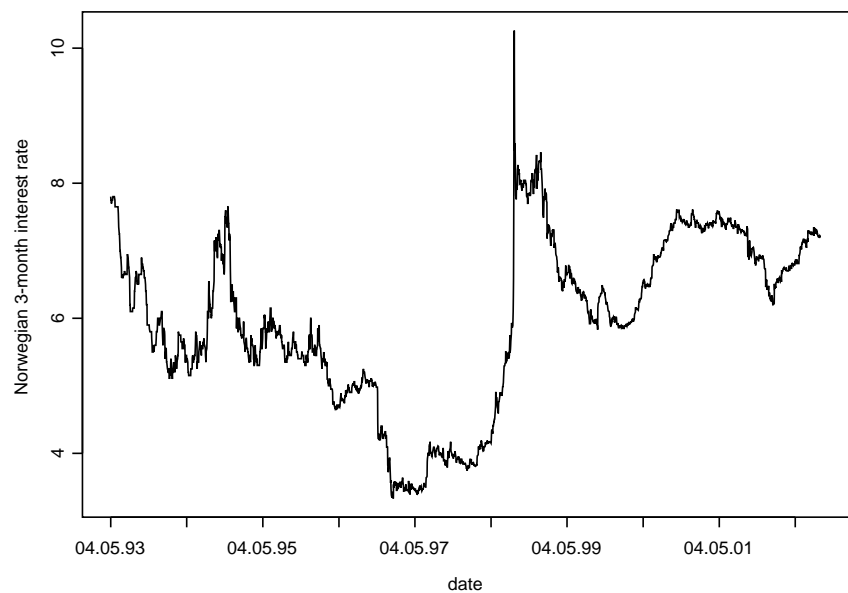


Statistical modelling of financial time series: An introduction



SAMBA/08/04

Kjersti Aas
Xeni K. Dimakos

8th March 2004

Title: Statistical modelling of financial time series: An introduction

Date: 8th March 2004

Year: 2004

Note no.: SAMBA/08/04

Author: Kjersti Aas <kjersti.aas@nr.no>
Xeni K. Dimakos <xeni.dimakos@nr.no>

Abstract: This note is intended as a summary of a one-day course in quantitative analysis of financial time series. It offers a guide to analysing and modelling financial time series using statistical methods, and is intended for researchers and practitioners in the finance industry.

Our aim is to provide academic answers to questions that are important for practitioners. The field of financial econometrics has exploded over the last decade. The intention of this course is to help practitioners cut through the vast literature on financial time series models, focusing on the most important and useful theoretical concepts.

Keywords: Statistical modelling, time series, stationarity, GARCH, correlation, copula

Target group: STAR

Availability: Open

Project: GB-BFF

Project no.: 220194

Research field: Finance, insurance and power market

No. of pages: 37

Contents

Introduction	1
1 Concepts of time series	2
1.1 Arithmetic and geometric returns	3
1.2 Aspects of time	5
2 Models	7
2.1 Random walk model	7
2.2 Autoregressive models	8
2.3 Stationarity	9
2.4 Autocorrelation function	10
3 Volatility	12
3.1 GARCH-models	13
3.2 Estimation of GARCH-models	15
3.3 Goodness-of-fit of a GARCH model	16
4 Return distributions	18
4.1 Marginal and conditional return distribution	18
4.2 Testing for normality	19
4.3 The scaled Students-t distribution	20
4.4 The Normal Inverse Gaussian (NIG) distribution	20
4.5 Extreme value theory	22
5 Multivariate time series	23
5.1 Vector Autoregression models	23
5.2 Multivariate GARCH models	24
5.2.1 The Diagonal-Vec model	24
5.2.2 The Constant Conditional Correlation model	24
5.3 Time-varying correlation	24
5.4 Multivariate return distributions	25
5.4.1 The multivariate Gaussian distribution	26
5.4.2 The multivariate normal inverse Gaussian (NIG) distribution	26
5.5 Copulas	27

