Revision Moment for the Retail Decision-Making System

Master’s Thesis in Financial Mathematics

Agnieszka Beata Juszczuk and Evgeniya Tkacheva
Revision Moment for the Retail Decision-Making System

Agnieszka Beata Juszczuk and Evgeniya Tkacheva

Halmstad University
Project Report IDE1026

Master’s thesis in Financial Mathematics, 15 ECTS credits
Supervisor: Ph.D. Mikhail Nechaev
Examiner: Prof. Ljudmila A. Bordag
External referees: Prof. Mikhail Babich

October 26, 2010

Department of Mathematics, Physics and Electrical Engineering
School of Information Science, Computer and Electrical Engineering
Halmstad University
Preface

We would like to thank our supervisor Mikhail Nechaev for very interesting theme for the research and for his useful comments during our work on this thesis. Also we would like to thank the programme coordinator prof. Ljudmila Bordag for the opportunity to study at the programme of Financial Mathematics and for her advices during our study. Especial thanks to Ilya Tkachev for the help with this thesis project, and also we would like thank parents and friends for support, that they give us everyday.
Abstract

In this work we address to the problems of the loan origination decision-making systems. In accordance with the basic principles of the loan origination process we considered the main rules of a clients parameters estimation, a change-point problem for the given data and a disorder moment detection problem for the real-time observations. In the first part of the work the main principles of the parameters estimation are given. Also the change-point problem is considered for the given sample in the discrete and continuous time with using the Maximum likelihood method. In the second part of the work the disorder moment detection problem for the real-time observations is considered as a disorder problem for a non-homogeneous Poisson process. The corresponding optimal stopping problem is reduced to the free-boundary problem with a complete analytical solution for the case when the intensity of defaults increases. Thereafter a scheme of the real time detection of a disorder moment is given.
Contents

1 Introduction .................................................. 1
   1.1 A Logistic Regression model ............................ 2
   1.2 A posteriori change-point problem for a given sample .. 2
   1.3 The Poisson disorder problem .......................... 2

2 The Multiple Logistic regression .......................... 5
   2.1 The real data ............................................ 5
   2.2 The multiple logistic regression ......................... 6
   2.3 The fitting procedure to the multiple Logistic Regression model 7
   2.4 The testing procedure for the significance of the model ... 8
   2.5 The area under the ROC curve .......................... 9

3 The A-posteriori change-point problem for a given sample 13
   3.1 The Maximum likelihood method ......................... 13
   3.2 The formulation of the problem in the discrete time ...... 14
   3.3 An exact test of to have no change against to have change .. 15
   3.4 The estimation of \( \tau \) .................................. 17
   3.5 The change-point problem in the model of the logistic regression 20
   3.6 The procedure for the known initial parameters and unknown final parameters ........................................ 21
   3.7 The results of our data .................................... 25
   3.8 The formulation of the problem in the continuous time .... 27
   3.9 The estimation of the change-point ....................... 29

4 The Poisson disorder problem ............................. 31
   4.1 The description of the problem .......................... 31
   4.2 The statement of the problem ............................ 32
   4.3 The a posteriori probability process ...................... 34
   4.4 The stochastic equation for the process \( (\varphi_t)_{t \geq 0} \) ... 37
   4.5 The stochastic equation for the process \( (\pi_t)_{t \geq 0} \) ....... 39
Chapter 1

Introduction

Loan origination decision-making systems are usually based on some kind of clients data analysis. By this analysis a financial institution tries to find the most relevant characteristics of the customers and construct the decision-making process in accordance to them. The major part of the clients characteristics includes non-numerical data (e.g. marital status, education, etc.), to represent these data in the numerical form credit scorecards are used.

Credit scorecards are mathematical models which provide a quantitative measurement of the likelihood that a customer will present a defined behaviour, with respect to the information about this customer. The output of a credit scorecard is the probability that the client will default. Using this method, from the set of the customers which apply for loans only clients with the probability of default less or equal to the stated cut-off rate will be accepted.

But even after such procedure some of the accepted clients default. A process of clients defaults can be considered as a stochastic process with some intensity of jumps. This intensity can change at any (unknown) time moment. In this work we consider the case when the intensity of defaults arises. The problem is to find a time moment which is as close as possible to the moment when the intensity of defaults arises. When such time moment is found we need some a posteriori method which allows to estimate the moment of the change of the intensity more accurately to obtain a homogeneous sample of the client’s data for the estimation of a new model. After that the model of estimation of the clients characteristics must be corrected for further using.

In the first part of this work in accordance to the described problem we consider the logistic regression model to estimate the probability of the client’s default. Further we provide methods for finding the change-point in the both cases of the discrete and continuous time. Finally the problem of
the real time detection of a disorder moment in the case when the intensity of defaults increases is solved. Below we give our motivation in more details.

1.1 A Logistic Regression model

Regression methods become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. The outcome variable in the logistic regression is binary or dichotomous.

The logistic regression is used for a prediction of the probability of an occurrence of an event by fitting data to a logistic function, in our case the probability that clients default.

The first part of the work on this project consists of the main rules of the parameters estimation for the logistic regression and corresponding quality characteristics such as a ROC curve, a sensitivity and a specificity.

1.2 A posteriori change-point problem for a given sample

The field of statistical research has many solutions reflecting different approaches to the main question: is the sample of observations homogeneous in a statistical sense? If the answer is no, the next question: what segments of homogeneity exist in a given sample?

In our case, the process of the data acquisition is completed when the homogeneity hypothesis is checked, so it is a posteriori change-point problem. In what follows the change-point problem in the discrete time, some tests estimation and results for our real data.

1.3 The Poisson disorder problem

In the second part of this work we solve the problem of the real time detection of a disorder. For this purpose we consider the process of clients defaults as a Poisson process. This is justified by the following natural assumptions which we made based on the logics of the loan origination process:

- the credit organization provides loans every time moment;
- the volume of the credit portfolio is sufficiently large;
• the clients applying for the loans independently of each other;
• the clients default at a random time moments;
• the clients default independently of each other.

Since we consider the process of the clients defaults as a Poisson process, we assume that this process has some admissible intensity in the beginning of the observation. The problem of a real-time detection of a disorder moment is equivalent to the problem of finding of a time moment which is as close as possible to the moment of increasing in the value of intensity of the clients defaults. Such problem is called a Poisson disorder problem which was studied by many authors (e.g. see [11], [2], [3], [1]). In this work we base on the paper by Peskir and Shiryaev [11]. They considered the disorder problem for the homogeneous Poisson process (i.e. in the case of the known constant intensities). To apply results of the paper [11] to the loan origination problem we assume that the intensity of defaults is time dependent. So it is natural to suppose that the process' intensity depends on the density of clients which is assumed to be a known function of time. Due to this reason we assume that the considered process of defaults is a non-homogeneous Poisson process. Therefore we can not use results of the previous studies directly, but we can use the approach which was provided by Peskir and Shiryaev [11] for the Poisson disorder problem. In this work we formulate the optimal stopping problem, then this problem is reduced to the free-boundary problem with using of the a posteriori probability process and an analytical solution for the free-boundary problem in the case when the intensity of defaults increases is found.
Chapter 1. Introduction
Chapter 2

The Multiple Logistic regression

In Chapter 2 we consider the logistic model in the case of more than one independent variable. Central to the consideration of the multiple logistic model is the estimation of the coefficients in the model and testing for their significance.

Using the Receiver Operating Characteristic (ROC) curve analysis accuracy of a test to discriminate defaulted cases from the normal cases is evaluated.

2.1 The real data

We consider a data set of clients who took the loans from some bank. The data contains 10231 loans. Among this loans 3216 (31,43\%) of the loans are defaulted ones.

Description of variables:

1. A dependent variable. The dependent variable is the log odds ratio of a default. This binary variable takes the value 1 if the loan is defaulted and 0 otherwise.

2. An independent variable. We have five independent variables: age, education, occupation, loan amount, extime, which is a number how many years client has been working till now.
Chapter 2. The Multiple Logistic Regression

2.2 The multiple logistic regression

We have \( p \) independent variables denoted by the vector \( \mathbf{x}' = (x_1, x_2, ..., x_p) \).

We can denote the conditional probability that our outcome is '1' by the following equation:

\[
P(Y = 1 | x) = \pi(x).
\]

The logit of the multiple logistic regression model is presented by

\[
g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p
\]

and the logistic regression model is

\[
\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}
\]

In our data the independent variables 'education' and 'occupation' have \( k > 2 \) distinct values. It is not proper to include them in the model as if they are interval scale variables. They do not have a numerical significance. Therefore, we will form a set of design variables to represent the categories of the variable. In our case, one of the independent variables is education, which has been coded as 'Professional', 'Standard' and 'High'. Therefore we need two design variables \( Ed_1 \) and \( Ed_2 \).

One possible coding strategy (see [10]) is that when the client is 'Professional', \( Ed_1 \) and \( Ed_2 \), both be equal zero, when client has 'Standard' education \( Ed_1 \) is equal one while \( Ed_2 \) zero; when client has 'High' education \( Ed_1 = 0 \) and \( Ed_2 = 1 \). The Table 2.1 and Table 2.2 illustrate this coding of the design variables.

Generally, if a nominal scaled variable has \( k \) possible values, then \( k-1 \) design variables will be needed.

Let us consider that the \( j^{th} \) independent variable \( x_j \) has \( k_j \) levels. The \( k_j - 1 \) design variables will be denoted as \( Ed_{jl} \) and the coefficient for these design variables will be denoted as \( \beta_{jl} \), \( l = 1, 2, ..., k_j - 1 \). The logit for a model with \( p \) variables and the \( j^{th} \) variable being discrete would be, (see [10])

\[
g(x) = \beta_0 + \beta_1 x_1 + ... + \sum_{l=1}^{k_j-1} \beta_{jl} D_{jl} + \beta_p x_p.
\]
Table 2.1: The coding of the design variables for education, coded at three levels

<table>
<thead>
<tr>
<th>Education</th>
<th>Ed₁</th>
<th>Ed₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prof</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Std</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>High</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2.2: The coding of the design variables for occupation, coded at three levels

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Oc₁</th>
<th>Oc₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transport</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Industry</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>State</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

2.3 The fitting procedure to the multiple Logistic Regression model

The data was collected on 10231 clients, who took the loans from the same bank. The value $n_1 = 3216$ is the number of clients who defaulted and $n_2 = 7015$ who didn’t default. The five variables are significant: age, education, occupation, extime and amount of the loan. We use the coding of the design variables for 'education' and 'occupation' from Table 2.1 and Table 2.2. The results of the fitting procedure of the logistic regression model to these data are presented in Table 2.3.

The estimated logit is given by the following expression

$$
\hat{g}(x) = -0.20 - 0.049 \ast extime - .05 \ast age + 9.47e - 06 \ast amount \ (2.3)
+ 0.9 \ast ed1 - 0.77 \ast ed2 + 1.11 \ast oc1 + 0.43 \ast ed2
$$

The fitted values are obtained using the estimated logit $\hat{g}(x)$. 
Chapter 2. The Multiple Logistic regression

Table 2.3: The estimated coefficients for a Multiple Logistic Regression Model.

|        | Coef.  | Std.Err. | z     | P>|z|  | 95% Conf. Interval |
|--------|--------|----------|-------|------|-------------------|
| default| -.0487731 | .0063161 | -7.72 | 0.000 | -.0611525 -.0363938 |
|        | -.0498521 | .0065691 | -7.59 | 0.000 | -.0627274 -.0369768 |
| amount | 9.47e-06  | 7.13e-07 | 13.28 | 0.000 | 8.07e-06  .0000109 |
| ed1    | .8961842  | .056516  | 15.86 | 0.000 | .7854149  1.006954 |
| ed2    | -.7740963 | .0739351 | -10.47| 0.000 | -.9190065 -.6291862 |
| oc2    | .4314866  | .1146536 | 3.76  | 0.000 | .2067696  .6562036 |
| oc1    | 1.110505  | .0564563 | 19.67 | 0.000 | .999853   1.221158 |
| cons   | -.2029166 | .1803296 | -1.13 | 0.260 | -.5563561 .150523 |

2.4 The testing procedure for the significance of the model

The first step in the process of a model assessment is usually devoted to evaluation of the significance of the variables in the model. The test is based on the statistic $G$ given by the equation

$$ G = -2 \ln \left[ \frac{L_2}{L_1} \right], $$

where $L_1$ is a likelihood with the independent variable, while $L_2$ is without. Let us consider the fitted model whose estimated coefficients are given in Table 2.3. The log likelihood of the fitted model is $L_1 = -5085,2597$. The log likelihood for the constant only model may be obtained by the method which yields the log likelihood $L_2 = -6369,0871$. Thus the value of the likelihood ratio test is equal to

$$ G = -2 \left[ (-6369,0871) - (-5085,2597) \right] = 2567,6548 $$

and the $p$–value for the test is $P \left( \chi^2(7) > 2567,6548 \right) = 0,00$ which is significant for the given $\alpha = 0,05$. We reject the null hypothesis and perhaps all $p$ coefficients are different from zero.

Now we look at the univariate Wald test statistics,

$$ W_j = \hat{\beta}_j / \hat{SE}(\hat{\beta}_j), $$

These values are given in the fourth column in Table 2.3. Under the hypothesis that every coefficient is zero, these statistics follow the standard normal
distribution. In the fifth column we have p-values and by these values we can conclude that all variables are significant with the level of significance \( \alpha = 0.05 \).

### 2.5 The area under the ROC curve

Specificity and sensitivity are statistical measures of a binary classification test. The sensitivity is a proportion of actual positives that has correctly identified as such (in our case the percentage of the people who default and who are correctly identified as that who is having this condition). While the specificity is a proportion of the negatives that are correctly identified (the percentage of the people who is not default to the people who are correctly identified as not having the condition). We can consider the people who wanted to take a loan. The test outcome can be positive (default) or negative (not default), while the actual status of the people may be different. In that setting:

- **True positive**: People who are defaulted correctly diagnosed as defaulted
- **False positive**: People who didn’t default are incorrectly identified as defaulted
- **True negative**: People who didn’t default are correctly identified as didn’t default
- **False negative**: People who defaulted are incorrectly identified as didn’t default

Then we obtain

\[
\text{sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}
\]

\[
\text{specificity} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}}
\]

Both of them sensitivity and specificity are closely related to the concept of the error of the first kind and the error of the second kind. From the theoretical point of view, the optimal prediction is the 100 percent sensitivity (i.e. all people from the default group identified as default) and the 100 percent specificity (i.e. not predict anyone from the not default group as to be default).
A better description of the classification accuracy is given by the area under the ROC (Receiver Operating Characteristics) curve. This curve shows how the receiver operates the existence of a signal in the presence of noise. It plots the probability of detecting the true signal (sensitivity) and the false signal (1-specificity) for all ranges of possible cutpoints. The area under the ROC curve describes the complete classification accuracy and it is in the interval from zero to one.

