Modeling and monitoring of the price process of Credit Default Swaps

Master's Thesis in Financial Mathematics

Anna Loshkina and Elena Malysheva
Modeling and monitoring of the price process of Credit Default Swaps

Anna Loshkina and Elena Malysheva

Halmstad University
Project Report IDE0837

Master’s thesis in Financial Mathematics, 15 ECTS credits
Supervisor: Ph.D. Eric Järpe
Examiner: Prof. Ljudmila A. Bordag
External referees: Prof. Vladimir Roubtsov, Prof. Lioudmila Vostrikova

November 1, 2008

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

We especially would like to thank our supervisor, Eric Järpe, who strongly helped us during this difficult time of working on our master thesis. We also want say a lot of thanks for our examiner, prof. Ljudmila A. Bordag, and other teachers at Halmstad University, who helped us during this year.
Abstract
Credit derivatives are very popular on financial markets in recent days. The most liquid credit derivative is a credit default swap (CDS). In this research we investigate methods for modeling and monitoring of the price process of CDS. We study Hull and White model to calculate CDS spread and have data for our analysis. We consider different methods for monitoring of the price process of CDS. In particular we study CUSUM method. And we calculate more commonly used performance measures for this method.
## Contents

1 Introduction ............................................. 1

2 Methods .................................................. 3
   2.1 Introduction to CDS ................................. 3
   2.2 Payoffs of CDS ........................................ 4
   2.3 Pricing and valuation of CDS ......................... 5
   2.4 ARARCH model ........................................ 11
   2.5 Change-point problem ................................. 12
   2.6 Performance measures ................................. 13
   2.7 Change-point detection methods ..................... 14
      2.7.1 Shewhart method ................................. 14
      2.7.2 The cumulative sum method (CUSUM) ... 15
      2.7.3 Shiryaev-Roberts method ....................... 15
      2.7.4 EWMA method ..................................... 15

3 Results .................................................... 17
   3.1 Plots of CDS spread ................................. 17
   3.2 Parameter estimation ................................. 18
   3.3 Recursive algorithm ................................. 21
   3.4 Calculating the Performance $ARL^0$ and $ARL^1$ for CUSUM ... 23

4 Conclusions ............................................... 25

Bibliography ................................................. 27

Appendix ..................................................... 29
Chapter 1

Introduction

In Russia first operations with derivatives started in 1992. From 1994 they became to be regular on Moscow Central Stock Exchange. However, a high level of risk on such a difficult market (as derivatives) led to a decrease of trading activity. There was a high level of uncertainty on Russian market, state regulation of operations with derivatives was very low. Financial crisis in Russia in 1998 was the last point in existence of derivatives in that moment. However, after that crisis derivatives began to take part in Russian market again. The actual problem in Russia now is to build a civilized market of derivatives. This is confirmed by experience from the past. The necessity to develop of derivative’s market is conditioned by deficiency in managing risks instruments in the Russian financial system. The prospective direction in development of derivatives in Russia is spreading of a rather new type – credit derivatives – for management of credit risks. They were registered by the International Derivatives Association in 1992. Credit derivatives are very popular on financial markets in recent days. The most liquid credit derivative is a credit default swap (CDS). The first CDS were developed in 1995. In 2007 new methods for evaluating CDS were introduced with using other instruments. According to the report of the Bank for International Settlements in 2007 the notional amount on outstanding over-the-counter CDS is around $ 42.6 trillion. Some dealers in Europe announce a value of CDS spread every day. Broad using of credit derivatives in Russia may make the work of the bank system more effective. So in consequence of all what is said before, one part of our research is dedicated to describing of CDS.

Market of securities is changeable, therefore management of a portfolio is necessary. It means that is necessary:

1. to save originally invested means;
2. to reach a maximum level of the income.

There are two ways of management of a portfolio: active and passive. The active model of management offers careful tracking and immediate purchase of tools
which satisfy to the investment purposes of a portfolio. Passive management assumes creation of diversification portfolios with in advance certain risk level.

Monitoring is a basis both active, and passive ways of management. The essence of a monitoring problem consists in following. Suppose we have process of identical independent distributed (i.i.d) observations. The distribution of this process till some change-point is known, but after this change-point the distribution of process became another. Suppose we are making observations online, i.e. one by one in time. Our problem consists of stopping this collection of observations as soon as the change has occurred. Such a question can be treated by method of statistical process control (SPC). The most important tools of SPC are control charts. There are many different control charts for monitoring: CUSUM control chart, EWMA control chart, Shiryaev-Robert’s control chart for instance. We will consider in detail CUSUM method and later we will compare CUSUM method with other methods.
Chapter 2

Methods

2.1 Introduction to CDS

A CDS, in its most general definition, is alike an insurance contract and protects the buyer against credit risks. One party – the protection buyer – pays a recurring fee. Another party – the protection seller – promises to pay some specified payment if credit event occurs. Simply, the holder of the security gives his risk of default to the seller of the swap. In other words, the seller pays an amount to cover losses if a credit event occurs during the life of the CDS. The basic mechanisms of a CDS are illustrated by Figure 2.1

CDS can be used for hedging, or insuring against for example default on a debt obligation or speculating. For hedging owners of a corporate bond protect themselves by purchasing a CDS. CDS are useful for speculating on changes in credit spread. A speculator has possibility to make a great profit from changes in a company’s credit quality. Usually CDS have from one to ten years maturity but the most traded is the 5-year CDS.

The regulator of CDS contracts is the International Swaps and Derivatives Association, in particular Credit Derivatives Definitions published in 2003. The confirmation typically specifies:

- reference entity (a company which has debt outstanding);
- reference obligation (bond issued by reference entity);
- effective date (the period of default protection);
- credit events (default by company);
- deliverable obligation characteristics (limit of obligation number that protection buyer can deliver on a credit event).

CDS notional principal is the total par value of the bond.
2.2 Payoffs of CDS

In this study we will consider 2 types of default swaps: standard and exotic (for more details see [1]). Let us describe the cash flows in standard default swaps. We will use following notation:

\( T \) – the maturity date;

\( L \) – face value of defaultable zero-coupon bond;

\( U \) – the maturity date of defaultable zero-coupon bond, \( U \geq T \);

\( \tau \) – the default time when the contingent payment is triggered;

\( T_i \) – time of credit swap premium.

