Stockholms universitet

Statistical Analysis on the Convergence to the Heavy-Tailed Stationary Distribution of a GARCH Process

Carlos Guevara

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2018:5
Matematisk statistik
Juni 2018

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2018:5**
http://www.math.su.se

Statistical Analysis on the Convergence to the Heavy-Tailed Stationary Distribution of a GARCH Process

Carlos Guevara*

June 2018

Abstract

The aim of this thesis is to study the convergence to the stationary distribution of the time series model GARCH(1,1). This model is often used when modelling volatility and can be written in terms of a so-called stochastic recurrence equation. The main result when studying the solution of this equation, is that it obtains a distribution which is of power-law type or heavy-tailed. Performing simulation studies and applying methods from extreme value theory (EVT), the theoretical results regarding the stationary distribution of the process have been successfully verified. Moreover we study the time it takes for the distribution of the process to converge to a heavy-tailed distribution. The convergence appears to happen faster given high values of todays volatility. The stationary distribution of the process is obtained implementing numerical computations and it can be used to calculate risk measurements such as value-at-risk (VaR), which is of importance when assessing the magnitude of risk of extreme events in finance and insurance.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: cargu224@outlook.com. Supervisor: Filip Lindskog.

Acknowledgements

I wish to express my sincerest gratitude to my supervisor Filip Lindskog at Stockholm University, for his guidance and time throughout the work of this thesis.

I would further like to thank my classmate Adam for feedback. In addition, thanks to my family and my girlfriend Vilma for their unconditional support during this work.

To the memory of my mother Esther del Rocío Guerrero Montenegro.

Contents

1 Introduction

Time series analysis is crucial for financial institutions and insurance companies when modelling volatility and computing risk measurements, such as value-at-risk (VaR) and expected shortfall (ES). The large increase of traded assets has contributed to more volatile markets, thus the need for adequate methods to forecast extreme events is of interest in quantitative risk management.

Non-linear time series models, such as the generalised autoregressive conditional heteroscedastic (GARCH) models, have been successfully used during the last decades to model volatility. The reason why such models have been used by risk managers and quantitative analysts is because they take into account the well-known phenomena of volatility clustering observed in financial markets, that is periods of high volatility are followed by periods of less volatility. One disadvantage of this approach is that the assumption of conditional normality does not seem to hold for real financial data and the magnitude of risk tend to be underestimated. Another approach to address for this problem is the estimation of extreme quantiles in extreme value theory (EVT), based on pure statistical theory and methods that allow to extrapolate beyond the tail of the stationary or marginal distribution of the stochastic process. In other words, we estimate the tail distribution of the exceedances, or extreme observations, beyond a given threshold.

Studying the stationary distribution of the model GARCH(1,1) is of interest in this thesis. In [7] it is shown that this model can be written as a stochastic recurrence equation, which has attracted a lot of attention because of its wide spectrum of applications in finance and insurance.The most surprising result is that the tail of the solution to the stochastic recurrence equation is of power-law type, that is the marginal or stationary distribution of the process is asymptotically heavy-tailed. Hence obtaining the stationary distribution for the process, give us an explicit formula to compute the probability of extreme events, meaning that it is be possible to compute a quantile representing a risk measurement.

In this thesis the daily log-returns for the Swedish stock index OMXS30 are used to estimate the parameters of the volatility model using quasi-maximum likelihood estimation (QMLE). Numerical computations are performed to verify theoretical conditions and to compute the theoretical tail index of the stationary distribution of the process.

Moreover, the theoretical results are verified implementing simulation studies where data samples for the solution of the stochastic recurrence equation are used to estimate the tail index of the distribution using extreme value theory, for instance applying the peaks over threshold (POT) method, we estimate the parameters of the generalised Pareto distribution (GPD). This to obtain an estimate of the tail index. Each simulation sample studied represents a GARCH process at some time t, hence we study how the tail index of the distribution varies over time. Doing inference we assess if the convergence value of the tail index correponds to the theoretical tail index value. We also study the kurtosis of some simulation samples to verify that the distribution of the process converges to a stationary heavy-tailed distribution. In addition we compare how the results change when assuming different distributions for the innovations of the volatility model, such as the standard normal distribution and the standardised t distribution.

It is also of interest to study the speed of convergence of the process to a stationary distribution, that is we study the time it takes for the GARCH model to obtain a heavy-tailed stationary distribution. We compute the convergence times of the tail index performing simulation studies for different start values, where the start values represent different possible values of todays volatility.

2 Theoretical framework

Time series analysis consist of methods for analysing time series data in order to get statistical information and characteristics of the data. It is of interest to study the dynamic structure of such series. This section gives a glimpse of the models and theory of interest for this thesis.

2.1 Return series

In financial studies it is often of interest to analyze asset returns instead of asset prices. The reason is that for the average investor, return of a certain asset is a complete and scale-free summary of an investment opportunity. Furthermore, return series are easier to handle than price series and provides more attractive statistical properties [9].

Let P_t be the price of an asset at time t. The one-period simple gross return between time $t+1$ and t is defined as

$$1 + R_t = \frac{P_{t+1}}{P_t} \tag{1}$$

Taking the natural logarithm of (1) gives the log-return or continuously compounded return X_t,

$$X_t = \log(1 + R_t) = \log\left(\frac{P_{t+1}}{P_t}\right) = \log P_{t+1} - \log P_t.$$

2.2 Autoregressive (AR) model

The autoregressive model is often used to deal with linear dependence in time series. Autocorrelation is the linear dependence of a variable with itself. An autoregressive model of order 1, AR(1), is defined as [9], using our notation for the log-returns,

$$X_t = \phi_0 + \phi_1 X_{t-1} + Z_t,$$

where $\{Z_t\}$ is assumed to be a white noise series with mean zero and variance σ_z^2.

2.3 Conditional Heteroscedastic Models

Conditional heteroscedastic models are of main interest in options pricing and quantitative risk management. The reason to this is that such models take into account heteroscedaticity, i.e. when the variance is not constant over time.

Volatilty can be defined as the standard deviation of the series and it has some characteristics commonly seen in asset returns, such as volatility clusters, periods of high volatility tend to be followed by changes of high volatility and periods of low volatility tend to be followed by small volatility changes. Other aspect is that volatility jumps are rare, it evolves over time continuously. Volatility does not diverge to infinity and it seems to react differently to a big price increase or a big price drop [9].

There is a variety of these models, but for the purpose of this thesis we will introduce the autoregressive conditional heteroscedastic (ARCH) model introduced by Engle (1982) and the generalised autoregressive conditional heteroscedastic (GARCH) model by Bollerslev (1986).

2.3.1 ARCH model

We start by introducing the conditional mean μ_t and conditional variance σ_t^2 of X_t given information available F_{t-1}, at time $t-1$,

$$\mu_t = \mathbb{E}[X_t|F_{t-1}], \quad \sigma_t^2 = \mathrm{Var}(X_t|F_{t-1}) = \mathbb{E}[(X_t-\mu_t)^2|F_{t-1}]$$

The autoregressive conditionally heteroscedatic process of order p, ARCH(p), is defined as [7] p. 2,

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_t X_{t-i}^2, \quad t \in \mathbb{Z},$$

where Z_t is an iid sequence of random variables with mean zero and and variance 1, $\alpha_0 > 0$ and $\alpha_p \geq 0$ for some $p \geq 0$. The quantity σ_t is the volatility of X_t.

Then the ARCH(1) model is

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2, \quad t \in \mathbb{Z}.$$

Often it is assumed that Z_t follows the standard normal or standardised t distribution, i.e. the t distribution is scaled to have mean zero and variance one, see Section 7.3.

Notice that large past squared shocks X_{t-1}^2 imply large conditional variance, thus the experienced volatility clustering in asset returns is taken into account.

The idea of an ARCH model is that the log-return X_t is serially uncorrelated, but dependent. The dependance of X_t can be described by a simple quadratic function of its lagged values. For a detailed explanation of this model and its properties see ch. 3 in [9].

2.3.2 GARCH model

To adequately describe the volatility process of asset returns with an ARCH(p), one may need to have many parameters, i.e. $p > 1$. To deal with this the generalised autoregressive conditional heteroscedastic model GARCH(p,q) is proposed, with $p = q = 1$ we get

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad t \in \mathbb{Z}, \tag{2}$$

where $\alpha_0, \alpha_1, \beta_1 > 0$ and as before Z_t is a white noise series with mean zero and variance one, see [7] p. 17.

2.4 Stochastic Recurrence Equation

This section is dedicated to the main theoretical approach studied in this thesis. The stochastic recurrence equation studied by Kesten (1973), see ch. 1 in [7],

$$X_t = A_t X_{t-1} + B_t \quad t \in \mathbb{Z}, \tag{3}$$

has been of interest over the last decades because of its applications in finance and insurance. One aspect of main importance is that, under general conditions, the tails of X_t are of power-law-type: light-tailed input variables (A_t, B_t) in (3) may cause heavy-tailed output X_t [7]. The power-law tail behaviour of the marginal distribution of the GARCH processes is of interest among experts in extreme value theory and time series analysis.

We verify that (2) can be written in terms of (3). Since $X_t = \sigma_t Z_t$,

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\ &= \alpha_0 + \alpha_1 \sigma_{t-1}^2 Z_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \\ &= (\alpha_1 Z_{t-1}^2 + \beta_1)\sigma_{t-1}^2 + \alpha_0.\end{aligned}$$

Thus $\sigma_t^2 = A_t\sigma_{t-1}^2 + B_t$ as in (3) with $A_t = \alpha_1 Z_{t-1}^2 + \beta_1$ and $B_t = \alpha_0$.

It is stated in [9] p. 23, that the foundations of time series analysis is stationarity and a time series $\{X_t\}$ is said to be strictly stationary if the joint distribution of $(X_{t_1}, ..., X_{t_k})$ is identical to that of $(X_{t_1+t}, ..., X_{t_k+t})$ for all t, where k is an arbitrary positive integer and $(t_1, ..., t_k)$ is a collection of k positive integers. In other words, strict stationarity requires that the joint distribution of $(X_{t_1}, ..., X_{t_k})$ is invariant under time shift. This condition is hard to verify empirically, thus a weaker condition of stationarity is often assumed. A time series $\{X_t\}$ is weakly stationary if both the mean of the series and the covariance between X_t and $X_{t-\ell}$ are time-invariant. More specifically, $\{X_t\}$ is weakly stationary if $\mathbb{E}[X_t] = \mu$, which is constant, and $\mathrm{Cov}(X_t, X_{t-\ell}) = \gamma_\ell$, which only depends on ℓ. In the condition of weak stationarity it is assumed that the first two moments of the time series are finite.