5.5.1	Definition	27
5.5.2	Parametric copulas	27
6	Applications	30
6.1	Prediction	30
6.1.1	Forecasting with an AR(1)-model	30
6.1.2	Forecasting with a GARCH(1,1)-model	31
6.2	Risk management	32
6.2.1	Simulations	32
	References	34

Introduction

Financial time series are continually brought to our attention. Daily news reports in newspapers, on television and radio inform us for instance of the latest stock market index values, currency exchange rates, electricity prices, and interest rates. It is often desirable to monitor price behaviour frequently and to try to understand the probable development of the prices in the future. Private and corporate investors, businessmen, anyone involved in international trade and the brokers and analysts who advise these people can all benefit from a deeper understanding of price behaviour. Many traders deal with the risks associated with changes in prices. These risks can frequently be summarised by the variances of future returns, directly, or by their relationship with relevant covariances in a portfolio context. Forecasts of future standard deviations can provide up-to-date indications of risk, which might be used to avoid unacceptable risks perhaps by hedging.

There are two main objectives of investigating financial time series. First, it is important to understand how prices behave. The variance of the time series is particularly relevant. Tomorrow's price is uncertain and it must therefore be described by a probability distribution. This means that statistical methods are the natural way to investigate prices. Usually one builds a model, which is a detailed description of how successive prices are determined.

The second objective is to use our knowledge of price behaviour to reduce risk or take better decisions. Time series models may for instance be used for forecasting, option pricing and risk management.

Concepts of time series

A time series $\{Y_t\}$ is a discrete time, continuous state, process where $t; t = 1, \dots, T$ are certain discrete time points. Usually time is taken at equally spaced intervals, and the time increment may be everything from seconds to years. Figure 1.1 shows nearly 2500 consecutive daily values of the Norwegian 3-month interest rate, covering the period May, 4th, 1993 to September, 2nd, 2002.

[lai suat](#)

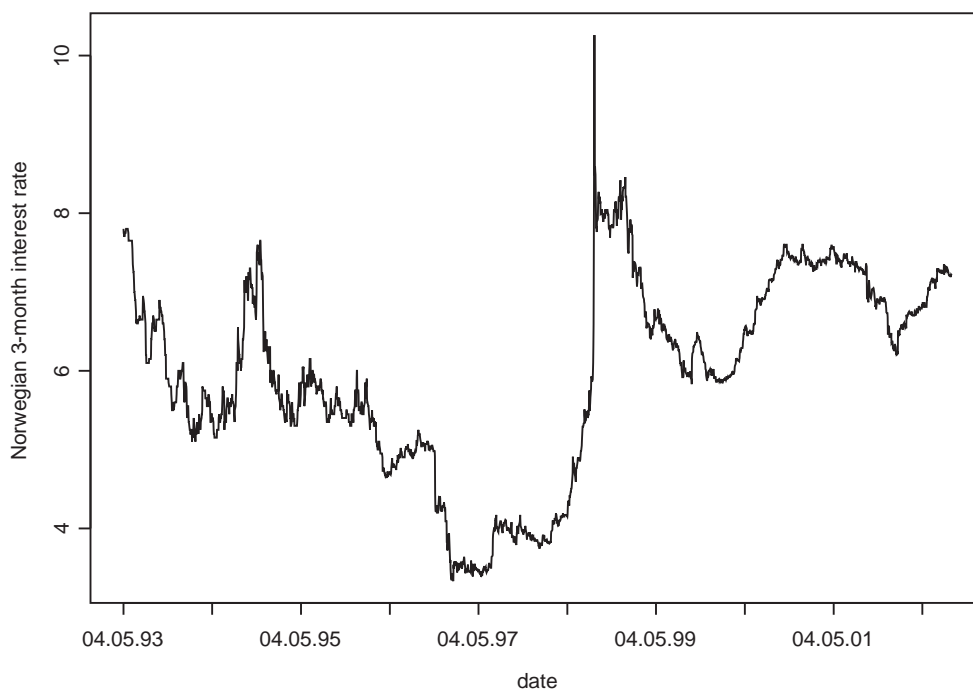


Figure 1.1: *NIBOR 3-month interest rates.*

1.1 Arithmetic and geometric returns

Loi nhuan

Nhung gia lien tiep tuong quan vs nhau rat cao.