Then payoff at default for standard swap is given in following way:

\[
(L - D(\tau, U))1_{\{\tau \leq T\}} = (L - D(\tau, U))^+ 1_{\{\tau \leq T\}},
\]

(2.1)

where \( D(\tau, U) \) is the value of the defaultable bond immediately after the bond has defaulted (post-default value).

The credit swap premium are payed by protection buyer at times \( T_i, i = 1, \ldots, m \) until default or maturity date (depend what happens first).

We can consider the alternative agreements for the recovery payments in following way:

\[
(LB(\tau, U) - D(\tau, U))1_{\{\tau \leq T\}},
\]

or

\[
(D(\tau -, U) - D(\tau, U))1_{\{\tau \leq T\}},
\]

where \( B(\tau, U) \) is the price at default time \( \tau \) of a \( U \)-maturity default-free bond with the face value 1. In the first case, the buyer is compensated for the loss.
of the defaultable bond’s value relative to the Treasury bond’s value at default time. In the second case, the protection buyer is compensated immediately after the default relative to the bond’s value instantly before the default.

The common expression 2.1 can be given in a more explicit form. For example, let $\delta$ be the fixed recovery rate. For Treasury value with fractional recovery formula 2.1 becomes:

$$L(1 - \delta B(\tau, U))1_{\{\tau \leq T\}}.$$

For par value with fractional recovery we will get the formula:

$$L(1 - \delta)1_{\{\tau \leq T\}}.$$

And for market value with fractional recovery formula 2.1 becomes

$$(L - \delta D(\tau - , U))1_{\{\tau \leq T\}}.$$

Exotic default swaps consider the agreements determining the triggering of credit event and/or agreements determining the amount of insurance payoff. There are different kinds of exotic CDS: digital, basket, contingent and dynamic CDS. For example, the payment at default is a pre-specified fixed amount in digital CDS. In the dynamic CDS the notional amount is the marked to market value of a portfolio of CDS. In the basket CDS the payment happens if the first one of a group of entities defaults. In the contingent contract payoff needs both the underlying credit event to occur and an extra trigger (for example, credit event with respect to other mention defaultable claim).

### 2.3 Pricing and valuation of CDS

There are two theories for the pricing of CDS. The first is by working with probabilities. Then the present value of CDS is the sum of cash flows weighted by their probability of non-default. The common generalization is that CDS should have a lower spread than corporate bonds. In the second approach non-arbitrage theory is used. It is proposed by Duffie and Hull and White.

Let us consider the first theory. In the probability model one uses the issue premium, the recovery rate, the credit curve for the reference entity and the LIBOR curve. LIBOR curve (or swap curve) is a type of yield curve which shows the relation between LIBOR rate (London Interbank Offered Rate) and the time to maturity. In case that default events never happen, the price of a CDS should be the sum of the discounted premium payments. So CDS pricing models consider case that default occur some time from the effective date to maturity date of the CDS contract. Let us consider a one year CDS with effective date $t_0$ with quarterly premium payments at times $t_i$, $i = 1, \ldots, n$.

We will use following notation:
**Chapter 2. Methods**

Figure 2.2: Tree diagram

$N$ – the nominal for the CDS,

c – the issue premium.

Then expression $\frac{Nc^4}{4}$ represents the quarterly premium payments. For simplicity we suppose that defaults can happen only on one of the payment dates. With this assumption there are following ways for end of contract:

- there is no default at all, the premium payments are made and the contract survives until the maturity date;

- a default occurs on the payment date,

For pricing the CDS we now need to find probabilities to all possible outcomes. Then the present value of CDS is the sum of the payoffs weighted by their probabilities. This is shown by Diagram 2.2 for case with $n=4$.

At each payment date:

- either the contract has a default event (shown in red on Figure 2.2; in this case a payment is equal $N(1 - r)$, where $r$ is recovery rate).

- or, there is no default (shown in blue; in this case a premium payment is equal to $\frac{Nc^4}{4}$).

Let the probability that there is no default on the interval from $t_{i-1}$ to $t_i$ be $p_i$ and in other case $1 - p_i$. The probability of default at time $t_i$ is $1 - p_i$ and in
this case of default at this first occasion there is no premium payment. Thus the
present value of default payment is
\[ N(1 - r)\delta_1. \]
The probability that default happens at time \( t_i, i = 2, \ldots, n \) is
\[ p_1p_2 \ldots p_{i-1}(1 - p_i) \]
then the present value of premium payment is
\[ -\frac{Nc}{n}(\delta_1 + \delta_2 + \ldots + \delta_{i-1}) \]
and the present value of default payment is
\[ N(1 - r)\delta_i. \]
The probability that there is no default at all is \( p_1p_2p_3 \ldots p_n \) then the present
value of premium payment is
\[ -\frac{Nc}{n}(\delta_1 + \delta_2 + \delta_3 + \ldots + \delta_n) \]
and there is no default payment. In all cases \( \delta_i \) are discount factors.
The probabilities \( p_i, i = 1, \ldots, n \) can be found from the credit spread curve. The probability that there is no default during time from \( t \) to \( t + \Delta t \) is commonly modeled by the formula: \( p = \exp(-s(t)\Delta t) \), where \( s(t) \) is the credit spread zero curve at time \( t \).

Then the total present value of the CDS is given by:
\[
PV = (1 - p_1)N(1 - r)\delta_1 + \\
+ \sum_{i=2}^{n} p_1p_2 \ldots p_{i-1}(1 - p_i)[N(1 - r)\delta_i - \frac{Nc}{n}(\delta_1 + \delta_2 + \delta_3 + \ldots + \delta_i)] + \\
- p_1p_2p_3 \ldots p_n(\delta_1 + \delta_2 + \delta_3 + \ldots + \delta_n)\frac{Nc}{n}.
\]

(2.2)

A fundamental assumption in the no-arbitrage approach is that there is no risk free way to make a profit. Let us consider the Hull and White model for valuation CDS (for more details see [5]). There are some assumptions in this model such as: no counterparty default risk; default probabilities, interest rates, and recovery rates are independent. First of all for valuation of CDS we should estimate the probability of default at different future times. As a source of data for parameter estimation we can use bond prices of reference entity.
Let us assume that if there is a possibility of default then a corporate bond sells with lower price than a similar treasury bond. We can write this relationship as:

\[ V_{TB} - V_{CB} = PV_d, \]

where \( V_{TB} \) is value of treasury bond, \( V_{CB} \) – value of corporate bond, \( PV_d \) – present value of cost of default. Using this representation and assumptions about recovery rates we can estimate the probability of default at different times. If we haven’t enough data from reference entity (meaning that there is an insufficient amount of traded bonds), we can use data from another company with the same risk of default (i.e. the same credit rating and in the same industry). Let defaults happen on any of the maturity dates of bonds.

