A Cross-Entropy Method for Optimal Stopping Problems

By

Thilini Dulanjali Kularatne
44792700

A thesis submitted to Macquarie University
for the degree of Master of Research
Department of Statistics
June 2017

MACQUARIE
University

44792700, 2017.

Typeset in LaTeX 2_ε.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Thilini Dulanjali Kularatne
44792700

Acknowledgements

I would like to extend my sincere gratitude to following persons who guided me and helped me to the successful completion of my Master of Research study. Their guidance, support and encouragement throughout this year was invaluable.

First, I want to express my sincere gratitude to my supervisor Dr. Gerogy Sofronov. His supervision and guidance throughout this year made my research life so easy. I highly appreciate his useful comments, remarks and continuous mentorship through this year. Thank you so much for your generous support in every manner and it made the past year extremely enjoyable and fruitful.

Next I would like to thank Dr. Thomas Fung for being my MRes advisor. Your support and guidance helped me a lot in many ways. Further I gratefully acknowledge all the academic and non-academic staff of Department of Statistics for providing their fullest assistance throughout this year. I would also like to thank Macquarie University for providing me necessary financial support through the International Macquarie University Research Excellence Scholarship (iMQRES). Further, I would like to thank Dr. Madawa Jayewardena for his support in our collaborative work. I would like to thank all my colleges in the department and specially Kasun Rathnayake for guiding me to apply for the research study at Macquarie University under Dr. Georgy Sofronov. Finally I wish to give my heartiest thank to my husband and my little daughter for giving me strength throughout this year.

List of Conference Presentations and Publications

- Thilini Dulanjali Kularatne and Georgy Sofronov. *A Cross-Entropy method for an optimal stopping problem.*
  (Poster Presentation at AustMS 2016. Held from December 5, 2016 - December 8, 2016 at Australian National University, Canberra, Australia)

- Thilini Dulanjali Kularatne and Georgy Sofronov. *Estimation of thresholds in optimal stopping problems via a Cross-Entropy method.*
  (Abstract Submitted for the International Conference on Robust Statistics 2017 (ICORS 2017). The conference is scheduled to be held at University of Wollongong, Australia from July 3, 2017 - July 7, 2017)

- Thilini Dulanjali Kularatne, Georgy Sofronov and Madawa W. Jayawardana. *A Cross-Entropy method for sequential decision problems.*
  (Paper submitted for 56^{th} IEEE Conference on Decision and Control. The conference is scheduled to be held in Melbourne, Australia from December 12, 2017 - December 15, 2017)

Abstract

There are frequent situations when observations are recorded consecutively over a period of time, for an example, daily values of currency exchange rates. Sequential observations appear one by one, so data are analysed as they are collected without fixing the sample size in advance. Further sampling may be terminated according to a pre-defined stopping rule. There are situations where we need to make decisions considering the observations which we are already having while future observations are not known yet. Sequential data analysis has a variety of applications in a wide range of fields including industrial quality control, econometrics, analysis of financial systems among many others. In this thesis, we develop several versions of a Cross-Entropy method to find an approximate optimal stopping rule. Here we have considered cases of both independent and dependent observations. We have carried out a simualtion study, which has shown the accuracy of the proposed algorithm.

Contents

Acknowledgements **iv**

List of Conference Presentations and Publications **v**

Abstract **vi**

List of Figures **ix**

List of Tables **x**

1 Introduction **1**

1.1 Background and Aims 1

1.2 Statement of Problem 4

1.2.1 Backward Induction 5

1.2.2 Optimal Stopping Problem for Uniformly Distributed Random Variables 6

1.3 Sequential Change-Point Problem 7

1.3.1 Sequential Change-Point Problem as an Optimal Stopping Problem . 7

1.4 Thesis Outline 8

2 Literature Review **10**

3 Theory and Methodology **15**

3.1 The Cross-Entropy Method 15

3.1.1 Importance Sampling 17
3.1.2 Kullback-Leibler Distance 18
3.2 General CE Algorithm for Optimization 19
3.2.1 CE Method for Noisy Optimization 21
3.3 Statistical Distributions to Simulate Samples for the CE Method 24
3.3.1 Normal Distribution 24
3.3.2 Truncated Normal Distribution 25
3.3.3 Beta Distribution 25
3.4 Convergence Properties of the CE Algorithm 26

4 Simulation Study **27**
4.1 The Cross-Entropy Method for Independent Random Variables 27
4.1.1 The Algorithm for Non-ordered Thresholds 28
4.1.2 The Algorithm for Ordered Thresholds 30
4.2 The Cross-Entropy Method for Dependent Random Variables 35
4.2.1 The Algorithm for Estimating the Coefficients of the Polynomial . . . 36

5 Discussion and Future Research **39**
5.1 Summary 39
5.2 Discussion 41
5.3 Future Work 42

A Appendix - R Code **45**
A.1 Cross-Entropy Method for Independent Random Variables 46
A.1.1 Algorithm for Non-ordered Thresholds 46
A.1.2 Algorithm for Ordered Thresholds 51
A.2 Cross-Entropy Method for Dependent Random Variables 52

References **55**

List of Figures

4.1 Values of the estimated thresholds using the CE method for non-ordered thresholds 31
4.2 Density plots for average RMSE by the CE method for non-ordered thresholds 32
4.3 Values of the estimated thresholds using CE method for ordered thresholds . 33
4.4 Density plots for average RMSE by CE method for ordered thresholds 34
4.5 Estimated thresholds using CE method for polynomials 37

List of Tables

4.1 True values and the average estimates of the thresholds for each time point produced by the CE method for non-ordered thresholds 30
4.2 True values and the average of the estimated gains for each time point by CE method for ordered thresholds 32
4.3 Average RMSE and the standard deviations of RMSE values of the estimates using two algorithms 33
4.4 Average values obtained for true value of game and the estimated value of game using two algorithms 34
4.5 The iteration count and the running time for the CE method considering all distributions for both algorithms 35

1

Introduction

1.1 Background and Aims

A sequential data set is a collection of records which are ordered with respect to time. There are frequent situations where data are sequentially collected over time such as rainfall measurements on successive days, daily values of currency exchange rates etc.

Sequential observations appear one by one, and the data are analysed as they are collected without fixing the sample size in advance. Therefore, there are situations when we need to make decisions on already obtained information while future observations are not known yet. The important question is, therefore, how a decision can be made when the data get updated in real-time.

Problems like this are faced regularly in many areas including industrial quality control (detecting changes in the manufacturing process) [63], environmental applications (detecting changes in ecological systems) [3, 34], signal processing (structural analysis of electroencephalographic signals)[47, 48], epidemiology (timely detection and prevention of various types of diseases)[19, 39] and finance (buying or selling an asset)[13, 17, 57].

The following are examples of some specific applications of sequential data analysis.s

1. **In finance**: Consider a problem of buying an asset and selling it in a later time within a specific time period. For this problem we need to assume the market prices are random observations from a known distribution. If $X_1, X_2, \ldots, X_n$ are the market prices of our asset for n time periods, we observe X_{n+1} and decide whether we are going to buy this asset at time $n+1$ (if it has been already purchased, we observe X_{n+1} and decide whether to sell the asset at time $n+1$ for a value of X_{n+1}). Our objective should be to maximize our profit while taking two decisions sequentially to stop and buy the asset and to stop and sell the asset for a given sequence of market prices [57]. Apart from parametric models, non-parametric techniques can be used for such problems [13].

2. **In industrial quality control**: Suppose we sequentially observe a series of products and we estimate the quality of those products. If observe $X_1, X_2, \ldots$ and we want to detect the time of which the distribution of the observations possibly has changed at some unknown time point. With sequential data analysis we can develop an on-line quality controlling procedure for the manufacturing process which will detect a change in the distribution as soon as possible with a low false alarm rate [63].

3. **In environmental applications**: Consider the process of changing the structure of ecological community over time. The change happens continuously over time and sometimes it is important to detect these changes. With sequential data analysis we can detect the changes, which can be used to prevent changes associated with loss of essential ecosystem functionality [2].

We observe a sequence of random variables, which can be interpreted as the value of an asset at a particular time. We have to decide when we must stop, given that no recall allowed, that is, a random variable once rejected cannot be chosen later on. Our decision to stop depends on the observations already made, but does not depend on the future observations which are not yet known. Our objective is to maximize an expected reward. There is an extensive literature on optimal stopping problem. For previous work related to this topic see [12, 18]. The main issue with finding an optimal stopping rule is the necessity to evaluate complex integrals. Thus efficient evaluation is not straightforward and requires the development of numerical methods, for example, methods based on Monte-Carlo (MC) integration.

A simple least squares approach was discussed in [36] to approximate the value American options using least squares MC method. The fitted values of the model were used to estimate the optimal stopping rule for the option. Linear and non-linear optimization techniques were used in [5, 9, 26] for optimal stopping problems for Markov decision processes. An adaptive learning version of regression MC to solve optimal stopping was proposed in [22]. The regression MC was considered as a sequential stochastic optimization design for a discrete time optimal stopping problem with a finite horizon. The problem of finding an optimal stopping rule for a buying and selling problem with independent observations was discussed in [57] to maximize the expected gain. The problem was discussed considering two stops to buy and to sell an asset. The further extensions and applications of multiple optimal stopping rule can be found in [40, 46, 56, 59, 60].

Sequential change-point problem (also known as the quickest change-point problem) can be considered as an optimal stopping problem. A procedure to detect sequential change-point using a Cross-Entropy (CE) method was discussed in [58]. The study was carried out considering an objective of maximizing the average detection delay with restrictions on average run length to false alarm. Further applications of the CE method in optimal stopping problem can be found in [61].

It is clear that the major focus of sequential data analysis is finding stopping rules (exact or approximate), that is, selecting a time for an action that would maximize the expected gain. Many problems including the problems we discussed above considered within the framework of the theory of optimal stopping rules for both dependent and independent observations. This study aims to develop numerical procedures that will allow us to make the best possible decision based on the data available. In this thesis, we focus on constructing a CE method to find an approximate optimal stopping rule using MC simulation. The study is developed considering independent random observations developing the stopping rule to maximize the expected gain. The CE method is further extended to dependent observations with the aim of applying to the real data sets. As our future research, we can consider real data related to stock market to develop an optimal stopping rule to buy or sell an asset by maximizing the expected gain. Further, we can also develop numerical techniques for a sequential change-point problem.

1.2 Statement of Problem

In this thesis, we consider an optimal stopping problem for finite discrete time stochastic sequences. We observe $X_1 = x_1, X_2 = x_2, \ldots, X_i = x_i$ sequentially and we must decide either to stop and accept the value x_i at time i $(i = 1, 2, \ldots, N-1)$ or continue and observe X_{i+1}. We must stop and accept the final value x_N as our gain if we did not stop earlier. We consider the case of "sampling without recall" for our problem, therefore, if we pass any observation without stopping, we cannot return to that object again in later time and accept the value as the gain. The random variable X_i can be interpreted as the value of an asset at time i.

At any stage of the sampling process, suppose n is the number of observations remained to be observed. After at least one observation has been considered for sampling, our position is considered by the value of x, the most recent observation taken. We denote $V_n(x)$ as the gain from the optimal procedure. Therefore, if we observe $X_1 = x_1, X_2 = x_2, \ldots, X_i = x_i$, the gain from the optimal procedure can be denoted as $V_{N-i}(x_i)$. Since we must stop and

accept the final observed value x as our gain as no further observations are available, it gives $V_0(x) = x$.

Let v_n be the expected gain with n observations remained, $n = 1, 2, \ldots, N$. For this problem we are sampling our observations using sampling without recall from a known distribution F. The value v_n depends only on the number of observations remaining to be considered for sampling and it does not depend on the observations which have already been taken for sampling. According to [18], the expected gain v_n can be expressed as follows

$$v_n = E[V_{n-1}(X)] = \int_{-\infty}^{\infty} V_{n-1}(x) dF(x). \tag{1.1}$$

The gain $V_n(x)$ for $n = 1, 2, \ldots, N-1$ from the optimal procedure is the maximum of the gain x obtaining if we stop and the the expected gain v_n by continuing. Therefore,

$$V_n(x) = \max\{x, v_n\}. \tag{1.2}$$

Since we have considered $V_0(x) = x$, the functions $V_1, V_2, \ldots$ and the values $v_1, v_2, \ldots$ can be determined using equations (1.1) and (1.2). Our aim is to find a stopping rule which maximizes the expected gain v_N.

1.2.1 Backward Induction

We have a finite horizon N for the number of observations for sampling. We apply backward induction method to obtain the optimal solution for the stopping problem. This means that we start with the final stage of observation and work backward to the first stage of observations [12, 18].