As stated in [7] p. 17, the conditions

$$\alpha_0 > 0 \quad \text{and} \quad \mathbb{E}[\log(\alpha_1 Z_{t-1}^2 + \beta_1)] < 0 \tag{4}$$

are necessary and sufficient for the existence of a non-vanishing a.s. unique causal strictly stationary solution to the equation

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\sigma_{t-1}^2, \quad t \in \mathbb{Z}. \tag{5}$$

With the statement above it is meant that conditions (4) are necessary and sufficient for the existence of a strictly stationary solution to equation (5), which is nonzero everywhere and with probability one it is unique causal, i.e. σ_t^2 is a causal solution to the stochastic recurrence equation of the GARCH(1,1) model because it is a function of past and present innovations $\{Z_t\}$, see Section 7.9. In other words conditions for strict stationarity of $\{X_t\}$ depend on the distribution of the innovations $\{Z_t\}$.

Moreover, using Jensen's inequality and since, by assumption $\mathbb{E}[Z_t^2] = 1$,

$$\mathbb{E}[\log(\alpha_1 Z_{t-1}^2 + \beta_1)] \le \log(\mathbb{E}[\alpha_1 Z_{t-1}^2 + \beta_1]) = \log(\alpha_1 + \beta_1),$$

satisfying (4) if $\alpha_0 > 0$ and $\alpha_1 + \beta_1 < 1$. Taking expectations in (5),

$$\begin{aligned}
\mathbb{E}[\sigma_t^2] &= \alpha_0 + \mathbb{E}[(\alpha_1 Z_{t-1}^2 + \beta_1)\sigma_{t-1}^2] \\
&= \alpha_0 + \alpha_1\mathbb{E}[Z_{t-1}^2\sigma_{t-1}^2] + \beta_1\mathbb{E}[\sigma_{t-1}^2] \\
&= \alpha_0 + (\alpha_1 + \beta_1)\mathbb{E}[\sigma_{t-1}^2].
\end{aligned}$$

Then for $\alpha_1 + \beta_1 < 1$, it follows that,

$$\mathbb{E}[\sigma_t^2] = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)} < \infty, \tag{6}$$

which ensures not only strict but also second-order stationarity of $\{\sigma_t\}$ and $\{X_t\}$, while the same procedure yields $\mathbb{E}[\sigma_t^2] = \infty$ for $\alpha_1 + \beta_1 \geq 1$.

2.5 Power-Law tails

In this section we present the definition of univariate regular variation as in [7] p. 273. Then we present Breiman's lemma as in [7] p. 275, which is used to prove and obtain the asymptotical results about the power-law tails of the solution to the stochastic recurrence equation. This was studied by Kesten (1973) and Goldie (1991) and we write this result found in Theorem 2.4.7 in [7] p. 61, in terms of the GARCH(1,1) model. Then we explain how the tail index for the volatility is obtained.

2.5.1 Univariate Regular Variation

A positive measurable function f on $(0, \infty)$ is to require that for any $\lambda > 0$,

$$\lim_{y \to \infty} \frac{f(\lambda y)}{f(y)} = \lambda^{\alpha}.$$

We say that a random variable Y is regularly varying with index $\alpha \geq 0$ if the function $f(y) = \mathbb{P}(|Y| > y)$ is regularly varying with index $-\alpha$ and a tail-balance condition holds: there exist constants $p, q \geq 0$ such that $p + q = 1$ and

$$\lim_{y \to \infty} \frac{\mathbb{P}(Y > y)}{\mathbb{P}(|Y| > y)} = p \quad \text{and} \quad \lim_{y \to \infty} \frac{\mathbb{P}(Y \leq -y)}{\mathbb{P}(|Y| > y)} = q.$$

In other words, Y has a regularly varying right-tail if

$$\lim_{y \to \infty} \frac{\mathbb{P}(Y > \lambda y)}{\mathbb{P}(Y > y)} = \lambda^{-\alpha} \quad \text{for all} \quad \lambda > 0.$$

2.5.2 Breiman's Lemma

Assume that X and Y are nonnegative independent random variables, Y is regularly varying with index $\alpha > 0$ and one of the following conditions hold

1. $\mathbb{E}[X^{\alpha+\varepsilon}] < \infty$ for some $\varepsilon > 0$.
2. $\mathbb{P}(Y > x) \sim c_0 x^{-\alpha}$ as $x \to \infty$ for some $c_0 > 0$ and $\mathbb{E}[X^\alpha] < \infty$.

Then

$$\mathbb{P}(XY > z) \sim \mathbb{E}[X^\alpha]\mathbb{P}(Y > z), \quad z \to \infty.$$

2.5.3 The Kesten-Goldie Theorem Applied to GARCH(1,1)

Assume that

1. $A = \alpha_1 Z_{t-1}^2 + \beta_1 > 0$ almost surely and the law of log(A) conditioned on $\{A > 0\}$ is not supported on any of the sets $a\mathbb{Z}, a \geq 0$.
2. There exists $\alpha \geq 0$ such that $\mathbb{E}[A^\alpha] = 1, \mathbb{E}[|B|^\alpha] < \infty$ and $\mathbb{E}[A^\alpha \max(\log A, 0)] < \infty$.
3. $\mathbb{P}(Ax + B = x) < 1$ for every $x \in \mathbb{R}$.

Then equation (5) has a stationary solution σ_∞^2 which satisfies

$$\sigma_\infty^2 \stackrel{d}{=} A\sigma_\infty^2 + B,$$

where σ_∞^2 is independent of (A, B) and $\stackrel{d}{=}$ means equality in distribution. Moreover there exists a positive constant c_+ such that

$$\mathbb{P}(\sigma_\infty^2 > x) \sim c_+ x^{-\alpha}, \quad x \to \infty, \tag{7}$$

where the constant c_+ is defined as

$$c_+ = \frac{1}{2\alpha \mathrm{m}_\alpha}\mathbb{E}[|A\sigma_{t-1}^2 + B|^\alpha - |A\sigma_{t-1}^2|^\alpha], \tag{8}$$

and $\mathrm{m}_\alpha = \mathbb{E}[|A|^\alpha \log|A|] > 0$.

2.5.4 Tail index for the stationary distribution of the volatility

To summarise this section we study the tail distribution of the squared log-returns in order to explain how the tail index of the stationary distribution of the volatility σ_t is obtained.

Assume that σ_t^2 is independent of Z_t^2 and $\alpha \geq 0$. Then

$$\begin{aligned} \mathbb{P}(X_t^2 > x) &= \mathbb{P}(\sigma_t^2 Z_t^2 > x) \\ &\sim \mathbb{E}[(Z_t^2)^\alpha]\mathbb{P}(\sigma_t^2 > x) \\ &\sim (2\alpha - 1)!!\ c_+ x^{-\alpha} =: c x^{-\alpha}, \quad x \to \infty, \end{aligned} \tag{9}$$

where $c = (2\alpha - 1)!!c_+$. In the second last step in (9) we used the fact that when $Z \sim N(0, 1)$, the central moment of order k is $\mathbb{E}[Z^k] = (k - 1)!!$, for k even and 0 otherwise. The notation $(2\alpha-1)!!$ means the *double factorial* of $(2\alpha-1)$, i.e. the product of all numbers from $(2\alpha-1)$ to 1 that has the same parity as $(2\alpha - 1)$. Thus the tail distribution for the conditional variance is

$$\mathbb{P}(\sigma_t^2 > x) \sim c_+ x^{-\alpha}, \quad x \to \infty. \tag{10}$$

Rewriting (10) yields

$$\mathbb{P}(\sigma_t > \sqrt{x}) \sim c_+ (\sqrt{x})^{-2\alpha}, \quad x \to \infty,$$

and since $\sqrt{x}$ can be any positive real number, we obtain that the volatility σ_t is regularly varying with index 2α, i.e.

$$\mathbb{P}(\sigma_t > x) \sim c_+ x^{-2\alpha}, \quad x \to \infty. \tag{11}$$

Then the log-returns X_t has an asymptotic tail distribution

$$\mathbb{P}(X_t > x) \sim c x^{-2\alpha}, \quad x \to \infty.$$

This means that the log-returns are regularly varying with index 2α and the constant c is in this case $c = (\alpha - 1)!!\ c_+$, when the innovations Z_t are assumed to be standard normal random variables.

When the innovations are standardised t distrbuted variables with ν degrees of freedom, Z_t is regularly varying with index ν. From Breiman's lemma we notice that Z_t must have higher order finite moments than σ_t, thus $\alpha < \nu/2$ is required to apply Breiman's lemma. If $\alpha > \nu/2$, then the tail behaviour of $\sigma_t Z_t$ is determined by that of Z_t rather than σ_t.

2.6 Extreme Value Theory (EVT)

Statistical analysis of extremes is vital for many risk management problems. Catastrophic events such as natural disasters or financial crises can result in astronomical losses for the banking and insurance industry.

In classical probability theory and statistics most of the results for insurance and finance are based on sums of random variables, where often approximations like the central limit theorem (CLT) are assumed. As the CLT is important when modelling sums of random variables, EVT plays a fundamental role when studying the asymptotic behaviour of extreme observations. The key idea of EVT is to consider the distribution of block maxima (or minima), this to focus on the tails of the distribution rather than the center.

The method to analyse the asymptotic tail behaviour of the distribution in this thesis is the Peaks over threshold (POT) method. This involves estimating the conditional distribution of exceedances beyond some threshold u, where the exceedances $X_k - u$ are assumed to belong to the generalised Pareto distribution (GPD).

Consider iid random variables $X_1, ..., X_n$ with distribution function F_X. The conditional excess distribution function for X over some threshold u is defined as in [6] p. 7,

$$F_u(x) = \mathbb{P}(X - u \leq x | X > u) = \frac{F(x+u) - F(u)}{1 - F(u)}, \quad x \geq 0,$$

where $F_u(x)$ is the probability that a loss exceeds u by no more than x given that the threshold is exceeded.