We will use following notation:

\[ N \] – number of bonds;
\[ t_i \] – the maturity of \( i \)-th bond;
\[ B_j \] and \( G_j \) – prices of the \( j \)-th bond and Treasury bond with the same cash flows as \( j \)-th bond today;
\[ F_j(t) \] – forward price of \( j \)-th bond for a forward contract with maturity \( t \), \( t < t_j \);
\[ v(t) \] – present value of $1 received at time \( t \) with certainty;
\[ C_j(t) \] – claim of the \( j \)-th bond if there is a default at time \( t \), \( t < t_j \);
\[ r_j(t) \] – recovery rate of the \( j \)-th bond at default at time \( t \), \( t < t_j \);
\[ \alpha_{ij} \] – present value of the loss, relative to the bond value if there were not default, from a default on the \( j \)-th bond at time \( t_i \);
\[ p_i \] – the risk-neutral probability of default at time \( t_i \).

Let \( t_1 < t_2 < ... < t_N \).

Assume that recovery rates and claims are known. Also assume that interest rates are deterministic. It follows that the price of the no-default value of the \( j \)-th bond is \( F_j(t) \) at time \( t \). In case of default holder makes recovery on claim of \( C_j(t) \) at rate \( r_j(t) \). Now for \( \alpha_{ij} \) we can write that:

\[ \alpha_{ij} = v(t_i)[F_j(t_i) - r_j(t_i)C_j(t_i)]. \]  

(2.3)

Let \( p_i \) be the probability of the loss \( \alpha_{ij} \) being incurred. Therefore, we can write a total present value of the losses on \( j \)-th bond:

\[ G_j - B_j = \sum_{i=1}^{j} p_i \alpha_{ij}. \]  

(2.4)
From last equation we can find \( p_j \):

\[
p_j = \frac{G_j - B_j - \sum_{i=1}^{j-1} p_i \alpha_{ij}}{\alpha_{jj}}.
\]  

(2.5)

As far as default events, *treasury interest rates* and recovery rates are independent, Equations (2.3) and (2.4) are correct for stochastic interest rates, uncertain recovery rates and uncertain default probabilities. If there is no systematic risk in recovery rate it follows that expected recovery rate in the real world is the same as expected recovery rates in the risk-neutral world. So we can estimate expected recovery rate from historical data.

Now we are ready to consider the valuation of CDS. Assume that a CDS has $1 notional principal. We will use following notation:
- \( T \) – life of CDS (years);
- \( p_i \) – risk-neutral probability of default at time \( t_i \);
- \( \hat{r} \) – expected recovery rate on the reference obligation in a risk-neutral world;
- \( u(t) \) – present value of payment at the rate of $1 per year on payments dates from zero to time \( t \);
- \( e(t) \) – present value of a payment at time \( t = t^* \) (\( t^* \) – payment date immediately before time \( t \));
- \( v(t) \) – present value of $1 receives at time \( t \);
- \( w \) – payments of CDS buyer per year;
- \( s \) – value of \( w \) when CDS has a zero value;
- \( \pi \) – the risk-neutral probability that there is no credit event until end of CDS life;
- \( A(t) \) – accrued interest on the reference obligation as a percent of face value at time \( t \).

For calculation of \( \pi \) we can use following formula:

\[
\pi = 1 - \sum_{i=1}^{N} p_i.
\]

Next representation fits to present value of payments:

\[
PVP = w \sum_{i=1}^{N} [u(t_i) + e(t_i)] p_i + w \pi u(T).
\]

The risk-neutral expected value of the reference obligation in case of a credit event occurring at time \( t_i \) is given by following formula:

\[
[1 + A(t_i)]\hat{r}.
\]

Then the risk-neutral expected payoff from that credit default swap is given by:

\[
1 - [1 + A(t_i)]\hat{r} = 1 - \hat{r} - A(t_i)\hat{r}
\]
so the present value of the expected payoff from the CDS is given by:

$$PVEP = \sum_{i=1}^{N} [1 - \hat{r} - A(t_i)\hat{r}] p_i v(t_i).$$

Now the value of the CDS to the buyer is:

$$V = PVEP - PVP = \sum_{i=1}^{N} [1 - \hat{r} - A(t_i)\hat{r}] p_i v(t_i) - w \sum_{i=1}^{N} [u(t_i) + e(t_i)] p_i + w\pi u(T).$$

From this representation we can solve the equation $V = 0$ with respect to the $w$ which in this case is called credit default swap spread and denoted as $s$.

$$s = \frac{\sum_{i=1}^{N} [1 - \hat{r} - A(t_i)\hat{r}] p_i v(t_i)}{\sum_{i=1}^{N} [u(t_i) + e(t_i)] p_i + w\pi u(T)}. \quad (2.6)$$

CDS spread is the total of the payments per year, as a percent of the notional principal.

Earlier we assumed that default can happen only on maturity dates of bonds. Now we consider case when it can happen at any time (for more details see [6]). Define $q(t)$ as the default probability density then $q(t)\Delta t$ is the probability of default time from $t$ to $t + \Delta t$ as seen at time zero. Suppose that $q(t)$ is constant and equal to $q_i$ for $t \in [t_{i-1}, t_i]$. Much as previous discussion for discrete case:

$$\beta_{ij} = \int_{t_{i-1}}^{t_i} v(t)[F_j(t) - \hat{r}C_j(t)]dt, \quad (2.7)$$

$$q_j = \frac{G_j - B_j - \sum_{i=1}^{j-1} q_i \beta_{ij} \beta_{jj}}{\beta_{jj}}. \quad (2.8)$$

In this case for valuation of a CDS we use the default probability density instead of probabilities of default at times $t_i$. Then value of $\pi$ is:

$$\pi = 1 - \int_0^T q(t)dt.$$ 

If default happens earlier than $T$ then the present value of the payments is given by $w[u(t) + e(t)]$. If there is not default until $T$ then the present value of the payment is given by $wu(T)$.