Suppose that we were interested in predicting the outcome of each client. These results are shown in Table 2.4. For example, we took a cutpoint 0.85, where the sensitivity is only 2.92% while the specificity 99.77%. The same can be done for any other choice of the cutpoint.
<table>
<thead>
<tr>
<th>Cutpoint</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>1-Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>99.22%</td>
<td>11.35%</td>
<td>88.65%</td>
</tr>
<tr>
<td>0.10</td>
<td>96.33%</td>
<td>27.78%</td>
<td>72.22%</td>
</tr>
<tr>
<td>0.15</td>
<td>92.07%</td>
<td>41.67%</td>
<td>58.33%</td>
</tr>
<tr>
<td>0.20</td>
<td>87.38%</td>
<td>52.32%</td>
<td>47.68%</td>
</tr>
<tr>
<td>0.25</td>
<td>80.78%</td>
<td>61.44%</td>
<td>38.56%</td>
</tr>
<tr>
<td>0.30</td>
<td>74.63%</td>
<td>69.48%</td>
<td>30.52%</td>
</tr>
<tr>
<td>0.35</td>
<td>68.07%</td>
<td>75.41%</td>
<td>24.59%</td>
</tr>
<tr>
<td>0.40</td>
<td>60.42%</td>
<td>80.81%</td>
<td>19.19%</td>
</tr>
<tr>
<td>0.45</td>
<td>53.89%</td>
<td>84.99%</td>
<td>15.01%</td>
</tr>
<tr>
<td>0.50</td>
<td>47.17%</td>
<td>88.38%</td>
<td>11.62%</td>
</tr>
<tr>
<td>0.55</td>
<td>40.52%</td>
<td>91.53%</td>
<td>8.47%</td>
</tr>
<tr>
<td>0.60</td>
<td>32.77%</td>
<td>94.13%</td>
<td>5.87%</td>
</tr>
<tr>
<td>0.65</td>
<td>26.12%</td>
<td>96.22%</td>
<td>3.78%</td>
</tr>
<tr>
<td>0.70</td>
<td>20.09%</td>
<td>97.49%</td>
<td>2.51%</td>
</tr>
<tr>
<td>0.75</td>
<td>13.31%</td>
<td>98.77%</td>
<td>1.23%</td>
</tr>
<tr>
<td>0.80</td>
<td>7.58%</td>
<td>99.47%</td>
<td>0.53%</td>
</tr>
<tr>
<td>0.85</td>
<td>2.92%</td>
<td>99.77%</td>
<td>0.78%</td>
</tr>
</tbody>
</table>

Table 2.4: The summary of sensitivity, specificity, and 1-specificity parameters for the classification tables based on the Logistic Regression Model in Table 2.3 using the Cutpoints of 0.05 to 0.85 in the increments of 0.05.

The plot of the sensitivity versus the 1-specificity over all possible cutpoints is shown in Figure 2.1. The curve generated by these points is called ROC-curve.

As a general rule (see[10]) holds

- If ROC=0.5 then this means no discrimination
- If 0.7≤ROC<0.8 then we have an acceptable discrimination
- If 0.8≤ROC<0.9 then it means the excellent discrimination
- If ROC≥0.9 then it is an outstanding discrimination

In practical point of view it is unusual to observe areas under the ROC Curve greater than 0.9.
The area under the ROC curve in Figure 2.1 is 0.7942 (counted by stata). Therefore we assume that for our data it is almost excellent discrimination achieved.
Chapter 2. The Multiple Logistic regression

Figure 2.1: Plot of the sensitivity and specificity versus all possible cutpoints.

Figure 2.2: The ROC curve.

In the next chapter we find the change-point in the data, which divides it in two subsamples. We need this because in our data one subsample works properly and the second not (more and more clients default), and then we must change our logistic regression model.
Chapter 3

The A-posteriori change-point problem for a given sample

In Chapter 2 we considered the method of the estimation of the probability of the clients defaults. To use this method we need a homogeneous sample of data. In our study we assume that there is just one change point in the data which divides it in two subsamples.

In Chapter 3 we formulate the change-point problem as well as in the discrete and the continuous time. Also we do an estimation of a change-point. For the real data we present the corresponding result.

3.1 The Maximum likelihood method

Suppose we have $x_1, x_2, ..., x_n$ as an observed sequence of independent random variables with one sudden change point at the moment $1 \leq t_0 \leq N$. The density function $f(x_t|\theta)$ defined as

$$f(x_t|\theta) = \begin{cases} f(x_t|\theta_1), & t = 1, ..., t_0 - 1, \\ f(x_t|\theta_2), & t = t_0, ..., N, \end{cases}$$

for $\theta_1 \neq \theta_2$.

The value $\theta_1$ as well as $\theta_2$ are a priori known and also the change point $t_0$ is uniformly distributed on $[1, ..., N]$.

To estimate the change-point $t_0$, we should test $N+1$ hypotheses: i-th hypothesis $H_i$ means that $t_0 = i$.

The maximum likelihood estimate of $t_0$ is: $H_k$ is accepted when (see [4])

$$l_k(x) \geq l_j(x), \forall j, j \neq k,$$
Chapter 3. The A-posteriori change-point problem for a given sample

\[ l_m(x) = \sum_{i=1}^{m-1} \ln f(x_i|\theta_1) + \sum_{i=m}^{N} \ln f(x_i|\theta_2), \]

where \( l_m(x) \), \( m = 1, ..., N \) is the likelihood function corresponding to the hypothesis \( H_m \).

In the case when \( \theta_1 \) and \( \theta_2 \) are a priori known, the method above is equivalent to the CUSUM procedure:

\[ \hat{t}_{CUSUM} = \inf (k : S_{k-1} \geq S_j, j = 2, ..., N), \]

\[ S_k = \sum_{i=1}^{k} [\ln f(x_i|\theta_1) - \ln f(x_i|\theta_2)]. \]

3.2 The formulation of the problem in the discrete time

When a client defaults our outcome is one, while if not it is zero, so we have a sequence of independent zero-one random variables, which has a change in the distribution at some unknown point.

Suppose that \( X_1, X_2, ..., X_T \) are independent random variables with

\[ P(X_i = 1) = 1 - P(X_i = 0) \quad P(X_i = 1) = \theta_0 \quad (i \leq \tau) \]

\[ P(X_i = 1) = \theta_1 \quad (i > \tau; 1 \leq \tau \leq T) \]

(3.1)

The value \( \tau \) is our change point. When \( 1 \leq \tau < T \), it has been a change in the sequence, but when \( \tau = T \) there has been no change.

We consider a statistic which can be used to make inferences about \( \tau \), while \( \theta_0 \) and \( \theta_1 \) are unknown.

Suppose, that we have a sequence \( (U_t; t = 1, ..., T) \), which is defined by \( U_t = TS_t - tS_T \), where, (see[13])

\[ S_t = \sum_{i=1}^{t} X_i, (t = 1, ..., T). \]

As an alternative to this method can be computed by using the cumulative sums of \( (X_i - T^{-1}S_T) \) that is,

\[ T^{-1}U_t = \sum_{i=1}^{t} (X_i - T^{-1}S_T). \]
If $\theta_0 > \theta_1$, then we estimate $\tau$ by considering $t_0$ which maximizes $U_t$ and take as the estimate $\tau$ the smallest value satisfying the condition, when the estimate $\hat{\tau}$ is given by

$$\hat{\tau} = \inf_{t_0} \{ t_0 : U_{t_0} \geq U_t; t = 1, \ldots, T \}.$$ \hfill (3.2)

### 3.3 An exact test of to have no change against to have change

In the case, when there is no change, the distribution of $(U_t|S_T = m)$ for $t = 1, \ldots, T - 1$ is independent of $\theta_0$ \cite{13} because $X_i$ are independent Bernoulli random variables with the constant probability $\theta_0$ of the value one. Firstly for testing to have no change against to have change we consider, $\theta_0 > \theta_1$, i.e.

$$V^+_m = (U_{t_0} : U_{t_0} \geq U_t; t = 1, \ldots, T|S_T = m).$$ \hfill (3.3)

In the case, when there is no change it is easy to show \cite{12}, that $V^+_m$ has the same distribution as the null distribution of $m(T - m)D^+_{m,T-m}$, a multiple of the Kolmogorov-Smirnov two sample statistic \cite{12}. Similarly

$$V^-_m = (U_{t_0} : U_{t_0} \leq U_t; t = 1, \ldots, T|S_T = m),$$ \hfill (3.4)

$$V^m = (U_{t_0} : |U_{t_0}| \geq |U_t|; t = 1, \ldots, T|S_T = m),$$

have the same distributions as $m(T - m)D^-_{m,T-m}$ and $m(T - m)D^-_{m,T-m}$ respectively. Of course the statistics $D^+_{m,n}$ and $D^-_{m,n}$ have the same null distributions. So this test can be made by statistic $V^+_m, V^-_m$ or $V^m$ and the null distribution of the corresponding Kolmogorov-Smirnov statistic. Tests about $\tau$ are based directly on an estimate, $\hat{\tau}$ are not good, therefore greater efficiency is expected for the tests based on the value of $U_t$ at the estimated change-point, \cite{9}.

On the other hand we can compare the statistic above with the likelihood ratio statistic for to have no change versus change. The maximum likelihood estimate of $\tau$, $\tau<T$, when $\theta_0$ and $\theta_1$ are unknown, will be find by maximizing the log likelihood. Firstly we define the likelihood function of $(X_1, \ldots, X_T)$ under the model (3.1), see \cite{8}

$$L(t) = \prod_{i=1}^{\tau} \theta_0^{X_i}(1 - \theta_0)^{1-X_i} \prod_{i=\tau+1}^{T} \theta_1^{X_i}(1 - \theta_1)^{1-X_i},$$ \hfill (3.5)

then the log likelihood conditional on $\tau = t$ with $\theta_0$ and $\theta_1$ known, can be
Chapter 3. The A-posteriori change-point problem for a given sample

written as

\[ M(t) = \sum_{i=1}^{t} \left[ X_t \log(\frac{\theta_0}{\theta_1}) + (1 - X_t) \log(\frac{1-\theta_0}{1-\theta_1}) \right] \]
\[ + \sum_{i=1}^{T} \left[ X_t \log(\tau_1) + (1 - X_t) \log(1 - \tau_1) \right]. \]

So the maximum likelihood estimate \( \hat{\tau} \) is the value which maximizes the sequence

\[ R_t = \sum_{i=1}^{t} \left[ X_t \log(\tau_0) + (1 - X_t) \log(\tau_1) \right], \]
\[ (t = 1, ..., T - 1), \]

where we assume that at least one variable, \( X_T \), have mean \( \theta_1 \).

In our case the marginal likelihood of \( \tau \) is the likelihood (3.5) maximized over \( \theta_0 \) and \( \theta_1 \), so that (3.7) is replaced by

\[ R_t = \sum_{i=1}^{t} \{ X_t \log(\frac{\theta_0}{\tau_0}) + (1 - X_t) \log(1 - \frac{\theta_0}{\tau_0}) \} \]
\[ + \sum_{i=1}^{T} \{ X_t \log(\frac{\tau_0}{\tau_1}) + (1 - X_t) \log(1 - \frac{\tau_0}{\tau_1}) \}, \]

where

\[ \bar{X}_t = t^{-1} \sum_{i=1}^{t} X_i, \quad \bar{X}_{t'} = (T - t)^{-1} \sum_{i=t+1}^{T} X_i, \]

over \( 1 \leq t < T \), are conditional estimates of \( \theta_0 \) and \( \theta_1 \).

When we have no change the log likelihood is given by

\[ R_T = \sum_{i=1}^{T} X_t \log(\bar{X}_T) + (1 - X_t) \log(1 - \bar{X}_T). \]

So a test of to have a no change versus to have a change can be based on

\[ R_T - \sup_{1 \leq t < T} R_t, \]

or \( R_T - R_\hat{\tau} \), while \( \hat{\tau} \) is the maximum estimator of \( \tau \). The null distribution of \( R_T - R_\hat{\tau} \) appears completely intractable.

To compare the likelihood ratio test with the conditional test based on \( V_m \) a small Monte Carlo study was undertaken, [12]. Different values of \( \theta_0 \), \( \theta_1 \), \( \tau \) and \( T \) were used. For \( V_m \) the test for the conditional null distribution was approximated using the asymptotic distribution of the two-sample Kolmogorov-Smirnov statistic, (see [12]) for further details. This case is only good for \( T \geq 50 \), but it is unchanging in that it gives a test of smaller size than the nominal size. There is a very small difference between powers of this two tests at the same approximate significance level, and for \( \tau = \frac{1}{2} T \) the \( V_m \) test was more powerful than the likelihood ratio test.

The \( V_m \) test seems better, because the null distribution of the likelihood ratio statistic is very difficult and the statistic hard to calculate, [13].
3.4 The estimation of $\tau$

Now we will consider the estimation of $\tau$ based on $U_t$. The procedure is to let $T \to \infty$, $\tau \to \infty$ such that $\tau T^{-1} \to \lambda$, the log likelihood $R_t$ of (3.8) satisfies

$$c(R_{t+1} - R_t) = X_{t+1} - \alpha(1 + \alpha)^{-1},$$

where $c$ is a constant independent of $t$ and $\alpha = \log[(1 - \theta_1)/(1 - \theta_0)]/\log(\theta_0/\theta_1)$. If $\theta_0 > \theta_1$ that is $\alpha > 0$, the statistic $\hat{\tau} - \tau$, where $\hat{\tau}$ is the maximum likelihood estimator, is then asymptotically equivalent in the distribution to the position of the larger of the two random walk maxima, where we defined the random walks

$$W = (0, R_{\tau-1} - R_\tau, ..., R_1 - R_\tau),$$
$$W' = (0, R_{\tau+1} - R_\tau, ..., R_{T-1} - R_\tau).$$

Thus $W$ and $W'$ represent the log likelihood function of $t$, relative to that of the true value $\tau$, for $t \leq \tau$ and $t \geq \tau$. Without loss of generality assume $\theta_0 > \theta_1$ and define

$$Y_i = -X_{\tau-i} + \phi(1 - X_{\tau-i}), \quad (i = 1, ..., \tau - 1),$$
$$Y'_i = X_{\tau+i} - \phi(1 - X_{\tau+i}), \quad (i = 1, ..., T - \tau - 1),$$

where

$$\phi = \log \left( \frac{1 - X_t}{1 - X_t^*} \right) / \log \left( \frac{X_t}{X_t^*} \right) > 0.$$

Then from (3.7) we have

$$Y_i = (R_{\tau-i} - R_{\tau-i+1}) / \log \left( \frac{X_t}{X_t^*} \right),$$
$$Y'_i = (R_{\tau+i} - R_{\tau+i-1}) / \log \left( \frac{X_t}{X_t^*} \right),$$

so that apart from a positive factor $\log \left( \frac{X_t}{X_t^*} \right)$, (3.9) becomes

$$W = \left( 0, Y_1, Y_1 + Y_2, ..., \sum_{j=1}^{\tau-1} Y_j \right),$$
$$W' = \left( 0, Y'_1, Y'_1 + Y'_2, ..., \sum_{j=1}^{T-\tau-1} Y'_j \right).$$

The random walks $W$ and $W'$ are independent, each with the independent increments, with distributions, by (3.10)

$$P[Y_i = (1 + \alpha)^{-1}] = \theta_0, \quad P[Y_i = -\alpha(1 + \alpha)^{-1}] = 1 - \theta_0,$$
$$P[Y'_i = (1 + \alpha)^{-1}] = \theta_1, \quad P[Y'_i = -\alpha(1 + \alpha)^{-1}] = 1 - \theta_0.$$
where the $Y_i$ and $Y'_i$ are independent.