The conditional excess distribution can be approximated by

$$F_u(x) \approx G_{\xi,\beta}(x), \quad u \to \infty,$$

where $G_{\xi,\beta}(x)$ is the cumulative distribution function for the GPD and is given by

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi} & \text{if} \quad \xi \neq 0, \\ 1 - \exp(-x/\beta) & \text{if} \quad \xi = 0, \end{cases}$$

where $\beta > 0$, the support is $x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ when $\xi < 0$. The case $\xi > 0$ corresponds to the heavy-tailed distributions whose tails decay like power functions such as the Pareto, Student's t, Cauchy, Burr, loggamma and Fréchet distributions. The case $\xi = 0$ corresponds to the thin-tailed distributions like the normal, exponential, gamma and lognormal where the tails decay faster than $x^{-\beta}$ for every $\beta > 0$. Lastly when $\xi < 0$ corresponds to the short-tailed distributions with a finite right endpoint like the uniform and beta distributions.

The tail distribution function for X can be expressed as

$$\begin{aligned}
\mathbb{P}(X > x) &= \mathbb{P}(X > u)\mathbb{P}(X > x|X > u) \\
&= \mathbb{P}(X > u)\mathbb{P}(X - u > x - u|X > u) \\
&= (1 - F(u))(1 - F_u(x - u)) \\
&\approx (1 - F(u))(1 - G_{\xi,\beta}(x - u)), \quad x > u,
\end{aligned}$$

where $1 - F(u)$ can be estimated by the empirical distribution function, i.e.

$$\bar{F}(u) = 1 - F(u) \approx \frac{k}{n},$$

where n is the sample size and k the number of observations above the threshold. In this thesis the threshold level is computed as $k = nu$ for some $u \in (0, 1)$.

Then the tail estimator can be written as

$$\mathbb{P}(X > x) = 1 - F(x) = \frac{k}{n}\left[1 + \hat{\xi}\left(\frac{x - u}{\hat{\beta}}\right)\right]^{-1/\hat{\xi}}, \quad x > u$$

In paper [6] p. 9, it is discussed that using this approach to estimate the tail distribution of the residuals can be viewed as a special case of assuming conditional t distribution when fitting a GARCH modell. It can be shown that assuming a conditional t distribution with ν degrees of freedom yields that the value of ξ in the limiting GPD is the reciprocal of the degrees of freedom, i.e. $\nu = 1/\xi$.

2.7 Value at Risk (VaR)

As stated in [5] p. 165-166, the value-at-risk (VaR) at level $p \in (0, 1)$ of a portfolio with value X at time 1 is

$$\mathrm{VaR}_p(X) = \min\{m : \mathbb{P}(mR_0 + X < 0) \leq p\},$$

where R_0 is the percentage return of a risk-free asset. In words, the VaR of a position with value X at time 1 is the smallest amount of money that if added to the position now and invested in the risk-free asset ensures that the probability of a strictly negative value at time 1 is not greater than p.

In statistical terms $\mathrm{VaR}_p(X)$ is the $(1-p)$-quantile of the discounted portfolio loss $L = -X/R_0$. The u-quantile of a random variable L with distribution function F_L is defined as

$$F_L^{-1}(u) = \min\{m : F_L(m) \geq u\},$$

and F_L^{-1} is the inverse if F_L is strictly increasing. If F_L is both continuous and strictly increasing, then $F_L^{-1}(u)$ is the unique value m such that $F_L(m) = u$. For a general F_L, the quantile value $F_L^{-1}(u)$ is obtained by plotting the graph of F_L and setting $F_L(m) = u$ to be the smallest value m for which $F_L(m) \geq u$. Then

$$\mathrm{VaR}_p(X) = F_L^{-1}(1-p).$$

3 Methodology

Figure 1 shows a so-called candlestick chart for the daily prices of OMXS30 from 16th February 2009 to 4th April 2018. Each bar in the candlestick chart represent the open, close, high and low prices of the index at each trading day. When the bar is red it indicates that during the tradig day the price closed lower than its open price and the bar is green in the other case. It seems that in overall the performance of the index is bullish, or has an upward trend, but notice that late 2011 and between 2015 and 2016 there are some remarkable price-falls. This kind of dramatical price changes are of interest in extreme value theory.

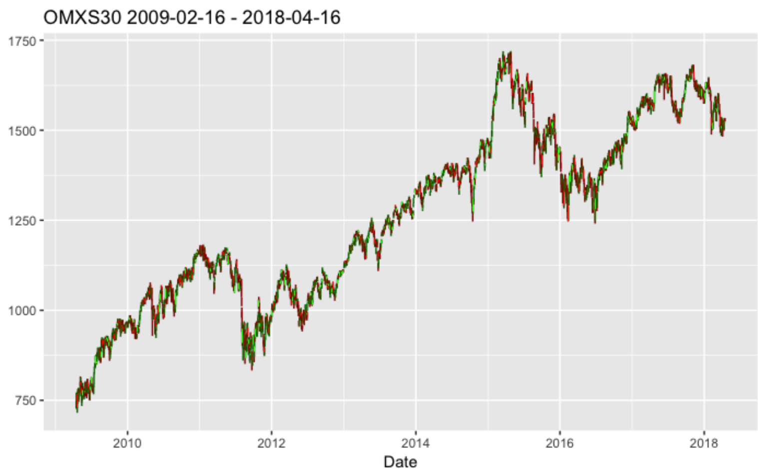

Figure 1: *Index prices for OMXS30 (SEK) between 2009-02-16 - 2018-04-16.*

In Section 3.1 we study the log-returns data and check if modelling the volatility with a GARCH model is adequate. Furthermore we estimate the parameters of the GARCH(1,1) model and fit a distribution to the residuals of the process using EVT. In Section 3.2 we present the recursive simulation procedure to obtain samples for the GARCH process. In Section 3.3, numerical computations are performed to compute the theoretical tail index of the stationary distribution and we verify these results with help of simulated data. In Section 3.4 we study how the tail index of the stationary distribution of the GARCH process behaves over time. In particular we compute the time it takes for the tail index to become stationary. Lastly in Section 3.5 we compute the constant c_+ in (11) and obtain an explicit formula for the stationary distribution of the volatility.

3.1 GARCH modelling

Before modelling the volatility of the daily log-returns, there are some features that need to be taken in consideration, such as the assumption of independent and identically distributed (iid) innovations and the well known phenomena of volatility clustering. Lets start by looking at a plot for the log-returns, see Figure 2. Here we notice that there are indeed some periods where the price changes are small and other periods where extreme values or big price changes are present.

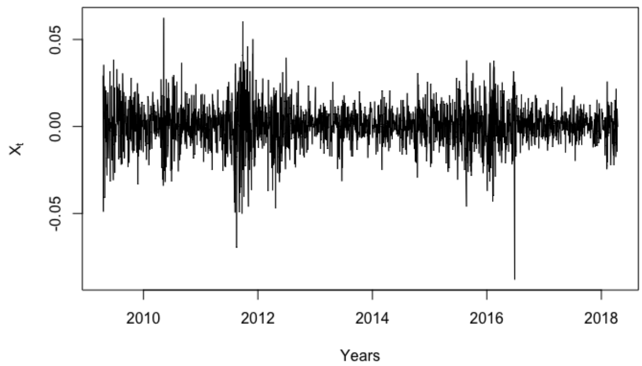

Figure 2: *Log-returns of OMXS30. Number of observations is 2260.*

Furthermore we present descriptive statistics of the log-returns data in Table 1. Notice that the skewness of the data is negative, which implies that the distribution of the data is weighted to the left. This is illustrated in Figure 10. Notice that there are observations in the leftmost part of the histogram. Moreover, the excess kurtosis (E.K.) is larger than 0, indicating that the distribution of the log-returns have heavier tails than the normal distribution. See the definition of skewness and kurtosis in Section 7.7.

Series	Min	Max	Median	Mean	Std.	Skeweness	E.K.
X_t	-0.088	0.062	0.00063	0.00031	0.0121	-0.3302	3.54

Table 1: *Descriptive statistics of log-returns.*

An autocorrelation plot is an useful tool to determine the presence of autocorrelation for the log-returns and heteroscedasticity for the squared values of the log-returns. This plot is based on the autocorrelation function (ACF), see Section 7.4. The horisontal dashed lines in Figure 3 indicates the 95% confidence interval for the estimators. If more than 5% of the spikes exceeds the confidence thresholds the ACF suggest the presence of autocorrelation.

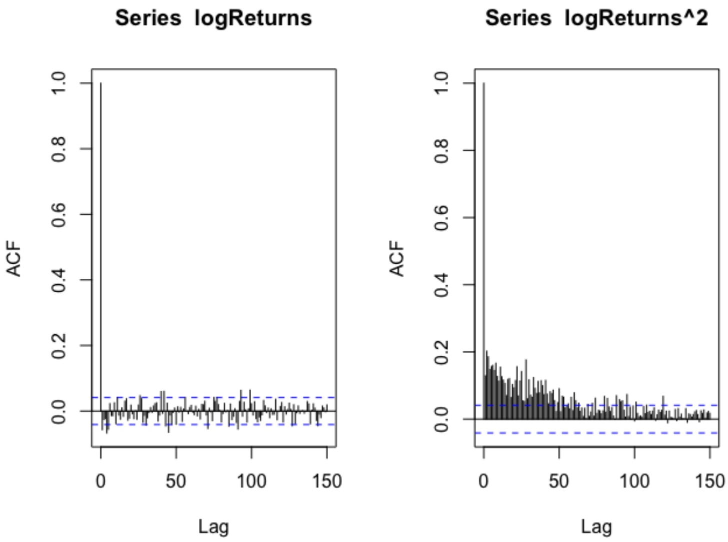

Figure 3: *Autocorrelation samples for daily log-returns. Left: log-returns* X_t*. Right: Squared log-returns* X_t^2*.*

Figure 3 left panel shows the ACF for the log-returns, here it seems that the log-returns series do not have any serial correlations since there are only a few spikes that extend beyond the significance levels. The right panel shows the ACF for the squared values of the log-returns. Notice the big spikes that exceeds the threshold levels, such spikes suggest that some ARCH effects are present in the series. Recall Section 2.3, where one of the objectives of ARCH-GARCH models is to adress for the need of many parameters in linear time series models in order to fit data well.