Direct statistical analysis of financial prices is difficult, because consecutive prices are highly correlated, and the variances of prices often increase with time. This makes it usually more convenient to analyse changes in prices. Results for changes can easily be used to give appropriate results for prices. Two main types of price changes are used: *arithmetic* and *geometric* returns (Jorion, 1997). There seems to be some confusion about the two terms, in the literature as well as among practitioners. The aim of this section is to explain the difference.

Daily arithmetic returns are defined by

$$r_t = (Y_t - Y_{t-1})/Y_{t-1},$$

where Y_t is the price of the asset at day t . Yearly arithmetic returns are defined by

$$R = (Y_T - Y_0)/Y_0,$$

where Y_0 and Y_T are the prices of the asset at the first and last trading day of the year respectively. We have that R may be written as

$$R = \frac{Y_T}{Y_0} - 1 = \frac{Y_T}{Y_{T-1}} \frac{Y_{T-1}}{Y_{T-2}} \cdots \frac{Y_1}{Y_0} - 1 = \prod_{t=1}^T \frac{Y_t}{Y_{t-1}} - 1,$$

that is, it is not possible to describe the yearly arithmetic return as a function, or a sum of, daily arithmetic returns!

Daily geometric returns are defined by

$$d_t = \log(Y_t) - \log(Y_{t-1}),$$

while yearly geometric returns are given by

$$D = \log(Y_T) - \log(Y_0).$$

We have that D may be written as

$$D = \log \left(\prod_{t=1}^T \frac{Y_t}{Y_{t-1}} \right) = \sum_{t=1}^T \log \left(\frac{Y_t}{Y_{t-1}} \right) = \sum_{t=1}^T d_t,$$

which means that yearly geometric returns are equal to the sum of daily geometric returns.

In addition to the fact that compounded geometric returns are given as sums of geometric returns, there is another advantage of working with the log-scale. **If the geometric returns are normally distributed, the prices will never be negative. In contrast, assuming that arithmetic returns are normally distributed may lead to negative prices,** which is economically meaningless.

The relationship between geometric and arithmetic returns is given by

$$D = \log(1 + R).$$

Hence, D can be decomposed into a Taylor series as

$$D = R + \frac{1}{2} R^2 + \frac{1}{3} R^3 + \cdots,$$

which simplifies to R if R is small. Thus, when arithmetic returns are small, there will be little difference between geometric and arithmetic returns. In practice, this means that if the volatility of a price series is small, and the time resolution is high, geometric and arithmetic returns are quite similar, but when volatility increases and the time resolution decreases, the difference grows larger.

Figure 1.2 shows historical arithmetic and geometric annual returns for the Norwegian and American stock market during the period 1970 to 2002. For the American market, all the points lie on a straight line, showing that the difference between the arithmetic and geometric returns is quite small. For the Norwegian market, however, there is significant deviation from the straight line, indicating greater differences. This can be explained by the historical annual volatilities in the two markets, which have been 18% and 44% for the American and Norwegian market, respectively, during this time period.