To estimate $\tau$ using (3.2) we consider $U_{t+1} - U_t = TX_{t+1} - ST$; after dividing by $T$ we have

$$T^{-1}U_{t+1} - T^{-1}U_t = X_{t+1} - ST^{-1},$$

and noting that $ST^{-1} \to \lambda \theta_0 + (1-\lambda)\theta_1$ in probability and that $\text{var}(ST^{-1}) = O(T^{-1})$, [13], and then asymptotically we have

$$T^{-1}U_{t+1} - T^{-1}U_t = X_{t+1} - \psi,$$

where

$$\psi = \lambda \theta_0 + (1-\lambda)\theta_1.$$

If we assume that $\theta_0 > \theta_1$, the statistic $(\hat{\tau} - \tau)$ is equivalent in distribution to the position of the larger of the two random walk maxima, where the random walk are defined by (3.13) and (3.14), but

$$P(Y_i = 1 - \psi) = \theta_0, \quad P(Y_i = \psi) = 1 - \theta_0,$$

$$P(Y'_i = 1 - \psi) = \theta_1, \quad P(Y'_i = \psi) = 1 - \theta_1.$$

We can note that [13]:

1. the asymptotic distributions of $\hat{\tau}$ and $\tilde{\tau}$ differ through the values $\psi$ and $\alpha(1+\alpha)^{-1}$;

2. if $\lambda = \frac{1}{2}$ and $\theta_0 = 1 - \theta_1$, then $\psi = \alpha(1+\alpha)^{-1}$ and the two estimates $\hat{\tau}$ and $\tilde{\tau}$ are asymptotically equivalent;

3. if $\theta_0 > \theta_1$, then $\theta_0 > \alpha(1+\alpha)^{-1} > \theta_1$.

Therefore when $\tau$ is near to $\frac{1}{2}T$ then there is a small difference between the maximizing $R_t$ and $U_t$ also $\hat{\tau}$ and $\tilde{\tau}$ are almost equivalent asymptotically.

In the case without the change, after considering the equivalence with the one-sided Kolmogorov-Smirnov statistic and using the limit theorems we can (see Takacs, 1970, p. 381):

- assume that $\theta_0 > \theta_1$,

- estimate $\tau$ using (3.2) to give $\hat{\tau}$,

- then, asymptotically, $\hat{\tau}T^{-1}$ has an uniform distribution over $(0,1)$.

When there is no change we would prefer $\hat{\tau}T^{-1} \to 0$ or 1 with the probability one.

It might be better to choose as our estimate of $\tau$ the value which maximizes
\[ [t(T - t)]^{-\frac{1}{2}} U_t \] since \( \text{var}(U_t) = t(T - t)T\theta(1 - \theta_0) \) on the hypothesis of no change. We denote this estimate by \( \hat{\tau}_N \), which using the normalized version of \( U_t \). One can appeal to the limit theorems involving the empirical distribution function to the reason that on the hypothesis of no change \( \hat{\tau}_N^{-1} \rightarrow 0 \) or 1 with the probability one, \[13\]. In particular the Law of the Iterated Logarithm indicates this result to be true. Consider \( T \rightarrow \infty \), \[13\], then

\[
[t(T - t)][(t + 1)(T - t - 1)]^{-\frac{1}{2}} U_{t+1} - [t(T - t)]^{-\frac{1}{2}} U_t
\]

\[
= [\lambda(1 - \lambda)]^{-\frac{1}{2}} \left[ T^{-1}U_{t+1} - T^{-1}U_t - \frac{1 - 2\nu}{2\nu (1 - \nu)} T^2 U_{t+1} \right]
\]

where \( \nu = tT^{\frac{1}{2}} \). While \( \tau - t = O(T) \), we have

\[
E\left( T^{-2} U_t \right) = \lambda(1 - \lambda)(\theta_0 + \theta_1), \quad \text{var}(T^{-2} U_t) = O(T^{-1}).
\]

If \( t \) is near \( \tau \) and \( T \rightarrow \infty \) we have

\[
[t(T - t)][(t + 1)(T - t - 1)]^{-\frac{1}{2}} U_{t+1} - [t(T - t)]^{-\frac{1}{2}} U_t,
\]

behaving similar to \( [\lambda(1 - \lambda)]^{-\frac{1}{2}} (X_{t+1} - \psi_1) \), where \( \psi_1 = \frac{1}{2}(\theta_0 + \theta_1) \). Considering the random walk ideas which we described above, we see that \( \hat{\tau}_N \) and \( \hat{\tau} \) are almost equivalent asymptotically, for \( \psi_1 \) close to \( \alpha(1 + \alpha)^{-1} \).

Canner(1975) has studied this statistic using the Monte Carlo method and gives some percentage points for \( n = m \), which could be used for \( n \neq m \) to provide approximate percentage points. Using these techniques, the test of no change against change was generally better than test considered in this section \[13\].

As a conclusion from this section we obtain that the distribution of \( (\hat{\tau} - \tau) \) is asymptotically independent on \( \tau \), so that in the theory is possible to find the asymptotic distribution and use it to find the confidence regions \[8\].
3.5 The change-point problem in the model of the logistic regression

In the previous sections we considered a simple cumulative sum type statistic for the change-point with the zero-one observations. We compared a conditional test of to have no change against to have change with a likelihood ratio test. Using the simple statistic we also considered an estimation of the change-point.

Now we consider a generalized maximum likelihood asymptotic power one test which aim to detect a change-point in the logistic regression when the alternative specifies that a change occurred in parameters of the model. We demonstrate applications of the test and the maximum likelihood estimation of the change-point using an actual problem, that was encountered with a real data.

Let's take a finite sequence of independent observations \((Y_i, x_i), \ i \geq 1,\) where

\[ x_i = [x_{i1}, \ldots, x_{id}]^T, \]

are fixed \(d\)-vector explanatory variables and

\[ Y_i = [Y_{i1}, \ldots, Y_{im_i}]^T, \ i \geq 1 \]

are Bernoulli variates. Because the exact distribution depends on the predictor \(x_i\) we can assume that,

\[
P(Y_{ij} = 1|x_i) = \left(1 + \exp(-x_i^T \beta_0)\right)^{-1} I(i < \nu) + \left(1 + \exp(-x_i^T \beta_1)\right)^{-1} I(i \geq \nu) \tag{3.16}
\]

\[ \beta_k = [\beta_{k1}, \ldots, \beta_{kd}]^T, \ k = 0, 1, \ j = 1, \ldots, m_i, \ i = 1, \ldots, n, \ \beta_0 \neq \beta_1, \]

where \(I\{\cdot\}\) is the indicator function and \(\beta_0, \ \beta_1, \ \nu\) are parameters of the model. In our situation up to the change-point \(\nu > 0\) our observations satisfy the logistic regression model with parameter \(\beta_0\) and after this change point \((\nu \leq n)\) satisfy the logistic regression with parameter \(\beta_1\).

Now we consider hypothesis testing, where

\[
H_0 : \nu \notin [1, n], \tag{3.17}
\]

\[
H_1 : 1 \leq \nu \leq n;
\]

\(\nu > 0\) is unknown.
3.6 The procedure for the known initial parameters and unknown final parameters

In our case we know $\beta_0$, but $\beta_1$ is unknown. We denote,

$$\Lambda_k^n(\beta_0, \hat{\beta}|x_k, ..., x_n) = \prod_{i=k}^{n} \prod_{j=1}^{m_i} \left( \frac{\exp(x_i^T \beta_{0} Y_{ij})}{1 + \exp(x_i^T \beta_0)} \right)^{-1} \frac{\exp(x_i^T \hat{\beta}^{(i+1,n)} Y_{ij})}{1 + \exp(x_i^T \hat{\beta}^{(i+1,n)})},$$

(3.18)

where $\hat{\beta}^{(i+1,n)} \in \mathcal{I} \{(Y_r, x_r)_{r=i+1}^{n}\}$, $1 \leq i \leq n$ is any estimator of $\beta_1$ in the $\sigma$-algebra generated by $\{(Y_r, x_r)_{r=i+1}^{n}\} \hat{\beta}^{(i+1,n)}$ based upon $\{(Y_r, x_r)_{r=i+1}^{n}\}$, where $\hat{\beta}^{(n+1,n)} = \beta_0$. Therefore $\Lambda_k^n(\beta_0, \hat{\beta}|x_k, ..., x_n)$ is the estimator of the likelihood ratio $\Lambda_k^n(\beta_0, \hat{\beta}|x_k, ..., x_n)$ from the following equation

$$\Lambda_k^n(\beta_0, \beta_1|x_k, ..., x_n) = \frac{\prod_{i=k}^{n} \prod_{j=1}^{m_i} P_{\nu=k}(Y_{ij}|x_i)}{\prod_{i=k}^{n} \prod_{j=1}^{m_i} P_{H_0}(Y_{ij}|x_i)} = \frac{\prod_{i=k}^{n} \prod_{j=1}^{m_i} P_{\nu=i}(Y_{ij}|x_i)}{\prod_{i=k}^{n} \prod_{j=1}^{m_i} P_{H_0}(Y_{ij}|x_i)}$$

(3.19)

$$= \exp \left( \sum_{i=k}^{n} \sum_{j=1}^{m_i} (x_i^T \beta_1 - x_i^T \beta_0) Y_{ij} \right) \prod_{i=k}^{n} \left( \frac{1 + \exp(x_i^T \beta_{0})}{1 + \exp(x_i^T \beta_1)} \right)^{m_i},$$

where

$$P_{\nu=k}(Y_{ij}|x_i) = P_{\nu=k}(Y_{ij} = 1|x_i) Y_{ij} (1 - P_{\nu=k}(Y_{ij} = 1|x_i))^{-Y_{ij}},$$

$$P_{H_0}(Y_{ij}|x_i) = P_{H_0}(Y_{ij} = 1|x_i) Y_{ij} (1 - P_{H_0}(Y_{ij} = 1|x_i))^{-Y_{ij}}.$$  

We consider the test: reject $H_0$ iff

$$\max_{1 \leq k \leq n} \Lambda_k^n(\beta_0, \hat{\beta}|x_k, ..., x_n) > C$$

(3.20)

**Significance level of the test.** The generalized maximum likelihood ratio test for the problem (3.17) in (3.20) was proposed in [7]. From the literature devoted to the change-point problem we known that such tests have high power. Therefore the evaluation of its significance is a major issue. The results on the significance level of the generalized maximum likelihood ratio test are asymptotic ($n \rightarrow \infty$). The guaranteed non-asymptotic upper bound for the significance level of the test was proposed in the following proposition, [7]
Proposition 1. [7] The significance level $\alpha$ of the test satisfies:

$$\alpha = P_{H_0} \left( \max_{1 \leq k \leq n} \Lambda_k^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) \geq C \right) \leq \frac{1}{C}. $$

Proof of the Proposition 1.

Lemma 1. [7] Under $H_0$ the sequence $\left[ \Lambda_k^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) \right]_{k=1}^n$ is a non-negative reverse martingale with respect to $[(Y_k, x_k)]_{k=1}^n$ and $E_{H_0} \Lambda_k^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) = 1$ for all $k = 1, 2, \ldots, n$.

By (3.18) we have for all $k = 1, \ldots, n$

$$E_{H_0} \left( \Lambda_k^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) | (Y_n, x_n), \ldots, (Y_{k+1}, x_{k+1}) \right) = \prod_{i=k+1}^{n} \prod_{j=1}^{n_i} \frac{\exp(x_i^T \hat{\beta}(k+1, n)) Y_{ij}}{1 + \exp(x_i^T \beta_0)} \frac{\exp(x_j^T \hat{\beta}(k+1, n)) Y_{kj}}{1 + \exp(x_j^T \beta_0)}$$

and

$$E_{H_0} \Lambda_k^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) = E_{H_0} E_{H_0} \left( \Lambda_{k+1}^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) | [(Y_r, x_r)]_{r=k+1}^n \right)$$

$$= E_{H_0} \Lambda_{k+1}^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) = E_{H_0} E_{H_0} \left( \Lambda_{k+1}^n(\beta_0, \hat{\beta} | x_k, \ldots, x_n) | [(Y_r, x_r)]_{r=k+2}^n \right)$$

$$= \ldots = 1,$$

Hence from Lemma 1, follows the theorem.

Theorem 1. [14](The Doob’s Submartingale Inequality). Let $Z$ be a non-negative submartingale. Then for $C > 0$ and $n \in \mathbb{Z}^+$,

$$CP \left( \sup_{k \leq n} Z_k \geq C \right) \leq E \left( Z_n; \sup_{k \leq n} Z_k \geq C \right) \leq E(Z_n).$$

Proof of the Theorem 1. The proof shows with the assumption that is not needed for the first inequality that $Z$ is non-negative, because the number of the steps $n$ does not feature directly in the last two expressions, what gives the result its power.

Let $F := \{ \sup_{k \leq n} Z_k \geq C \}$. Then $F$ is a disjoint union

$$F = F_0 \cup F_1 \cup \ldots \cup F_n$$
where $F_0 := \{ Z_0 \geq C \}$.
$F_k := \{ Z_0 < C \} \cap \{ Z_1 < C \} \cap \ldots \cap \{ Z_{k-1} < C \} \cap \{ Z_k \geq C \}$.

Now, $F_k \in \mathcal{F}_k$, and $Z_k \geq C$ on $F_k$. So

$$E(Z_n; F_k) \geq E(Z_k; F_k) \geq CP(F_k).$$

Summing over $k$ yields the results. Now by the Lemma 1 and the Doob’s theorem for a non-negative submartingale we have

$$\alpha = P_{H_0}(M_n > C) = P_{H_0} \left( \Lambda_{N_n} (\hat{\beta} | x_{N_n}, \ldots, x_n) > C \right)$$
$$\leq E_{H_0} \Lambda_{N_n} (\hat{\beta} | x_{N_n}, \ldots, x_n) = E_{H_0} \Lambda_{N_n} (\hat{\beta} | x_n) = \frac{1}{C}.$$

**Power of the test:** We define for all $1 \leq i \leq n$,

$$a_i \equiv E_\nu \ln \left( \prod_{j=1}^{m_i} \frac{P_{H_0}(Y_{ij}|x_i)}{P_{H_0}(Y_{ij}|x_i)} \right)$$
$$m_i \left( (x_i^T \beta_1 - x_i^T \beta_0) \frac{\exp(x_i^T \beta_1)}{1+\exp(x_i^T \beta_1)} + \ln \left( \frac{1+\exp(x_i^T \beta_0)}{1+\exp(x_i^T \beta_1)} \right) \right),$$

if $i < \nu$ then $a_i \leq 0$ and if $i \geq \nu$ then $a_i \geq 0$.

The next result gives the expression for the power of the test (3.19).