It follows that the expected present value of the payment is:

$$EPVP = w \int_0^T q(t)[u(t) + e(t)]dt + w\pi u(T).$$

Formula for the risk-neutral expected payoff from the credit default swap is given by:

$$1 - [1 + A(t)]\hat{r} = 1 - \hat{r} - A(t)\hat{r}.$$
And the present value of the expected payoff from the credit default swap is:

\[
PVEP = \int_0^T [1 - \hat{r} - A(t)\hat{r}]q(t)v(t)dt.
\]

Then the value of the CDS to the buyer is:

\[
V = PVEP - EPVP = \int_0^T [1 - \hat{r} - A(t)\hat{r}]q(t)v(t)dt - w\int_0^T q(t)[u(t) + e(t)]dt - \pi u(T).
\]

From this representation we can calculate credit default swap spread, \(s\). It is equal to \(w\) when this representation is zero:

\[
s = \frac{\int_0^T [1 - \hat{r} - A(t)\hat{r}]q(t)v(t)dt}{\int_0^T q(t)[u(t) + e(t)]dt + \pi u(T)}.
\]  

(2.9)

2.4 ARARCH model

For modeling the CDS spread process we consider the ARARCH(p,q) model in the special case when \(p = 1\) and \(q = 1\).

Suppose we observe the price process of the credit derivative \(X_t\).

**Theorem 1** The process \(\{X_t\}\) is a Markov process.

Let us consider a decomposition of the observation process \(\{X_t\}\) into a pure ARARCH part \(\{h_t\}\) and a shift process part \(\{\mu_t\}\) as

\[
X_t = h_t + \mu_t,
\]

(2.10)

where

\[
h_t = a_0X_{t-1} + \epsilon_t\sqrt{a_1 + a_2X_{t-1}^2},
\]

(2.11)

\(a_0 \in \mathbb{R}, \ a_1 > 0, \ a_2 > 0\) and the shift process \(\{\mu_t\}\) is specified below (see [2]). In Equation (2.11) the \(\epsilon_t\) is a sequence of independent and identically distributed random variables \(\epsilon_t \sim N(0,1)\).

Suppose we are considering the price process of CDS’s under the discrete time and monitoring a sudden change of the distribution of this process. For the distribution of \(x_t\) given \(x_{t-1}\) \(F(x_t|x_{t-1})\) the change-point at the time \(t\) can be a random time-point \(\theta\), which indicates that the change occurs at time \(\theta\).

\[
F(x_t|x_{t-1}) = \begin{cases} 
F_0(x_t|x_{t-1}), & t < \theta; \\
F_1(x_t|x_{t-1}), & t \geq \theta.
\end{cases}
\]
In Equation (2.10) the $\mu_t$ is the shift process

$$
\mu_t = \begin{cases} 
a, & t < \theta; \\
a + \alpha(t - \theta + 1), & t \geq \theta.
\end{cases}
$$

where $a$ is a some constant level, $t$ is the time of observation, $\theta$ is the change point and $\alpha$ is the slope. This means that, at times before the change-point, the expectation is constant, but in and after the change-point the expectation is linearly increasing.

Therefore we can transform the Equation (2.10) in the following form:

$$X_t = a_0X_{t-1} + \epsilon_t\sqrt{a_1 + a_2X^2_{t-1} + a + \alpha(t - \theta + 1)}1_{\{t \geq \theta\}}. \tag{2.12}$$

According to Theorem 1 the process of the observations $X_t = \{x_t : t \geq 1\}$ satisfy the Markov property. Therefore we can write the conditional joint density function of $x_t, x_{t-1}, \ldots$ with respect to $\theta$ as

$$f(x_t, x_{t-1}, \ldots | \theta) = f(x_0) \prod_{u=1}^{t} f(x_u|x_{u-1}). \tag{2.13}$$

For the monitoring of the price process of CDS’s we want to have a good idea whether for each time $t$ has the change occurred or not?

### 2.5 Change-point problem

Suppose we sequentially observe the random process $X_t = \{x_1, x_2, \ldots, x_{\theta-1}, x_{\theta}, \ldots\}$ with respect to the filted probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, where $\Omega$ is a sample space, $\mathcal{F}$ is some $\sigma$-algebra of $\Omega$, $\{\mathcal{F}_t\}$ is a filtration, i.e. a sequence of $\sigma$-algebras such that $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \ldots$ and $P$ is a probability measure. On $[0, \theta)$ the distribution of this process is one, but on $[\theta, \infty)$ the distribution of the process becomes another. For the distribution of $x_t$ given $x_{t-1}$ $F(x_t|x_{t-1})$ the change-point at the time $t$ can be a random time-point $\theta$, which indicates that the change occurs at time $\theta$

$$F(x_t|x_{t-1}) = \begin{cases} 
F_0(x_t|x_{t-1}), & t < \theta; \\
F_1(x_t|x_{t-1}), & t \geq \theta.
\end{cases}$$

The part of the process before the change is said to be ”in control” and the part of the process after the change is said to be ”out of control”.

The change-point problem is to detect the change quickly and accurately. To deal with this problem we construct a stopping rule

$$\tau = \inf\{t \geq 1 : a(t) > d\},$$

where $a(t) = f(x_t, x_{t-1}, \ldots, x_1)$ is a specific for each method and $d$ is a control limit.
2.6 Performance measures

A desirable property of any method is that detects a change quickly and accurately. For this purpose there are measures of the false alarm properties as well as measures of delay of motivated alarm and credibility of alarm (for more details see [4]). The most commonly used measure is the average run length. Denote by $E_{\theta}(.)$ the expectation when the change is at $\theta$. When $\theta = \infty$ there is no change in the system under observation.

$$ARL^0 = E_{\infty}(\tau).$$

$ARL^0$ is usually called the average in-control run length. This is the commonly used measure of false alarms rate. When $\theta = 1$ the change-point is at the beginning

$$ARL^1 = E_1(\tau).$$

$ARL^1$ is usually called the average out-of-control run length.

When the change occurs it is important to minimize the expected delay of detection. In 1963 Shiryaev (see [8]) suggested measures of the expected value of the delay. The expected delay from the time of change, $\theta = t$, to the time of alarm, $\tau$, is denoted by

$$ED(\nu) = E(\tau - \theta|\tau \geq \theta).$$

It is easier to evaluate the conditional expected delay

$$CED(t) = E(\tau - \theta|\tau \geq \theta = t).$$

When $\theta = 1$ the value of the expected delay equal the value of $ARL^1 - 1$. The Shiryaev-Roberts method has worse $ARL^1$ than the Shewhart method. The CUSUM and the Shewhart methods are alike for $\theta = 1$. For finite time the CUSUM method is alike to the Shewhart method but not to the Shiryaev-Roberts method.