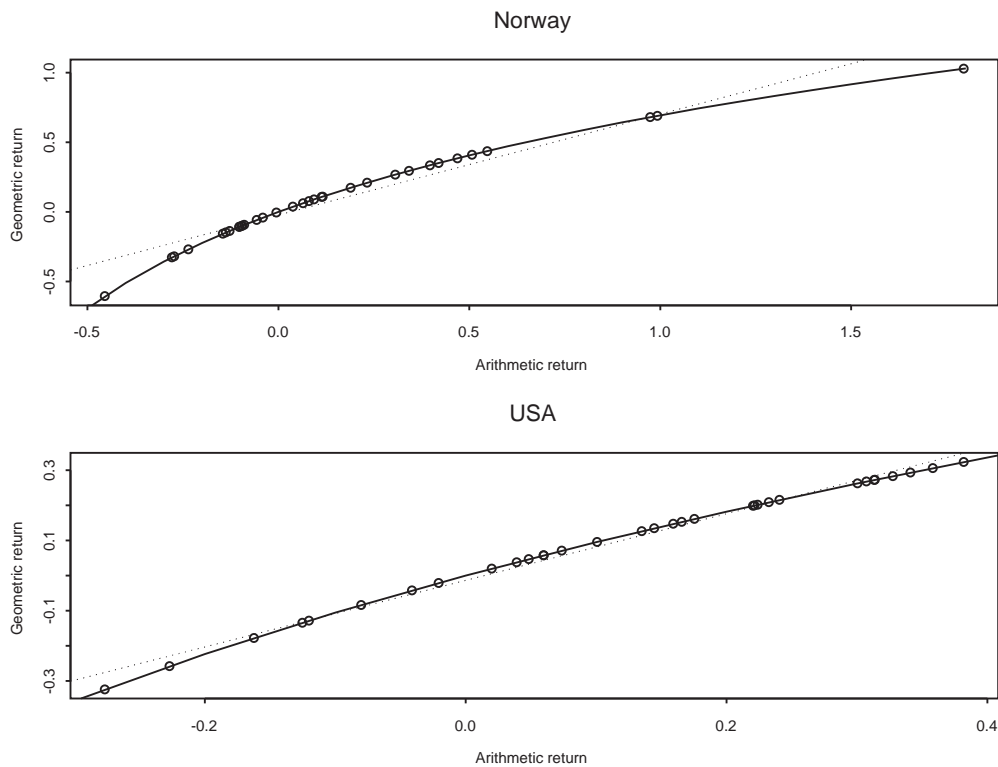


Figure 1.2: Arithmetic and geometric annual returns for the Norwegian and American stock market during the time period 1970 to 2002.

It is very common to assume that geometric returns are normally distributed on all time resolutions, Black & Scholes (Black and Scholes, 1973) formula for option pricing is for instance based on this assumption. If geometric returns are normally distributed, it follows from the relationship between arithmetic and geometric returns specified above, that arithmetic returns follow a lognormal distribution. The famous Markowitz portfolio theory (Markowitz, 1952), use the fact that the return of a portfolio can be written as a weighted average of component asset returns. It can be shown that this implies that arithmetic returns are used (the weighted average of geometric returns is not equal to the geometric return of a portfolio). Markovitz framework (commonly denoted the mean-variance approach) further assumes that

the risk of the portfolio can be totally described by its variance. This would be true if the arithmetic returns were multivariate normally distributed, but not necessarily otherwise. If the arithmetic returns are lognormally distributed, as implied by the Black & Scholes formula, and dependent, we don't even have an explicit formula for the distribution of their weighted sum.

1.2 Aspects of time

In practice, financial models will be influenced by time, both by **time resolution** and **time horizon**. The concept of resolution signifies how densely data are recorded. In applications in the finance industry, this might vary from seconds to years. The finer the resolution, the heavier the tails of the return distribution are likely to be. For intra-daily, daily or weekly data, failure to account for the heavy-tailed characteristics of the financial time series will undoubtedly lead to an underestimation of portfolio Value-at-Risk (VaR). Hence, market risk analysis over short horizons should consider heavy-tailed distributions of market returns. For longer time periods, however, many smaller contributions would average out and approach the normal as the lag ahead expands¹. This is illustrated by Figure 1.3, which shows the distributions of daily and monthly geometric returns for the time series in Figure 1.1. As can be seen from the figure, the first distribution is peaked and heavy-tailed, while the other is closer to the normal.