To confirm the above visual analysis we perform a Ljung-Box test, see Section 7.5. The null hypothesis is that the log-returns do not have any serial correlations, i.e. no autocorrelation. The null hypothesis for the squared log-returns is that heteroscedasticity is not present in the series. The result is shown in Table 3.

Series	P-value
X_t	0.0069
X_t^2	9.029e-10

Table 2: *Ljung-Box test for log-returns* X_t *and squared log-returns* X_t^2*. Test suggest rejection of* H_0 *for both series on* 5% *confidence level. This indicates that linear dependency and heteroscedasticity are present in the series.*

From the result in Table 3 we conclude that a non-linear time series model seems appropriate for modelling the volatility of the daily log-returns since heteroscedasticity is present in the series. But on the other hand we find evidence against the iid assumption.

Moreover we study the residuals of the GARCH fit, i.e.

$$\hat{Z} = (z_{t-n+1}, \cdots, z_t) = \left(\frac{x_{t-n+1} - \hat{\mu}}{\hat{\sigma}_{t-n+1}}, \dots, \frac{x_t - \hat{\mu}}{\hat{\sigma}_t} \right),$$

where the conditional standard deviation series $(\hat{\sigma}_{t-n+1}, \dots, \hat{\sigma}_t)$ are computed recursively from (5). It may be appropriate to model the conditional mean series $(\hat{\mu}_{t-n+1}, \dots, \hat{\mu}_t)$ with an AR model to filter the autocorrelation observed in the data. This because performing a two-sided t-test where the null hypothesis that the mean of the log-returns is equal to 0, is not rejected on 5% significance level (P-value is 0.2137). But for the purpose of this thesis modelling the conditional mean of the series properly is not of major interest. Hence the conditional mean $\hat{\mu}$ is estimated with the mean value of the log-returns.

The residuals are of interest when studying the heavy-tailed structure of the series and to check adequacy of the model. In Figure 11 the residuals are plotted against the theoretical quantiles of the normal distribution. The residuals appears to have a leptokurtic, heavy-tailed, distribution. This contradicts the assumption of conditional normality for the residuals. Moreover we perform a Ljung-Box test where the null hypothesis is that the residuals do not have any serial correlations, i.e. no autocorrelation. The same test is performed for the squared residuals where the null hypothesis is that heteroscedasticity is not present in the series.

Series	P-value
$\hat{Z}$	0.0433
$\hat{Z}^2$	0.3221

Table 3: *Ljung-Box test for residuals and their squared values. Test suggest rejection of H_0 for residuals on 5% confidence level, this indicates that the iid-hypothesis for the residuals does not seem to hold. On the other hand the test suggest presence of heteroscedacity for the squared residuals.*

Even though the iid assumption does not seem to hold for neither raw data nor the residuals, for the purpose of this thesis, we assume that the GARCH(1,1) is a proper volatility model for the log-returns of OMXS30. This assumption is made because it is of interest to study the explicit formula of the stochastic recurrence equation according to [7] in order to implement simulation procedures to study the stationary distribution of the GARCH process.

To estimate the coefficients in (5) we use Quasi-Maximum Likelihood estimation (QMLE). This method is particularly relevant for GARCH models because it provides consistent and asymptotically normal estimators for strictly stationary GARCH processes [3]. See Section 7.8 for a detailed explanation of the theory behind this optimisation method. The GARCH(1,1) is optimised to obtain the vector of parameter estimates $\hat{\theta} = (\hat{\mu}, \hat{\alpha_0}, \hat{\alpha_1}, \hat{\beta_1})^T$. This thesis concerns the study of equation (5), hence coefficients $(\hat{\alpha_0}, \hat{\alpha_1}, \hat{\beta_1})^T$ are of main interest. From the GARCH fit we obtain parameter estimates and the stochastic recurrence equation of the GARCH(1,1) model is

$$\sigma_t^2 = 0.0000023 + (0.096 Z_{t-1}^2 + 0.89)\sigma_{t-1}^2. \tag{12}$$

Notice that $\hat{\alpha}_1 + \hat{\beta}_1 \approx 0.985 < 1$, hence condition (6) is satisfied and ensures that the process is strictly stationary.

Lastly we fit a distribution for the residuals and the log-returns using EVT in order to take into account the leptokurtic structure of the tails. For this we estimate the tail index $1/\xi$ of the GPD using the POT method from EVT as in [6] p. 8. Let the ordered residuals $z_{(1)} > \cdots > z_{(n)}$ and fix a number of the data in the tail to be $k \ll n$, i.e. $k = nu$ for $u = (0.05, 0.1)$. This gives a random threshold at the $(k+1)$th order statistic. The GPD with parameters ξ and β is then fitted to the excess amounts $(z_{(1)} - z_{(k+1)}, \ldots, z_{(k)} - z_{(k+1)})$ over the threshold for all residuals exceeding the threshold. The result is given in Table 4.

Sample	u	$1/\hat{\xi}$	$\hat{\beta}$
$\hat{Z}^2$	0.05	3.43	1.6687
X_t^2	0.05	2.24	0.0003
$\hat{Z}^2$	0.10	6.58	1.9194
X_t^2	0.10	2.90	0.0003

Table 4: *Tail index estimation of GPD using POT method for squared residuals and log-returns. This indicates that the distribution of the residuals is less heavy-tailed than for the log-returns.*

3.2 Recursive simulation

It is of main interest to verify the theoretical results about the heavy-tailed stationary distribution of the GARCH process. Especially we want to verify that the solution of the stochastic recurrence equation (13) has a heavy-tailed distribution asymptotically. Moreover it is of interest to study how the tail index of the stationary distribution behaves over time and to obtain the time it takes for the tail index to converge to a certain value.

To obtain samples representing $X_t = \sigma_t Z_t$, the following recursive simulation procedure is performed:

1. Create a N by T matrix, each column represents a time-lag $t = 0, ..., T-1$ and N is the number of simulations.
2. Choose a start value σ_0 and compute recursively $\sigma_1, ..., \sigma_{T-1}$ according to (13), where the innovations Z_t are standard normal or standardised t distributed variables.
3. Iterate the simulation N times.
4. Extract for each column t, samples $\sigma_t Z_t$ or $\sigma_t^2 Z_t^2$, to get the desired simulation sample at time t.

3.3 Tail index

The result from Section 2.5 shows that the stationary distribution of σ_t^2 in equation (5) has a power-law tail asymptotically, i.e.

$$\mathbb{P}(\sigma_t^2 > x) \sim c_+ x^{-\alpha^*}, \quad x \to \infty.$$

We are interested in finding the theoretical tail index α^*, recall from Section 2.5 that one of the conditions of (11) is that $\mathbb{E}[A^{\alpha^*}] = 1$. Using the parameter estimates of GARCH(1,1) we apply the following computation to find the solution of α^* numerically.

1. Define the sequence $\{\alpha\}$ ranging between 0 and τ with 0.1 steps, where $\tau \in [0, \infty)$.
2. Iterate $A_i = \hat{\alpha_1} \cdot Z_i^2 + \hat{\beta_1}$ for $i = 1, ..., N$, for large N. The innovations Z_i are either random standard normal or standardised t distributed variables.
3. Approximate $\mathbb{E}[A^\alpha]$ with $\frac{1}{N}\sum_{i=1}^{N} A_i^\alpha$, according to the law of large numbers.
4. Iterate for each value in the sequence $\{\alpha\}$.

5. Stop the iteration when $\mathbb{E}[A^{\alpha}] \approx 1$ to get α^*.

Figure 4 illustrates the numerical computation performed when $N = 10000$. When the innovations are random standard normal variables, the theoretical tail index α^* is approximately equal to 2.45.

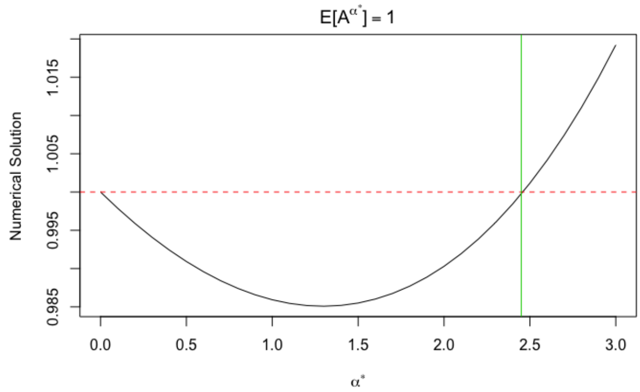

Figure 4: *Numerical solution to* $\mathbb{E}[A^{\alpha^*}] = 1$. *Vertical green line shows that* $\alpha^* \approx 2.45$.

We want to estimate, using the POT method from EVT, the tail index for simulated data representing $X_t^2 = \sigma_t^2 Z_t^2$ as t goes to infinity. To verify if the convergence value of the tail index corresponds to the value of the theoretical tail index α^* we implement the simulation procedure described in Section 3.2 for $T = 1000$ and $N = 100000$. The simulation is performed with start value σ_0 set to be equal to the empirical standard deviation of the log-returns data. The innovations are random standard normal variables. Each column in the simulation matrix represents a sample $\sigma_t^2 Z_t^2$. Recall from Section 2.5 that analysing the tail behaviour of the distribution of $\sigma_t^2 Z_t^2$ corresponds to analysing the samples σ_t^2 asymptotically. For $\sigma_t^2 Z_t^2$ the tail index should converge to a value that is lower than for σ_t^2, but both of these values should correspond to the theoretical tail index value. Thus it is also of interest to study the convergence value of the tail index for samples σ_t^2.

For each column in the obtained simulation matrix, the parameter $1/\xi_t$ of the GPD corresponding to the tail index of the stationary distribution is estimated using the POT method from EVT where $u = 0.05$. We get a vector of tail index estimates, i.e. $1/\hat{\xi}_1, ..., 1/\hat{\xi}_{1000}$. The convergence value of the tail index is approximated with the mean value of the parameter estimates sample according to the law of large numbers, see Section 7.1, i.e. we get the tail index estimator $\hat{\delta} = \frac{1}{T}\sum_{t=1}^{T} 1/\hat{\xi}_t$. The tail index estimator is tested to assess if it is statistical significant equal to the theoretical tail index α^* performing a one sample two-sided t-test on 95% confidence level, i.e. the null hypothesis being $H_0 : \hat{\delta} = 2.45$ and the alternative hypothesis

$H_a : \hat{\delta} \neq 2.45$. The result is shown in Table 5.