**Proposition 2.** Assume for some $0 < \rho < 1, N > 0$ and $(n - \nu) > N : (1 - \rho) \sum_{i=\nu}^n a_i > \ln(C)$ and

$$\frac{((1 - \rho) \sum_{i=\nu}^n a_i - \ln(C))^2}{\sum_{i=\nu}^n (m_i | x_i^T (\beta_1 - \beta_0))} \to \infty,$$

as $(n - \nu) \to \infty$,

where $a_i > 0$ was defined above. Then

$$0 \leq 1 - P_\nu \left\{ \max_{1 \leq k \leq n} \Lambda_{k} (\beta_0, \hat{\beta} | x_k, \ldots, x_n) > C \right\}$$
$$\leq \exp \left( - \frac{(1-\rho) \sum_{i=\nu}^n a_i - \ln(C))^2}{2 \sum_{i=\nu}^n (m_i | x_i^T (\beta_1 - \beta_0))} \right) + \tau_{\nu,n},$$

where

$$\tau_{\nu,n} \equiv P_\nu = \left\{ \sum_{i=\nu}^n 2m_i | x_i^T (\hat{\beta}^{(i+1,n)} - \beta_1) | \geq \rho \sum_{i=\nu}^n a_i \right\}.$$
Chapter 3. The A-posteriori change-point problem for a given sample

Proof of the Proposition 2.

By definitions (3.18), (3.21) and we define for all $1 \leq i \leq n$

$$\lambda_i \equiv \ln \left( \prod_{j=1}^{m_i} \frac{P_\nu(Y_{ij}|x_i)}{P_{H_0}(Y_{ij}|x_i)} \right) - a_i$$

(3.23)

$$= (x_i^T \beta_1 - x_i^T \beta_0) \sum_{j=1}^{m_i} \left( Y_{ij} - \frac{\exp(x_i^T \beta_1)}{1 + \exp(x_i^T \beta_1)} \right),$$

where $a_i$ defined by (3.21), as well as a Taylor series expansion and some rearranging of terms, the probability of the type II error for the test (3.20) is bounded by

$$P_\nu \left\{ \max_{1 \leq k \leq n} \Lambda_k^n(\beta_0, \hat{\beta}|x_k, ..., x_n) \leq C \right\} \leq P_\nu \left\{ \ln \Lambda_k^n(\beta_0, \hat{\beta}|x_k, ..., x_n) > \ln(C) \right\}$$

$$= P_\nu \left\{ \sum_{i=\nu}^{n} \sum_{j=1}^{m_i} x_i^T (\beta_1 - \beta_0) Y_{ij} + m_i \ln \left( \frac{1 + \exp(x_i^T \beta_0)}{1 + \exp(x_i^T \beta_1)} \right) \right\}$$

$$+ \sum_{i=\nu}^{n} \left( \sum_{j=1}^{m_i} x_i^T (\hat{\beta}^{(i+1,n)} - \beta_1) Y_{ij} + m_i \ln \left( \frac{1 + \exp(x_i^T \beta_1)}{1 + \exp(x_i^T \hat{\beta}^{(i+1,n)})} \right) \right) \leq \ln(C) \right\},$$

$$P_\nu \left\{ \sum_{i=\nu}^{n} \lambda_i + \sum_{i=\nu}^{n} a_i + \sum_{i=\nu}^{n} x_i^T (\hat{\beta}^{(i+1,n)} - \beta_1) \left( \sum_{j=1}^{m_i} Y_{ij} - m_i \frac{\exp(\theta_i)}{1 + \exp(\theta_i)} \right) \leq \ln(C) \right\}$$

$$\leq P_\nu \left\{ \sum_{i=\nu}^{n} \lambda_i + (1 - \rho) \sum_{i=\nu}^{n} a_i \leq \ln(C) \right\} + \tau_{\nu,n},$$

(3.24)

where $\theta_i \in \left( x_i^T \hat{\beta}^{(i+1,n)}, x_i^T \beta_1 \right)$. Noting that, by Lemma 1 and (3.19)

$$0 \leq 1 - P_\nu \left\{ \max_{1 \leq k \leq n} \Lambda_k^n(\beta_0, \hat{\beta}|x_k, ..., x_n) > C \right\} \leq P_\nu \left\{ \sum_{i=\nu}^{n} (\lambda_i + a_i) \leq \ln(C) \right\}$$

$$= P_\nu \left\{ \sum_{i=\nu}^{n} (-\lambda_i) \geq \sum_{i=\nu}^{n} a_i - \ln(C) \right\},$$
where, by (3.21) and (3.23) $E_{\nu} \lambda_i = 0$ and $|\lambda_i| \leq m_i |x_i^T \beta_1 - x_i^T \beta_0|, \ i \geq \nu$. Therefore, by Hoeffding(1963) we have

$$0 \leq 1 - P_{\nu} \left\{ \max_{1 \leq k \leq n} \Lambda_k^\nu (\beta_0, \hat{\beta}|x_k, ..., x_n)>C \right\}$$

$$\leq \exp \left( \frac{((1-\rho) \sum_{i=\nu}^{n} a_i - \ln(C))^2}{2 \sum_{i=\nu}^{n} (m_i |x_i^T (\beta_1 - \beta_0)|)^2} \right) + \tau_{\nu,n},$$

where $\tau_{\nu,n}$ by (3.22), thereby completing the proof.

The maximum likelihood estimation of a change point: If $H_0$ is rejected, then the following estimator of $\nu$ is considered, [7]

$$\hat{\nu}_n = \arg \max \prod_{1 \leq k \leq n} \prod_{i=k}^{n} \prod_{j=1}^{m_i} \frac{\exp(x_i^T \beta_0 Y_{ij})}{1 + \exp(x_i^T \beta_0)} \frac{1}{\exp(x_i^T \hat{\beta}(k,n) Y_{ij}) + 1 + \exp(x_i^T \hat{\beta}(k,n))},$$

(3.25)

where $\hat{\beta}(k,n)$ is the MLE of $\beta_1$ in the $\sigma$-algebra generated by $(Y_k, x_k), ..., (Y_n, x_n)$.

3.7 The results of our data

Our clients took the loans, everyone at some unknown time moment. Based on the information from the data sample we do not know the exact time moment when the client took the loan, but we know the order of the clients. We would like to know on which client the change-point has occurred.

We have 10231 clients who took the loans, so $i \in [1, 10231]$ and $Y_i$ is an outcome for every client. Because of it for every $i$ we have $m_i = 1$, i.e $j$ always take just one value, $j = 1$. Therefore:

$$Y_{ij} = Y_i = \begin{cases} 1, & \text{if } i\text{-th client default} \\ 0, & \text{if not} \end{cases}$$

For every $i$ we have a vector for explanatory variables $x_i$, which consists of all data from clients (age, education, occupation, extime and loan amount). The change point which we will find $\nu \in [1, 10231]$. We consider hypothesis testing where:

$$H_0 : \nu \notin [1, n],$$

$$H_1 : 1 \leq \nu \leq n;$$

(3.26)
In our case \( n = 10231, m_i = 1 \) for all \( i \in [1, 10231] \). Therefore from (3.18) we have:

\[
\Lambda^n_k(\beta_0, \hat{\beta}|x_k, ..., x_n) = \prod_{i=k}^{10231} \left( \frac{\exp(x_i^T \beta_0 Y_{ij})}{1 + \exp(x_i^T \beta_0)} \right)^{-1} \frac{\exp(x_i^T \hat{\beta}^{(i+1,10231)} Y_{ij})}{1 + \exp(x_i^T \hat{\beta}^{(i-1,10231)})}.
\]

Our model assumes \( \beta_0 = [-0.20, -0.049, -0.05, 9.47e-06, 0.9, -0.77, 1.11, 0.43]^T \).

So

\[
\ln \left[ \max_{1 \leq k \leq 10231} \Lambda^{10231}_k(\beta_0, \hat{\beta}|x_k, ..., x_n) \right] = \ln[\Lambda^{10231}_{6665}(\beta_0, \hat{\beta}|x_k, ..., x_n)] = 2126
\]

where \( \hat{\beta} \) is the MLE of \( \beta_1 \) (\( \hat{\beta}^{(l,n)} = \beta_0, n - l \leq 3567 \)). Based on the test (3.20) and Proposition 3.1 we reject \( H_0 \) with a \( p-value \leq \exp(-2126) \).

By applying (3.25) we can say that estimator of the change point is \( \hat{\nu} = 6665 \).

So after 6664 steps the change has occurred. That is why we can say from 6665 client our logistic regression model is not working properly. So our model must be corrected for further using.
3.8 The formulation of the problem in the continuous time

In the previous section we considered the case in the discrete time. This model is quite good and give an idea about finding the change-point but it was assumed that the credit company provides a large number of loans. Therefore it is natural to extend our model to the continuous time case. Based on the motivation from the introduction we consider the process of clients defaults as a Poisson process.

**Definition 1.** The stochastic process $N = (N_t)_{t \geq 0}$ which admits following properties:

1. the process starts at zero $N_0 = 0$;
2. $N = (N_t)_{t \geq 0}$ has independent increments: for any $t_1 < \ldots < t_n$, $N_{t_n} - N_{t_{n-1}}, \ldots, N_{t_2} - N_{t_1}, N_{t_1}$ are independent random variables;
3. for any $t > 0$, $N_t$ follows a Poisson distribution with parameter $\lambda t$, i.e.

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \forall n \in \mathbb{N}$$

is called Poisson process with an intensity $\lambda$.

**Definition 2.** The stochastic process $L = (L_t)_{t \geq 0}$ is a continuous Markov chain if for all $s, t > 0$ and non-negative integers $i, j, l_u, 0 \leq u < s$ holds:

$$\mathbb{P}\{L_{s+t} = j | L_s = i, L_u = l_u, 0 \leq u < s\} = \mathbb{P}\{L_{s+t} = j | L_s = i\}.$$

Suppose we observe a trajectory of the Poisson process $N = (N_t)_{t \geq 0}$, and an intensity is changed from $\lambda_1$ to $\lambda_2$ at some unknown time $\theta$. Based upon the information which represented by the trajectory of $N$, our problem is to detect this change point as fast as possible by the means of the point process observations.

Now, let $N^{\lambda_1} = (N_{t}^{\lambda_1})_{t \geq 0}$, $N^{\lambda_2} = (N_{t}^{\lambda_2})_{t \geq 0}$ and $L = (L_t)_{t \geq 0}$ be independent stochastic processes defined on a probability space $(\Omega, F; P, \pi \in [0, 1])$ such that:

1. $N^{\lambda_1}$ is a Poisson process with the intensity $\lambda_1$;
2. $N^{\lambda_2}$ is a Poisson process with the intensity $\lambda_2$;
3. $L$ is a continuous Markov chain with two states $\lambda_1$ and $\lambda_2$. 


So $P_\pi(L_1 = \lambda_2) = 1 - P_\pi(L_1 = \lambda_1) = \pi$, and for $L_1 = \lambda_1$, there is a single passage of $L$ from $\lambda_1$ to $\lambda_2$ at a random time $\theta > 0$ satisfying $P_\pi(\theta \leq t) = t$. The process $N = (N_t)_{t \geq 0}$ is given by:

$$N_t = \int_0^t I(L_s = \lambda_1)dN_s^{\lambda_1} + \int_0^t I(L_s = \lambda_2)dN_s^{\lambda_2}, \quad (3.28)$$

and $F^N_1 = \sigma(N_s : 0 \leq s \leq 1)$.

We consider the change-point problem for the Poisson process in the case when the intensities $\lambda_1$ and $\lambda_2$ are constants and the change-point $\tau$ is unknown. Then we can easily transfer the problem to the problem about a sequence of independent exponentially distributed random variables, because the intervals between jumps in the Poisson process are exponentially distributed independent random variables. In our case the likelihood function can be expressed as:

$$L(t) = \ln \left( \prod_{i=1}^{N_t} \lambda_1 e^{-\lambda_1 \tau_i} \prod_{i=N_t+1}^{N_T} \lambda_2 e^{-\lambda_2 \tau_i} \right), \quad (3.29)$$

where $\tau_i$ is the time between jump $i-1$ and $i$.

We see that

$$L(t) = \ln \left( \lambda_1^{N_t} e^{-\lambda_1 \sum_{i=1}^{N_t} \tau_i} \lambda_2^{N_T-N_t} e^{-\lambda_2 \sum_{i=N_t+1}^{N_T} \tau_i} \right)$$

$$= \ln \left( \left( \frac{\lambda_1}{\lambda_2} \right)^{N_t} \lambda_2^{N_T} e^{-\lambda_1 \sum_{i=1}^{N_t} \tau_i - \lambda_2 (N_T-N_t)} \right).$$

So we have

$$L(t) = \ln \left( \lambda_2^{N_T} e^{-\lambda_2 T} \left( \frac{\lambda_1}{\lambda_2} \right)^{N_t} e^{-(\lambda_1 - \lambda_2) t} \right), \quad (3.30)$$

Thus the function $L(t)$ depends only on the term

$$K_t = \ln \left( \frac{\lambda_1}{\lambda_2} \right) N_t - (\lambda_1 - \lambda_2) t. \quad (3.31)$$

Hence the maximum likelihood estimate $\hat{\tau}$ is the value of $t$ which maximizes the sequence above.
3.9 The estimation of the change-point

From the other side we can use other method of the change-point detection. To use this method we will define the a posteriori probability:

$$\pi_t = P_\pi (\theta \leq t | F_1^N) = E_\pi \left[ I_{(\theta \leq t)} | F_1^N \right], \quad (3.32)$$

for $t \geq 0$.

For the detection of the change-point we can use $E_\pi (\theta | F_1^N)$, which is an a posteriori estimator of $\theta$.

By the standard arguments we get the expression:

$$E_\pi (\theta | F_1^N) = \int_0^1 (1 - \pi_t) dt = 1 - \int_0^1 \pi_t dt,$$

for the forecast of $\theta$ based on the observation.

To calculate $\pi_t$ we define the likelihood ratio process as

$$\varphi_t = \frac{\pi_t}{1 - \pi_t}, \quad (3.33)$$

By the Bayes formula we can see that

$$\pi_t = \int_0^t \frac{dP^s}{dP_\pi} | F_1^N ds, \quad (3.34)$$

where $P^s(.) = P(.) | \theta = s$ is the probability law of the process $N$ given that $\theta = s$ for $s \in [0, 1]$, $P(.) = \int_0^1 P^s(.) ds + (1 - t) P_t(.)$. And $\frac{dP^s}{dP_\pi}$ is a Radon-Nikodym density of the measure $P^s | F_1^N$ with respect to the measure $P_\pi | F_1^N$, so similarly

$$1 - \pi_t = (1 - t) \frac{dP_\pi}{dP_\pi} | F_1^N.$$

(3.35)

After a substitution of the expressions for $\pi_t$ and $1 - \pi_t$ we get that the likelihood ratio process $\varphi_t$ is given by

$$\varphi_t = \frac{1}{1 - t} \int_0^t \frac{dP^s}{dP_\pi} | F_1^N ds = \frac{Z_1}{1 - t} \int_0^t \frac{1}{Z_s} ds, \quad (3.36)$$

where the process $Z_t$ equals to

$$Z_t = \frac{dP^0}{dP^\infty} (t, N) = \frac{d(P^0 | F_0^N)}{d(P^\infty | F_0^N)} = \exp \left( \log \frac{\lambda_2}{\lambda_1} N_t - (\lambda_2 - \lambda_1) t \right) \quad (3.37)$$
Chapter 3. The A-posteriori change-point problem for a given sample

Now the process $\varphi_t$ is expressed in terms of the observed process

$$E_{\pi}(\theta|\mathcal{F}_1^N) = \int_0^1 (1 - \pi_t) dt = \int_0^1 \frac{1}{1 + \varphi_t} dt,$$

(3.38)

where $\varphi_t = \frac{Z_t}{1 - t} \int_0^1 \frac{1}{Z_s} ds$.