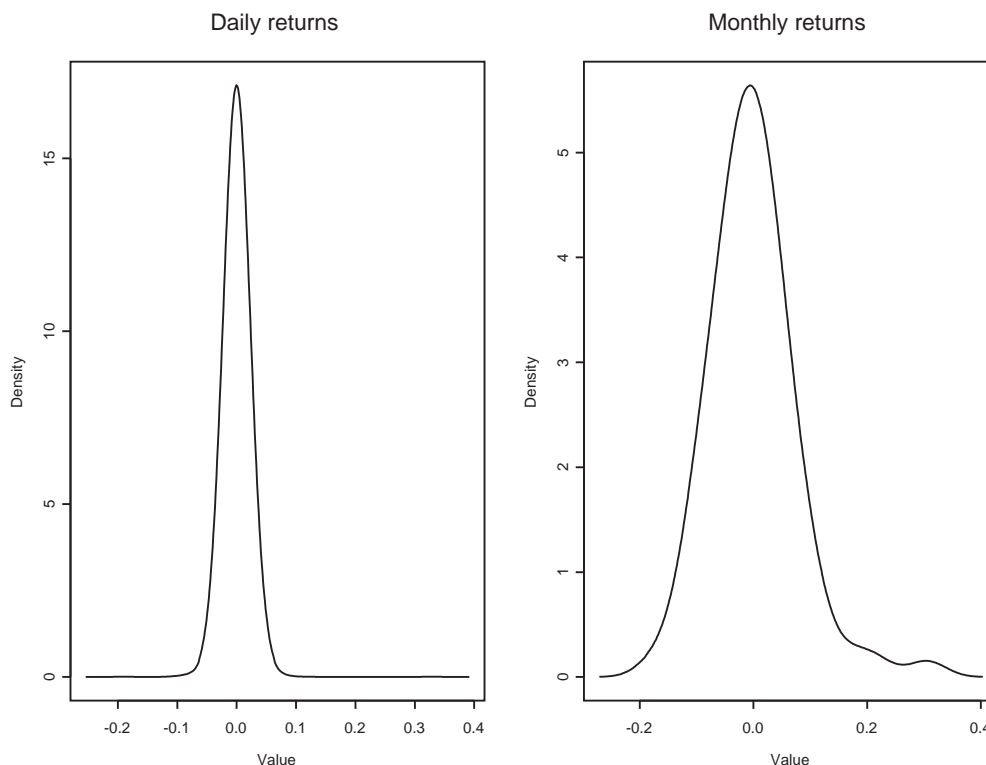


Figure 1.3: *Daily and monthly geometric returns for the time series in Figure 1.1.*

¹If the daily return distribution has very heavy tails, and/or the daily returns are dependent, the convergence to the normal distribution might be quite slow (Bradley and Taqqu, 2003)

It is also important to employ a statistical volatility or correlation model that is consistent with the horizon of the forecast/risk analysis. To forecast a long-term average volatility, it makes little sense (it may actually give misleading results) to use a high-frequency time-varying volatility model. On the other hand, little information about short-term variations in daily volatility would be forthcoming from a long-term moving average volatility model. As the time horizon increases, however, one encounters a problem with too few historical observations for estimating the model. One then might have to estimate the model using data with a higher resolution and aggregate this model to the correct resolution.

Financial prices are determined by many political, corporate, and individual decisions. A *model* for prices is a detailed description of how successive prices are determined. A good model is capable of providing simulated prices that behave like real prices. Thus, it should describe the most important of the known properties of recorded prices. In this course we will discuss two very common models, *the random walk model* and the *autoregressive model*.

2.1 Random walk model

A commonly used model in finance is the random walk, defined through

$$Y_t = \mu + Y_{t-1} + \epsilon_t,$$

bien dong

where μ is the drift of the process and the increments $\epsilon_1, \epsilon_2, \dots$ are serially independent random variables. Usually one requires that the sequence $\{\epsilon_t\}$ is identically distributed with mean zero and variance σ^2 , but this is not a necessary assumption. The variance of the process at time t is given by

$$\text{Var}(Y_t) = t \sigma^2,$$

i.e. it increases linearly with time. Figure 2.1 shows a simulated random walk where $\mu = 0.00022$ and $\sigma = 0.013$.