Sample	$\sigma_t^2 Z_t^2$	σ_t^2
$\hat{\delta}$	**2.42** (2.37, 2.47)	**2.60** (2.20, 2.99)

Table 5: *Text in bold indicates that values are statistically significant equal to the theoretical tail index value $\alpha^* = 2.45$. In parenthesis are given the 95% confidence intervals for the t-test. Notice that the tail index convergence value is lower for $\sigma_t^2 Z_t^2$ indicating that the stationary distribution is more heavy-tailed for $\sigma_t^2 Z_t^2$ than for σ_t^2.*

For the simulation sample $\sigma_t^2 Z_t^2$ the null hypothesis that the tail index estimator $\hat{\delta}$ is equal to the theoretical value α^*, is accepted on 5% significance level (P-value is 0.2587). In the same way the null hypothesis is accepted for sample σ_t^2 (P-value is 0.4581), see Table 5. This means that the tail index α^* is the true convergence value of the stationary distribution of the GARCH(1,1) model when the innovations of the process are assumed to be standard normal variables. Hence we can say that when assuming standard normal innovations, $\sigma_t^2 Z_t^2$ is regularly varying with index 2.42 and σ_t^2 is regularly varying with index 2.60, and both of these values are statistically significant equal to the theoretical index value $\alpha^* = 2.45$. This confirms Breiman's lemma, see 2.5.

Moreover we assess if σ_t is regularly varying with index $2\alpha^*$. Appling the POT method from EVT to samples σ_t we obtain that $\hat{\delta} = 5.83$. We perform the following two-sided t-test

$$H_0 : \hat{\delta} = 2\alpha^* = 4.9, \quad H_a : \hat{\delta} \neq 4.9,$$

and obtain that on 5% significance level the null hypothesis is accepted (P-value is 0.101) where the 95% condidence interval for the test is $(4.72, 6.93)$. Hence σ_t is regularly varying with index $2\alpha^*$. This is what we wanted to verify in Section 2.5.

We perform the above analysis assuming that the innovations are standardised t distributed variables with ν degrees of freedom. In the same way we verify if the the tail index estimator $\hat{\delta}$ is statistically significant equal to its theoretical value α^*. The result can be found in Section 7.10 Table 9.

3.4 Tail index convergence

To further analyse the stationary distribution of the GARCH process, we study how the tail index of the distribution behaves over time. For this we use the simulation procedure described in Section 3.2 for $T = 1000$ and $N = 100000$ when the innovations are standard normal variables. In the same way as before the simulation is performed with start value σ_0 equal to the empirical standard deviation of the stock index log-returns. For each simulation sample $\sigma_t^2 Z_t^2$, the parameter $1/\xi_t$ of the GPD, corresponding to the tail index of the stationary distribution at time t, is estimated using the POT method from EVT. We set $u = 0.05$ and obtain a random threshold for the POT method $k = 0.05 \cdot n$. Here we get a vector of tail index estimates, i.e. $1/\hat{\xi}_1, ..., 1/\hat{\xi}_{1000}$. The tail index estimates are plotted against time, see Figure 5. Here we see that for both simulation samples the tail index converges to a stationary value as time goes to infinity. From the plots we notice that the convergence appears to occur between times 0 and 200.

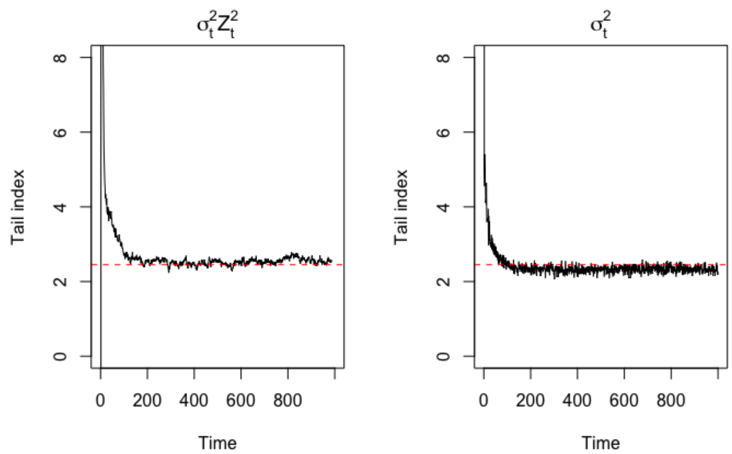

Figure 5: *Tail index behaviour over daily time-lags for samples $\sigma_t^2 Z_t^2$ (left) and samples σ_t^2 (right) assuming that the innovations $Z_t \sim N(0,1)$. Horisontal red dashed line represents value of theoretical index $\alpha^* = 2.45$.*

To obtain the time it takes for the distribution of $\sigma_t^2 Z_t^2$ to reach stationarity we study the simulation samples $\sigma_t^2 Z_t^2$ at times $t = 0, ..., t^*, ..., 1000$, where t^* is the convergence time. For each simulation sample, the tail index $1/\xi_t$ is estimated using the POT method from EVT. Then we check the first time the tail index obtains a value that is less than or equal to the theoretical tail index value. When the start value of the simulation σ_0 is equal to the empirical standard deviation of the log-returns, we obtain that $t^* = 78$, see Figure 6.

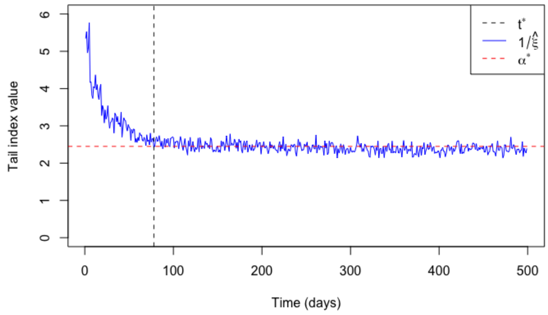

Figure 6: *Convergence time of tail index is* $t^* = 78$.

Figure 7 illustrates how the distribution of some simulation samples $\sigma_t Z_t$ changes over time. Here the sample quantiles are plotted against the theoretical quantiles of the normal distribution. Notice that when $t = 0$, the simulation sample looks like a straight line since we assumed standard normal distribution for the innovations. As time increases the tails for the distribution are more leptokurtic. Notice the light-blue line representing the sample at time 78, it looks very similar to the line at time 1000. Hence time 78 seems to be the actual convergence time to stationarity of the distribution for the GARCH process.

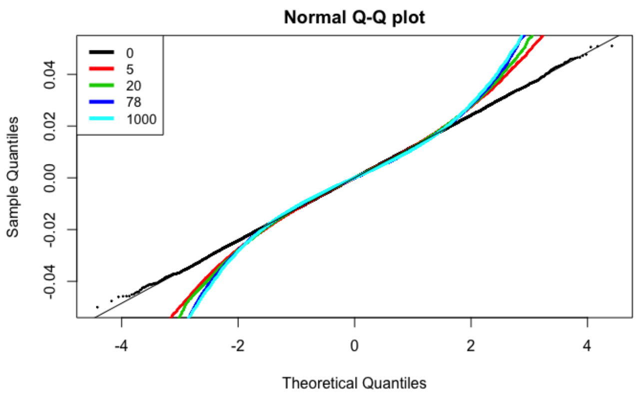

Figure 7: *Sample quantiles of* $\sigma_t Z_t$ *at times* $t = 0, 5, 20, 78, 1000$*, are plotted against theoretical quantiles of the normal distribution. As time increases the tails of the GARCH process are more leptokurtic.*

In Table 6 we present descriptive statistics for the distribution of samples $\sigma_0 Z_0$, $\sigma_{78} Z_{78}$ and $\sigma_{1000} Z_{1000}$. Notice that since we assume standard normal innovations, the excess kurtosis for $\sigma_0 Z_0$ is almost zero. As time increases the distribution is weighted to the right, this is confirmed by the positive value of skewness. More important is that the excess kurtosis is much higher for samples $\sigma_{78}Z_{78}$ and $\sigma_{1000}Z_{1000}$, indicating a heavy-tailed distribution. As seen from Table 6, the kurtosis and skewness are not the same for samples $\sigma_{78}Z_{78}$ and $\sigma_{1000}Z_{1000}$. This inidcates that our result regarding the convergence time 78 may not that reliable.

Series	Min	Max	Median	Mean	Std.	Skewness	E.K.
$\sigma_0 Z_0$	-0.0500	0.0509	3.3e-5	-4.7e-7	0.0121	-0.0016	-0.0062
$\sigma_{78}Z_{78}$	-0.1869	0.2177	-4.4e-5	-3.7e-5	0.0134	0.0205	5.13
$\sigma_{1000}Z_{1000}$	-0.2034	0.2112	3.6e-5	6.4e-5	0.0136	0.0131	6.75

Table 6: *Descriptive statistics for samples $\sigma_0 Z_0$, $\sigma_{78}Z_{78}$ and $\sigma_{1000}Z_{1000}$.*

Furthermore we study how the tail index behaves over time when instead of normality, we assume that the innovations are standardised t distributed. In the same way as in Section 3.3 we compute the value of the theoretical tail index when the innovations are t distributed with $\nu = 5$ degrees of freedom and obtain that $\alpha^* \approx 1.69$. Notice that condition $\alpha^* < \nu/2$ is satisfied. Figure 13 in Section 7.10 illustrates how the tail index behaves over time in this case. In the right plot we study the simulation sample σ_t^2. Here the tail index appears to converge to the theoretical tail index value 1.69. A one sample two-sided t-test confirms this, i.e. the null hypothesis that the convergence value of the tail index is equal to the theoretical tail index α^* is not rejected on 5% significance level (P-value is 0.20). The 95% confidence interval for the test is $(1.46, 1.74)$. On the other hand we notice that for sample $\sigma_t^2 Z_t^2$, the convergence value of the tail index is lower than the value of the theoretical tail index, see the left plot in Figure 13. Recall Section 3.3 where all the tests performed, assessing if the tail index estimator $\hat{\delta}$ is equal to the theoretical value of the tail index α^*, are rejected on 5% confidence level, see Table 9. Intuitively this is because σ_t^2 has already a heavy-tailed distribution since we assumed t distributed innovations for the recursive simulation. Multiplying this by another t distributed variable Z_t^2 makes the output more heavy-tailed. Therefore we do not compute the convergence time when the innovations are t distributed variables.