Often used the probability of the difference does not exceed a fixed limit. The fixed limit, for example $l$, is the time available for successful detection. Frisen (see [3]) suggested the probability of successful detection of a measure of the performance. The probability of successful detection is

$$PSD(\theta, l) = Pr(\tau - \theta \leq l|\tau \geq \theta).$$

The $PSD$ is worse for the Shewhart method than for the CUSUM method for $l = 3$.

The Shiryaev-Roberts method and the LR method have a worse probability of detection of a change if the change happens early but if the change happens late these methods are good.
When an alarm is we need to know the alarm is true or not. The risk of false alarms and the risk of delay we must consider. The probability that a change has occurred is considered as a time-dependent predictive value

\[ PV(t) = P(\tau \leq t | \tau = t). \]

PV specifies, whether there is a large probability or not that change has occurred. The Shiryaev-Roberts method has a relatively constant predicted value therefore the same kind of action is appropriate both for early and late alarms.

There is one more measure, the name of this measure is false alarm probability

\[ PFA = P(\tau < \theta). \]

It is probability, that the alarm occurs before the change. The properties differ between the LR methods on one hand and the Shewhart, the CUSUM, the EWMA and the Shiryaev-Roberts methods on the other.

### 2.7 Change-point detection methods

Now we will consider the most commonly used methods (for more details see [9]). All methods will be based on the likelihood ratio between the change having occurred at time \( s < t \) and the change having not occurred by time \( t \). This is

\[
L(s, t) = \log \frac{f(x_0, x_1, \ldots, x_t | \theta = s)}{f(x_0, x_1, \ldots, x_t | \theta < s = t)}
= \log \frac{f_0(x_0) \prod_{u=1}^{s-1} f_0(x_u | x_{u-1}) \prod_{u=s}^{t} f_1(x_u | x_{u-1}, \theta = s)}{f_0(x_0) \prod_{u=1}^{s} f_0(x_u | x_{u-1})}
= \sum_{u=s}^{t} lr(s, t),
\]

where

\[
lr(s, t) = \log \frac{f_1(x_t | x_{t-1}, \theta = s)}{f_0(x_t | x_{t-1})}
\]

and \( f_0 \) and \( f_1 \) are the conditional density functions.

#### 2.7.1 Shewhart method

The first method was invented by Walter A. Shewhart in the 1924. The Shewhart method is the simplest and most widely used method for observations. Denoted the process of observations by \( \{X_t : t = 1, 2, \ldots\} \) where \( x_t = \{x_1, \ldots, x_t\} \) the observation at time \( t \). At each observable time \( s \) we want to know the change
has occurred or not? Let \( X_s = \{X_t : t \leq s\} \). If we use the likelihood ratio then alarm is trigger at

\[
\tau = \min\{t \geq 1 : L(t, t) > c\},
\]

(2.15)

where \( c \) is a constant called threshold value. The Shewhart method are sensitive to large process shift.

### 2.7.2 The cumulative sum method (CUSUM)

The CUSUM method was developed by Page (1954) (see [7]). This method is one of the most famous methods used in the change point problem, because CUSUM method is more sensitive to small process shifts. The alarm function of the CUSUM method can be express by the likelihood ratios as

\[
\tau = \min\{t \geq 1, \max(s \leq t \sum_{s=1}^{t} \exp(L(s, t)) : s = 1, 2, \ldots) > c\},
\]

(2.16)

where \( c \) is a constant called threshold value.

### 2.7.3 Shiryaev-Roberts method

Shiryaev (1963) and Roberts (1966) suggested the method in which an alarm is triggered at the time

\[
\tau = \min\{t \geq 1, \sum_{s=1}^{t} \exp(L(s, t)) > c\},
\]

(2.17)

where \( c \) is a constant called threshold value.

### 2.7.4 EWMA method

The exponentially weighted moving average (EWMA) method was introduced by Roberts (1959). This method is very popular in quality control applications. The stopping rule of the EWMA rules is defined as

\[
\tau = \min\{t \geq 1; \lambda^{t-s}L(s, t) : s = 1, 2, \ldots, t \geq c\},
\]

(2.18)

where \( c \) is a constant called threshold value and \( \lambda \in (0, 1) \). If \( \lambda \to 0 \) we obtain a CUSUM method, if \( \lambda = 1 \) we recover a Shewhart method.
Chapter 2. Methods
Chapter 3

Results

3.1 Plots of CDS spread

Now we want to define what kind of processes is CDS spread process. To figure it out we plotted hypothetical changes of CDS spread during 5 years (shown on Figure 3.1).

![Figure 3.1: The changing of CDS spreads during time: Mobile TeleSystem Open Joint Stock Company](image)

Figure 3.1: The changing of CDS spreads during time: Mobile TeleSystem Open Joint Stock Company
Since CDS are over-the-counter derivatives we haven’t available data for our research. That’s why we calculated it with using real historical data of underlying bond and Hull and White model for pricing of CDS considered in Section 2.3. In particular we used available quotes of bond of Mobile TeleSystems Open Joint Stock Company. For first view we suppose that it is ARARCH process. To be more sure about type of process we simulated ARARCH(1,1) (shown of Figure 3.2). In particular we simulated volatility of our process.

![Figure 3.2: The simulation of ARARCH(1,1)](image)

As we can see on Figure 3.2 behavior of both plots are very similar. It gives us reasons to consider this type of processes in our case. But if it is unit case and for other bonds it is not true? To be sure let us consider bond’s plots for other two companies (shown on Figure 3.3).

As we can see in these cases we have change-points too. So we will assume that this kind of process and this change-point analysis applies to more than one realisation of one process.