For a sequential random sample $X_1, X_2, \ldots, X_N$, we observe $X_1, X_2, \ldots, X_{N-1}$ and consider whether it is better to select the final observation X_N. We start to make decisions by considering backwards. At the next stage, since we know the values $X_1, X_2, \ldots, X_{N-2}$, we can now decide whether it is better to select X_{N-1}. Similarly, since we know the optimality of selecting observations at each point, knowing each possible value of X_{n-1} for $n = 1, 2, \ldots, N$,

we can make our decisions about future observations. Therefore, we can compare the risk of observing X_1 and continuing in an optimal pattern after that with taking the immediate decision without observing any further variables. This method of comparing risks in early stages of decision making gives an optimal sequential decision procedure.

1.2.2 Optimal Stopping Problem for Uniformly Distributed Random Variables

Consider a sequential random sample $X_1, X_2, \ldots, X_N$ taken from a uniform distribution on the interval (a, b). If the sampling process is terminated after observing $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ for $(n = 1, 2, \ldots, N)$, the gain is x_n. For $n = 1, 2, \ldots,$ if $v_0 = a$, $v_1 = (a + b)/2$, then we can use equations (1.1) and (1.2) to obtain v_n.

$$v_2 = E[V_1(X)],$$
$$V_1(X) = \max\{x, v_1\}.$$

Then we can write

$$v_2 = E[V_1(X)] = \int_a^b \max\{x, v_1\} \frac{1}{(b-a)} dx$$

since

$$\max\{x, v_1\} = \begin{cases} x, & x \geq v_1 \\ v_1, & x < v_1. \end{cases}$$

We have

$$\begin{aligned} v_2 &= \int_a^{v_1} v_1 \frac{1}{(b-a)} dx + \int_{v_1}^b x \frac{1}{(b-a)} dx \\ &= \frac{v_1 x}{(b-a)}\Big|_a^{v_1} + \frac{x^2}{2(b-a)}\Big|_{v_1}^b \\ &= \frac{1}{2(b-a)}[b^2 - 2v_1 a + {v_1}^2]. \end{aligned}$$

Therefore, in general we can write the value of the threshold at $(n+1)$-th time as

$$v_{n+1} = \frac{1}{2(b-a)}[b^2 - 2v_n a + {v_n}^2]. \tag{1.3}$$

For our problem we are considering a sequential random sample $X_1, X_2, \ldots$ taken from a uniform distribution on the interval $(0, 1)$. Therefore, using equation (1.3), the value of the threshold at $(n+1)$-th time can be written as

$$v_{n+1} = \frac{1}{2}[1 + {v_n}^2]. \tag{1.4}$$

1.3 Sequential Change-Point Problem

In the sequential change-point problem (or the quickest change-point problem) we consider sequence of observations appear one by one in time and we do not have any information on future observations. The objective is to identify the change-point in the sequence of observations as quickly as possible while avoiding "false alarms". There have been developed several sequential detection rules for the quickest change-point problem. These include cumulative sum (CUMSUM) procedure [42], exponentially weighted moving average (EWMA) procedure [49] and Shiryaev-Roberts (SR) procedure [54, 55].

Suppose we observe independently and identically distributed (iid) random variables $Y_1, Y_2, \ldots, Y_N$ sequentially. Consider the sequence is in the "control state" initially, that is, the random variables are distributed according to a common probability function F_0 for $t = 1, 2, \ldots, \tau - 1$ with parameter θ_0 and at unknown time τ a change happens in the distribution and the process becomes "out of control". After the change has occurred, we observe random variables with a different probability function $F_1 (F_0 \neq F_1)$ with parameter $\theta_1 (\theta_0 \neq \theta_1)$.

1.3.1 Sequential Change-Point Problem as an Optimal Stopping Problem

The sequential change-point can be considered as an optimal stopping problem [12]. Suppose that a change occurs in the process at a time τ and we have the aim of detecting the occurrence of the change. Assume that τ is a non-negative integer-valued random variable

with probabilities

$$P(\tau = 0) = \pi,$$
$$P(\tau = k|\tau > 0) = r_k, \qquad k = 1, 2, \ldots \quad \Big(\sum_{1}^{\infty} r_k = 1\Big).$$

As above, if the change occurres at time τ then we observe $y_1, \ldots, y_{\tau-1}, y_\tau{}', y_{\tau+1}{}', \ldots$ in which $y_1, y_2, \ldots$ are iid with a known distribution F_0 and $y_1{}', y_2{}', \ldots$ are iid with a known distribution F_1.

Suppose we stop the process at time n. Then we will loose

$$c \quad \text{if} \quad \tau > n,$$
$$n - \tau \quad \text{if} \quad \tau \leq n,$$

where c is considered as a fixed inspection cost. Our aim is to maximize the expected gain (the loss multiplied by -1)

$$x_n = -c(1 - \pi_n) - \sum_{i=0}^{n-1}(n - i){p_i}^n \quad n = 0, 1, 2, \ldots$$

with

$${p_i}^n = P(\theta = i|\mathcal{F}_n) \quad (i, n \geq 0),$$
$$\pi_n = P(\theta \leq n|\mathcal{F}_n) \quad (n \geq 0),$$

where $\mathcal{F}_n$ is the σ-algebra generated by the first n random variables.

1.4 Thesis Outline

The remainder of this thesis is organized as follows. Chapter 2 discusses the literature related to the optimal stopping problem. It contains detailed information on previously proposed numerical method to solve optimal stopping problems. The numerical methods discussed

in this chapter cover parametric and non-parametric methods. Chapter 3 covers the theory and methodology used in this thesis. It provides detailed information about the CE method including its modification for noisy optimization. The chapter further describes the convergence properties of the CE method.

Chapter 4 displays the results of the simulation study which we have performed by implementing the proposed versions of the CE method for sequences of independent and dependent random variables. Chapter 5 summarises the thesis outcomes and discusses the scope of the future extensions related to the topic.

2

Literature Review

In this chapter of the thesis, we discuss the literature related to optimal stopping problem. We give detailed information about some numerical techniques developed for optimal stopping problem.

Secretary Problem

According to [21], the secretary problem or the marriage problem (or the dowry problem) was first introduced by Martin Gardner in early 1960's in a mathematical games column issued by Scientific American. The basic form of the problem was discussed like following.

> "Imagine a company wants to hire a secretary for a one secretarial position available out of n number of applicants. The applicants are interviewed one-by-one

in a random order. The decision about the applicant is made as soon as the interview is finished for that particular applicant. The applicants are ranked from best to worst without any ties and the decision is based on those ranks. When assigning the ranks the interviewer has no idea about the unseen applicants. If an applicant is rejected then that applicant cannot be recalled later. The objective should be to maximize the probability of selecting the best applicant via an optimal strategy."

The above problem was then extended by many mathematicians and statisticians. Dynamic programming principles were introduced for this sequential decision problem by Lindley in 1961 [35], where the author showed that the dynamic programming methods which can be used as computational techniques in the problem of statistical decision making. He extended the secretary problem of assigning a rank arbitrary for the applicant selected and then he particularly considered the problem of minimizing the expected rank of the selected applicant. However, the author could not address some issues like finding asymptotic optimal strategies for large n and relative ranks. Those problems were solved by Chow *et al.* in 1964 [10]. They discussed the secretary problem as an optimal selection problem based on relative ranks. The authors developed the theories related to the problem considering relative ranks. They solved the problem by having the objective to find an optimal stopping rule which minimizes the expectation of the absolute rank of the person selected.

Optimal Stopping Problem

With the introduction of the secretary problem, extensions and developments have been made on the basic problem which was extended to the problem of optimal stopping [11]. Applications of the optimal stopping problem can be found in many areas; Griffeath and Snell in 1974 [23] consider a class of optimal stopping problems for conditioned random walk with selling strategies in the stock market. They aim to demonstrate the results obtained by Boyce in 1970 [6], on how stock prices adopt to Brownian motion, using a simple process of discrete random walk. Authors were able to develop maximum entropy methods for the optimal stopping problem in selling strategies in the stock market. Also, the authors used

the maximum entropy measures to formulate discrete solutions for the stock market results in Boyce (1970) [6].

The applications of the optimal stopping problem in option pricing can be found in the literature. Jacka (1991) showed that the problem of pricing American put option is equivalent to solving an optimal stopping problem [30]. Author established the optimal stopping problem by considering the martingale measures and the fact that the fair price of the option could be regarded as a function of the present stock price and the time horizon. The connection between the option pricing and the optimal stopping problem was discussed by generalizing the Black-Scholes option formula. Further applications of optimal stopping problem in option pricing can be found in [24, 36, 44].

Optimal stopping problems can further modified by introducing reward functions to the problem. In 2005, Novikov and Shiryaev presented a method to obtain an effective solution of the optimal stopping problem when a reward function is an integer power function of a random walk on an infinite time interval [41]. The authors have used the results obtained by Darling *et al.* (1972) [16] when generalizing the proposed method to check the optimality of the stopping time with integer power function rewards. The optimal stopping problem was extended to a case when more than one decision should be made. For example, Sofronov *et al.* (2006) consider a problem which involves two stops: the first stop is to buy an asset and the second stop for selling the asset [57]. For further extensions and applications of multiple optimal stopping rules see [40, 43, 56].

Numerical Methods on Optimal Stopping Problem

Several numerical methods were developed considering the optimal stopping problem. Longstaff and Schwartz (2001) used the least squares Monte-Carlo (LSM) method to approximate the value of American options [36]. The idea of the method is to compare the immediate exercise value of an option with the expected cash flow from continuing to develop the optimal strategy. The conditional expectations of the cash flows were approximated using the least

squares applying backward induction [12, 18]. The proposed LSM algorithm was illustrated using cross sectional information of an American put-option on a share of non-dividend-paying stock. A set of basis functions was used to identify the conditional expectation using simple regression. One possible choice of the basis functions is the set of weighted Laguerre polynomials.

$$L_n(X) = \exp(-X/2)\frac{e^X}{n!}\frac{d^n}{dX^n}(X^n e^{-X}).$$

Then the fitted value of regression was obtained using

$$F(\omega; t_{K-1}) = \sum_{j=0}^{\infty} a_j L_j(X)$$

for first $M < \infty$ basis functions. The fitted values accurately estimate the optimal stopping rule for the option. The authors have discussed the convergence properties of the proposed method and illustrated the applicability of the method considering examples related to American put options and American-Bermuda-Asian options. Further applications of regression Monte-Carlo to solve optimal stopping problems through simulations can be found in [22].

Yu and Bertsekas (2006) proposed another algorithm based on the least squares methodology [64]. The methodology was developed as a new Q-learning algorithm to approximate cost evaluation of optimal stopping problems. The authors consider optimal stopping of Markovian decision problems for discrete time with large state space where dynamic programming solutions cannot be implemented practically. The authors use linear function approximation to the optimal cost functions (Q-factors) [62].

Linear programming and non-linear programming techniques have also been proposed to solve optimal stopping problems. Cho and Stockbridge (2002) formulated the optimal stopping problem using an infinite-dimensional linear programming (LP) technique [9]. The optimal stopping problem was described considering a process X and a reward $R(x)$ which can be obtained when the process is stopped at time τ in state x (when $X(\tau) = x$). The objective is to find the stopping rule which maximizes the expected reward $E[R\{X(\tau)\}]$.

The authors describe the LP formulation for the optimal stopping problem and establish the equivalence between the stochastic process and the LP formulation. Moreover, they develop the equivalence between optimal stopping problems and control problems. To illustrate the accuracy of the LP formulation, the authors have used an example of one-dimensional Brownian motion.

Helms (2002) considers both liner and non-linear programming techniques to solve optimal stopping problems [26]. The author has considered a linear programming method for optimal stopping problems with unimodal function of a threshold value and another method which is applicable to more general stopping problems. For the later the author has used both linear and non-linear optimization techniques. Both methods were illustrated using Shiryaev's quickest detection problem [55] for Brownian process and the author suggests that the LP techniques are convenient and easy when analysing decision problems. For more applications of the linear programming formulation in the optimal stopping problem for Markov decision process see [5].

The Cross-Entropy (CE) method [50, 52] is another numerical techniques which can be used to solve the optimal stopping problem effectively. Sofronov *et al.* (2013) applied the CE method for finding an approximate solution of the multiple best choice problem [61]. The CE method was used to develop the optimal stopping rule and the value of the game. Further applications of the CE method in optimal stopping problems can be found in [45].

The sequential change-point detection problem can be treated as an optimal stopping problem. There is an extensive literature on this topic and for previous work related, see, for example, [20, 58].