Lastly we study the convergence times of the tail index when the simulations are implemented for different start values σ_0. Here we assume as before that the innovations are standard normal random variables. The start value σ_0 represents todays volatility. The results can be found in Table 7. The convergence tends to be faster for higher start values σ_0 than for lower, meaning that for high values of todays volatility, the GARCH process obtains a heavy-tailed distribution faster.

σ_0	0.007	0.009	0.011	0.013	0.015	0.017	0.019	0.021
t^*	116	91	78	57	57	57	52	52
σ_0	0.023	0.025	0.027	0.029				
t^*	52	52	52	52				

Table 7: *Convergence times t^* when simulations are performed for different start values σ_0 and innovations are assumed to be standard normal variables. This indicates that the convergence is faster given higher start values. The mean convergence time is $t^* = 65$.*

3.5 The stationary distribution

We verified in Section 3.3 that the tail index α^* converges to a value which corresponds to the tail index of the stationary distribution of the GARCH process. To compute the constant c_+ of the stationary distribution in (11) numerically, we implement the result from equation (8) as follows:

1. Let $\alpha = \alpha^*$ where α^* is the theoretical tail index.
2. Choose a start value σ_0 and compute recursively $\sigma_1, ..., \sigma_{T-1}$ according to (13), where the innovations Z_i are standard normal random variables.
3. Compute $\mathbb{E}[|A_i\sigma_{t-1}^2 + B_i|^\alpha - |A_i\sigma_{t-1}^2|^\alpha]$, where $A_i = \hat{\alpha}_1 Z_i + \hat{\beta}_1$ and $B_i = \hat{\alpha}_0$ and σ_{t-1} from step 2. Iterate this for $i = 1, ..., N$, for large N. Then approximate the expected value with the mean according to the law of large numbers.
4. Compute $\frac{1}{2\alpha \mathrm{m}_\alpha}$ where $\mathrm{m}_\alpha = \mathbb{E}[|A_i|^\alpha \log |A_i|]$, by first iterating $|A_i|^\alpha \log |A_i|$, $i = 1, ..., N$, for large N and A_i as in step 3. Then approximate the expected value with the mean.
5. Lastly compute $c_+ = \frac{1}{2\alpha \mathrm{m}_\alpha}\mathbb{E}[|A\sigma_{t-1}^2 + B|^\alpha - |A\sigma_{t-1}^2|^\alpha]$.

From the computation described above we obtain that $c_+ \approx 1.72 \cdot 10^{-10}$. Since $\alpha^* = 2.45$, the stationary distribution of σ_t is

$$\mathbb{P}(\sigma_t > x) \sim 1.72 \cdot 10^{-10} x^{-4.9}, \quad x \to \infty. \tag{13}$$

To verify if this result seems reasonable we study the empirical distribution of σ_t when $t = 78$. This because in Section 3.4 we obtained that at time 78, the tail index obtains a stationary value. Hence we study the stationary distribution of σ_{78}. Using the same recursive simulation procedure described in Section 3.2 we study σ_{78} for $N = 10000$.

Let $\overline{F}_{\sigma_t}(x)$ be the empirical distribution of the simulation sample σ_t. We want to assess if

$$\overline{F}_{\sigma_t}(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}\{\sigma_t > x\} \approx \mathbb{P}(\sigma_t > x). \tag{14}$$

Figure 8 shows a histogram for the distribution of σ_{78}. The vertical dashed lines represent the 0.90, 0.95, 0.975, 0.99 and 0.995 quantiles of the sample data. These are the x values that we want to compare in (15). To compute the corresponding probabilities for the empirical distribution we use the indicator function $\mathbb{I}\{\sigma_t > x\}$ to count the proportion of values that exceeds the given x quantile or threshold value, see Table 8.

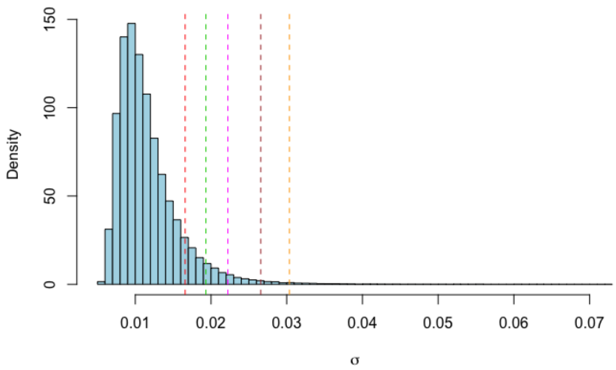

Figure 8: *Histogram of* $\{\sigma_{78}\}$. *Vertical dashed lines represent the 0.90, 0.95, 0.975, 0.99 and 0.995 quantiles of the sample data.*

x	0.017	0.019	0.022	0.027	0.030
$\overline{F}_{\sigma_{78}}(x)$	0.100	0.050	0.025	0.010	0.005
$\mathbb{P}(\sigma_{78} > x)$	0.091	0.043	0.022	0.009	0.005

Table 8: *Probabilities for stationary vs the empirical distribution of* $\{\sigma_{78}\}$.

Notice from Table 8 that in overall the computed probabilities with the stationary distribution (14) seem to correspond to the probabilities obtained from the empirical distribution.

3.6 Software

The software used in this thesis is R. R is a free software environment for statistical computing and graphics, see https://www.r-project.org/ for more information.

The packages used throutout this thesis are *tidyquant*, *quantmod*, *xts*, *dplyr*, *ggplot2*, *fitdistrplus*, *fGarch*, *QRM* and *fExtremes*. The latter three are aimed for volatility modelling and extreme value theory. The other packages are used for data management, fitting distributions and explanatory data analysis. For a detailed explanation of each package consult https://cran.r-project.org/web/packages/available_packages_by_name.html.

The simulation procedures and numerical computations used in this thesis have also been implemented in R.

4 Results

In Section 3.3 we studied the asymptotic results for the power-law tail of the non-linear time series model GARCH(1,1). More specific we looked at the solution of the stochastic recurrence equation (5). First we computed the theoretical tail index α^* numerically using condition $\mathbb{E}[\alpha^*] = 1$ from Theorem 2.4.7 in [7] page 61. We first assumed that the innovations are standard normal random variables. Figure 4 illustrates the numerical solution.

Moreover we implemented a recursive simulation procedure to obtain samples $\sigma_t^2 Z_t^2$ at times $t = 0, ..., T-1$, see Section 3.2. This because it was of main interest to study how the tail index behaves over time. For each simulation sample, we estimated $1/\hat{\xi}_t$ from the GPD using the POT method from EVT. Then we introduced the tail index estimator $\hat{\delta}$, which is simply the mean value of the tail index estimates. A one sample two-sided t-test was used to test if the tail index estimator is statistically significant equal to the theoretical tail index value on 5% confidence level, i.e.

$$H_0 : \hat{\delta} = \frac{1}{T} \sum_{t=0}^{T-1} 1/\hat{\xi}_t = \alpha^* = 2.45, \quad H_a : \hat{\delta} \neq 2.45.$$

Here the null hypothesis for the test was accepted for both $\sigma_t^2 Z_t^2$ and σ_t^2, see Table 5. The fact that the null hypothesis was accepted confirms that studying the distribution of $\sigma_t^2 Z_t^2$ corresponds to analysing the distribution of σ_t^2 asymptotically. This was discussed in Section 2.5 when we applied Breiman's lemma to derive the stationary distribution of σ_t. We also notice that the convergence value of the tail index is lower for $\sigma_t^2 Z_t^2$ than for σ_t^2 indicating that the distribution is more heavy-tailed for $\sigma_t^2 Z_t^2$ than for σ_t^2. This result is illustrated in Figure 5.

The tail index estimator $\hat{\delta}$ computed for $\sigma_t^2 Z_t^2$, obtained a value equal to 2.42 when $Z_t \sim N(0, 1)$. Hence $\sigma_t^2 Z_t^2$ is regularly varying with index 2.42. Moreover is σ_t regularly varying with index 5.83. This value was statistically significant equal to the theoretical $2\alpha^* = 4.9$. From this we conclude that when the input variables (A_t, B_t) of stochastic recurrence equation (3) are light-tailed, corresponding to the case when the innovations Z_t are standard normal random variables, it generates an output which is heavy-tailed since the tail index value is positive. This means that after some time the distribution of the stochastic process will converge to a heavy-tailed stationary distribution.

Furthermore we implemented the same recursive simulation as before to obtain samples $\sigma_t^2 Z_t^2$ and σ_t^2, when instead of normality, we assumed that the innovations are standardised t distributed with ν degrees of freedom.

We computed the tail index estimator $\hat{\delta}$ and assessed if its value is equal to the theoretical tail index α^*. The tests failed for samples $\sigma_t^2 Z_t^2$ and were accepted for most of the tests performed when studying samples σ_t^2, see Table 9. How the tail index behaves over time is illustrated in Figure 13.

When the innovations are t distributed with $\nu = 5$ degrees of freedom, we obtained $\alpha^* = 1.69$. The tail index estimator for $\sigma_t^2 Z_t^2$ was in this case $\hat{\delta} = 1.223$. This implies that σ_t is regularly varying with index $2\hat{\delta} \approx 2.45$ asymptotically. But since $2\alpha^* = 3.38 \neq 2\hat{\delta}$, means that when we assume t distribution for the innovations, the output generated from the GARCH process is more heavy tailed and the tail index of the distribution for σ_t does not correspond to the tail index of $\sigma_t Z_t$. This result seems logical since we assume t distribution for the innovations, the t distribution is heavier-tailed than the normal distribution, thus the input variables (A_t, B_t) of the stochastic recurrence equation generates an output which is even more heavier-tailed.

Moreover we studied the speed of convergence of the tail index. For this we checked the time when the tail index obtains a value that is less than or equal to the theoretical tail index. Graphically it seems to correspond to the actual convergence time of the process, i.e. when the innovations are assumed to be standard normal variables the convergence time is 78, see Figure 6.