Using the calculation from above we find the value of $\pi_t$ and based on this we find the $E(\theta|\mathcal{F}_1^N)$.

Remark. The point $\hat{t} = \varphi_{\hat{t}} = 1$ provides a median for the distribution of $\theta$ given by $\mathcal{F}_1^N$. Therefore it can be used as another estimation of $\theta$. 
Chapter 4

The Poisson disorder problem

Now we begin to consider the problem of the real time detection of a disorder moment. Before we start the theoretical considerations let us recall our motivation for the disorder problem. We assume that a credit organization provides a large number of loans based on some model of estimation of clients characteristics. After the loan origination some of the clients will default. The moment when the intensity of defaults increases significantly we consider as a moment of the model reconciliation. The problem can be formulated as a disorder problem for a non-homogeneous Poisson process.

In this Chapter we give a basic description and a statement of the Poisson disorder problem as well as the theoretical background for its solution.

4.1 The description of the problem

To give a basic description for the disorder problem we need the following definition of the non-homogenous Poisson process

**Definition 3.** A stochastic process $N = (N_t)_{t \geq 0}$ which admits following properties:

1. the process starts at zero $N_0 = 0$;
2. $N = (N_t)_{t \geq 0}$ has independent increments: for any $t_1 < \ldots < t_n$, $N_{t_n} - N_{t_{n-1}}, \ldots, N_{t_2} - N_{t_1}, N_{t_1}$ are independent random variables;
3. for any $t > 0$, $N_t$ follows a Poisson distribution with parameter $\Gamma(t) = \int_0^t \gamma(s)ds, \gamma(s) > 0$:

$$\mathbb{P}(N_t = n) = e^{-\Gamma(t)} \frac{(\Gamma(t))^n}{n!}, \forall n \in \mathbb{N}$$

is called **non-homogenous Poisson process** with an intensity $\gamma(t)$. 
Chapter 4. The Poisson disorder problem

We assume that we are observing in the real time a non-homogeneous Poisson process and that there can be a single passage at a random (unknown) time $\theta$ in the intensity of this process from $\gamma^1_t = \alpha_1 f(t)$ to $\gamma^2_t = \alpha_2 f(t)$, where $f(t)$ is a function which describes the density of the clients (i.e. the number of an active clients at the moment $t$), $\alpha_1$ and $\alpha_2$ are coefficients which define the share of defaulted clients. Then an expected number of defaults for the time period $[0, t]$ is given by

$$
\Gamma_t = \alpha_1 \int_0^{t \wedge \theta} f(s) \, ds + \alpha_2 \int_{t \wedge \theta}^t f(s) \, ds.
$$

Here we assume that the intensity of defaults depends on the density of clients.

The problem was formulated in [11] for the case of the constant intensities and for the present settings it can be given in the same way. Suppose at the time $t = 0$ we begin to observe a Poisson process $N = (N_t)_{t \geq 0}$ whose intensity changes from $\gamma^1_t$ to $\gamma^2_t$ at a some random (unknown) time $\theta$ with the following distribution

$$
P(\theta = 0) = \pi,
$$

$$
P(\theta > t | \theta > 0) = e^{-\lambda t},
$$

so $\theta$ is assumed to take the value 0 with a probability $\pi$, and to be exponentially distributed with the parameter $\lambda$ given that $\theta > 0$. Based upon the information which is continuously updated through the observation of the process $N$, the problem is to terminate the observation (and declare the alarm) at the time $\tau$, which is as close as possible to $\theta$ in accordance to the given criteria.

4.2 The statement of the problem

In this section is given a formulation of the disorder problem for a non-homogeneous Poisson process. We use the notion of a continuous Markov chain defined in Chapter ??.

Let $N^\gamma_t = (N^\gamma_t^1)_{t \geq 0}$, $N^\gamma_t = (N^\gamma_t^2)_{t \geq 0}$ and $L = (L_t)_{t \geq 0}$ be independent stochastic processes defined on the probability space $(\Omega, \mathcal{F}, P, \pi, \pi \in [0, 1])$ such that:

1. $N^\gamma_t$ is a Poisson process with the intensity $\gamma^1_t > 0$;
2. $N^\gamma_t$ is a Poisson process with the intensity $\gamma^2_t > 0$;
3. $L = (L_t)_{t \geq 0}$ is a continuous Markov chain with two states $\alpha_1$ and $\alpha_2$, initial distribution $[1 - \pi, \pi]$, and the transition probability matrix $[e^{-\lambda t}, 1 - e^{-\lambda t}; 0, 1]$ for $t > 0$ where $\lambda > 0$.

Thus $P_{\pi}(L_0 = \alpha_2) = 1 - P_{\pi}(L_0 = \alpha_1) = \pi$, and given that $L_0 = \alpha_1$, there is a single passage from $\alpha_1$ to $\alpha_2$ at a random time $\theta > 0$ satisfying $P_{\pi}(\theta > t) = (1 - \pi) e^{-\lambda t}$ for all $t > 0$. 

The process $N = (N_t)_{t \geq 0}$ is given by

$$N_t = \int_0^t \mathbb{I}(L_{s-} = \alpha_1) dN_s^1 + \int_0^t \mathbb{I}(L_{s-} = \alpha_2) dN_s^2,$$

and we set a natural filtration of the process $N$ as $\mathcal{F}_t^N = \sigma(N_s : 0 \leq s \leq t)$ for $t \geq 0$. Denoting $\theta = \inf\{t \geq 0 : L_t = \alpha_2\}$ we see that

$$P_\pi(\theta = 0) = \pi,$$

and

$$P_\pi(\theta > t|\theta > 0) = e^{-\lambda t}$$

for all $t > 0$.

It is assumed that the disorder time $\theta$ is unknown (i.e. it can not be observed directly from the trajectory of $N$). Now we formulate the optimal stopping problem and at first we need the following definition.

**Definition 4.** We call a non-negative random value $\tau = \tau(\omega)$ a stopping time or a random variable independent of the “future” if

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t^N \text{ and } P\{\tau < \infty\} = 1$$

for each $t \geq 0$. This means that at any moment $t$ we can see if the stopping happened or not.

A stopping time of the process $N$ means a stopping time adopted to the natural filtration $(\mathcal{F}_t^N)_{t \geq 0}$ generated by the process $N$. The Poisson disorder problem is equivalent to the problem of the finding of a stopping time $\tau^*$ of $N$ which is “as close as possible” to the disorder time $\theta$ as a solution of the following optimal stopping problem

$$V(\pi) = \inf_{\tau} \left( P_\pi(\tau < \theta) + cE_\pi(\tau - \theta)^+ \right),$$

(4.2)

where $P_\pi(\tau < \theta)$ is interpreted as a probability of the “false alarm”, $E_\pi(\tau - \theta)^+$ is interpreted as the “average” delay in the detection of the occurrence of the “disorder”, $c > 0$ is a given constant which is used to balance the errors of both types, and the infimum in (4.2) is taken over all stopping times $\tau$ of $N$. 

4.3 The a posteriori probability process

Following [11] we can reformulate the problem (4.2) as an optimal stopping problem for the Markov process. To do this we define the a posteriori probability process by

\[ \pi_t = P_\pi(\theta \leq t | \mathcal{F}_t^N) = E_\pi(\mathbb{I}_{\{\theta \leq t\}} | \mathcal{F}_t^N), \quad (4.3) \]

for \( t \geq 0 \). With use of the dominated convergence theorem we get

\[ P_\pi(\tau < \theta) = E_\pi(\mathbb{I}_{\{\tau < \theta\}}) = E_\pi(1 - \mathbb{I}_{\{\tau \geq \theta\}}) = E_\pi(1 - \pi_\tau), \]

for all stopping times \( \tau \) of \( N \), also

\[ E_\pi(\tau - \theta)^+ = E_\pi\left(\int_0^\tau \mathbb{I}_{\{\theta \leq t\}} dt\right) = E_\pi\left(\int_0^\tau P_\pi\{\mathbb{I}_{\{\theta \leq t\}} \mid \mathcal{F}_t^N\} dt\right) = E_\pi\left(\int_0^\tau \pi_t dt\right) \]

for all stopping times \( \tau \) of \( N \). So (4.2) can be rewritten as

\[ V(\pi) = \inf_\tau E_\pi\left(\left(1 - \pi_\tau\right) + c \int_0^\tau \pi_t dt\right), \quad (4.4) \]

where the infimum is taken over all stopping times \( \tau \) of \( (\pi_t)_{t \geq 0} \).

In order to get an explicit formula for the process \( \pi_t \) we use the likelihood ratio process defined as

\[ \varphi_t = \frac{\pi_t}{1 - \pi_t}. \quad (4.5) \]

According to [11]

\[ \pi_t = \pi \frac{dP^0}{dP_\pi}(t, N) + (1 - \pi) \int_0^t \frac{dP^s}{dP_\pi}(t, N) \lambda e^{-\lambda s} ds, \quad (4.6) \]

where \( P^s(N \in \cdot) = P_\pi(N \in \cdot | \theta = s) \) is the distribution law of the Poisson process which changes the intensity from \( \gamma_1 \) to \( \gamma_2 \) at a time \( s \geq 0 \), and \( \frac{dP^s}{dP_\pi}(t, N) \) is a Radon-Nikodym density of the measure \( P^s | \mathcal{F}_t^N \) with respect to the measure \( P_\pi | \mathcal{F}_t^N \). Similarly

\[ 1 - \pi_t = (1 - \pi)e^{-\lambda t} \frac{dP^t}{dP_\pi}(t, N) = (1 - \pi)e^{-\lambda t} \frac{dP^\infty}{dP_\pi}(t, N), \quad (4.7) \]
where $P^\infty$ is the probability measure of the process $N^\gamma^1$.

In order to get the formula for the likelihood ratio process $\varphi_t$ we substitute expressions for $\pi_t$ and $(1 - \pi_t)$ into (4.5), then

$$\varphi_t = e^{\lambda t} Z_t \left( \varphi_0 + \lambda \int_0^t e^{-\lambda s} \frac{Z_s}{Z_t} ds \right),$$

(4.8)

where

$$Z_t = \frac{dP^0}{dP^\infty}(t, N) = \exp \left\{ (\alpha_2 - \alpha_1) \int_0^t f(s) ds + \log \left( \frac{\alpha_2}{\alpha_1} \right) N_t \right\}. \quad (4.9)$$

Note, that the process $N$ has the intensity $\gamma_1^1$ under the measure $P^\infty$ and $\gamma_2^2$ under $P^0$.

Based on [10] we derive the properties which the process $Z_t$ must satisfy for the case of a non-homogeneous Poisson process.

**Proposition 3.** Let $N$ be a Poisson process. The process $Z_t$ is a Radon-Nikodym density which is given by (4.9), then $Z_t$ satisfies following properties

1. $E^\infty[Z_t] = 1$;
2. $Z_t$ is a martingale with respect to $P^\infty$;
3. under the measure $P^0$, $N$ is a Poisson process with the intensity $\gamma_2^2$.

**Proof.** 1. Show that $E^\infty[Z_t] = 1$. Under the measure $P^\infty$ for a given $t$ the process $N_t$ is a Poisson process with the intensity $\gamma_1^1 = \alpha_1 f(t)$. The corresponding moment generating function equals to

$$\Psi_t(v) = E^\infty \left[ e^{vN_t} \right] = \exp \left\{ (e^v - 1)\alpha_1 \int_0^t f(s) ds \right\}. \quad \left(\frac{\alpha_2}{\alpha_1}\right)$$

If $\ln \left( \frac{\alpha_2}{\alpha_1} \right)$ is denoted by $v$ then

$$E^\infty[Z_t] = E^\infty \left[ \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s) ds + v N_t \right\} \right] = \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s) ds \right\} \Psi_t(v).$$
We can check that
\[ \Psi_t \left( \ln \left( \frac{\alpha_2}{\alpha_1} \right) \right) = \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s) ds \right\} \]
and therefore
\[ E^\infty[Z_t] = \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s) ds \right\} \exp \left\{ (\alpha_2 - \alpha_1) \int_0^t f(s) ds \right\} = 1. \]

2. Let us show that \( Z_t \) is a martingale under \( P^\infty \), i.e. \( E^\infty[Z_t|\mathcal{F}_u] = Z_u \).

\[ E^\infty[Z_t|\mathcal{F}_u] = E^\infty \left[ \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s) ds + u N_t \right\} \right| \mathcal{F}_u \] =
\[ \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s) ds \right\} e^{u N_u} E^\infty \left[ e^{v(N_t-N_u)} \right| \mathcal{F}_u \].

We use the fact that \( N_t - N_u \sim \text{Poi} \left( \alpha_1 \int_u^t f(s) ds \right) \), then the corresponding moment generating function is given by
\[ \hat{\Psi}_t(v) = E^\infty[e^{v(N_t-N_u)}] = \exp \left\{ (\alpha_2 - \alpha_1) \int_u^t f(s) ds \right\}. \]

Therefore
\[ E^\infty[Z_t|\mathcal{F}_u] = \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s) ds \right\} e^{u N_u} \hat{\Psi}_t(v) =
\[ \exp \left\{ (\alpha_1 - \alpha_2) \int_0^u f(s) ds \right\} \exp \left\{ \log \frac{\alpha_2}{\alpha_1} N_u \right\} = Z_u. \]

3. To prove this fact we calculate the moment generating function of \( N \) under the measure \( P^0 \) using the fact that \( E^0[e^{u N_t}] = E^\infty[e^{u N_t} Z_t] \). We see
\[ E^0[e^{u N_t}] = \hat{\Psi}_t(u) = \exp \left\{ (e^u - 1) \alpha_2 \int_0^t f(s) ds \right\}, \]
on the other hand

\[ E^\infty[e^{uN_t}Z_t] = E^\infty \left[ \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s)ds + (v + u)N_t \right\} \right] = \]

\[ \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s)ds \right\} E^\infty [e^{N_t(u+v)}], \]

where

\[ E^\infty [e^{N_t(u+v)}] = \Psi_t(v + u) = \exp \left\{ (e^{v+u} - 1)\alpha_1 \int_0^t f(s)ds \right\}. \]

Then we obtain

\[ E^\infty[e^{uN_t}Z_t] = \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s)ds \right\} \exp \left\{ (e^u - 1)\alpha_2 \int_0^t f(s)ds \right\} = \]

\[ \exp \left\{ (e^u - 1)\alpha_2 \int_0^t f(s)ds \right\} = E^0[e^{uN_t}]. \]

4.4 The stochastic equation for the process 

\((\varphi_t)_{t \geq 0}\).