3

Theory and Methodology

In this section of the thesis, we provide detailed information on the theory and methodology used in the subsequent chapter corresponding to our simulation study. The theory and methods have discussed in detail to deliver an overview of the scope of the thesis.

3.1 The Cross-Entropy Method

The Cross-Entropy (CE) method is an adaptive Monte Carlo (MC) approach introduced by Reuven Y. Rubinstein in 1999 [50]. By extending the work done in [51], the method was developed as an adaptive algorithm for rare event simulation using variance minimization technique. Then the method was further modified not only for estimating rare event probabilities but also for solving complex combinatorial, continuous and multi-extremal optimization

problems [32, 50, 52]. The CE method has been employed to solve many complicated optimization problems like travelling salesman problem (TSP), the max-cut problem and the quadratic assignment problem (QAP) [52]. Also, the CE method has been applied to many research problems like buffer allocation problem [1, 29], sequential change-point detection [58], vehicle routing problem [8], decision problem [37, 61], control and navigation [27], noisy optimization etc.

The CE method can be applied to both *estimation* and *optimization* problems. It uses Kullback-Leibler [33] (or cross-entropy) distance, which is considered as a fundamental concept of modern information theory. For the estimation problems, the CE method uses the Kullback-Leibler distance to measure the difference between two sampling distributions. The CE method works together with the importance sampling (IS) technique to solve rare event estimation problems.

The CE method is a fast adaptive procedure with asymptotic convergence properties [28]. With the optimization problem, first the CE method translates the problem into a rare estimation problem. The translated estimation problem is called as an associate stochastic problem (ASP) and it typically involves rare event estimation.

The CE method is an iterative procedure, which can be summarized using the following three steps.

- **Step 1:** Generate a random sample of objects (vector of parameters) according to a statistical distribution.

- **Step 2:** Obtain the best performing sample of objects using the performance (objective) function.

- **Step 3:** Update the parameters of the statistical distribution in Step 1, using the sample from Step 2, to produce a better sample at the next iteration.

The CE method can be applied for both deterministic and stochastic combinatorial optimization problems. For stochastic optimization, the objective function should be estimated through simulations. Stochastic optimization involves optimizing *noisy* objective functions. For this case it is assumed that the objective function is corrupted with some additional noise and this can be frequently found in simulation-based problems. Examples of noisy optimization include vehicle routing problem [4, 8], buffer allocation problem [1, 29] and simulated annealing [31].

3.1.1 Importance Sampling

Importance sampling (IS) is considered as a variance reduction technique and can be seen as more efficient approach for simulation. It is considered as a most important pre-requisite for sequential Monte-Carlo. IS can be used to study one distribution while sampling from another. Suppose we want to find l, the expected performance of a stochastic system. Let

$$l = \mathbb{E}_f H(X) = \mathbb{E}_f \Psi\Big(S(X);\gamma\Big) = \int \Psi\Big(S(x);\gamma\Big) f(x)\mu(dx).$$

Here
S is the sample performance function.
$\Psi\Big(\cdot;\gamma\Big)$ is the real-valued function of the sample performance which depends on γ.
f is the density of X with respect to some measure μ.

Let g be another probability density such that H is dominated by g. Using g, we can write

$$l = \int H(x)\frac{f(x)}{g(x)} g(x)\mu(dx) = \mathbb{E}_g H(X)\frac{f(X)}{g(X)}.$$

Here g is the importance sampling density and f is the nominal density. The importance sampling estimator of l is

$$\hat{l} = \frac{1}{N}\sum_{i=1}^{N} H(X_i)\frac{f(X_i)}{g(X_i)}.$$

Here $X_1, X_2, \ldots, X_N$ are iid random vectors with density g.

3.1.2 Kullback-Leibler Distance

Kullback-Leibler (KL) distance, the fundamental concept of information theory [33], is a natural distance function between the true probability distribution and the target probability distribution. The KL distance is also called KL divergence, relative entropy and discrimination information. The principles of KL divergence is used in developing the CE method.

For two probability distributions P and Q with probability distribution functions p and q on the sample sample space $\mathcal{X}$, the KL divergence is defined as

$$\mathcal{D}_{KL}(P,Q) = \mathbb{E}_P \ln\Big(\frac{p(X)}{q(X)}\Big).$$

The KL divergence can be defined for discrete and continuous distributions separately as follows.

$$\mathcal{D}_{KL}(P,Q) = \begin{cases} \sum_{i=1}^{N} p(x_i) \ln\Big(\frac{p(x_i)}{q(x_i)}\Big), & \text{if } P \text{ and } Q \text{ are discrete distributions} \\ \int_{-\infty}^{\infty} p(x) \ln\Big(\frac{p(x)}{q(x)}\Big) dx, & \text{if } P \text{ and } Q \text{ are continuous distributions} \end{cases}$$

The following are some properties of the KL divergence.

1. KL divergence is always non-negative

$$\mathcal{D}_{KL}(P,Q) \geq 0.$$

2. In general $\mathcal{D}_{KL}(P,Q) \neq \mathcal{D}_{KL}(Q,P)$. Therefore, it is not symmetric and does not verify the triangular inequality.

3. The equality of the KL divergence reached when $P(X) = Q(X)$.

4. KL divergence is a convex function on the domain of probability distributions.

3.2 General CE Algorithm for Optimization

Consider $\mathcal{X}$, an arbitrary set of states and a real valued performance function $\mathbb{S}$ on $\mathcal{X}$. If there is only one corresponding state x^* where our maximum is attained, then we can denote our maximum as γ^*

$$\gamma^* = \max_{x \in \mathcal{X}} \mathbb{S}(x). \tag{3.1}$$

When the state space $\mathcal{X}$ is finite, our optimization problem (3.1) is referred as a discrete or combinatorial optimization problem. To apply the CE method, first our deterministic optimization problem is translated into an estimation problem, so-called an associate stochastic problem (ASP). To develop the ASP for our deterministic optimization problem, we consider a collection of indicator functions $I_{\{\mathbb{S}(X) \geq \gamma\}}$ on $\mathcal{X}$ for levels $\gamma \in \mathbb{R}$. We define family of probability density functions (pdfs) $\{f(\cdot; u), u \in \mathcal{U}\}$ on the set $\mathcal{X}$. For a certain $u \in \mathcal{U}$, associating (3.1), the following estimation problem

$$l(\gamma) = \mathbb{P}_u(\mathbb{S}(X) \geq \gamma) = \mathbb{E}_u I_{\{\mathbb{S}(X) \geq \gamma\}} \tag{3.2}$$

is considered as a rare event estimation problem. Here X is a random vector with pdf $f(\cdot; u)$ for some $x \in \mathcal{X}$, $\mathbb{P}_u$ is a probability measure associated with $f(\cdot; u)$, $\mathbb{E}_u$ is the expectation operator and γ is a known or an unknown parameter.

For a certain γ which will be closer to γ^*, estimating l is a rare event estimation problem. For that, the CE method uses the KL divergence in a multi-level algorithm. It constructs a sequence of levels $\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_t$ and reference parameters (vectors) $\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_t$ corresponding to the state x^* that gives the highest performance such that $\hat{\gamma}_t$ converges to the optimal γ^* and $\hat{u}_t$ converges to the optimal reference vector u^* [53]. That is, for $\gamma = \gamma^*$, $l(\gamma)$ in equation (3.2) can be estimated using a log-likelihood estimator with reference parameter u^*.

$$u^* = \arg\max_u \ \mathbb{E}_u I_{\{\mathbb{S}(X) \geq \gamma\}} \ln f(X, u).$$

This parameter u^* can be estimated by

$$\hat{u}^* = \arg\max_u \frac{1}{M} \sum_{i=1}^{M} I_{\{\mathbb{S}(X_i) \geq \gamma\}} \ln f(X_i, u).$$

Here $X_1, X_2, \ldots, X_M$ are generated from pdf $f(\cdot; u)$.

The general CE algorithm can be described as follows.

Algorithm 1: General CE algorithm

---

1. **Initialization:** Define an initial parameter vector $\hat{u_0}$ of the statistical distribution. Set the iteration counter $t = 1$.

2. **Adaptive updating of γ_t:** Generate $X_1, X_2, \ldots, X_M$ from the density $f(\cdot; u)$ and calculate the performances $\mathbb{S}(X_1), \mathbb{S}(X_2), \ldots, \mathbb{S}(X_M)$. Sort the values in an increasing order $\mathbb{S}_{(1)} \leq \mathbb{S}_{(2)} \leq \cdots \leq \mathbb{S}_{(M)}$, where $\mathbb{S}_{(i)}$ is the i-th order statistic of the performances $\mathbb{S}(X_1), \mathbb{S}(X_2), \ldots, \mathbb{S}(X_M)$.

   Let γ_t be the $(1-\rho)$- quantile of $\mathbb{S}(X)$ under u_{t-1} which satisfies

$$\mathbb{P}_{u_{t-1}}(\mathbb{S} \geq \gamma_t) \geq \rho,$$
$$\mathbb{P}_{u_{t-1}}(\mathbb{S} \leq \gamma_t) \geq 1 - \rho.$$

   The estimate of γ_t is

$$\hat{\gamma}_t = \mathbb{S}_{(\lceil (1-\rho)M \rceil)}.$$

3. **Adaptive updating of u_t:** For fixed γ_t and u_{t-1}, obtain estimates for u_t applying the CE method to the optimization problem

$$\hat{u} = \arg\max_u \mathbb{E}_{u_{t-1}} I_{\{\mathbb{S}(X) \geq \gamma_t\}} \ln f(X, u).$$

Update the parameters of the statistical distribution using the best performing (elite) sample.

$$\hat{u} = \arg\max_{u} \frac{1}{M} \sum_{i=1}^{M} I_{\{\mathbb{S}(X_i) \geq \hat{\gamma}_t\}} \ln f(X_i, u). \tag{3.3}$$

4. **Smoothed updating of u_t:** Instead of updating the parameter vector $\hat{u}_t$ directly, the following smoothed updating procedure can be used.

$$\hat{u}_t = \alpha \tilde{u}_t + (1 - \alpha)\hat{u}_{t-1}$$

Here $\tilde{u}_t$ is a vector derived using equation (3.3) and α is the smoothing coefficient ($0 \leq \alpha \leq 1$). For best results it is empirically found that α should be $0.4 \leq \alpha \leq 0.9$. If there are two or more optimal solutions the smoothed updating of u_t helps CE method to prevent from converging to sub-optimal solutions [15]. If $\alpha = 1$, the updating will not be smoothed.

5. If the specified stopping criterion is met, obtain the solution which maximizes the performance function. Otherwise, increase $t = t + 1$, and repeat steps 2 to 4 until the stopping criterion is met.

To complete specification of the algorithm, initial parameters $\hat{u}_0, M, \rho$ and a stopping criterion should be supplied.

3.2.1 CE Method for Noisy Optimization

The CE method is an effective method to handle noisy optimization, that is, when the objective function is corrupted with noise. For example, when the objective function value obtained via simulations [1, 14]. Noisy optimization may also be called stochastic optimization. The general CE algorithm can easily be modified to be used for noisy optimization

problems [52].

For the maximization problem in (3.1), consider our performance function $\mathbb{S}(x)$ corrupted with noise. For example, we can consider another random vector Y with a pdf which may depend on x. Then we can write the noisy optimization function

$$\hat{\mathbb{S}}(x) = \hat{\mathbb{S}}(x, Y).$$

Our aim is to solve the maximization problem

$$\max_{x \in \mathcal{X}} \mathbb{E}\,\hat{\mathbb{S}}(x, Y).$$

When $\mathbb{S}(x) = \mathbb{E}\,\hat{\mathbb{S}}(x)$ is not available, we can use the sample value $\hat{\mathbb{S}}(x)$ as an unbiased estimate of $\mathbb{E}\,\hat{\mathbb{S}}(x)$. Therefore $\mathbb{S}(x) = \mathbb{E}\,\hat{\mathbb{S}}(x) = \mathbb{E}\,\hat{\mathbb{S}}(x, Y)$ for all x and $\mathbb{S}(x)$ has an optimal value of γ^* corresponding to the state x^* that gives the highest performance. Therefore the steps for updating γ_t and u_t in the noisy optimization algorithm can be considered as follows.