### 3.2 Parameter estimation

For calculating the performance $ARL^0$ we have to find the formula for estimates values of ARARCH(1,1) model $a_1$ and $a_2$. Before the change in the
distribution we consider a normal density function
\[
f(x_t|x_{t-1}) = \frac{1}{\sigma_t\sqrt{2\pi}} \exp\left(-\frac{(x_t - \mu)^2}{2\sigma_t^2}\right), \tag{3.1}
\]
where
\[
\sigma_t^2 = a_1 + a_2x_{t-1}^2
\]
and \(\mu_t\) is the shift process such that
\[
\mu_t = a + \alpha(t - \theta + 1)1_{t \geq \theta}.
\]

Let us first consider the case then \(t < \theta\), then \(\mu_t = a\). Therefore for the ARARCH(1,1) model the distribution is
\[
f(x_t|x_{t-1}) = \frac{1}{\sigma_t\sqrt{2\pi}} \exp\left(-\frac{(x_t - a)^2}{2\sigma^2}\right). \tag{3.2}
\]

Let’s write the log-likelihood function as a function of \(a\), \(a_1\) and \(a_2\)
\[
L(a, a_1, a_2) = \sum_{t=2}^{n} \log f_0(x_t|x_{t-1}) \tag{3.3}
\]
Chapter 3. Results

and using the Equation (3.2) we will transform the Equation (3.3) in the form of

$$L(a, a_1, a_2) = \sum_{t=2}^{n} \log \left( \frac{1}{\sigma_t \sqrt{2\pi}} \exp \left( -\frac{(x_t - a)^2}{2\sigma_t^2} \right) \right).$$

(3.4)

Using the property of logarithm function we obtain

$$L(a, a_1, a_2) = \sum_{t=2}^{n} \left( \log(2\pi)^{-\frac{1}{2}} + \log(\sigma_t^2)^{-\frac{1}{2}} - \frac{(x_t - a)^2}{2\sigma_t^2} \right)$$

(3.5)

and we will remove a brackets

$$L(a, a_1, a_2) = -\frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^{n} \log(\sigma_t^2) - \frac{1}{2} \sum_{t=2}^{n} \frac{(x_t - a)^2}{\sigma_t^2}.$$  (3.6)

We know that $\sigma_t^2 = a_1 + a_2 x_{t-1}^2$, hence

$$L(a, a_1, a_2) = -\frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^{n} \log(a_1 + a_2 x_{t-1}^2) - \frac{1}{2} \sum_{t=2}^{n} \frac{(x_t - a)^2}{a_1 + a_2 x_{t-1}^2}.$$  (3.7)

We can estimate the value of $a, a_1, a_2$ by taking the partial derivative. Consider the partial derivative of log-likelihood with respect to $a_1$

$$\frac{\partial L(a, a_1, a_2)}{\partial a_1} = -\frac{1}{2} \sum_{t=2}^{n} \frac{1}{a_1 + a_2 x_{t-1}^2} + \frac{1}{2} \sum_{t=2}^{n} \frac{(x_t - a)^2}{(a_1 + a_2 x_{t-1}^2)^2}. $$

(3.8)

We have

$$\frac{\partial L(a, a_1, a_2)}{\partial a_1} = \sum_{t=2}^{n} \frac{(x_t - a)^2 - a_1 - a_2 x_{t-1}^2}{2(a_1 + a_2 x_{t-1}^2)^2}.$$ 

(3.9)

Now we consider the partial derivative of the log-likelihood with respect to $a_2$

$$\frac{\partial L(a, a_1, a_2)}{\partial a_2} = -\frac{1}{2} \sum_{t=2}^{n} \frac{x_{t-1}^2}{a_1 + a_2 x_{t-1}^2} + \frac{1}{2} \sum_{t=2}^{n} \frac{(x_t - a)^2 x_{t-1}^2}{(a_1 + a_2 x_{t-1}^2)^2}. $$

(3.10)

We have

$$\frac{\partial L(a, a_1, a_2)}{\partial a_2} = \sum_{t=2}^{n} \frac{(x_t - a)^2 x_{t-1}^2 - a_1 x_{t-1}^2 - a_2 x_{t-1}^4}{2(a_1 + a_2 x_{t-1}^2)^2}. $$

(3.11)

And consider the partial derivative of the log-likelihood with respect to $a$

$$\frac{\partial L(a, a_1, a_2)}{\partial a} = \frac{1}{2} \sum_{t=2}^{n} \frac{2(x_t - a)}{a_1 + a_2 x_{t-1}^2}. $$

(3.12)
We can find all coefficients by solving following system of equations:

\[
\begin{align*}
\frac{\partial L(a,a_1,a_2)}{\partial a} &= 0, \\
\frac{\partial L(a,a_1,a_2)}{\partial a_1} &= 0, \\
\frac{\partial L(a,a_1,a_2)}{\partial a_2} &= 0.
\end{align*}
\]

3.3 Recursive algorithm

Now we will construct the recursive algorithm for CUSUM method. For the monitoring of the price process of CDS’s we must know for each time \(t\) has the change occurred or not? We know that the problem of detection of the change in the distribution of the process are based on a stopping rule

\[
\tau = \inf \{t : a(t) > c\},
\]

where \(a(t) = f(x_t, x_{t-1}, \ldots, x_1)\) is a specific for each method and \(a(t)\) is called alarm function and \(c\) threshold.

**Recursive algorithm for CUSUM method**

Let’s consider the CUSUM alarm function \(a(t)\) as a function of the observations \(x_t, x_{t-1}, \ldots\) and \(a(t - 1)\). For CUSUM method we have that

\[
a(t) = \max_{0 \leq s \leq t} L(s, t).
\]

Now we will define the conditional density function \(f_0(x_t|x_{t-1})\) before the change.

By definition

\[
f_0(x_t|x_{t-1}) = \frac{d}{dx_t} P(X_t \leq x_t|X_{t-1} = x_{t-1}),
\]

where

\[
X_t = h_t + \mu_t,
\]

\[
h_t = a_0 X_{t-1} + \epsilon_t \sqrt{a_1 + a_2 X_{t-1}^2}
\]

and

\[
\mu_t = a + \alpha(t - \theta + 1)1_{t\geq\theta}.
\]

Since we consider the case before the change, therefore \(\mu_t = a\). Therefore, we can write

\[
f_0(x_t|x_{t-1}) = \frac{d}{dx_t} P(a_0 X_{t-1} + \epsilon_t \sqrt{a_1 + a_2 X_{t-1}^2} + a \leq x_t|X_{t-1} = x_{t-1})
\]

\[
= \frac{d}{dx_t} P\left(\epsilon_t \leq \frac{x_t - a - a_0 X_{t-1}}{\sqrt{a_1 + a_2 X_{t-1}^2}}|X_{t-1} = x_{t-1}\right).
\]
Since $\epsilon_t$ is a sequence of independent and identically distributed random variables $\epsilon_t \sim N(0, 1)$ we can write

$$f_0(x_t|x_{t-1}) = \frac{d}{dx_t} \Phi \left( \frac{x_t - a_0x_{t-1} - a}{\sqrt{a_1 + a_2x_{t-1}^2}} \right)$$

$$= \frac{1}{\sqrt{a_1 + a_2x_{t-1}^2}} \cdot \phi \left( \frac{x_t - a_0x_{t-1} - a}{\sqrt{a_1 + a_2x_{t-1}^2}} \right).$$

