

λ_0	β_0	a	$\hat{\beta}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$	MSE
1	0.6	1	0.60542	0.00542	0.009033333	0.0000974
1	0.6	1	0.60206	0.00206	0.003433333	0.000026
1	0.6	2	0.6079	0.0079	0.013166667	0.0008
1	0.6	2	0.60188	0.00188	0.003133333	0.00000525
2	0.8	1	0.7996	0.0004	0.0005	0.0000048
2	0.8	1	0.8027	0.0027	0.003375	0.0000342
1	0.8	2	0.7958	0.0042	0.00525	0.0000402
1	0.8	2	0.80298	0.00298	0.003725	0.0000426
1	0.8	1	0.799	0.001	0.00125	0.0000406
1	0.8	1	0.80148	0.00148	0.00185	0.0000535
2	0.8	2	0.8011	0.0011	0.001375	0.00000825
2	0.8	2	0.7994	0.0006	0.00075	0.0000254

Table B.14: Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ using the numeric solutions (5 runs)

λ_0	β_0	a	$\hat{\beta}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$	MSE
1	0.6	1	0.6062	0.0062	0.010333333	0.000109
1	0.6	1	0.608	0.008	0.013333333	0.000122
1	0.6	2	0.5966	0.0034	0.005666667	0.0000638
1	0.6	2	0.6027	0.0027	0.0045	0.00000885
2	0.8	1	0.8008	0.0008	0.001	0.0000052
2	0.8	1	0.8035	0.0035	0.004375	0.0000404
1	0.8	2	0.7964	0.0036	0.0045	0.0000352
1	0.8	2	0.8032	0.0032	0.004	0.0000444
1	0.8	1	0.7994	0.0006	0.00075	0.0000458
1	0.8	1	0.8018	0.0018	0.00225	0.0000566
2	0.8	2	0.802	0.002	0.0025	0.0000123
2	0.8	2	0.8001	1E-04	0.000125	0.0000236

Table B.15: Model 4: Estimated values, bias, relative error, MSE for $\hat{\beta}$ using the analytic approximation (5 runs).

λ_0	β_0	a	$\hat{\beta}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$	MSE
1	0.6	1	0.6072	0.0072	0.012	0.000202
1	0.6	1	0.6057	0.0057	0.0095	0.0000765
1	0.6	2	0.5996	0.0004	0.000666667	0.0000512
1	0.6	2	0.602	0.002	0.003333333	0.0000192
2	0.8	1	0.79776	0.00224	0.0028	0.0000354
2	0.8	1	0.8032	0.0032	0.004	0.0000484
1	0.8	2	0.7944	0.0056	0.007	0.0000496
1	0.8	2	0.7988	0.0012	0.0015	0.0000556
1	0.8	1	0.7972	0.0028	0.0035	0.0000928
1	0.8	1	0.8001	1E-04	0.000125	0.0000319
2	0.8	2	0.8024	0.0024	0.003	0.0000556
2	0.8	2	0.80046	0.00046	0.000575	0.0000547

Table B.16: Model 4: Estimated values, bias, relative error, MSE for $\hat{\beta}$ when $r(T)$ is known (5 runs).

λ_0	β_0	a	$\hat{\beta}$	$\hat{\lambda}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\lambda}-\lambda_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$	$\lvert\hat{\lambda}-\lambda_0\rvert/\lambda_0$	MSE $\hat{\beta}$	MSE $\hat{\lambda}$
1	0.6	1	0.60788	1.5227	0.00788	0.5227	0.01313	0.5227	8.88E-05	0.552
1	0.6	1	0.605	1.0656	0.005	0.0656	0.00833	0.0656	4.58E-05	0.037
1	0.6	2	0.6082	1.5562	0.0082	0.5562	0.01366	0.5562	0.00014	0.633
1	0.6	2	0.6028	1.606	0.0028	0.606	0.00466	0.606	0.0002	0.146
2	0.8	1	0.7958	2.73	0.0042	0.73	0.00525	0.365	7.34E-05	2.590
2	0.8	1	0.7982	1.92	0.0018	0.08	0.00225	0.04	1.95E-05	0.265
2	0.8	1	0.8006	2.2894	0.0006	0.2894	0.00075	0.1447	6.38E-05	1.291
1	0.8	2	0.8002	1.2004	0.0002	0.2004	0.00025	0.2004	0.000027	0.067
1	0.8	2	0.8038	1.4676	0.0038	0.4676	0.00475	0.4676	0.000119	0.582
1	0.8	1	0.8008	1.444	0.0008	0.444	0.001	0.444	6.44E-05	0.734
1	0.8	1	0.7994	1.4346	0.0006	0.4346	0.00075	0.4346	0.000011	0.917
2	0.8	2	0.7982	1.9012	0.0018	0.0988	0.00225	0.0494	7.46E-05	0.529
2	0.8	2	0.7968	2.3888	0.0032	0.3888	0.004	0.1944	5.40E-05	0.723
2	0.8	2	0.7996	2.885	0.0004	0.885	0.0005	0.4425	1.28E-05	1.745

Table B.17: Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ and $\hat{\lambda}$ using numeric solutions (5 runs).