1. **Adaptive updating of γ_t:** Let γ_t be a $(1-\rho)$-quantile of $\hat{\mathbb{S}}(x)$ under u_{t-1}, for a fixed u_{t-1}. Then the estimate of γ_t is

   $$\hat{\gamma}_t = \hat{\mathbb{S}}_{(\lceil (1-\rho)M \rceil)}.$$

   Here $\hat{\mathbb{S}}_{(i)}$ is the i^{th} order statistic of the performances $\hat{\mathbb{S}}(X_1), \hat{\mathbb{S}}(X_2), \ldots, \hat{\mathbb{S}}(X_M)$ for a random sample of $X_1, X_2, \ldots, X_M$ from $f(\cdot; u_{t-1})$ (Similar to Step 2 in Algorithm 1).

2. **Adaptive updating of u_t:** For fixed γ_t and u_{t-1}, obtain estimates for u_t applying CE method to

   $$\hat{u} = \arg\max_{u} \mathbb{E}_{u_{t-1}} I_{\{\hat{\mathbb{S}}(X) \geq \gamma_t\}} \ln f(X, u).$$

   For fixed $\hat{\gamma}_t$ and $\hat{u}_{t-1}$ update the parameters of the statistical distribution for the elite sample using

$$\hat{u} = \arg\max_{u} \frac{1}{M} \sum_{i=1}^{M} I_{\{\hat{\mathbb{S}}(X_i) \geq \hat{\gamma}_t\}} \ln f(X_i, u).$$

CE Method for Optimal Stopping Problem

Consider a problem of finding a set of thresholds that maximizes the value of the game for an optimal stopping problem. The maximization problem was developed considering random variables $Y_1, Y_2, \ldots, Y_N$ from the standard uniform distribution $U(0,1)$. We can write the maximization problem to obtain the set of thresholds $x^* = ({x_1}^*, {x_2}^*, \ldots, {x_N}^*)$ as

$$\max_{x \in \mathcal{X}} \mathbb{E}\, \hat{\mathbb{S}}(x, Y).$$

Here $\mathcal{X} = \{x = (x_0, x_1, \ldots, x_N),\ x_j \in (0,1),\ j = 1, 2, \ldots, N,\ x_0 = 0\}$, N is the number of thresholds and $\hat{\mathbb{S}}(x, Y)$ is an estimate of the value of the game (expected gain). We can describe the CE algorithm for the above problem as follows. In this algorithm we have considered the normal distribution to generate samples of X.

Algorithm 2: CE algorithm for optimal stopping problem considering normal distribution

---

1. **Initialize:** Choose initial values for

   $$\mu^{(0)} = \left({\mu_1}^{(0)}, {\mu_2}^{(0)}, \ldots, {\mu_N}^{(0)}\right) \quad \text{and} \quad \left(\sigma^2\right)^{(0)} = \left(\left({\sigma_1}^2\right)^{(0)}, \left({\sigma_2}^2\right)^{(0)}, \ldots, \left({\sigma_N}^2\right)^{(0)}\right).$$

   Set $t = 0$.

2. **Repeat:** Steps 3 to 7 until $\max_j({\sigma_j}^{(t)}) < \varepsilon$.

3. **Draw:** Increase t by 1. Generate M random vectors $X_1, X_2, \ldots, X_M$ from the multivariate normal distribution with parameters $\left(\mu^{(t-1)}, \left(\sigma^2\right)^{(t-1)}\right)$, that is, for all $j = 1, 2, \ldots, N$, draw X_{ij} from $\mathsf{Normal}\left({\mu_j}^{(t-1)}, \left({\sigma_j}^2\right)^{(t-1)}\right), i = 1, 2, \ldots, M$.

4. **Evaluate:** Evaluate the performance of each $X_i, i = 1, 2, \ldots, M$: $\hat{\mathbb{S}}(X_1), \hat{\mathbb{S}}(X_2), \ldots, \hat{\mathbb{S}}(X_M)$.

5. **Select:** Let I be the indices of the N^{elite} best performing (elite) samples. Let $N^{elite} = \rho M$ be the size of the elite sample.

6. **Update:** for all $j = 1, 2, ..., N$, estimate $\tilde{\mu}_j^{(t)}$ and $(\tilde{\sigma}_j^2)^{(t)}$ using elite sample

$$\tilde{\mu}_j^{(t)} = \sum_{i \in I} X_{ij} / N^{elite}$$

and

$$(\tilde{\sigma}_j^2)^{(t)} = \sum_{i \in I} (X_{ij} - {\mu_j}^{(t)})^2 / N^{elite}.$$

7. **Smooth:**

$$\mu^{(t)} = \alpha \tilde{\mu}^{(t)} + (1 - \alpha)\mu^{(t-1)},$$

$$(\sigma^2)^{(t)} = \alpha (\tilde{\sigma^2})^{(t)} + (1 - \alpha)(\sigma^2)^{(t-1)}.$$

3.3 Statistical Distributions to Simulate Samples for the CE Method

When using the CE method to obtain thresholds in the optimal stopping problem, we need to generate random samples according to a statistical distribution. In our study we have considered three distributions to generate random samples with the aim of comparing the performance of the proposed CE algorithms. They are normal, truncated normal and beta distributions.

3.3.1 Normal Distribution

A random variable X has a normal (or Gaussian) distribution whose probability density function is defined for any point of $x\,(-\infty < x < \infty)$ by

$$f(x|\mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\Big[\frac{-(x-\mu)^2}{2\sigma^2}\Big].$$

Here $\mu\,(-\infty < \mu < \infty)$ is the mean and $\sigma^2\,(\sigma^2 > 0)$ is the variance of the the distribution.

3.3.2 Truncated Normal Distribution

A truncated normal distribution, which is defined on the range (L, U) is a modified version of the normal distribution. According to [7], depending on the truncated range we can consider four cases of the distribution.

1. Nontruncated case when $L = -\infty$ and $U = \infty$.
2. Lower truncated case when $L > -\infty$ and $U = \infty$.
3. Upper truncated case when $L = -\infty$ and $U < \infty$.
4. Doubly truncated case when $L > -\infty$ and $U < \infty$.

For our study, we have considered the fourth case which is the doubly truncated case within the range $(0, 1)$. The truncated normal probability density function is denoted by $f(x|\mu, \sigma^2, L, U)$. Here, μ and σ^2 are the parameters for the mean and the variance of the distribution whereas, L and U give the lower and the upper bound of the truncated range. The probability density function for the truncated normal distribution can be given by

$$f(x|\mu,\sigma^2,L,U) = \begin{cases} 0, & \text{if } L \geq x \\ \dfrac{(2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[\frac{-(x-\mu)^2}{(2\sigma^2)}\right]}{\Phi\left(\frac{U-\mu}{\sigma}\right) - \Phi\left(\frac{L-\mu}{\sigma}\right)}, & \text{if } L < x < U \\ 0, & \text{if } x \geq U \end{cases}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

3.3.3 Beta Distribution

A random variable X has a beta distribution with parameters α and β with a probability density function is defined by

$$f(x|\alpha,\beta) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & \text{for } 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Here $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(\theta) = \int_0^\infty y^{\theta-1} e^{-y} dy, \qquad \theta > 0.$$

3.4 Convergence Properties of the CE Algorithm

The asymptotic convergence properties of the CE algorithm is discussed in detail in [15, 25, 38]. The idea and the technique was mainly developed in [25] considering two ant algorithms, and [38] has done several additional modifications to speed up the CE method and to improve the convergence. In [38], the author derives two graph-based CE algorithms which can be considered to be very close to ant algorithms in [25]. The two modifications were named as the conservative modification and the conservative modification with lower bound. Author further proved asymptotic convergence of these modified algorithms.

The theoretical results in convergence properties of the CE algorithm in discrete optimization was discussed in [15]. The authors have showed that the CE algorithm with a constant smoothing parameter has a significantly faster rate of convergence compare to the CE algorithm with decreasing smoothing schemes. They prove several theorems on the convergence of the CE method and illustrate them using numerical examples.

In this thesis, we develop CE algorithms for optimal stopping problems. These may be considered as discrete optimization problems. Therefore, we can refer to the convergence properties discussed in [15, 25, 38].

4

Simulation Study

In this chapter of the thesis, we perform a simulation study using the algorithms discussed in Chapter 3. All the data are simulated and analysed using the statistical software R.

4.1 The Cross-Entropy Method for Independent Random Variables

We consider an optimal stopping problem for a sequence of independent uniformly distributed random variable over the interval $(0, 1)$. For this study, we have considered $N = 7$ consecutive time points for $K = 1000$ simulation paths. We obtained the set of thresholds $v^* = ({v_1}^*, {v_2}^*, \ldots, {v_N}^*)$ at each time point and the value of the game v. The objective function is to maximise the expected gain; see equations (1.1) and (1.2).

We developed two versions the CE method for the optimal stopping problem. We use the following parametrisation of the algorithm: $M = 1000$, $\rho = 0.05$, $\alpha = 0.7$ and $\epsilon = 0.01$. The two versions of the CE algorithm are described below.

4.1.1 The Algorithm for Non-ordered Thresholds

To perform the CE algorithm for optimal stopping problem, we generate random samples from three different distributions for the step 3 in Algorithm 2, which we have discussed in Chapter 2 of this thesis.

1. **Normal distribution (CE-N):** As we have mentioned in algorithm 2, we have considered the normal distribution with initial parameter vectors

$$\mu^{(0)} = \left(\mu_1{}^{(0)}, \mu_2{}^{(0)}, \ldots, \mu_N{}^{(0)}\right) \quad \text{and} \quad (\sigma^2)^{(0)} = \left((\sigma_1{}^2)^{(0)}, (\sigma_2{}^2)^{(0)}, \ldots, (\sigma_N{}^2)^{(0)}\right),$$

   where $\mu_i{}^{(0)} = (\sigma_i{}^2)^{(0)} = 0.5$ for $i = 1, 2, \ldots, N$.

2. **Truncated normal distribution (CE-TN):** We have considered the truncated normal distribution with initial parameter vectors $\mu_i{}^{(0)} = (\sigma_i{}^2)^{(0)} = 0.5$ for $i = 1, 2, \ldots, N$ within the interval 0 and 1.

3. **Beta distribution (CE-B):** As the third distribution, we have considered beta distribution with initial shape parameters $\alpha^* = \beta^* = 1$. Then to initialize $\mu_i{}^{(0)}$ and $(\sigma_i{}^2)^{(0)}$ we have considered the relationships

$$\mu^{(0)} = \frac{\alpha^*}{\alpha^* + \beta^*}$$

   and

$$(\sigma_i{}^2)^{(0)} = \frac{\alpha^*\beta^*}{(\alpha^* + \beta^*)^2(\alpha^* + \beta^* + 1)}.$$

In each iteration we estimate the two shape parameters α^* and β^* for the elite sample using

$$\alpha^{*(t)} = \left[\frac{(1-\mu^{(t)})}{\sigma^{2(t)}} - \frac{1}{\mu^{(t)}}\right](\mu^{(t)})^2$$

and

$$\beta^{*(t)} = \alpha^{*(t)}\left[\frac{1}{\mu^{(t)}} - 1\right]$$

and update the current parameter set. Here t indicates the iteration count.

The R code for the CE algorithm for non-ordered thresholds considering all 3 distributions can be found in the appendix of the thesis.

The accuracy of the estimated set of thresholds and the value of the game was checked using the root mean squared error (RMSE). The RMSE for the estimated values can be obtained by using

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\text{true}_i - \text{estimate}_i\right)^2}{N}}. \tag{4.1}$$

The true values can be obtained using equation (1.4) in Chapter 1. The algorithm was performed for 50 repetitions and we obtained the distribution of the average RMSE values for the estimated gain at each point. Table 4.1 represents the true values for each time point and the average values of the estimated gain we obtained by using each distribution.

It is clear that the estimated thresholds are very close to the true values of the thresholds for each time point, which shows that the algorithm performs very well. We have represented the behaviour of the estimated thresholds compared with the true values in Figure 4.1. Figure 4.2 displays the density for average RMSE obtained for three distributions using

the CE method for non-ordered thresholds.

TABLE 4.1: True values and the average estimates of the thresholds for each time point produced by the CE method for non-ordered thresholds

Time Point n	True Value	Average of the Estimates		
		CE-N	CE-TN	CE-B
1	0.500	0.513	0.514	0.509
2	0.625	0.635	0.638	0.626
3	0.695	0.697	0.692	0.706
4	0.742	0.749	0.752	0.744
5	0.775	0.782	0.777	0.778
6	0.800	0.809	0.803	0.807
7	0.820	0.820	0.817	0.824

4.1.2 The Algorithm for Ordered Thresholds

The CE algorithm was then modified for ordered thresholds. That is, in each iteration the randomly generated samples X_i's for thresholds were sorted in an ascending order before evaluating performance of each X_i using step 4 in Algorithm 2. In order to compare the performance of this algorithm with the algorithm for non-ordered thresholds, we used normal (CE-No), truncated normal (CE-TNo) and beta (CE-Bo) distributions with same initial parameters ("o" in the abbreviations stands for "ordered"). Algorithm was performed for 50 repetitions for the same simulated data values which we have used in the algorithm for non-ordered thresholds. The R code for the CE algorithm for ordered thresholds considering all 3 distributions can be found in the appendix of the thesis.