Before we begin to find the differential equation for the process \(\pi_t\), we need to find such an equation for \(\varphi_t\). In order to find the stochastic differential equation for the process \((\varphi_t)_{t \geq 0}\) we introduce two jump processes, based on the process (4.8) we define

\[ \Phi_1(t) = e^{\lambda t}Z_t, \quad (4.10) \]

and

\[ \Phi_2(t) = \varphi_0 + \int_0^t \frac{e^{-\lambda s}}{Z_s} ds. \quad (4.11) \]
In accordance to Theorem 2, The Ito’s product rule for jump processes (see Appendix C) we have

$$\varphi_t = \Phi_1(0)\Phi_2(0) + \int_0^t \Phi_2(s) d\Phi_1(s) + \int_0^t \Phi_1(s) d\Phi_2(s) + [\Phi_1, \Phi_2](t). \tag{4.12}$$

We see that in the processes $\Phi_1(t)$ and $\Phi_2(t)$ there do not exist a diffusion component, moreover since the process $\Phi_2(t)$ is continuous, then its jump part is equal to zero. So in accordance to Theorem 3 (see Appendix C) we obtain

$$[\Phi_1, \Phi_2](t) = 0.$$

Taking into the account all arguments from above we write $\varphi_t$ in the following form

$$\varphi_t = \Phi_1(t)\Phi_2(t) = \varphi_0 Z_0 + \int_0^t \left( \varphi_0 + \lambda \int_0^s e^{-\lambda u} \frac{dZ_u}{Z_u} \right) d(e^{\lambda s} Z_s) +$$

or in the differential form

$$d\varphi_t = \left[ \varphi_0 + \lambda \int_0^t e^{-\lambda s} \frac{dZ_s}{Z_s} \right] d(e^{\lambda t} Z_t) + e^{\lambda t} Z_t \left( \varphi_0 + \lambda \int_0^t e^{-\lambda s} \frac{dZ_s}{Z_s} \right).$$

(4.13) (4.14)

To obtain the final formula for $d\varphi_t$ we need to calculate differentials in (4.14). At first we calculate $d(e^{\lambda t} Z_t)$

$$d(e^{\lambda t} Z_t) = Z_t \lambda e^{\lambda t} dt + e^{\lambda t} dZ_t,$$

here $dZ_t$ we find using Theorem 4 (see Appendix C). For our case we have

$$dZ_t = \left( \frac{\alpha_2}{\alpha_1} - 1 \right) Z_{t-} dM_t,$$

where the process $M_t$ is given by the following formula

$$M_t = N_t - \alpha_1 \int_0^t f(s) ds.$$
Now we can calculate the stochastic differential of the process $Z_t$

$$dZ_t = \left( \frac{\alpha_2}{\alpha_1} - 1 \right) Z_{t-} dN_t - (\alpha_2 - \alpha_1) Z_{t-} f(t) dt. \tag{4.15}$$

Using the formula (4.15) we get

$$d(e^\lambda Z_t) = e^\lambda (\lambda Z_t - (\alpha_2 - \alpha_1) Z_{t-} f(t)) dt + e^\lambda Z_{t-} \left( \frac{\alpha_2}{\alpha_1} - 1 \right) dN_t. \tag{4.16}$$

Now we calculate $d \left( \varphi_0 + \lambda \int_0^t \frac{e^{-\lambda s}}{Z_s} ds \right)$ in (4.14)

$$d \left( \varphi_0 + \lambda \int_0^t \frac{e^{-\lambda s}}{Z_s} ds \right) = \lambda \frac{e^{-\lambda s}}{Z_s} dt. \tag{4.17}$$

According to (4.16) and (4.17) we write equation (4.14) in the following form

$$d\varphi_t = \lambda e^\lambda Z_t \left[ \varphi_0 + \lambda \int_0^t \frac{e^{-\lambda s}}{Z_s} ds \right] dt + e^\lambda Z_{t-} \lambda \frac{e^{-\lambda s}}{Z_s} dt - \left( \alpha_2 - \alpha_1 \right) f(t) dt$$

$$- \left( \frac{\alpha_2}{\alpha_1} - 1 \right) e^\lambda Z_{t-} \left[ \varphi_0 + \lambda \int_0^t \frac{e^{-\lambda s}}{Z_s} ds \right] dN_t.$$

Simplifying the latter expression we obtain the stochastic differential equation for the process $\varphi_t$

$$d\varphi_t = \lambda (\varphi_t + 1) dt + \left( \frac{\alpha_2}{\alpha_1} - 1 \right) \varphi_{t-} (dN_t - \alpha_1 f(t) dt). \tag{4.18}$$

### 4.5 The stochastic equation for the process $(\pi_t)_{t \geq 0}$.

In order to find the stochastic differential equation for $(\pi_t)_{t \geq 0}$ we use the equality
\[ \pi_t = \frac{\varphi_t}{1 + \varphi_t} = g(\varphi_t), \]
then the derivative of the function \( g \) is calculated in the following way
\[ g'(\varphi_t) = \frac{1}{(1 + \varphi_t)^2}. \]
Also we need the value of the process \( \pi_t \) just before the jump expressed in the form
\[ \pi_{t-} = \frac{\varphi_{t-}}{1 + \varphi_{t-}}. \]
In accordance to Theorem 5, The Ito-Doeblin formula for the jump process (see Appendix C), we have the expression for \( \pi_t \)
\[ \pi_t = \pi + \int_0^t \frac{1}{(1 + \varphi_s)^2} d\varphi^c_s + \sum_{0 < s \leq t} \left[ \frac{\varphi_s}{1 + \varphi_s} - \frac{\varphi_{s-}}{1 + \varphi_{s-}} \right]. \quad (4.19) \]
The continuous part of the differential of \( \varphi_t \) has the form
\[ d\varphi^c_t = \lambda(\varphi_t + 1)dt - (\alpha_2 - \alpha_1)f(t)\varphi_t dt, \quad (4.20) \]
thus we can rewrite (4.19) as
\[ \pi_t = \pi + \int_0^t \frac{\lambda(\varphi_s + 1) - (\alpha_2 - \alpha_1)f(s)\varphi_s}{(1 + \varphi_s)^2} ds + \sum_{0 < s \leq t} \left[ \frac{\varphi_s}{1 + \varphi_s} - \frac{\varphi_{s-}}{1 + \varphi_{s-}} \right]. \]
To obtain this equality in the differential form, we should represent the latter sum as an integral with respect to some jump process. For this reason we find a dependence between \( \varphi_s \) and \( \varphi_{s-} \)
\[ \varphi_s = e^{\lambda s} Z_s \left( \varphi_0 + \lambda \int_0^s \frac{e^{-\lambda u}}{Z_u} du \right), \]
and
\[ \varphi_{s-} = e^{\lambda s} Z_{s-} \left( \varphi_0 + \lambda \int_0^s \frac{e^{-\lambda u}}{Z_u} du \right). \]
Now we get
\[ \varphi_s = \frac{Z_s}{Z_{s-}} \varphi_{s-}. \]
In order to find an explicit dependence between $\varphi_s$ and $\varphi_{s-}$ we simplify the following fraction

\[
\frac{Z_s}{Z_{s-}} = \exp\{\log \frac{\alpha_2}{\alpha_1} (N_s - N_{s-})\} = \begin{cases} 
\frac{\alpha_2}{\alpha_1}, & \text{if there is a jump at time } s, \\
1, & \text{if there is no jump at time } s.
\end{cases}
\]

Thus we obtain

\[
\varphi_s = \begin{cases} 
\frac{\alpha_2}{\alpha_1} \varphi_{s-}, & \text{if there is a jump at time } s, \\
\varphi_{s-}, & \text{if there is no jump at time } s.
\end{cases}
\]

It is clear, that

\[
\frac{\varphi_s}{1 + \varphi_s} - \frac{\varphi_{s-}}{1 + \varphi_{s-}} = 0,
\]

if there is no jump at the time $s$. Suppose now, that there is a jump, then $\varphi_s = \frac{\alpha_2}{\alpha_1} \varphi_{s-}$ and

\[
\frac{\varphi_s}{1 + \varphi_s} - \frac{\varphi_{s-}}{1 + \varphi_{s-}} = \frac{\frac{\alpha_2}{\alpha_1} \varphi_{s-}}{1 + \varphi_{s-}} - \frac{\varphi_{s-}}{1 + \varphi_{s-}} = \frac{\alpha_2 \varphi_{s-}}{\alpha_1 + \alpha_2 \varphi_{s-}} - \frac{\varphi_{s-}}{1 + \varphi_{s-}}
\]

\[
= \frac{\left(\frac{\alpha_2}{\alpha_1} - 1\right) \varphi_{s-}}{(1 + \varphi_{s-})(1 + \frac{\alpha_2}{\alpha_1} \varphi_{s-})}.
\]

Since $\Delta N_s = 1$ if it is a jump at $s$ and $0$ otherwise, we write

\[
\frac{\varphi_s}{1 + \varphi_s} - \frac{\varphi_{s-}}{1 + \varphi_{s-}} = \frac{\left(\frac{\alpha_2}{\alpha_1} - 1\right) \varphi_{s-}}{(1 + \varphi_{s-})(1 + \frac{\alpha_2}{\alpha_1} \varphi_{s-})} \Delta N_s. \quad (4.21)
\]

Now, we rewrite the sum of jumps as

\[
\sum_{0 < s \leq t} \left[ \frac{\varphi_s}{1 + \varphi_s} - \frac{\varphi_{s-}}{1 + \varphi_{s-}} \right] = \int_0^t \frac{\left(\frac{\alpha_2}{\alpha_1} - 1\right) \varphi_{s-}}{(1 + \varphi_{s-})(1 + \frac{\alpha_2}{\alpha_1} \varphi_{s-})} dN_s
\]

and we obtain an integral form for $\pi_t$

\[
\pi_t = \pi + \int_0^t \frac{\lambda(\varphi_s + 1) - (\alpha_2 - \alpha_1)f(s)\varphi_s}{(1 + \varphi_s)^2} ds + \int_0^t \frac{\left(\frac{\alpha_2}{\alpha_1} - 1\right) \varphi_{s-}}{(1 + \varphi_{s-})(1 + \frac{\alpha_2}{\alpha_1} \varphi_{s-})} dN_s.
\]

It has the following differential form

\[
d\pi_t = \frac{\lambda(\varphi_t + 1) - (\alpha_2 - \alpha_1)f(t)\varphi_t}{(1 + \varphi_t)^2} dt + \frac{\left(\frac{\alpha_2}{\alpha_1} - 1\right) \varphi_{t-}}{(1 + \varphi_{t-})(1 + \frac{\alpha_2}{\alpha_1} \varphi_{t-})} dN_t.
\]
After we replace $\varphi_t$ through $\frac{\pi_t}{1-\pi_t}$, we get

$$d\pi_t = \lambda(1-\pi_t)dt + \frac{(\alpha_2 - \alpha_1)\pi_t(1-\pi_t)}{\alpha_2\pi_t + \alpha_1(1-\pi_t)} \left(dN_t - (\alpha_2\pi_t + \alpha_1(1-\pi_t))f(t)dt\right).$$  

(4.22)

We see (as in [11]) that $\varphi_t$ and $\pi_t$ are strong Markov processes under $P_\pi$ with respect to the natural filtrations which coincide with $(F_t^N)_{t \geq 0}$ respectively. Thus the infimum in (4.4) may be viewed as taken over all stopping times $\tau$ of $\pi_t$ and the optimal stopping problem (4.4) falls into the class of optimal stopping problems for Markov processes. Thus we proceed by finding of the infinitesimal operator for the Markov process $\pi_t$. 
Chapter 5

The free-boundary problem

In this Chapter we represent the optimal stopping problem from Chapter 4 as a free-boundary problem. An analytical solution of this problem is given for the case when the intensity of defaults increases.

5.1 An infinitesimal operator for process \((\pi_t)_{t \geq 0}\).

For the formulation of a free-boundary problem we need to find an infinitesimal operator for the Markov process \((\pi_t, F^N_t, P_\pi)_{t \geq 0}\) for \(\pi \in [0, 1]\) using the equation (4.22). Let \(F = F(t, \pi_t)\), \(F \in C^1(\mathbb{R}^+ \times [0, 1])\), then to proceed we need Theorem 6, The two-dimensional Ito-Doeblin formula for the processes with jumps (see Appendix C), in accordance to which we obtain the following equality for \(F(t, \pi_t)\)

\[
F(t, \pi_t) = F(0, \pi_0) + \int_0^t F_t(s, \pi_s)ds + \int_0^t F_s(s, \pi_s)d\pi_s + \sum_{s \leq t} [F(s, \pi_s) - F(s, \pi_{s-})].
\]

We represent the latter sum as an integral with respect to some jump process, for this purpose we express \(\pi_t\) through \(\pi_{t-}\) if there is a jump at the time \(t\)

\[
\pi_t = \frac{\varphi_t}{1 + \varphi_t} = \frac{\alpha_2 \pi_{t-}}{\alpha_1 + (\alpha_2 - \alpha_1) \pi_{t-}},
\]

then

\[
\sum_{s \leq t} [F(s, \pi_s) - F(s, \pi_{s-})] = \sum_{s \leq t} \left[ F \left( s, \frac{\alpha_2 \pi_{s-}}{\alpha_1 + (\alpha_2 - \alpha_1) \pi_{s-}} \right) - F(s, \pi_{s-}) \right] \Delta N_s =
\]

\[
= \int_0^t \left[ F \left( s, \frac{\alpha_2 \pi_{s-}}{\alpha_1 + (\alpha_2 - \alpha_1) \pi_{s-}} \right) - F(s, \pi_{s-}) \right] dN_s.
\]
To obtain an infinitesimal operator for the process $\pi_t$ we need a compensated Poisson process which is a martingale under the measure $P_\pi$. Follow by [11] define the process $(\overline{N}_t)_{t \geq 0}$ as

$$\overline{N}_t = N_t - \int_0^t E(f(s)L_s|\mathcal{F}_s^N)ds = N_t - \int_0^t f(s)(\alpha \pi s_+ + \alpha_1(1 - \pi s_+))ds, \quad (5.2)$$

then

$$d\overline{N}_t = dN_t - f(t)(\alpha \pi t_+ + \alpha_1(1 - \pi t_+))dt.$$

**Proposition 4.** The process $\overline{N}_t$ defined by (5.2) is a martingale under the measure $P_\pi$.

**Proof.** To prove this proposition we consider the Poisson process $N$ which was defined in the following form

$$N_t = \int_0^t \mathbb{1}(L_s = \alpha_1)dN_t^{\gamma_1} + \int_0^t \mathbb{1}(L_s = \alpha_2)dN_t^{\gamma_2},$$

where $N^\gamma_1, N^\gamma_2$ are Poisson processes, and $L = (L_t)_{t \geq 0}$ is a continuous Markov chain. Then the compensated Poisson processes given by

$$\overline{N}_t^{\gamma_1} = N_t^{\gamma_1} - \alpha_1 \int_0^t f(s)ds$$

and

$$\overline{N}_t^{\gamma_2} = N_t^{\gamma_2} - \alpha_2 \int_0^t f(s)ds.$$

Following Theorem 7 (see Appendix C) from [16] we get that the processes $\overline{N}_t^{\gamma_1}$ and $\overline{N}_t^{\gamma_2}$ are martingales under the measure $P_\pi$. Using this fact we can rewrite the process $N$ as

$$N_t = M_t + \alpha_1 \int_0^t \mathbb{1}\{L_{s-} = \alpha_1\}f(s)ds + \alpha_2 \int_0^t \mathbb{1}\{L_{s-} = \alpha_2\}f(s)ds,$$

where

$$M_t = \int_0^t \mathbb{1}\{L_{s-} = \alpha_1\}d\overline{N}_t^{\gamma_1} + \int_0^t \mathbb{1}\{L_{s-} = \alpha_2\}d\overline{N}_t^{\gamma_2}.$$
is a martingale under the measure $P_{\pi}$. After all these arguments we write the process $N_t$ in the following form

$$N_t = M_t - \int_0^t f(s)(\alpha_2 \pi_s + \alpha_1(1 - \pi_s))ds + \int_0^t f(s)(\alpha_1 I\{L_s = \alpha_1\} + \alpha_2 I\{L_s = \alpha_2\})ds.$$