The obtained convergece time was studied using a QQ-plot where the sample quantile of $\sigma_t Z_t$ were plotted against the theoretical quantiles of the normal distribution. As seen in Figure 7, the distribution of $\sigma_{78} Z_{78}$ looks very similar the distribution of $\sigma_{1000} Z_{1000}$. Hence it seems reasonable to state that at time 78 the GARCH process obtains a heavy-tailed stationary distribution when the recursive simulation of the process has start value equal to the empirical standard deviation of the log-returns for OMSXS30, i.e. $\sigma_0 = 0.012$. In Table 6 we looked at the descriptive statistics for the distribution of samples $\sigma_0 Z_0$, $\sigma_{78} Z_{78}$, $\sigma_{1000} Z_{1000}$, here we noticed that for the latter two samples, the kurtosis is positive and larger than for $\sigma_0 Z_0$, indicating that as time increases the stationary distribution of the GARCH process converges to a heavy-tailed distribution, but since there is a small difference in the skewness and kurtosis values for $\sigma_{78} Z_{78}$ and $\sigma_{1000} Z_{1000}$, the obtained convergence value 78 may be questionable.

In addition we computed the convergence times when the recursive simulation of $\sigma_t^2 Z_t^2$ was implemented with different start values σ_0. This was done assuming standard normal innovations. The results are summarised in Table 7. Figure 9 illustrates this result. Observe that for higher values of σ_0 the convergence occurs faster.

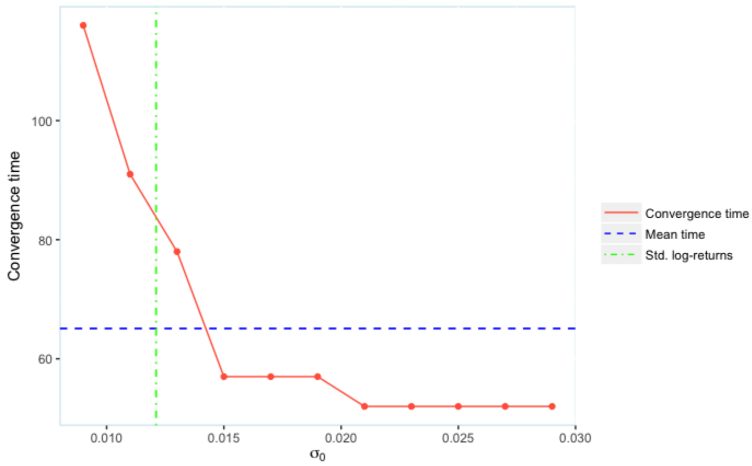

Figure 9: *Convergence times t^* plotted against start values σ_0 indicates that for higher start values, the convergence of tail index is faster.*

Lastly we computed the constat c_+ in (11), this to get the formula of the stationary distribution of σ_t, see (14). This result was verified successfully by studying the empirical distribution of σ_t when t was set to the convergence time of the tail index, i.e. $t = 78$. The computed probabilities with the formula of the stationary distribution (11) seemed to correspond to the proportion of exceedances beyond a threshold (quantile) x in the empirical distribution, see Table 8.

5 Conclusion

This thesis concerns a study on the convergence to the stationary distribution of the GARCH(1,1) process. In particular we study the tail index convergence to stationarity. We conclude that the process obtains a heavy-tailed distribution when the input variables (A_t, B_t) of the stochastic recurrence equation are light-tailed, i.e. when the innovations are assumed to be standard normal variables.

Using the log-returns of OMXS30 we fitted a GARCH(1,1) model and estimated the coefficients to the stochastic recurrence equation. Here we assumed that the GARCH(1,1) model fits data well. The log-returns data appeared to have significant serial correlations as discussed in Section 3.1. The conditional mean of the series could have been fitted using an AR model but since we wanted to use the explicit formula of the stochastic recurrence equation to implement simulation procedures, we assumed that the conditional mean of the series is constant.

To properly model the volatility of OMXS30, another non-linear time series model, i.e. a higher order model, can be fitted and analysed. This may be the case when modelling volatility of financial time series for predictive purposes. Fitting another AR-GARCH model for the series implies some major changes in the structure of the stochastic recurrence equation and the simulation procedures. This is left for further research.

We also studied the convergence times of the tail index and an interesting result is that given high values of todays volatility the convergence appears to be faster than in the case when todays volatility is assumed to be low, see Figure 9. Volatile markets imply that the risk of extreme events increase, hence the fact that the convergence to a heavy-tailed distribution of the process is faster for high volatility values seems reasonable. The mean convergence time when assuming different values of todays volatility is around time 65. The time-lags in this thesis represent days, this implies that at least two months of data is required to obtain a stationary distribution for the process. This is of course under the assumption that the GARCH(1,1) model fits the log-returns data well.

The importance of this result is that given the stationary distribution of the process and the amount of data needed to compute it, we can calculate VaR_p in a straightforward manner. Solving x from (11) by inverting the formula of the stationary distribution gives

$$x = \left(\frac{c_+}{p}\right)^{\frac{1}{2\alpha}} := \hat{x_p}, \tag{15}$$

where $p \in (0, 1)$ is the VaR level. Thus $\text{VaR}_p(X) \approx \hat{x}_p$.

This means that setting the computed values for the constant of the stationary distribution c_+, the estimated tail index α^* and the value-at-risk level p in (16), we can calculate a quantile value $\hat{x_p}$. This quantile value corresponds to the magnitude of risk for a potential loss in a portfolio, or in this case, the value-at-risk at level p for the log-returns of OMXS30.

Moreover one would need to backtest this approach of calculating VaR and compare its performance against other methods for tail estimation. See for instance [6], where the GPD-approximation appears to be preferable to model the leptokurtic tails of the residuals of an AR-GARCH process. This method outperforms various other tail estimation methods when computing risk measurements.

This study can be applied to multivariate time series models, see for instance ch. 4 in [7], this is left for further reasearch.

6 References

[1] Alm S.E., Brittom T. (2011). *Stokastik: sannolikhetsteori och statistikteori med tillämpningar.* Första upplagan, Liber AB.

[2] Brockwell P.J., Davis R.A. (2016) *Introduction to Time Series and Forecasting*, Springer.

[3] Francq C., Zakoïan J.-M. (2010) *GARCH Models - Structure, Statistical Inference and Financial Applications* John Wiley & Sons Ltd.

[4] Held L., Sabanés B.D. (2014) *Applied Statistical Inference.* Springer.

[5] Hult H., Lindskog F., Hammarlid O., Rehn C. J. (2014) *Risk and Portfolio Analysis: Principles and Methods.* Springer.

[6] McNeil A.J., Frey R. (2000) Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach. *Journal of Empirical Finance*, Vol 7(3-4): 271-300.

[7] Mikosch T., Buraczewski S., Damek E. (2010) *Stochastic Models with Power-Law Tails, The Equation* $X = AX + B$, Springer.

[8] Resnick S. I. (2007) *Heavy-tail Phenomena, Probabilistic and Statistical Modeling*, Springer.

[9] Tsay R.S. (2010) *Analysis of Financial Time Series, Third Edition.* John Wiley & Sons Inc.

7 Appendix

7.1 Law of large numbers (LLN)

Using the definition as in [4] p. 330, let $X_1, X_2, ...$ be a sequence of iid random variables with finite expectation μ. Then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu \quad \text{as} \quad n \to \infty.$$

7.2 Normal distribution

The density function of normal distributed random variable X, i.e. $X \sim N(\mu, \sigma^2)$, is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}.$$

The expected value is $\mathrm{E}[X] = \mu \in \mathbb{R}$ and the variance is $\mathrm{Var}(X) = \sigma^2 > 0$.

X is standard normal if $\mu = 0$ and $\sigma^2 = 1$, i.e. $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$. [4] p. 337.

7.3 Student's t distribution

The density function of a random variable X that is t distributed with $\nu > 0$ degrees of freedom, i.e. $X \sim t(\mu, \sigma^2, \nu)$, is given by

$$f(x) = C \cdot \left\{1 + \frac{1}{\nu\sigma^2}(x-\mu)^2\right\}^{-\frac{\nu+1}{2}}.$$

The variance is $\mathrm{Var}(X) = \sigma^2 \frac{\nu}{\nu-2}$ if $\nu > 2$, $\sigma^2 > 0$ and the expected value is $\mathrm{E}[X] = \mu$ if $\nu > 1$, $\mu \in \mathbb{R}$.

A t distributed random variable with ν degrees of freedom, i.e. $X \sim t(\mu, \sigma^2, \nu)$ converges in distribution to a normal random variable $Y \sim N(\mu, \sigma^2)$ as $\nu \to \infty$. [4] p. 338.

7.4 Autocorrelation function (ACF)

Consider a weakly stationary return series X_t. The correlations between a variable X_t and its past values $X_{t-1}, X_{t-2}, ..., X_{t-\ell}$ are referred to as serial correlations or autocorrelations. The lag-ℓ autocorrelation of X_t is defined as

$$\rho_\ell = \frac{\text{Cov}(X_t, X_{t-\ell})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t-\ell})}} = \frac{\text{Cov}(X_t, X_{t-\ell})}{\text{Var}(X_t)} = \frac{\gamma_\ell}{\gamma_0},$$

where the property $\text{Var}(X_t) = \text{Var}(X_{t-\ell})$ for a weakly stationary series is used. In general the lag-ℓ sample autocorrelation of X_t is defined as

$$\hat{\rho}_\ell = \frac{\sum_{t=\ell+1}^{T}(X_t - \bar{x})(X_{t-\ell} - \bar{x})}{\sum_{t=1}^{T}(X_t - \bar{x})^2}, \quad 0 \leq \ell < T - 1,$$

where $\bar{x} = \sum_{t=1}^{T} X_t/T$. If $\{X_t\}$ is an iid sequence with $\text{E}[X_t^2] < \infty$, then $\hat{\rho}_\ell$ is asymptotically normal with mean zero and variance $1/T$ for any fixed positive integer ℓ. [9] p. 24.

7.5 Ljung-Box test

Consider an iid sequence $X_1, ..., X_n$ with finite-variance . The Ljung-Box test can be used to test the significance of the autocorrelation coefficient. The test statistic, which is an improvement of *the portmanteau test*, is defined as

$$Q_{LB} = n(n+2)\sum_{t=1}^{h}\frac{\hat{\rho}_\ell^2}{n-t}, \tag{16}$$

where n is the length of the series and $\hat{\rho}_\ell$, $\ell = 1, ..., h$, is the autocorrelation coefficient. A large value of (17) suggests that the sample autocorrelation of the data are too large to be from an iid sample. We therefore reject the iid hypothesis at level α if $Q_{LB} > \chi^2_{1-\alpha}(h)$, where $\chi^2_{1-\alpha}(h)$ is the $1-\alpha$ quantile of the chi-squared distribution with h degrees of freedom. [2] p. 30-31.