Average of the estimated thresholds from 50 repetitions are recorded in Table 4.2. The RMSE for the estimated thresholds were calculated using equation (4.1). We can clearly see that the estimated thresholds converge to their true values. The behaviour of the estimated

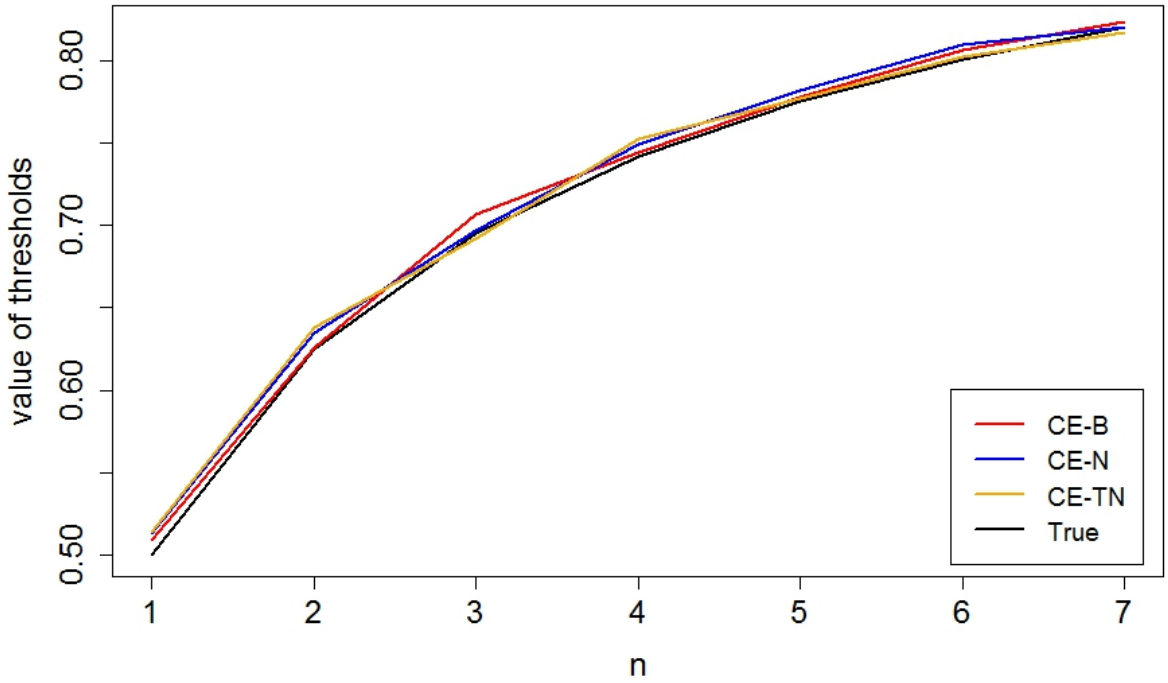

Figure 4.1: Values of the estimated thresholds using the CE method for non-ordered thresholds

thresholds with their true values are presented in Figure 4.3. The distribution of the average RMSE using the CE method for ordered thresholds is given in Figure 4.4.

To compare the convergence of the estimated thresholds by using the two algorithms, Table 4.3 summarizes the average RMSE and the standard deviations of RMSE values of the estimates obtained for the thresholds.

The value of the game v was obtained for each distribution for 50 repetitions. The true value for the value of game v is also calculated by considering each simulated data matrix in each repetition. The average of the true value of game is obtained as 0.836. The average values obtained for the estimated value of game considering both algorithms are included in Table 4.4.

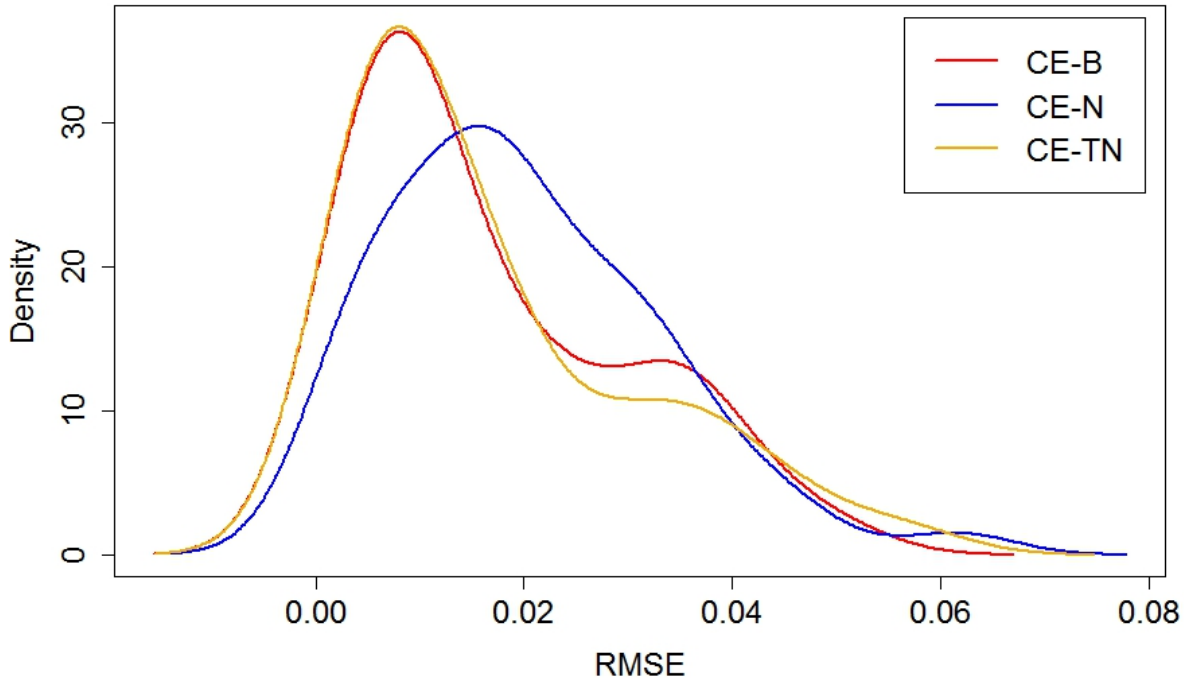

FIGURE 4.2: Density plots for average RMSE by the CE method for non-ordered thresholds

TABLE 4.2: True values and the average of the estimated gains for each time point by CE method for ordered thresholds

Time Point N	True Value	Average of the Estimates		
		CE-No	CE-TNo	CE-Bo
1	0.500	0.495	0.510	0.502
2	0.625	0.614	0.619	0.621
3	0.695	0.674	0.695	0.688
4	0.742	0.728	0.733	0.739
5	0.775	0.769	0.776	0.773
6	0.800	0.801	0.801	0.803
7	0.820	0.825	0.829	0.825

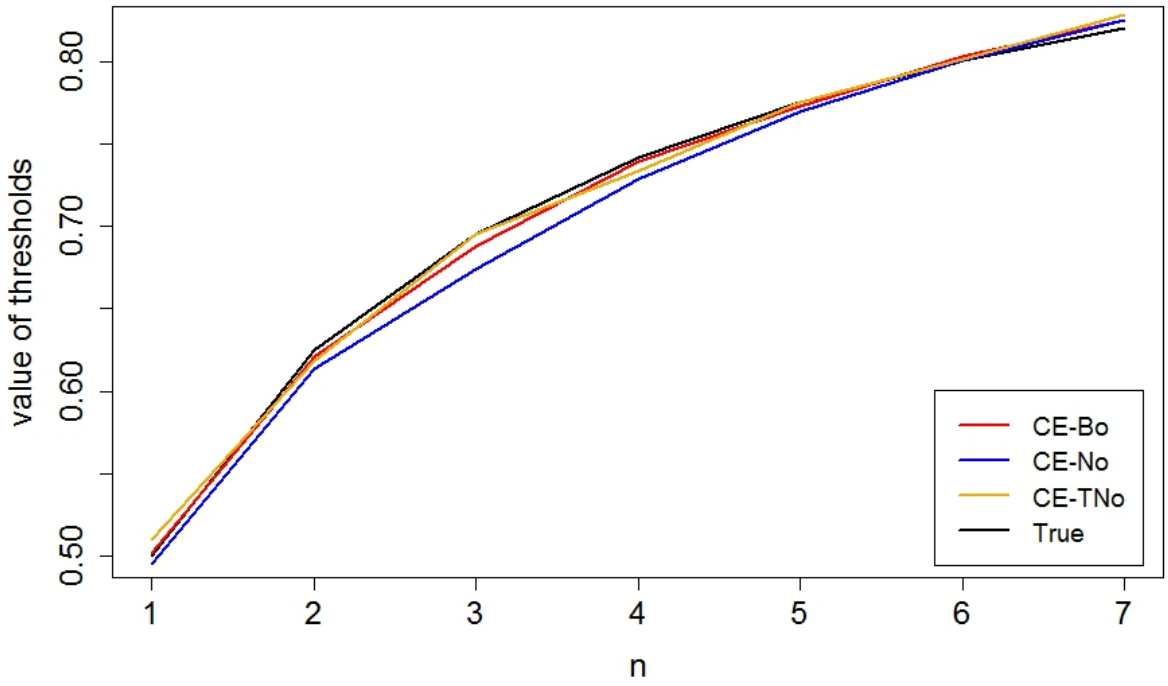

Figure 4.3: Values of the estimated thresholds using CE method for ordered thresholds

Table 4.3: Average RMSE and the standard deviations of RMSE values of the estimates using two algorithms

Algorithm	Distribution	Average RMSE	Standard Deviation of RMSE
Non-ordered thresholds	CE-N	0.020	0.013
	CE-TN	0.017	0.014
	CE-B	0.017	0.013
Ordered thresholds	CE-No	0.018	0.011
	CE-TNo	0.016	0.012
	CE-Bo	0.019	0.014

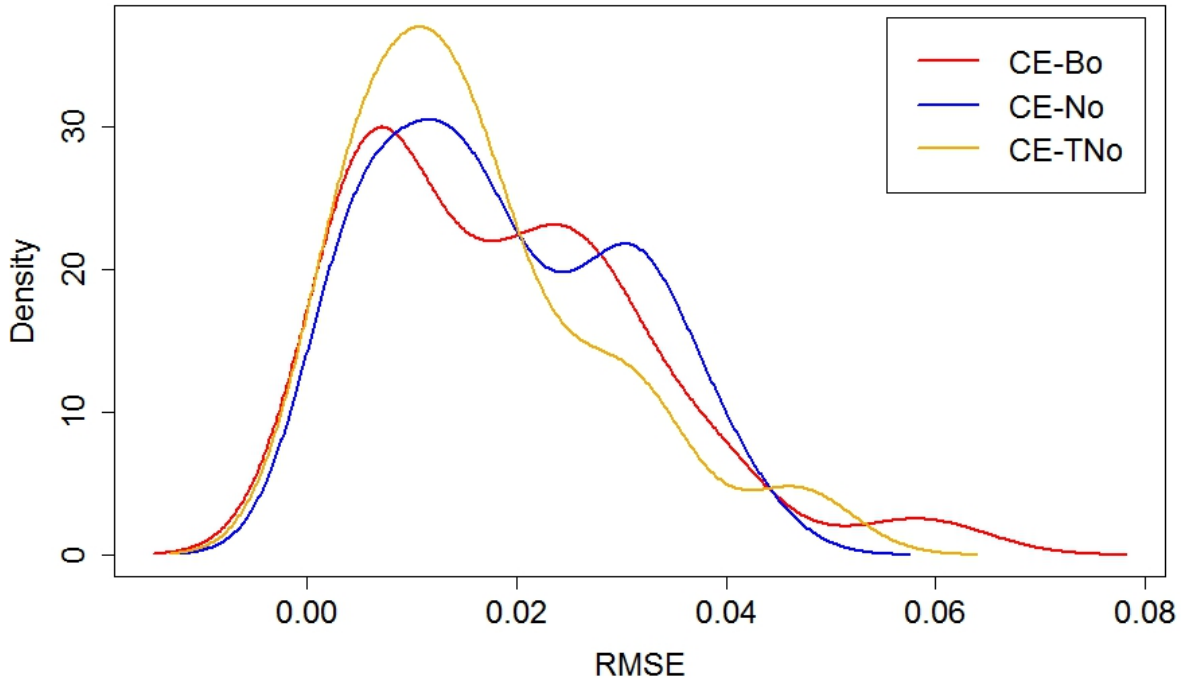

FIGURE 4.4: Density plots for average RMSE by CE method for ordered thresholds

TABLE 4.4: Average values obtained for true value of game and the estimated value of game using two algorithms

Algorithm	Distribution	Average of Estimated Value of Game
Non-ordered thresholds	CE-N	0.841
	CE-TN	0.842
	CE-B	0.842
Ordered thresholds	CE-No	0.839
	CE-TNo	0.841
	CE-Bo	0.840

Table 4.5 provides summary statistics for the iteration count and the processing time (in seconds) for the three distributions with respect to the proposed algorithms. The numerical results show that the algorithm developed for the CE method for ordered thresholds performs better than the non-ordered version of the algorithm. The processing time is relevant to a 64-bit OS, 3.60GHz Intel Core i7 processor with 16GB RAM.