To prove that $N_t$ is a martingale under the measure $P_{\pi}$ we show that the process $A_t$ defined as

$$A_t = \int_0^t f(s)(\alpha_1 I\{L_s = \alpha_1\} + \alpha_2 I\{L_s = \alpha_2\})ds - \int_0^t f(s)(\alpha_2 \pi_s + \alpha_1(1 - \pi_s))ds$$

is a martingale under the measure $P_{\pi}$. Recall that

$$\pi_s = P\{\theta \leq s | \mathcal{F}_s\} = E[I\{L_s = \alpha_2\} | \mathcal{F}_s],$$

and

$$1 - \pi_s = P\{\theta > s | \mathcal{F}_s\} = E[I\{L_s = \alpha_1\} | \mathcal{F}_s].$$

If we assume that $u < t$, then

$$E[A_t | \mathcal{F}_u] = \int_0^t f(s)(\alpha_1 E[I\{L_s = \alpha_1\} | \mathcal{F}_u] + \alpha_2 E[I\{L_s = \alpha_2\} | \mathcal{F}_u])ds - \int_0^t f(s)(\alpha_2 E[\pi_s | \mathcal{F}_u] + \alpha_2 E[(1 - \pi_s) | \mathcal{F}_u]) =$$

$$A_u + \int_u^t f(s)(\alpha_1 E[I\{L_s = \alpha_1\} | \mathcal{F}_u] + \alpha_2 E[I\{L_s = \alpha_2\} | \mathcal{F}_u])ds - \int_u^t f(s)(\alpha_2 E[\pi_s | \mathcal{F}_u] + \alpha_2 E[(1 - \pi_s) | \mathcal{F}_u]).$$

For any $s \in [u, t]$ we can see that

$$E[E[I\{L_s = \alpha_1\} | \mathcal{F}_s] | \mathcal{F}_u] = E[I\{L_s = \alpha_1\} | \mathcal{F}_u],$$
Chapter 5. The free-boundary problem

\[
E[E[I\{L_{s-} = \alpha_2\}|\mathcal{F}_s]|\mathcal{F}_u] = E[I\{L_{s-} = \alpha_2\}|\mathcal{F}_u]
\]

and using this we obtain

\[
E[A_t|\mathcal{F}_u] = A_u + \int_0^t f(s)(\alpha_1 E[I\{L_{s-} = \alpha_1\}|\mathcal{F}_u] + \alpha_2 E[I\{L_{s-} = \alpha_2\}|\mathcal{F}_u])ds - \int_0^t f(s)(\alpha_2 E[I\{L_{s-} = \alpha_2\}|\mathcal{F}_u] + \alpha_1 E[I\{L_{s-} = \alpha_1\}|\mathcal{F}_u])ds = A_u.
\]

Hence the process \(A_t\) is a martingale under the measure \(P_\pi\) and \(\overline{N}_t = M_t + A_t\) is a martingale too.

Using (5.2) we write (4.22) in the following form

\[
d\pi_t = \lambda(1 - \pi_t)dt + \frac{(\alpha_2 - \alpha_1)\pi_{t-}(1 - \pi_{t-})}{\alpha_2\pi_{t-} + \alpha_1(1 - \pi_{t-})} d\overline{N}_t. \tag{5.3}
\]

Based on (4.22) and the definition of the jump process we can write

\[
d\pi_t^c = (\lambda(1 - \pi_t) - f(t)(\alpha_2 - \alpha_1)\pi_{t-}(1 - \pi_{t-}))dt,
\]

so

\[
F(t, \pi_t) = F(0, \pi_0) + \int_0^t F_i(s, \pi_s)ds + \int_0^t F_{\pi}(s, \pi_s)(\alpha_2 - \alpha_1)\pi_{s-}(1 - \pi_{s-}))ds + \int_0^t \left[F\left(s, \frac{\alpha_2\pi_{s-}}{\alpha_1 + (\alpha_2 - \alpha_1)\pi_{s-}}\right) - F(s, \pi_{s-})\right] d\overline{N}_s - f(s)(\alpha_2\pi_{s-} + \alpha_1(1 - \pi_{s-}))ds.
\]

Due to the fact that the process \(\overline{N}_t\) is a martingale under \(P_\pi\) we have

\[
F(t, \pi_t) = F(0, \pi_0) + \int_0^t \left[F\left(s, \frac{\alpha_2\pi_{s-}}{\alpha_1 + (\alpha_2 - \alpha_1)\pi_{s-}}\right) - F(s, \pi_{s-})\right] d\overline{N}_s + \int_0^t (\mathbb{L}F)(s, \pi_s)ds,
\]
where

\[(\mathbb{L}F)(t, \pi) = F_t(t, \pi) + F_\pi(t, \pi)(\lambda(1 - \pi) - f(t)(\alpha_2 - \alpha_1)\pi(1 - \pi)) +
\]

\[f(t)(\alpha_2\pi + \alpha_1(1 - \pi)) \left[ F\left(t, \frac{\alpha_2\pi}{\alpha_1 + (\alpha_2 - \alpha_1)\pi}\right) - F(t, \pi) \right].\]

### 5.2 The free-boundary problem

From (4.22) we see that the distribution law of \((\pi_t)_{t \geq 0}\) depends on the current state of \((\pi_t)_{t \geq 0}\) and on the time \(t\). Hence we have a dependence on the time in the optimal stopping problem (4.2) and (4.4).

The map \((t, \pi) \rightarrow V(t, \pi)\) is concave and decreasing on \([0, 1]\). The stopping time \(\tau^* = \inf\{t \geq 0 : \pi_t \geq B_*(t)\}\) is optimal in the problem (4.2) and (4.4), where \(B_*(t)\) is the smallest \(\pi\) from \([0, 1]\) satisfying \(V(t, \pi) = 1 - \pi\).

Thus \(V(t, \pi) < 1 - \pi\) for all \(\pi \in [0, B_*(t)]\) and \(V(t, \pi) = 1 - \pi\) for all \(\pi \in [B_*(t), 1]\).

Based on the general optimal stopping theory of the Markov processes and using the preceding facts we can formulate the free-boundary problem for \((t, \pi) \rightarrow V(t, \pi)\) and \(B_*(t)\) defined above as follows:

\[(\mathbb{L}V)(t, \pi) = -c\pi, \quad (0 < \pi < B_*(t)), \quad (5.4)\]

\[V(t, \pi) = 1 - \pi, \quad (B_*(t) \leq \pi \leq 1), \quad (5.5)\]

\[V(t, B_*(t) -) = 1 - B_*(t). \quad (5.6)\]

We assume that \(\alpha_2 > \alpha_1\) and consider the drift term of \((\pi_t)_{t \geq 0}\) from (4.22)

\[\lambda(1 - \pi) - f(t)(\alpha_2 - \alpha_1)\pi(1 - \pi) = f(t)(\alpha_2 - \alpha_1)\left(\frac{\lambda}{f(t)(\alpha_2 - \alpha_1)} - \pi\right)(1 - \pi),\]

where we insert \(\hat{B}(t) = \frac{\lambda}{f(t)(\alpha_2 - \alpha_1)}\), so the drift term is

\[\lambda(1 - \pi) - f(t)(\alpha_2 - \alpha_1)\pi(1 - \pi) = f(t)(\alpha_2 - \alpha_1)\left(\hat{B}(t) - \pi\right)(1 - \pi).\]

The sign of the drift term is determined by the sign of \((\hat{B} - \pi)\), hence we deduce that \((\pi_t)_{t \geq 0}\) has a positive drift in \([0, \hat{B}]\), a negative drift in \([\hat{B}, 1]\) and a zero drift at \(\hat{B}\).
Now we consider the optimal stopping problem (4.4) taking into account that the process (5.2) is a martingale under \( P_\pi \) and that by (5.3)
\[
\pi_t = \pi + \lambda \int_0^t (1 - \pi_s) ds + M_t,
\]
where
\[
M_t = \int_0^t \frac{(\alpha_2 - \alpha_1) \pi_s (1 - \pi_s)}{\alpha_2 \pi_s - \alpha_1 (1 - \pi_s)} dN_s
\]
is a martingale under \( P_\pi \), \( M_0 = 0 \) and hence by the optional sampling theorem \( E[M_t] = 0 \). So from (4.4) we can find that
\[
V(\pi) = E_\pi \left( 1 - \pi_\tau + c \int_0^\tau \pi_t dt \right)
\]
\[
= E_\pi \left( 1 - \pi - \lambda \int_0^\tau (1 - \pi_t) dt + c \int_0^\tau \pi_t dt \right)
\]
\[
= (1 - \pi) + (\lambda + c) E_\pi \left( \int_0^\tau \left( \pi_t - \frac{\lambda}{\lambda + c} \right) dt \right)
\]
for all stopping times \( \tau \) of \( (\pi_t)_{t \geq 0} \). Let us denote
\[
\tilde{B} = \frac{\lambda}{\lambda + c}.
\]
The integral \( \int_0^\tau \left( \pi_t - \frac{\lambda}{\lambda + c} \right) dt \) in (5.7) decreases whenever \( \pi_t < \tilde{B} \). Suppose now, that \( \tilde{B} < \tilde{B}(t) \) for all \( t \geq 0 \), then if \( (\pi_t)_{t \geq 0} \) once leaves the interval \([0, \tilde{B}]\) it will never return. After \( \pi_t \) crosses the barrier \( \tilde{B} \) the value of \( \int_0^\tau \left( \pi_t - \frac{\lambda}{\lambda + c} \right) dt \) only increases. Since it is never optimal to stop until \( \pi_t < \tilde{B} \), as well as that the process \( (\pi_t)_{t \geq 0} \) should be stopped when it enters \([\tilde{B}, 1]\), these arguments prove that the optimal threshold is
\[
B_*(t) = \tilde{B},
\]
(5.10)
and the time

\[ \tau_\ast = \inf \{ t \geq 0 : \pi_t \geq \tilde{B} \} \]  

is an optimal stopping time for the problem (4.2) and (4.4).

**Remark 1.** We can see that the inequality in the supposition (5.9) can be rewritten in the following form

\[ \frac{\lambda}{\lambda + c} \leq \frac{\lambda}{f(t)(\alpha_2 - \alpha_1)}, \]

or

\[ f(t) \leq \frac{\lambda + c}{\alpha_2 - \alpha_1}. \]
Chapter 6

The scheme of a real time detection of a disorder moment

Now we consider how to use the obtained results in order to find an optimal stopping moment $\tau_*$ of the Poisson disorder problem. Suppose that we observe in the real time a Poisson process $N$, then to find the value of $\tau_*$ the following procedure can be applied:

1. we assume that the constants $c, \alpha_1, \lambda, \pi$ are given, the function $f(t)$ is known and the constant $\alpha_2$ is chosen in accordance to the change of the intensity which is significant for the credit organization;

2. we find the value of $\pi_t$ using the stochastic differential equation (4.22);

3. according to the solution of the considered free-boundary problem we terminate the observation of the process $N$ when the value of $\pi_t$ becomes greater or equal to $B_* = \frac{\lambda}{\lambda+c}$.

Now we describe how to find an exact value of $\pi_t$ based on the observations of the process $N$ and using the equation (4.22). We rewrite this equation in the following form

$$d\pi_t = (\lambda(1-\pi_t) - f(t)(\alpha_2 - \alpha_1)\pi_t(1-\pi_t)) \, dt + \frac{(\alpha_2 - \alpha_1)\pi_t(1-\pi_t)}{\alpha_2\pi_t + \alpha_1(1-\pi_t)} \, dN_t.$$  

(6.1)

From this equation we can see that the dynamics of $\pi_t$ consists of the drift term and the jump term. Between jumps the value of $\pi_t$ depends only on the drift term, as the jump term is equal to zero. So between jumps $\pi_t$ must solve the following ordinary differential equation

$$d\pi_t = (\lambda(1-\pi_t) - f(t)(\alpha_2 - \alpha_1)\pi_t(1-\pi_t)) \, dt.$$  

(6.2)
with initial conditions which should be changed after each jump.

**Proposition 5.** Equation (6.2) has an explicit solution

\[
\pi_t = 1 + \frac{1}{G_1(t)(G_2(t) + C_i)},
\]

(6.3)

where \(C_i\) is a constant which should be calculated after the \(i\)-th jump in accordance to the new initial conditions caused by this jump. Functions \(G_1(t)\) and \(G_2(t)\) admit following explicit formulas

\[
G_1(t) = \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s)ds + \lambda t \right\},
\]

\[
G_2(t) = \int_0^t (\alpha_1 - \alpha_2) f(u) \exp \left\{ -(\alpha_1 - \alpha_2) \int_0^u f(s)ds - \lambda t \right\} du.
\]

**Proof.** Rewrite equation (6.2) in the following form

\[
\frac{d\pi_t}{dt} = \pi_t^2 (\alpha_2 f(t) - \alpha_1 f(t)) + \pi_t (\alpha_1 f(t) - \alpha_2 f(t) - \lambda) + \lambda,
\]

we see that this is a Riccati equation with a partial solution \(\pi_t^* \equiv 1\). Using the substitution

\[
\pi_t = \pi_t^* + \frac{1}{y(t)} = 1 + \frac{1}{y(t)},
\]

where \(y(t)\) is a new unknown function, we reduce the Riccati equation to the following linear differential equation

\[
\frac{d(1 + 1/y(t))}{dt} = (1 + 1/y(t))^2 (\alpha_2 f(t) - \alpha_1 f(t)) + (1 + 1/y(t))(\alpha_1 f(t) - \alpha_2 f(t) - \lambda) + \lambda,
\]

from which we obtain a non-homogeneous linear differential equation in the form

\[
\frac{dy(t)}{dt} = y(t)(\alpha_1 f(t) - \alpha_2 f(t) + \lambda) - \alpha_2 f(t) + \alpha_1 f(t).
\]

Using the solution of the corresponding homogeneous equation we obtain

\[
y(t) = C(t) \exp \left\{ (\alpha_1 - \alpha_2) \int_0^t f(s)ds + \lambda t \right\} = C(t)G_1(t),
\]
where

\[ C(t) = \int_0^t (\alpha_1 - \alpha_2) f(u) \exp \left\{ -(\alpha_1 - \alpha_2) \int_0^u f(s) ds - \lambda t \right\} du + C_i = G_2(t) + C_i. \]

So

\[ y(t) = G_1(t)(G_2(t) + C_i), \]

and

\[ \pi_t = 1 + \frac{1}{G_1(t)(G_2(t) + C_i)}. \]

Now using the results of the latter proposition we can describe the procedure of finding of the value \( \pi_t \) from jump to jump.