λ_0	β_0	a	$\hat{\beta}$	$\hat{\lambda}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\lambda}-\lambda_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$	$\lvert\hat{\lambda}-\lambda_0\rvert/\lambda_0$	MSE $\hat{\beta}$	MSE $\hat{\lambda}$
1	0.6	1	0.6094	1.589	0.0094	0.589	0.0156	0.589	0.000113	0.6907
1	0.6	1	0.6062	1.0906	0.0062	0.0906	0.0103	0.0906	6.02E-05	0.045
1	0.6	2	0.6116	1.64	0.0116	0.64	0.0193	0.64	0.000212	0.747
1	0.6	2	0.6056	1.228	0.0056	0.228	0.0093	0.228	0.000241	0.215
2	0.8	1	0.7976	3.059	0.0024	1.059	0.003	0.5295	6.64E-05	4.029
2	0.8	1	0.7994	2.035	0.0006	0.035	0.0008	0.0175	0.000017	0.321
2	0.8	1	0.80054	2.1282	0.00054	0.1282	0.0007	0.0641	6.82E-05	1.887
1	0.8	2	0.8014	1.26972	0.0014	0.26972	0.0018	0.26972	2.90E-05	0.107
1	0.8	2	0.8052	1.632	0.0052	0.632	0.0065	0.632	0.000141	0.921
1	0.8	1	0.8022	1.5206	0.0022	0.5206	0.0027	0.5206	6.74E-06	0.918
1	0.8	1	0.7988	1.7204	0.0012	0.7204	0.0015	0.7204	3.32E-05	1.203
2	0.8	2	0.8002	2.2128	0.0002	0.2128	0.00025	0.1064	7.58E-05	1.021
2	0.8	2	0.8002	2.914	0.0002	0.914	0.00025	0.457	5.50E-05	2.064
2	0.8	2	0.802	3.8042	0.002	1.8042	0.0025	0.9021	2.36E-05	6.025

Table B.18: Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ and $\hat{\lambda}$ using analytic approximation (5 runs).

λ_0	β_0	a	$\hat{\beta}$	$\hat{\lambda}$	$\lvert\hat{\beta}-\beta_0\rvert$	$\lvert\hat{\lambda}-\lambda_0\rvert$	$\lvert\hat{\beta}-\beta_0\rvert/\beta_0$	$\lvert\hat{\lambda}-\lambda_0\rvert/\lambda_0$	MSE $\hat{\beta}$	MSE $\hat{\lambda}$
1	0.6	1	0.5952	1.1358	0.0048	0.1358	0.008	0.1358	0.000366	0.0266
1	0.6	1	0.6044	1.0418	0.0044	0.0418	0.0073	0.0418	7.12E-05	0.0038
1	0.6	2	0.596	1.1217	0.004	0.1217	0.0066	0.1217	7.00E-05	0.0216
1	0.6	2	0.6008	1.053	0.0008	0.05296	0.00133	0.05296	0.000024	0.0106
2	0.8	1	0.7942	2.219	0.0058	0.21996	0.00725	0.10998	0.000119	0.188
2	0.8	1	0.8003	2.004	0.0003	0.00442	0.000375	0.00221	6.51E-06	0.033
2	0.8	1	0.8004	2.117	0.0004	0.1167	0.0005	0.05835	6.24E-05	0.177
1	0.8	2	0.79609	1.01894	0.00391	0.01894	0.0048	0.01894	4.62E-05	0.001
1	0.8	2	0.7978	1.02972	0.0022	0.02972	0.00275	0.02972	6.30E-05	0.0038
1	0.8	1	0.7938	1.0004	0.0062	0.0004	0.00775	0.0004	0.000131	0.0059
1	0.8	1	0.79	1.06034	0.01	0.06034	0.0125	0.06034	0.000254	0.1693
2	0.8	2	0.8044	2.0017	0.0044	0.0017	0.0055	0.00085	3.00E-05	0.112
2	0.8	2	0.7955	2.14238	0.0045	0.14238	0.0056	0.07119	0.00013	0.0394
2	0.8	2	0.79562	2.35488	0.00438	0.35488	0.0055	0.17744	2.92E-05	0.212

Table B.19: Model 4: Estimated values, bias, relative error and MSE for $\hat{\beta}$ and $\hat{\lambda}$ when r(T) is known (5 runs).