Thus, the conditional density function before the change is equal

$$f_0(x_t|x_{t-1}) = \frac{1}{\sqrt{2\pi(a_1 + a_2x_{t-1}^2)}} \cdot \exp \left(\frac{-(x_t - a_0x_{t-1} - a)^2}{2 \cdot (a_1 + a_2x_{t-1}^2)}\right). \quad (3.14)$$

And now we will define conditional density function $f_1(x_t|x_{t-1})$ after the change. In this case we have

$$X_t = a_0X_{t-1} + \epsilon_t\sqrt{a_1 + a_2X_{t-1}^2} + a + \alpha(t - s + 1).$$

Therefore we can write that

$$f_1(x_t|x_{t-1}) = \frac{d}{dx_t} P(\epsilon_t\sqrt{a_1 + a_2X_{t-1}^2} + a_0X_{t-1} + a + \alpha(t - s + 1) \leq x_t | X_{t-1} = x_{t-1})$$

$$= \frac{d}{dx_t} P \left( \epsilon_t \leq \frac{x_t - a_0X_{t-1} - a - \alpha(t - s + 1)}{\sqrt{a_1 + a_2X_{t-1}^2}} | X_{t-1} = x_{t-1} \right).$$

Since $\epsilon_t$ is a sequence of independent and identically distributed random variables $\epsilon_t \sim N(0, 1)$ we can write

$$f_1(x_t|x_{t-1}) = \frac{d}{dx_t} \Phi \left( \frac{x_t - a_0x_{t-1} - a - \alpha(t - s + 1)}{\sqrt{a_1 + a_2x_{t-1}^2}} \right)$$

$$= \frac{1}{\sqrt{a_1 + a_2x_{t-1}^2}} \cdot \phi \left( \frac{x_t - a_0x_{t-1} - a - \alpha(t - s + 1)}{\sqrt{a_1 + a_2x_{t-1}^2}} \right).$$

Thus, the conditional density function after the change is equal

$$f_1(x_t|x_{t-1}) = \frac{1}{\sqrt{2\pi(a_1 + a_2x_{t-1}^2)}} \cdot \exp \left(\frac{-(x_t - a_0x_{t-1} - a - \alpha(t - s + 1))^2}{2 \cdot (a_1 + a_2x_{t-1}^2)}\right). \quad (3.15)$$

Since

$$lr(s, t) = \log \frac{f_1(x_t|x_{t-1}, \theta = s)}{f_0(x_t|x_{t-1})}$$

$$= \frac{1}{\sqrt{2\pi(a_1 + a_2x_{t-1}^2)}} \cdot \exp \left(\frac{-(x_t - a_0x_{t-1} - a - \alpha(t - s + 1))^2}{2 \cdot (a_1 + a_2x_{t-1}^2)}\right). \quad (3.16)$$
therefore we can write that
\[
lr(s, t) = \log \frac{1}{\sqrt{2\pi(a_1 + a_2 x_{t-1}^2)}} \cdot \exp \left( -\frac{(x_t - a_0 x_{t-1} - a - \alpha(t-s+1))^2}{2(a_1 + a_2 x_{t-1}^2)} \right)
\]
\[
= \log \exp \left( \frac{(x_t - a_0 \cdot x_{t-1} - a)^2}{2(a_1 + a_2 \cdot x_{t-1}^2)} - \frac{(x_t - a_0 \cdot x_{t-1} - a - \alpha(t-s+1))^2}{2(a_1 + a_2 \cdot x_{t-1}^2)} \right)
\]
\[
= \frac{2\alpha(x_t - a_0 x_{t-1} - a)(t-s+1) - \alpha^2(t-s+1)^2}{2(a_1 + a_2 x_{t-1}^2)}.
\]

Therefore we can write
\[
lr(s, t) = \frac{2\alpha(x_t - a_0 x_{t-1} - a)(t-s+1) - \alpha^2(t-s+1)^2}{2(a_1 + a_2 x_{t-1}^2)}.
\] (3.17)

Back to Equation (3.13) we can write that
\[
a(t) = \max_{1 \leq s \leq t} L(s, t)
\]
\[
= \max(lr(t, t), lr(t, t) + lr(t-1, t), \ldots, lr(t, t) + lr(t-1, t) + \ldots + lr(1, t))
\]
\[
= lr(t, t) + \max(0, lr(t-1, t), \ldots, lr(t-1, t) + lr(t-2, t) + \ldots + lr(1, t))
\]
\[
= lr(t, t) + \max(0, \max(lr(t-1, t), \ldots, lr(t-1, t) + lr(t-2, t) + \ldots + lr(1, t))
\]
\[
= lr(t, t) + \max(0, a(t-1)),
\] (3.18)

where
\[
\max(lr(t-1, t), \ldots, lr(t-1, t) + \ldots + lr(1, t)) = a(t-1).
\]

Using Equation (3.17) we can write that
\[
a(t) = \frac{2\alpha(x_t - a_0 x_{t-1} - a) - \alpha^2}{2(a_1 + a_2 x_{t-1}^2)} + \max(0, a(t-1)).
\] (3.19)

Since we know recursive algorithm for CUSUM method we can find the average in-control run length \(ARL^0\) and average out-of-control run length \(ARL^1\) for this method.

### 3.4 Calculating the Performance \(ARL^0\) and \(ARL^1\) for CUSUM

By definition of the stopping rule for CUSUM is
\[
\tau = \min\{t : a(t) > c\},
\] (3.20)
where
\[ a(t) = \max_{0 \leq s \leq t} L(s, t). \]

We know the recursive algorithm for calculating \( a(t) \). For finding of the average in-control run length we have to average value of \( t \) such that \( a(t) > c \), where \( c \) is threshold value.

The expected time for the first alarm denoted by \( ARL^0 \). For the calculating of the performance measure \( ARL^0 \) we fixed \( t \) runs for the first shift, then to calculate the threshold value to be used for comparing different shifts. We will consider the case then \( t = 100 \). For every 100 runs on average we have a false alarm. The expected delay to motivated alarm, denoted by \( ARL^1 \), is the most important. The \( ARL^1 \) indicates delay to alarm assumming the process change immediate monitoring start.