**Before the first jump.** We have an initial condition for (6.2) in the form \( \pi_0 = \pi \). We assume that \( \pi_t \) satisfies the formula (6.3) with a constant \( C_0 \)

\[ \pi_t = 1 + \frac{1}{G_1(t)(G_2(t) + C_0)} \]

until the first jump, where \( C_0 \) can be obtained from the initial condition, so

\[ C_0 = \frac{1}{G_1(0)(\pi - 1)} - G_2(0). \]

**At the time of the first jump.** Assume that the first jump occurs at the time \( t_1 \). Just before the jump we have

\[ \pi_{t_1-} = 1 + \frac{1}{G_1(t_1)(G_2(t_1) + C_0)}, \]

just after the jump

\[ \pi_{t_1} = \pi_{t_1-} + \frac{(\alpha_2 - \alpha_1)\pi_{t_1-}(1 - \pi_{t_1-})}{\alpha_2\pi_{t_1-} + \alpha_1(1 - \pi_{t_1-})}, \]

where \( \frac{(\alpha_2 - \alpha_1)\pi_{t_1-}(1 - \pi_{t_1-})}{\alpha_2\pi_{t_1-} + \alpha_1(1 - \pi_{t_1-})} \) is a jump term from equation (6.1). After the first jump \( \pi_t \) satisfies the same formula (6.3) but with an another constant \( C_1 \) instead of \( C_0 \). As we know the new initial condition \( \pi_{t_1} \), we can find the value of the constant \( C_1 \)

\[ C_1 = \frac{1}{G_1(t_1)(\pi_{t_1} - 1)} - G_2(t_1). \]
At the time of the i-th jump. Assume that the jump occurs at the time $t_i$. Just before the i-th jump we have

$$
\pi_{t_{i-}} = 1 + \frac{1}{G_1(t_i)(G_2(t_i) + C_{i-1})},
$$

just after the i-th jump

$$
\pi_{t_i} = \pi_{t_{i-}} + \frac{(\alpha_2 - \alpha_1)\pi_{t_{i-}}(1 - \pi_{t_{i-}})}{\alpha_2\pi_{t_{i-}} + \alpha_1(1 - \pi_{t_{i-}})}.
$$

After the i-th jump $\pi_t$ satisfies the same formula (6.3) but with a constant $C_i$ instead of $C_{i-1}$. As we know the initial condition $\pi_{t_{i-}}$ at $t_i$, we find the value of a constant $C_i$

$$
C_i = \frac{1}{G_1(t_i)(\pi_{t_i} - 1) - G_2(t_i)}.
$$

Remark 2. We proceed the observation of the process $N$ until the value of $\pi_t$ becomes greater or equal to $B_\ast$. As it happens we terminate the observation of the process $N$ and declare that the disorder has already occurred.
Chapter 7

Conclusions

In this work we addressed the problems of the loan origination decision-making systems. The goal of this work was to estimate the disorder moment, i.e. the moment when the intensity of defaults increases significantly. The problem was solved using two approaches: the change-point detection for a given sample and the real time detection of the disorder moment.

For the first approach results are following. We obtained the estimated logit and by the $p$-values, we can conclude that all variables are significant in our model. Also after looking for the area under ROC curve that for our data it is almost excellent discrimination.

We considered a simple cumulative sum type statistic for the change-point with the zero-one observations. We compared a conditional test of no change against the change with a likelihood ratio test. Using the simple statistic we also considered an estimation of the change-point.

Then after results in the change point problem in the discrete time in the model of a logistic regression we assume that from 6665 clients we need to change our regression model to work properly for further using. Also using calculation from our thesis we detect the change point problem in the continuous time.

For the problem of the real time detection it was assumed that the process of clients defaults is a non-homogeneous Poisson process with a single passage in the intensity at some unknown time. Based on the paper by Peskir and Shiryaev [11] the problem of finding a passage time was solved as a disorder problem (4.2) for the non-homogeneous Poisson process. Using the a-posteriori probability process we reduced the optimal stopping problem (4.4) to the free-boundary problem (5.4)-(5.6).

The case of interest is when the intensity of defaults increases, in this case the complete analytical solution for the free-boundary problem was obtained. Moreover using these results there was represented the scheme of the real time
detection of the disorder moment and given an explicit formula (6.3) for the a posteriori probability process.
Chapter 8

Further research

In this work we were focused on the solution of the disorder problem in the case when the intensity of the process of defaults increases. Due to the initial problem of the loan origination it is natural also to consider the case when the intensity of defaults decreases, because in this situation an initial model of the clients parameters estimation need to be corrected as well. In the latter case the solution of the free-boundary problem (5.4)-(5.6) will differ from the solution introduced in this work. While considering the case when the intensity of defaults decreases an infinitesimal operator from the section 5.1 will be used directly. It is important to mention that the scheme of the practical realization from Chapter 6 will be the same and the only thing that will be changed is a free-boundary $B_*(t)$.

One other way for the extension of this research is to consider the disorder problem for the compound Poisson process with multiple hypotheses about the values of intensities and distributions of jumps. Such study can be based on the paper by Dayanik, Poor and Sezer [6].
Notation

\( N = (N_t)_{t \geq 0} \) Poisson process.

\( (\mathcal{F}_t^N)_{t \geq 0} \) filtration generated by \( N \).

\( \gamma_t \) intensity of a non-homogeneous Poisson process.

\( \theta \) disorder time.

\( L = (L_t)_{t \geq 0} \) continuous Markov chain.

\( \pi_t \) a posteriori probability process.

\( P_\pi \) initial measure.

\( \Gamma_t \) cumulative intensity function.

\( \mathbb{1} \) indicator function.

\( \varphi_t \) likelihood ratio process.

\( Z_t \) likelihood process.

\( E_\pi \) mathematical expectation with respect to measure \( P_\pi \).

\( \Psi_t(v) \) moment generating function.

\( \overline{N}_t = (\overline{N}_t)_{t \geq 0} \) compensated Poisson process.

\( V(t, \pi) \) loss function.

\( \mathbb{L} \) infinitesimal operator.

\( B_* \) free boundary.

\( \tau \) change point.

\( \hat{\tau} \) maximum estimator of \( \tau \).

\( L(t) \) likelihood function.

\( \alpha \) significance level.

\( W \) random walk.
Bibliography


_The problem of the Nile: Conditional solution to the change-point problem._ Biometrika 65, 243 – 251.


_Journal of Statistical Planning and Inference._ Change point problems in the model of logistic regression, 313 – 331.

_Inference about the change-point in a sequence of binomial variables._ Biometrika 57, 477 – 488.
Inference about the change-point from cumulative sum tests. Biometrika 53, 509–523.


### Appendix A

<table>
<thead>
<tr>
<th>N</th>
<th>Age</th>
<th>Education</th>
<th>Occupation</th>
<th>ExTime</th>
<th>Amount</th>
<th>Default</th>
<th>ed₁</th>
<th>ed₂</th>
<th>oc₁</th>
<th>oc₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39</td>
<td>Prof</td>
<td>Transport</td>
<td>12</td>
<td>115000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>Stnd</td>
<td>Transport</td>
<td>2</td>
<td>110000</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>High</td>
<td>State</td>
<td>23</td>
<td>140000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
<td>High</td>
<td>State</td>
<td>17</td>
<td>145000</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>Prof</td>
<td>Industry</td>
<td>12</td>
<td>65000</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>Prof</td>
<td>Industry</td>
<td>9</td>
<td>35000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>Stnd</td>
<td>Transport</td>
<td>13</td>
<td>105000</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>46</td>
<td>Prof</td>
<td>Transport</td>
<td>29</td>
<td>125000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>41</td>
<td>Stnd</td>
<td>Industry</td>
<td>14</td>
<td>95000</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>31</td>
<td>Stnd</td>
<td>Industry</td>
<td>14</td>
<td>80000</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>32</td>
<td>High</td>
<td>State</td>
<td>15</td>
<td>145000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>48</td>
<td>High</td>
<td>State</td>
<td>25</td>
<td>70000</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>49</td>
<td>High</td>
<td>Transport</td>
<td>26</td>
<td>120000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>23</td>
<td>Prof</td>
<td>Industry</td>
<td>6</td>
<td>35000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>49</td>
<td>Stnd</td>
<td>Transport</td>
<td>32</td>
<td>45000</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>43</td>
<td>Prof</td>
<td>Transport</td>
<td>20</td>
<td>50000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>36</td>
<td>High</td>
<td>Industry</td>
<td>19</td>
<td>65000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>23</td>
<td>High</td>
<td>State</td>
<td>6</td>
<td>60000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>Stnd</td>
<td>Industry</td>
<td>15</td>
<td>135000</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>24</td>
<td>Stnd</td>
<td>Industry</td>
<td>1</td>
<td>90000</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>50</td>
<td>High</td>
<td>State</td>
<td>27</td>
<td>150000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>25</td>
<td>Prof</td>
<td>Industry</td>
<td>2</td>
<td>65000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>21</td>
<td>High</td>
<td>Transport</td>
<td>0</td>
<td>65000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>47</td>
<td>Prof</td>
<td>Transport</td>
<td>20</td>
<td>125000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>47</td>
<td>High</td>
<td>State</td>
<td>20</td>
<td>95000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 8.1: Some examples of real statistical data.
Appendix B

In this Appendix we would present the basic theory related to jump processes which was used during the study. Following by Shreve [16] consider the stochastic process $X(t)$ of the following form

$$X(t) = X(0) + I(t) + R(t) + J(t), \quad (8.1)$$

where

- $X(0)$ is the non-random initial condition,
- $I(t) = \int_0^t \Gamma(s)dW(s)$ is the Ito integral of the adapted process $\Gamma(s)$ with respect to the Brownian motion relative to the considered filtration,
- $R(t) = \int_0^t \Theta(s)ds$ is the Riemann integral for some adapted process $\Theta(s)$, and
- $J(t)$ is a pure jump process or a jump part of $X(t)$.

The continuous part of $X(t)$ is defined as

$$X^c(t) = X(0) + I(t) + R(t) = X(0) + \int_0^t \Gamma(s)dW(s) + \int_0^t \Theta(s)ds.$$

The quadratic variation of this process is

$$[X^c, X^c](t) = \int_0^t \Gamma^2(s)ds.$$

**Definition 5.** A process $X(t)$ of the form (8.1), with the Ito integral part $I(t)$, Riemann integral part $R(t)$ and the pure jump part $J(t)$ as described above, is called a jump process.

The jump size of $X(t)$ at the time $t$ is denoted by

$$\Delta X(t) = X(t) - X(t-).$$
Appendix C

**Theorem 2.** (Corollary 11.5.5 from Shreve [16], Chapter 11, Section 11.5). Let $X_1(t)$ and $X_2(t)$ be jump processes. Then

$$X_1(t)X_2(t) = X_1(0)X_2(0) + \int_0^t X_2(s-)dX_1(s) + \int_0^t X_1(s-)dX_2(s) + [X_1, X_2](t).$$

**Theorem 3.** (Theorem 11.4.7 from [16], Chapter 11, Section 11.4). Let $X_1(t) = X_1(0) + I_1(t) + R_1(t) + J_1(t)$ be a jump process, where $I_1(t) = \int_0^t \Gamma_1(s)dW(s)$, $R_1(t) = \int_0^t \Theta_1(s)ds$, and $J_1(t)$ is a right continuous jump process. Then $X_1^c(t) = X_1(0) + I_1(t) + R_1(t)$, and

$$[X_1, X_1](t) = [X_1^c, X_1^c](t) + [J_1, J_1](t) = \int_0^t \Gamma_1^2(s)ds + \sum_{0<s\leq t} (\Delta J_1(s))^2.$$

Let $X_2(t) = X_2(0) + I_2(t) + R_2(t) + J_2(t)$ be a jump process, where $I_2(t) = \int_0^t \Gamma_2(s)dW(s)$, $R_2(t) = \int_0^t \Theta_2(s)ds$, and $J_2(t)$ is a right continuous jump process. Then $X_2^c(t) = X_2(0) + I_2(t) + R_2(t)$, and

$$[X_1, X_2](t) = [X_1^c, X_2^c](t) + [J_1, J_2](t) = \int_0^t \Gamma_1^2(s)\Gamma_2^2(s)ds + \sum_{0<s\leq t} \Delta J_1(s)\Delta J_2(s).$$

**Theorem 4.** (Lemma 11.6.1 from [16], Chapter 11, Section 11.6).

The process $Z(t) = e^{(\lambda - \tilde{\lambda})t} \left( \begin{array}{c} \lambda \\ \tilde{\lambda} \end{array} \right)^N(t)$ satisfies the condition

$$dZ(t) = \frac{\tilde{\lambda} - \lambda}{\lambda} Z(t-)dM(t).$$

In particular, $Z(t)$ is a martingale under $\mathbb{P}$ and $\mathbb{E}Z(t) = 1$ for all $t$. 

67
Theorem 5. (Theorem 11.5.1 from [10], Chapter 11, Section 11.5). Let $X(t)$ be a jump process and $f(x)$ be a function for which $f'(x)$ and $f''(x)$ are well defined and continuous. Then

$$f(X(t)) = f(X(0)) + \int_0^t f'(X(s))dX^c(s) + \frac{1}{2} \int_0^t f''(X(s))dX^c(s)dX^c(s) + \sum_{0<s\leq t} [f(X(s)) - f(X(s-))].$$

Theorem 6. (Theorem 11.5.4 from [10], Chapter 11, Section 11.5). Let $X_1(t)$ and $X_2(t)$ be jump processes, and let $f(t,x_1,x_2)$ be a function whose the first and second derivatives appearing in the following formula are well defined and continuous. Then

$$f(t, X_1(t), X_2(t)) = f(0, X_1(0), X_2(0)) + \int_0^t f_t(s, X_1(s), X_2(s))ds +$$

$$\int_0^t f_{x_1}(s, X_1(s), X_2(s))dX_1^c(s) + \int_0^t f_{x_2}(s, X_1(s), X_2(s))dX_2^c(s) +$$

$$\frac{1}{2} \int_0^t f_{x_1,x_1}(s, X_1(s), X_2(s))dX_1^c(s)X_1^c(s) + \int_0^t f_{x_1,x_2}(s, X_1(s), X_2(s))dX_1^c(s)X_2^c(s) +$$

$$\frac{1}{2} \int_0^t f_{x_2,x_2}(s, X_1(s), X_2(s))dX_2^c(s)X_2^c(s) + \sum_{0<s\leq t} [f(s, X_1(s), X_2(s)) - f(s, X_1(s-), X_2(s-))].$$

Theorem 7. (Theorem 11.2.4 from [10], Chapter 11, Section 11.2). Let $N(t)$ be a Poisson process with intensity $\lambda$. We define the compensated Poisson process by

$$M(t) = N(t) - \lambda t.$$

Then $M(t)$ is a martingale.