Table 4.5: The iteration count and the running time for the CE method considering all distributions for both algorithms

Distribution	Iteration Count				Running time			
	Min	Max	Average	Median	Min	Max	Average	Median
CE-B	14	48	21.96	21.00	33.88	122.39	54.82	53.09
CE-Bo	11	33	16.96	15.50	29.87	103.97	52.63	50.27
CE-N	14	33	19.94	20.00	33.53	90.44	50.29	48.94
CE-No	11	25	16.44	16.00	26.30	85.69	50.38	48.54
CE-TN	13	33	19.80	19.00	33.36	79.03	49.56	48.24
CE-TNo	11	33	15.64	14.00	30.54	86.75	48.49	45.78

4.2 The Cross-Entropy Method for Dependent Random Variables

In this section, we consider an optimal stopping problem for a sequence of dependent random variables. We are particularly focussed in an autoregressive (AR) model. For our simulation study we consider an AR(1) model:

$$X_n = \phi_1 X_{n-1} + \epsilon_n,$$

where $\phi_1 = 0.5$ and $\epsilon_n \backsim \mathsf{Normal}(0, 1)$. For our study we simulate $N = 5$ consecutive time points for $K = 1000$ simulation paths. For the dependent observations we estimate the

threshold value at each time point X_n considering the values at previous time point X_{n-1}. For that, we used a linear model

$$v_n = a_{0,n} + a_{1,n} X_{N-n} \tag{4.2}$$

We develop a CE algorithm to obtain the vectors of coefficients $a_n = (a_{0,n}, a_{1,n})$ for each $n = 1, \ldots, N$ with initial values $M = 1000, \rho = 0.05$, $\alpha = 0.7$ and $\epsilon = 0.03$. The proposed CE algorithm can be described as follows.

4.2.1 The Algorithm for Estimating the Coefficients of the Polynomial

To perform the CE algorithm for an optimal stopping problem considering dependent random variables, we generate random samples for the vector of coefficients $a_n = (a_{0,n}, a_{1,n})$ using a normal distribution in step 3, Algorithm 2 (see Chapter 2). We use the following initial parametrization:

$$\mu^{(0)} = \left(\mu_1{}^{(0)}, \ldots, \mu_N{}^{(0)}\right) \quad \text{and} \quad (\sigma^2)^{(0)} = \left((\sigma_1{}^2)^{(0)}, \ldots, (\sigma_N{}^2)^{(0)}\right)$$

where $\mu_i{}^{(0)} = 0$ and $(\sigma_i{}^2)^{(0)} = 1$ for $i = 1, 2, \ldots, N$.

We update the parameters at each iteration considering the elite sample obtained by the estimated thresholds values of the model (4.2). We try to obtain the best set of coefficients which can be used to find the thresholds in order to maximize our expected gain. When updating the parameters at each iteration for the coefficients, we maximize the expected gain as a performance function. Therefore, after reaching the stopping criterion the CE method provides the coefficients for the best fitted values of the threshold.

The R code of the proposed CE algorithm can be found in the appendix of this thesis.

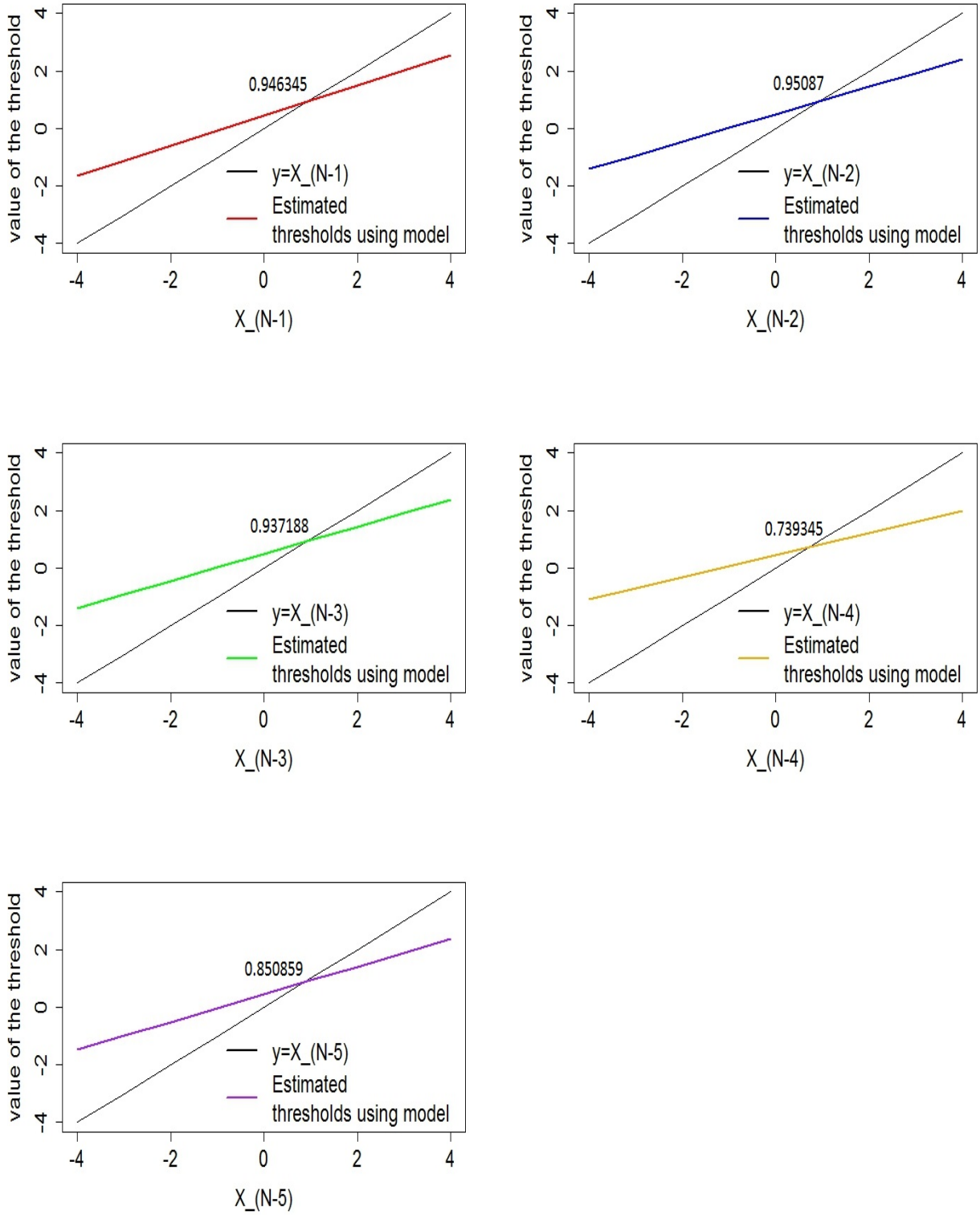

Figure 4.5: Estimated thresholds using CE method for polynomials

Figure 4.5 illustrates the estimated linear models for the thresholds. This means, for example, if we observe a value of X_{N-1} greater than 0.946345 (the intersection of the two lines), we will accept this value as our gain and stop. Otherwise, we will proceed to the next observation.

5

Discussion and Future Research

5.1 Summary

Chapter 1 of this thesis provides the general introduction and the background of our study. It discusses the nature of sequential data and the how we can make decisions using these data. It further provides details on the applicability of the sequential data analysis in many research areas and specifically how it is related with decision making. The optimal stopping problem was developed in this chapter considering discrete time stochastic sequences. It describes the theory to develop the objective function in order to maximize the expected gain. The chapter describes the backward induction method, which was used for uniformly distributed random variables to obtain the threshold values and the value of game. These results are later used in Chapter 4 to illustrate the accuracy of the estimated results. Lastly, the chapter provides brief theoretical information on sequential change-point detection, which

can be considered as an optimal stopping problem.

Chapter 2 gives a literature review on the topic. It also provides the historical background of the optimal stopping problem, which was initially discussed as the secretary problem appeared as a mathematical game in early 1960's. Then the chapter outlines applications of the the optimal stopping problem, mainly in the area of finance. It provides extensive literature on the optimal stopping problem considering various modifications and extensions in reward functions and problems involving more than one stops. It further describes several numerical methods that have recently been used to solve the optimal stopping problem. Finally, the chapter reviews literature on the Cross-Entropy (CE) method, which we are focussed at to solve the optimal stopping problem.

Chapter 3 covers the theory and methodology we have used in the thesis. It starts with the general introduction to the CE method, which was introduced by Reuvan Y. Rubinstein in 1999 [50] as an adaptive Monte-Carlo approach. The chapter gives information about the application of CE method to both estimation and optimization problems. It further describes importance sampling, a variance reduction techniques which is considered as a most important pre-requisite for sequential Monte-Carlo. Then the chapter gives details on Kullback-Leibler (KL) distance, which is used to develop the CE method. The description of the general CE algorithm for optimization is provided along with the corresponding algorithm for noisy optimization. Finally, the chapter discusses relevant literature on the convergence properties of the CE algorithm.

Chapter 4 provides the numerical results obtained from our simulation study. We considered an optimal stopping problem for independent random variables from the uniform distribution $\mathsf{Uniform}(0, 1)$. The chapter describes the method we have used to build up CE algorithms for non-ordered thresholds and ordered thresholds considering three separate distributions to generate random samples in the algorithms. We show how well the estimated values from the proposed algorithms perform when compared with the theoretical values.

Further, the chapter provides information on the speed of the performance of the two algorithms using the descriptive statistics on iteration count and the running time, relevant to a 64-bit OS, 3.60GHz Intel Core i7 processor with 16GB RAM. Finally, we consider a simulation study for dependent random variables. A CE method was developed for the coefficients for the simple model $v_n = a_{0,n} + a_{1,n}X_{N-n}$ considering an optimal stopping problem for an autoregressive model of order 1, AR(1).

5.2 Discussion

In this thesis, we applied the Cross-Entropy (CE) method as a numerical technique to solve optimal stopping problems for sequences of discrete random variables. Our objective is to maximize the expected gain. At each time point we compared the observed value with the threshold value and decided whether to stop or to continue to the next observation. We considered sequences of random variables from the uniform distribution over the interval $(0, 1)$. Using [18], we derived formulae for the thresholds at each time point and the value of the game. We carried out a simulation study considering 7 consecutive time points for 1000 simulated paths to estimate the thresholds at each time point and the value of the game using the CE method that we developed.

The first CE algorithm was developed considering non-ordered thresholds. That is, we generated random samples from a distribution with corresponding initial parameters. Then we evaluated the performance of our sample using the objective function. We selected the best performing sample as the elite sample and updated our initial parameters according to the elite sample until we met our stopping criterion. The algorithm was performed for 50 repetitions on different sequences of uniform distributed random variables. To generate the random samples, we considered, firstly, a normal distribution (CE-N); secondly, a truncated normal distribution (CE-TN) within the interval $(0, 1)$ and, thirdly, a beta distribution (CE-B). We used the Root Mean Squared Error (RMSE) to calculate the accuracy of the estimates.

For the second CE algorithm we modified the first algorithm using ordered thresholds. At each iteration the randomly generated samples were sorted in an ascending order and then we evaluated the performance of the sorted sample using the objective function. The algorithm was performed considering the same three distributions normal (CE-No), truncated normal (CE-TNo) and beta (CE-Bo). To check the accuracy of the estimates, the RMSE was used. When comparing the algorithm for non-ordered thresholds with the algorithm for ordered thresholds, it was clear that the later performs in a better way since this optimal stopping problem belongs to the monotonic case. The second algorithm provided smaller RMSE values for the estimated thresholds and for the value of the game. The iteration count and the processing time (in seconds) obtained for the three distributions with respect to the proposed algorithms further confirmed that the CE method for ordered thresholds performed better than the non-ordered version of the algorithm.