Now we will consider the graphs of \( ARL^0 \) and \( ARL^1 \) plotted using the our real data.

![Figure 3.4: Plots of \( ARL^0 \) and \( ARL^1 \)](image)

When we consider the \( ARL^1 \) it can be seen that CUSUM is very effective for detecting small shift. We can notice that the size of the shift is increasing when \( ARL^0 \) and \( ARL^1 \) is decreasing. Therefore we plotted graphs of \( ARL^0 \) and \( ARL^1 \) against the different volatility shift size.
Chapter 4
Conclusions

Credit Default Swap is an over-the-counter derivative so we haven’t available real data for analysis. We have studied Hull and White model for pricing of CDS and used this knowledges for calculating CDS spreads. As result we have had data for the change-point analysis. We have choosed ARARCH(1,1) model for our process, our choice have based on the analysis of graphs of hypothetical CDS spreads and simulated ARARCH(1,1).

In this paper we have analysed the monitoring of the price process of a Credit Default Swaps. For the monitoring this process we used the change-point analysis. Our problem consisted in detection of the change-point quickly and accurately. For solving this problem we considered Shewhart, CUSUM, Shiryaev-Roberts and EWMA methods. All methods differ with respect to the way the alarm limit change with decision time and the number parameters they depend on. Also we considered different performance measures. At the present time the performance measure for the methods of observation is becoming very useful. We have constructed recursive algorithm for CUSUM method and got following results:

\[ a(t) = lr(t,t) + \max(0, a(t-1)), \]

where \( a(t) \) is alarm function and

\[ lr(t,t) = \frac{2\alpha(x_t - a_0x_{t-1} - a) - \alpha^2}{2(a_1 + a_2x_{t-1}^2)}. \]

For CUSUM method was calculated the most commonly used measure average run length. The CUSUM method detects process shift very fast. We can see this from our results. If we consider the plot of ARL we can observe that we have change in ARL for a small size of shift. Detection of the change-point quickly and accurately is very important in financial markets. For further research it can be useful to consider observations for ARARCH process with a higher order. Also methods which are considered in this master thesis can be basis for a case when we bet for decreasing of a price process.
Chapter 4. Conclusions
Bibliography


Continuous Inspection Schemes. Biometrica, 41, pp. 100-114.


Appendix

**Definition 1** A sequence of random variable \( \{X_t\} \) is called a Markov chain if for any \( t = 1, 2, \ldots \)

\[
P(X_t = x_t | X_{t-1} = x_{t-1}, \ldots, X_1 = x_1, X_0 = x_0) = P(X_t = x_t | X_{t-1} = x_{t-1}).
\]

The Markov chain is a random process with the property that conditional on its present value, the future is independent of the past.

**Theorem 1** The process \( \{X_t\} \) is a Markov process.

**Proof**: We observe a random process \( \{X_t; t = 1, 2, \ldots\} \) where \( X_t \) are observations from an ARARCH(1,1) model. Therefore we can write that

\[
X_t = h_t + \mu_t,
\]

where

\[
h_t = a_0 X_{t-1} + \epsilon_t \sqrt{a_1 + a_2 X_{t-1}^2}
\]

and where \( a_0 \in \mathbb{R}, a_1 > 0 \) and \( a_2 > 0 \). In Equation 4.2 the \( \epsilon_t \) is a sequence of independent and identically distributed random variables \( \epsilon_t \sim N(0,1) \). The \( \mu_t \) is the shift process

\[
\mu_t = \begin{cases} a, & t < \theta; \\ a + \alpha(t - \theta + 1), & t \geq \theta. \end{cases}
\]

where \( t \) is the time of observation and \( \theta \) is the change point.

Let’s rewrite the Equation 4.1 in the next form:

\[
X_t = a_0 \cdot X_{t-1} + \epsilon_t \sqrt{a_1 + a_2 X_{t-1}^2} + a + \alpha(t - \theta + 1) I_{\{t \geq \theta\}}.
\]

For proof that \( \{X_t\} \) is a Markov process we must prove that \( \{X_t\} \) to satisfy a Markov property. Now we will consider a two cases: the first case then \( t < \theta \)
and the second case then \( t \geq \theta \). Let’s consider the first case \( (t < \theta) \), using of the definition of the joint density function:

\[
\begin{align*}
f(x_t|x_{t-1}, \ldots, x_0) & = \frac{d}{dx_t} P(X_t \leq x_t \mid X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) \\
& = \frac{d}{dx_t} P \left( a_0 X_{t-1} + \epsilon_t \frac{\sqrt{a_1 + a_2 X_{t-1}^2}}{a_1 + a_2 x_{t-1}^2} \leq x_t \mid X_{t-1} = x_{t-1}, \ldots, X_0 = x_0 \right) \\
& = \frac{d}{dx_t} P \left( \epsilon_t \leq \frac{x_t - a_0 X_{t-1} - a}{\sqrt{a_1 + a_2 x_{t-1}^2}} \mid X_{t-1} = x_{t-1}, \ldots, X_0 = x_0 \right) \\
& = \frac{d}{dx_t} P \left( \epsilon_t \leq \frac{x_t - a_0 x_{t-1} - a}{\sqrt{a_1 + a_2 x_{t-1}^2}} \mid X_{t-1} = x_{t-1} \right) \\
& = f(x_t|x_{t-1}),
\end{align*}
\]

where in the step \( \ast \) the independence of the sequence \( \epsilon_t \) is used therefore we have that \( X_t \) is depended only of the \( X_{t-1} \) and we can lose of the next parts \( X_{t-2} = x_{t-2}, \ldots, X_0 = x_0 \). Thus we have

\[
\frac{d}{dx_t} P(X_t = x_t \mid X_{t-1} = x_{t-1}) = f(x_t|x_{t-1}).
\]

When we consider of the second case \((t \geq s = \theta)\), we have what

\[
X_t = a_0 X_{t-1} + \epsilon_t \sqrt{a_1 + a_2 X_{t-1}^2} + a + \alpha(t - s + 1).
\]

Since we change only the distribution function, not the Markov property of the process \( \{X_t\} \). For this reason, similarly to the proof of the first case we is see that

\[
f(x_t|x_{t-1}, x_{t-2}, \ldots, x_0) = f(x_t|x_{t-1}).
\]

Thus, the ARARCH(1,1) model satisfies the Markov property.