7.6 T-test

If given a sample $x = \{x_i\}_{i=1}^n, n > 30$, we want to study a hypothetical mean value μ_0 by performing a t-test. The null hypothesis $H_0 : \mu_0 = \bar{x}$ is tested against the alternative hypothesis $H_a : \mu_0 \neq \bar{x}$ by computing the following test statistic

$$T = \frac{\bar{x} - \mu_0}{s_x^2/\sqrt{n}} \sim t_{\alpha/2}(n-1).$$

Here $\bar{x}$ is the sample mean, s_x^2 is the sample variance and n the sample size. The test statistic is t distributed with $n-1$ degrees of freedom and the null hypothesis is rejected on confidence level α if $|T| > t_{\alpha/2}(n-1)$. [1]

7.7 Skewness and Kurtosis

As stated in [9] p. 8, the third central moment of a random variable X, measures the symmetry of X with respect to its mean, whereas the 4th central moment measures the tail behaviour of X. In statistics, skewness (3rd normalized central moment) and kurtosis (4th normalized central moment), are often used to summarise the extent of asymetry and tail thickness.

Skewness is defined as

$$S(x) = \mathrm{E}\left[\frac{(X-\mu_x)^3}{\sigma_x^3}\right],$$

and kurtosis is defined as

$$K(x) = \mathrm{E}\left[\frac{(X-\mu_x)^4}{\sigma_x^4}\right].$$

The quantity $K(x) - 3$ is called excess kurtosis because $K(x) = 3$ for a normal distribution. Thus, the excess kurtosis of a normal random variable is zero. A distribution with positive excess kurtosis is sais to have hevy tails.

7.8 Quasi-maximum likelihood estimation (QMLE)

We write this optimising method as in ch. 7 [3], but using the notations in this thesis.

Assume that the observations $X_1, ..., X_n$ constitute a realisation of a GARCH(p,q) process,

$$\begin{cases} X_t = \sigma_t Z_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \quad \forall t \in \mathbb{Z}, \end{cases}$$

where Z_t is a sequence of iid variables with mean zero and variance 1, $\alpha_0 > 0, \alpha_i \geq 0$ $(i = 1, ..., p)$, and $\beta_j \geq 0$ $(j = 1, ..., q)$. The orders p and q are assumed to be known. The vector of parameters

$$\theta = (\theta_1, ..., \theta_{p+q+1})^T := (\alpha_0, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q)^T$$

belongs to a parameter space of the form

$$\Theta \subset (0, +\infty) \times [0, \infty)^{p+q}.$$

The true value of the parameter is unknown, and is denoted by

$$\theta_0 = (\alpha_0, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q)^T.$$

To write the likelihood of the model the Gaussian quasi-likelihood function is assumed for the innovation Z_t, which, conditionally on initial values $X_0, ..., X_{1-q}, \tilde{\sigma}_0^2, ..., \tilde{\sigma}_{1-p}^2$ to be specified below, coincides with the likelihood when Z_t are distributed as standard Gaussian or normal. The conditional Gaussian quasi-likelihood is given by

$$L_n(\theta) = L_n(\theta; X_1, ..., X_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\tilde{\sigma}_t^2}} \exp\left(-\frac{X_t^2}{2\tilde{\sigma}_t^2}\right)$$

where $\tilde{\sigma}_t^2$ are recursively defined, for $t \geq 1$, by

$$\tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\theta) = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i}^2 + \sum_{j=1}^{q} \beta_j \tilde{\sigma}_{t-j}^2.$$

A QMLE of θ is defined as any measurable solution $\hat{\theta}_n$ of

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} L_n(\theta).$$

Taking the logarithm, it is seen that maximizing the likelihood is equivalent to minimizing, with respect to θ,

$$\tilde{\mathbf{I}}_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \tilde{\ell}_t, \quad \text{where } \tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{X_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2.$$

A QMLE is thus a measurable solution of the equation

$$\hat{\theta}_n = \arg\min_{\theta \in \Theta} \tilde{\mathbf{I}}_n(\theta).$$

For more information about the choice of the initial values, likelihood equations and asymptotic properties, see ch. 7 [3].

7.9 Causality

As stated in [7] p. 10, the solution X_t to the stochastic recurrence equation 3 is *causal* or *non-anticipative*, if for every t, X_t is a measurable function of past and present noise variables $(A_s, B_s)_{s \leq t}$.

7.10 Figures and Tables

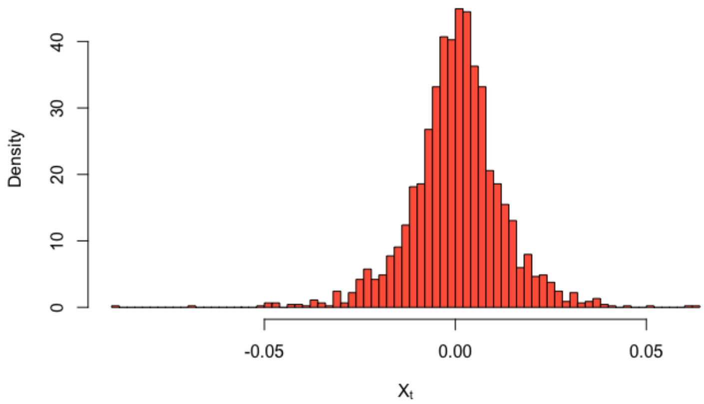

Figure 10: *Histogram for log-returns of OMXS30 illustrates that the data is weighted to the left, i.e. negative skeweness.*

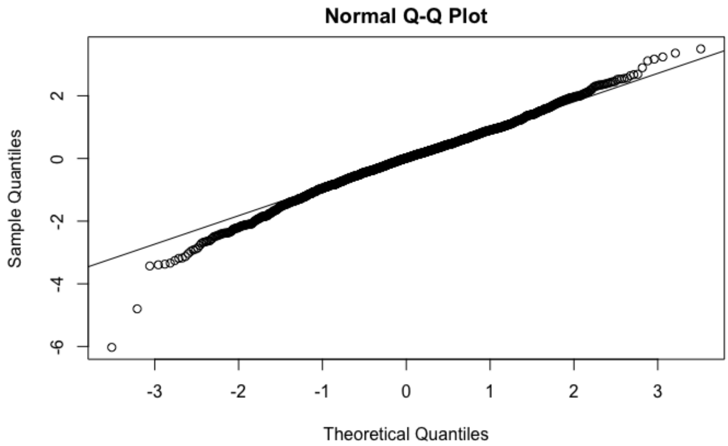

Figure 11: *Plot of residuals against the theoretical quantiles of the normal distribution shows that the distribution of residuals is leptokurtic.*

Sample		$\sigma_t^2 Z_t^2$	σ_t^2
Z-dist.	α^*	$\hat{\delta}$ (C.I.)	$\hat{\delta}$ (C.I.)
t_4	1.46	1.054 (1.053, 1.056)	1.431 (1.412, 1.450)
t_5	1.69	1.223 (1.220, 1.226)	**1.598** (1.455, 1.740)
t_6	1.76	1.367 (1.364, 1.370)	**2.034** (1.732, 2.335)
t_7	1.89	1.480 (1.477, 1.484)	**1.900** (1.746, 2.055)
t_8	1.85	1.572 (1.568, 1.577)	2.097 (1.856, 2.338)
t_9	1.94	1.650 (1.645, 1.655)	**2.135** (1.636, 2.633)

Table 9: *Text in bold indicates that values are statistically significant equal to the theoretical tail index value* α^**. In parenthesis are given the* 95% *confidence intervals for the t-test.*

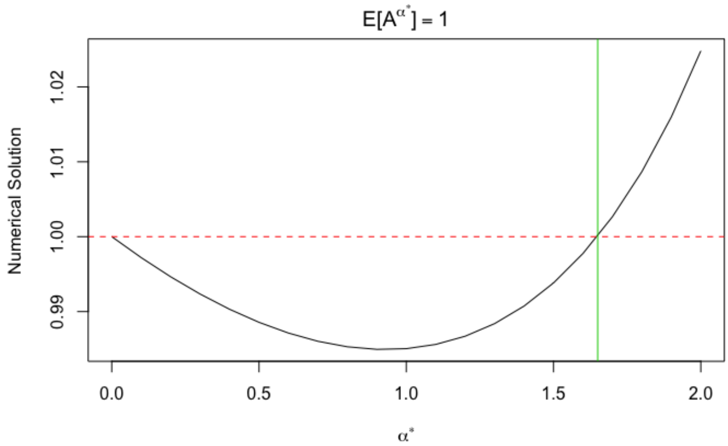

Figure 12: *Numerical solution to* $\mathbb{E}[A^{\alpha^*}] = 1$ *when* $Z_t \sim t_5$*. Vertical green line shows that* $\alpha^* \approx 1.65$*.*

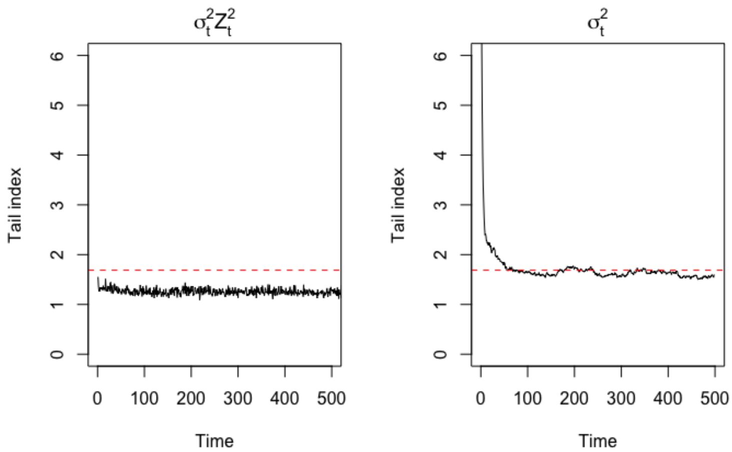

Figure 13: *Tail index behaviour over daily time-lags for sample* $\sigma_t^2 Z_t^2$ *(left plot) and* σ_t^2 *(right plot) assuming the innovations* $Z_t \sim t_5$*. Horisontal red dashed line represents value of theoretical tail index* $\alpha^* = 1.69$*.*