The next modification of the CE method was for optimal stopping problems with the dependent observations. We considered the optimal stopping problem for random data simulated from an AR(1) autoregressive model. Here we considered data for 5 consecutive time points and 1000 simulation paths. We considered a linear model $v_n = a_{0,n} + a_{1,n}X_{N-n}$ to estimate v_n, the value of threshold at time point $N - n$, using the observed value X_{N-n} as a covariate in the model. Our approach was somewhat similar to the method used by [36] but here we proposed to estimate the thresholds directly. We developed a CE method to obtain the best suitable coefficients for the model. We obtained the threshold values which we can be used to make decisions at each time point. It is clear that a linear model is the simplest model that depends on the observed value, so our method can be modified by using other basis functions like set of Laguerre polynomials discussed in [36].

5.3 Future Work

We have carried out a study on solving optimal stopping problems. The study presented in this thesis can be modified and generalised in many directions. The following are potential extensions and applications which we plan to develop.

Developing other numerical techniques

In this thesis we have presented only one specific numerical techniques in optimal stopping problems. We considered the Cross-Entropy (CE) method based on Monte-Carlo (MC) simulation technique for different types of stopping problems for both independent and dependent observations. We plan to extend our study to other numerical methods including Expectation-Maximization (EM) algorithm and Markov chain Monte-Carlo (MCMC). Also, we intend to apply numerical methods such as linear programming principles or dynamic programming methods to solve our problem (see, for example, [9, 26]).

Modifying gain function

For this study we have considered the observed value as a reward. As a future research, we will modify our gain function to solve the optimal stopping problems. For example, we can consider as a reward a more complicated function when obtaining the expected gain. Also, it will be possible to include cost functions at each time point when maximizing the reward.

Dependent case with more complicated polynomials

In the simulation study of this thesis, we discussed developing a CE method for a linear function to estimate threshold values in order to make the decision in our optimal stopping problem for dependent observations. For this study we considered observations from an AR(1) autoregressive model. In the future, we will consider data from some other distributions like Geometric Brownian Motion. While in this study we have started our analysis considering a simple linear function, we plan to use more complicated polynomials to estimate the thresholds in decision making. In future research we can include set of Laguerre polynomials, Hermite polynomials or Jacobi polynomials as our basis function in the model [36].

Multiple stopping problem

So far we have considered only one stop to make our decision. That is, we are comparing our value with a threshold value and we are deciding to stop at that point taking the observed value as a gain, or to proceed with the next observation. As future research, we can generalize our study to optimal stopping problems involving more than one stop (multiple stopping problems) [40, 59]. We will modify our proposed CE method and also the other numerical methods which we plan to develop for multiple stopping problems.

Application of sequential change-point detection problem

As we mentioned in Chapter 1 of this thesis, the sequential change-point problem (or the quickest change-point problem) can be considered as an optimal stopping problem (see, for example, [12]). We intend to further develop our numerical methods for the sequential change-point problem.

Developing an R package

With the extensions we have discussed above, we will be able to develop a new package within R software, in which the proposed methods and techniques will be directly used by simply inputting data into a function. The developed R package would be very useful for decision makers who are interested in optimal stopping problems.

A

Appendix - R Code

This appendix contains supplementary information on the algorithms we have developed using CE method for optimal stopping problems. We have discussed the methodology we have used to develop these algorithms and the results which we have obtained in Chapter 4 and Chapter 5.

A.1 Cross-Entropy Method for Independent Random Variables

A.1.1 Algorithm for Non-ordered Thresholds

Using Normal Distribution (CE-N)

```
library(msm)

n <- 8              # number of time points
Sim <- 1000         # number of simulated paths
a=0                 # lower value for the uniform distribution
b=1                 # upper value for the uniform distribution
d=n-1     # the length of x

M <- matrix(runif((n*Sim),min=a, max=b),ncol=n)

Gain <- function(x)  # the length of x is n-1
{
  x1 <- c(x[1:(n-1)],-Inf)
  MaxFirst <- function(y)
  {
    y[min(which(y>=x1))]
  }
  mean(apply(M,1,MaxFirst))
}

true=numeric()
true[d]=0.5
for(k in (d-1):1)
{
  true[k]=(b/2)*(1+(true[k+1]/b)^2)
}

------CE Method Using Normal Distribution-------

myfunctionCE=function(M)
{
  N = 1000 # number of observations generated for the algorithm
  rho =0.05
  Nelite =round(N*rho) # number of observations in the elite sample
  smooth=0.7
```

```
eps=0.01
d=n-1  # the length of x

mu_0_1=rep(0.5,d)
sd_0_1=rep(0.5,d)

t_1=1

mu_old_1=mu_0_1
sd_old_1=sd_0_1

t_max_1=200

ptm1_1 <- proc.time()

while ((max(sd_0_1)>eps)&(t_1<=t_max_1))
{

  Gain_est_1 <- array(0,dim=c(N,d))

  for (i in 1:d)
  {
    Gain_est_1[,i] <- matrix(rnorm(N,mean=mu_0_1[i],sd=sd_0_1[i]),ncol=1)
  }

  score_1 <- apply(Gain_est_1,1,Gain)

  score_sorted_1=sort(score_1,decreasing=TRUE,index.return=TRUE)

  elite_index_1=score_sorted_1$ix[1:Nelite]
  elite_sample_1= Gain_est_1[elite_index_1,]

  if (d>1)
  {
    mu_new_1=apply(elite_sample_1,2,mean)
    sd_new_1=apply(elite_sample_1,2,sd)

  }

  if (d==1)
  {
    mu_new_1=mean(elite_sample_1)
    sd_new_1=sd(elite_sample_1)
```

```
    }

    mu_0_1=smooth*mu_new_1+(1-smooth)*mu_old_1
    sd_0_1=smooth*sd_new_1+(1-smooth)*sd_old_1

    mu_old_1=mu_0_1
    sd_old_1=sd_0_1

    print(c(t_1,mu_0_1,sd_0_1))

    t_1=t_1+1
  }

  print(proc.time()-ptm1_1)

---------RMSE Calculation--------

  CE=as.data.frame(Gain_est_1)
  CE["RMSE"]<-NA

  SE_1=matrix(0,Sim,d)

  for(m in 1:d)
  {
    SE_1[,m]=(true[m]-Gain_est_1[,m])^2
  }

  for(m in 1:d)
  {
    for(l in 1:N)
    {
      CE$RMSE[l]=sqrt(sum(SE_1[l,m])/d)
    }
  }

  CE$Val_Game=Gain(apply(Gain_est_1,2,mean))

  return(apply(CE,2,mean))
}

------ Results for 50 Repetitions---------

getResults <- function(M)
{
```

```
  resCE <- as.numeric(myfunctionCE(M))

  return(resCE)

}

Ntimes <- 50
ResultCE=matrix(NA,nrow = Ntimes,ncol= 9)  # store results for 50
                                                    different M values
Resulttrue=matrix(NA,nrow = Ntimes,ncol= 1)

for (i in 1:Ntimes)
{
  M <- matrix(runif((n*Sim),min=a, max=b),ncol=n)

  ResultCE[i,] <- getResults(M)
}
```

Using Truncated Normal Distribution (CE-TN)

For the algorithm performed using truncated normal distribution, we changed only the `Gain_est_1` array using

```
Gain_est_2[,i] <- matrix(rtnorm(N,mean=mu_0_2[i],sd=sd_0_2[i],
                                                    lower=0,upper=1),ncol=1).
```

Using Beta Distribution (CE-B)

```
------CE Method Using Beta Distribution-------

myfunctionBE=function(M)
{
  N = 1000
  rho =0.05
  Nelite =round(N*rho)
  smooth=0.7
  eps=0.01

  alpha0=beta0=rep(1,d)
  mu0=1/(1+(beta0/alpha0))
  sd0=sqrt((alpha0*beta0)/((alpha0+beta0)^2)*(alpha0+beta0+1))

  t=1
```

```
alpha_old=alpha0
beta_old=beta0
mu_old=mu0
sd_old=sd0

t_max=200

ptm1 <- proc.time()

while ((max(sd0)>eps)&(t<=t_max))
{
  Gain_est <- array(0,dim=c(N,d))

  for (i in 1:d)
  {
   Gain_est[,i] <- matrix(rbeta(N,alpha0[i],beta0[i]),ncol=1)
  }

  score <- apply(Gain_est,1,Gain)

  score_sorted=sort(score,decreasing=TRUE,index.return=TRUE)

  elite_index=score_sorted$ix[1:Nelite]
  elite_sample= Gain_est[elite_index,]

  if (d>1)
  {
    mu_new=apply(elite_sample,2,mean)
    sd_new=apply(elite_sample,2,sd)
  }

  if (d==1)
  {
    mu_new=mean(elite_sample)
    sd_new=sd(elite_sample)
  }

  mu0=smooth*mu_new+(1-smooth)*mu_old
  sd0=smooth*sd_new+(1-smooth)*sd_old
  alpha0=(((1-mu0)/(sd0^2))-(1/mu0))*(mu0^2)
  beta0=alpha0*((1/mu0)-1)

  mu_old=mu0
  sd_old=sd0
```

```
    alpha_old=alpha0
    beta_old=beta0

    print(c(t,mu_0,sd_0))

    t=t+1
  }

  print(proc.time()-ptm1)
```

A.1.2 Algorithm for Ordered Thresholds

In the algorithm for ordered thresholds, we sorted the samples we are generating for `Gain_est` array in an ascending order before evaluating the performance using the performance function `Gain`. The modification we have done for each distribution can be represent like follows. We have only provided the modified part of the original code in this section.

```
Gain_est_sn <- array(0,dim=c(N,d))
    Gain_est_st <- array(0,dim=c(N,d))
    Gain_est_sb <- array(0,dim=c(N,d))

    -----Normal Distribution---------

    for (i in 1:d)
    {
      Gain_est_n[,i] <- matrix(rnorm(N,mean=mu_0_1[i],sd=sd_0_1[i]),ncol=1)
    }

    ----------Truncated Normal Distribution---------

    for (i in 1:d)
    {
      Gain_est_t[,i] <- matrix(rtnorm(N,mean=mu_0_2[i],sd=sd_0_2[i],
                                                    lower=0,upper=1),ncol=1)
    }

    -----------Beta Distribution-----------

    for (i in 1:d)
    {
     Gain_est_b[,i] <- matrix(rbeta(N,alpha0[i],beta0[i]),ncol=1)
    }
```

```
    ----------Sorting the generated samples in ascending order----------

    for (j in 1:N)
    {
      Gain_est_sn[j,]=sort(Gain_est_n[j,],decreasing = TRUE)
      Gain_est_st[j,]=sort(Gain_est_t[j,],decreasing = TRUE)
      Gain_est_sb[j,]=sort(Gain_est_b[j,],decreasing = TRUE)
    }
```

A.2 Cross-Entropy Method for Dependent Random Variables

```
library(msm)

n <- 5       # number of time points
K <- 1000    # number of simulated paths
d=n-1        # the length of x
p=0.5        # order of the AR model

M=matrix(NA,ncol=n,nrow=K)

for(k in 1:K)
{
  M[k,] <- arima.sim(model=list(ar=p),n=n)
}

Gain <- function(x)   # the length of x is n-1
{
  x1 <- c(x[1:(n-1)],-Inf)
  MaxFirst <- function(y)
  {
    y[min(which(y>=x1))]
  }
  mean(apply(M,1,MaxFirst))
}

----------- CE method for coefficients of the polynomial-----

myfunctionLA <- function(M)
{
N = 1000
rho =0.05
```

```
Nelite =round(N*rho)
smooth=0.7
eps=0.03
d=n-1        # the length of x
m=1          # order of the polynomial
l=m+1        # number of parameters in the polynomial

mu0_l=rep(0,n)
sd0_l=rep(1,n)

t_l=1

mu_old_l=mu0_l
sd_old_l=sd0_l

t_max_l=500

ptm1_l <- proc.time()

while ((max(sd0_l)>eps)&(t_l<=t_max_l))
{
  a_l <- matrix(NA, nrow=n, ncol=l)  # coefficient matrix
  fit_val <- matrix(NA,nrow = N, ncol=d)

  for(k in d:1)
  {
  for (i in 1:n)
  {
    for(j in 1:l)
    {
      a_l[i,j] <- matrix(rnorm(1,mean=mu0_l[j],sd=sd0_l[j]),nrow = 1)
    }
  }

  L=matrix(0,nrow=K,ncol=l)

  for(j in 1:l)
  {
    L[,j]=M[,k]^(j-1)
  }

    fit_val= L %*% t(a_l)
  }
```

```
  score_l <- apply(fit_val,1,Gain)

  score_sorted_l=sort(score_l,decreasing=TRUE,index.return=TRUE)

  elite_index_l=score_sorted_l$ix[1:Nelite]
  elite_sample_l= fit_val[elite_index_l,]

  if (d>1)
  {
    mu_new_l=apply(elite_sample_l,2,mean)
    sd_new_l=apply(elite_sample_l,2,sd)
  }

  if (d==1)
  {
    mu_new_l=mean(elite_sample_l)
    sd_new_l=sd(elite_sample_l)
  }

  mu0_l=smooth*mu_new_l+(1-smooth)*mu_old_l
  sd0_l=smooth*sd_new_l+(1-smooth)*sd_old_l

  mu_old_l=mu0_l
  sd_old_l=sd0_l

  print(c(t_l,mu0_l,sd0_l))

  t_l=t_l+1

}
proc.time()-ptm1_l
```

References

[1] Alon, G., Kroese, D. P., Raviv, T., and Rubinstein, R. Y. (2005). Application of the cross-entropy method to the buffer allocation problem in a simulation-based environment. *Annals of Operations Research*, 134(1):137–151.

[2] Andersen, T., Carstensen, J., Hernandez-Garcia, E., and Duarte, C. M. (2009). Ecological thresholds and regime shifts: approaches to identification. *Trends in Ecology & Evolution*, 24(1):49–57.

[3] Baker, M. E. and King, R. S. (2010). A new method for detecting and interpreting biodiversity and ecological community thresholds. *Methods in Ecology and Evolution*, 1(1):25–37.

[4] Bertsimas, D. J. (1992). A vehicle routing problem with stochastic demand. *Operations Research*, 40(3):574–585.

[5] Borkar, V. S., Pinto, J., and Prabhu, T. (2009). A new learning algorithm for optimal stopping. *Discrete Event Dynamic Systems*, 19(1):91–113.

[6] Boyce, W. M. (1970). Stopping rules for selling bonds. *The Bell Journal of Economics and Management Science*, pages 27–53.

[7] Burkardt, J. (2014). The truncated normal distribution. *Department of Scientific Computing, Florida State University*.

[8] Chepuri, K. and Homem-De-Mello, T. (2005). Solving the vehicle routing problem with stochastic demands using the cross-entropy method. *Annals of Operations Research*, 134(1):153–181.

[9] Cho, M. J. and Stockbridge, R. H. (2002). Linear programming formulation for optimal stopping problems. *SIAM Journal on Control and Optimization*, 40(6):1965–1982.

[10] Chow, Y., Moriguti, S., Robbins, H., and Samuels, S. (1964). Optimal selection based on relative rank (the âĂIJsecretary problemâĂİ). *Israel Journal of Mathematics*, 2(2):81–90.

[11] Chow, Y. S. and Robbins, H. (1963). On optimal stopping rules. *Probability Theory and Related Fields*, 2(1):33–49.

[12] Chow, Y. S., Robbins, H., and Siegmund, D. (1971). *Great expectations: The theory of optimal stopping*. Houghton Mifflin Co., Boston, Mass.

[13] Chun, Y. H., Plante, R. D., and Schneider, H. (2002). Buying and selling an asset over the finite time horizon: A non-parametric approach. *European Journal of Operational Research*, 136(1):106–120.

[14] Cohen, I., Golany, B., and Shtub, A. (2007). Resource allocation in stochastic, finite-capacity, multi-project systems through the cross entropy methodology. *Journal of Scheduling*, 10(3):181–193.

[15] Costa, A., Jones, O. D., and Kroese, D. (2007). Convergence properties of the cross-entropy method for discrete optimization. *Operations Research Letters*, 35(5):573–580.

[16] Darling, D. A., Liggett, T., and Taylor, H. M. (1972). Optimal stopping for partial sums. *The Annals of Mathematical Statistics*, pages 1363–1368.

[17] David, I. (1998). Explicit results for a class of asset-selling problems. *European journal of operational research*, 110(3):576–584.

[18] DeGroot, M. H. (1970). *Optimal statistical decisions*. McGraw-Hill Book Co., New York-London-Sydney.

[19] Elgersma, O. E., van Leersum, M., Buijs, P. C., van Leeuwen, M. S., van de Schouw, Y. T., Eikelboom, B. C., and van der Graaf, Y. (1998). Changes over time in optimal duplex threshold for the identification of patients eligible for carotid endarterectomy. *Stroke*, 29(11):2352–2356.

[20] Feinberg, E. A. and Shiryaev, A. N. (2006). Quickest detection of drift change for brownian motion in generalized bayesian and minimax settings. *Statistics & Decisions*, 24(4/2006):445–470.

[21] Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical science*, pages 282–289.

[22] Gramacy, R. B. and Ludkovski, M. (2015). Sequential design for optimal stopping problems. *SIAM Journal on Financial Mathematics*, 6(1):748–775.

[23] Griffeath, D. and Snell, J. L. (1974). Optimal stopping in the stock market. *The Annals of Probability*, pages 1–13.

[24] Guo, X. and Shepp, L. (2001). Some optimal stopping problems with nontrivial boundaries for pricing exotic options. *Journal of Applied Probability*, 38(03):647–658.

[25] Gutjahr, W. J. (2002). Aco algorithms with guaranteed convergence to the optimal solution. *Information processing letters*, 82(3):145–153.

[26] Helmes, K. (2002). Numerical methods for optimal stopping using linear and non-linear programming. In Pasik-Duncan, B., editor, *Stochastic Theory and Control: Proceedings of a Work-shop held in Lawrence, Kansas*, pages 185–203. Springer, Berlin Heidelberg, Berlin, Heidelberg.

[27] Helvik, B. E. and Wittner, O. (2001). Using the cross-entropy method to guide/govern mobile agentâĂŹs path finding in networks. In *International Workshop on Mobile Agents for Telecommunication Applications*, pages 255–268. Springer.

[28] Homem-de Mello, T. and Rubinstein, R. Y. (2002). Estimation of rare event probabilities using cross-entropy. In Yücesan, E., Chen, C. H., Snowdon, J. L., and MCharnes, J., editors, *Proc. 2002 Winter Simulation Conf., San Diego, CA*, pages 310–319.

[29] Hui, K.-P., Bean, N., Kraetzl, M., and Kroese, D. P. (2005). The cross-entropy method for network reliability estimation. *Annals of Operations Research*, 134(1):101.

[30] Jacka, S. (1991). Optimal stopping and the american put. *Mathematical Finance*, 1(2):1–14.

[31] Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., et al. (1983). Optimization by simulated annealing. *science*, 220(4598):671–680.

[32] Kroese, D. P., Porotsky, S., and Rubinstein, R. Y. (2006). The cross-entropy method for continuous multi-extremal optimization. *Methodology and Computing in Applied Probability*, 8(3):383–407.

[33] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

[34] Lambin, E. F. (1997). Modelling and monitoring land-cover change processes in tropical regions. *Progress in physical Geography*, 21(3):375–393.

[35] Lindley, D. V. (1961). Dynamic programming and decision theory. *Applied Statistics*, pages 39–51.

[36] Longstaff, F. A. and Schwartz, E. S. (2001). Valuing american options by simulation: a simple least-squares approach. *Review of Financial studies*, 14(1):113–147.

[37] Mannor, S., Rubinstein, R. Y., and Gat, Y. (2003). The cross entropy method for fast policy search. In *ICML*, pages 512–519.

[38] Margolin, L. (2005). On the convergence of the cross-entropy method. *Annals of Operations Research*, 134(1):201–214.

[39] Mateuca, R., Lombaert, N., Aka, P., Decordier, I., and Kirsch-Volders, M. (2006). Chromosomal changes: induction, detection methods and applicability in human biomonitoring. *Biochimie*, 88(11):1515–1531.

[40] Nikolaev, M. and Sofronov, G. Y. (2007). A multiple optimal stopping rule for sums of independent random variables. *Discrete Mathematics and Applications dma*, 17(5):463–473.

[41] Novikov, A. A. and Shiryaev, A. N. (2005). On an effective solution of the optimal stopping problem for random walks. *Theory of Probability & Its Applications*, 49(2):344–354.

[42] Page, E. (1954). Continuous inspection schemes. *Biometrika*, 41(1/2):100–115.

[43] Pearce, C. E., Szajowski, K., and Tamaki, M. (2012). The duration problem with multiple exchanges. *Numer. Algebra Control Optim.*, 2(2).

[44] Pham, H. (1997). Optimal stopping, free boundary, and american option in a jump-diffusion model. *Applied mathematics & optimization*, 35(2):145–164.

[45] Polushina, T. V. (2009). Estimating optimal stopping rules in the multiple best choice problem with minimal summarized rank via the cross-entropy method. In *Evolutionary Computation, 2009. CEC'09. IEEE Congress on*, pages 1668–1674. IEEE.

[46] Porosinski, Z. (1987). The full-information best choice problem with a random number of observations. *Stochastic processes and their applications*, 24(2):293–307.

[47] Preiss, M., Gray, D., and Stacy, N. (2003). Change detection in repeat pass interferometric synthetic aperture radar. In *Radar Conference, 2003. Proceedings of the International*, pages 98–102. IEEE.

[48] Radke, R. J., Andra, S., Al-Kofahi, O., and Roysam, B. (2005). Image change detection algorithms: a systematic survey. *IEEE transactions on image processing*, 14(3):294–307.

[49] Roberts, S. (1959). Control chart tests based on geometric moving averages. *Technometrics*, 1(3):239–250.

[50] Rubinstein, R. (1999). The cross-entropy method for combinatorial and continuous optimization. *Methodology and computing in applied probability*, 1(2):127–190.

[51] Rubinstein, R. Y. (1997). Optimization of computer simulation models with rare events. *European Journal of Operational Research*, 99(1):89–112.

[52] Rubinstein, R. Y. and Kroese, D. P. (2004). *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer Science & Business Media.

[53] Rubinstein, R. Y. and Kroese, D. P. (2016). *Simulation and the Monte Carlo method*. John Wiley & Sons.

[54] Shiryaev, A. (1961). The problem of the most rapid detection of a disturbance in a stationary process. In *Soviet Math. Dokl*, volume 2.

[55] Shiryaev, A. N. (1963). On optimum methods in quickest detection problems. *Theory of Probability & Its Applications*, 8(1):22–46.

[56] Sofronov, G. (2013). An optimal sequential procedure for a multiple selling problem with independent observations. *European Journal of Operational Research*, 225(2):332–336.

[57] Sofronov, G., Keith, J. M., and Kroese, D. P. (2006a). An optimal sequential procedure for a buying-selling problem with independent observations. *Journal of applied probability*, pages 454–462.

[58] Sofronov, G., Polushina, T., and Priyadarshana, M. (2012). Sequential change-point detection via the cross-entropy method. In *Neural Network Applications in Electrical Engineering (NEUREL), 2012 11th Symposium on*, pages 185–188. IEEE.

[59] Sofronov, G. Y. (2016). A multiple optimal stopping rule for a buying–selling problem with a deterministic trend. *Statistical Papers*, 57(4):1107–1119.

[60] Sofronov, G. Y., Kroese, D., Keith, J., Nikolaev, M., et al. (2006b). Simulations of thresholds in multiple best choice problem. *Obozr. Prikl. Prom. Mat.*, 13(6):975–982.

[61] Sofronov, G. Y. and Polushina, T. V. (2013). Evaluating optimal stopping rules in the multiple best choice problem using the cross-entropy method. In Klement, E. P., Borutzky, W., Fahringer, T., Hamza, M. H., and Uskov, V., editors, *MIC 2013: proceedings of the IASTED International Conference on Modelling, Identification and Control : 11-13 February 2013, Innsbruck, Australia*, pages 205–212.

[62] Tsitsiklis, J. N. and Van Roy, B. (1999). Optimal stopping of markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives. *IEEE Transactions on Automatic Control*, 44(10):1840–1851.

[63] Yakir, B. (1994). Optimal detection of a change in distribution when the observations form a markov chain with a finite state space. *Lecture Notes-Monograph Series*, pages 346–358.

[64] Yu, H. and Bertsekas, D. P. (2006). A least squares q-learning algorithm for optimal stopping problems. *Lab. for Information and Decision Systems Report*, 2731.