In finance, the random walk model is commonly used for equities, and it is usually assumed that it is the geometric returns of the time series that follows this model. As we will discuss in Lesson 3, **the variance σ might be dependent of the time t** . The assumption of serially independent increments of the series can be motivated as follows. If there were correlation between different epochs, smart investors could bet on it and beat the market. However, in the process they would then destroy the basis for their own investment strategy, and drive the correlations they utilised back to zero. Hence, the (geometric) random walk model assumes that it at a given moment is impossible to estimate where in the business cycle the economy is, and utilise such knowledge for investment purposes.

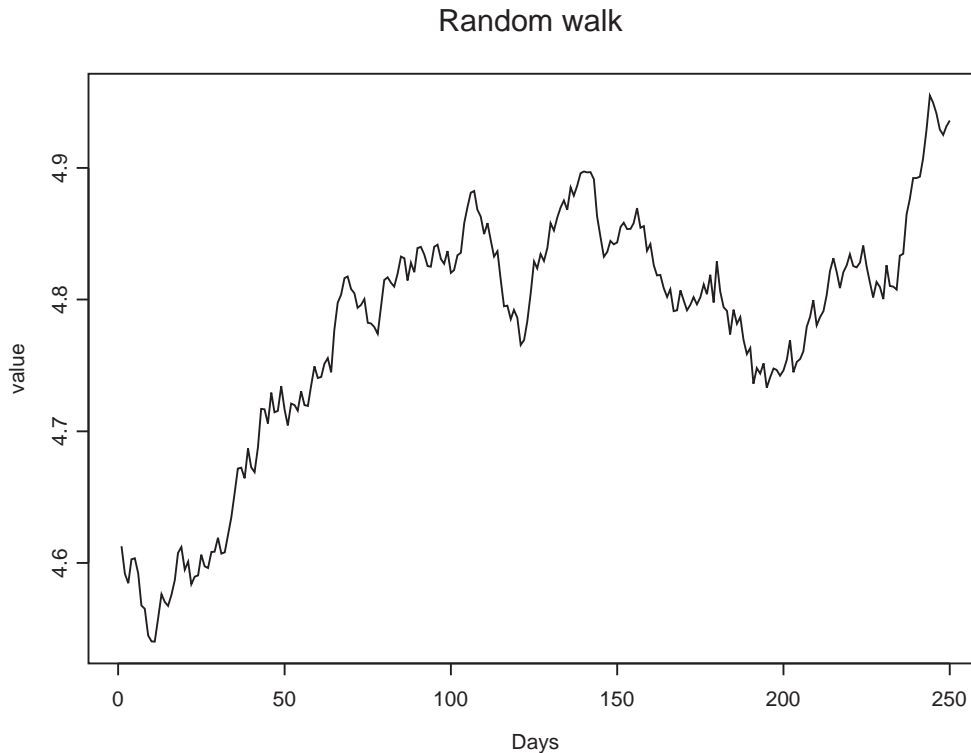


Figure 2.1: *Simulated random walk with $\mu = 0.00022$ and $\sigma = 0.013$.*

2.2 Autoregressive models

Random walk models cannot be used for all financial time series. Interest rates, for instance, are influenced by complicated political factors that make them difficult to describe mathematically. However, if a description is called for, the class of autoregressive models is a useful candidate. We shall only discuss the simplest first order case, the AR(1)-model:

$$Y_t = \mu + \alpha Y_{t-1} + \epsilon_t,$$

where $|\alpha| < 1$ is a parameter and $\epsilon_1, \epsilon_2, \dots$ are serially independent random variables. As for the random walk model we assume that the random terms have mean 0 and variance σ^2 . We are back to the random walk model if $\alpha = 1$. This autoregressive process is important, because it provides a simple description of the stochastic nature of interest rates that is consistent with the empirical observation that interest rates tend to be mean-reverting. The parameter α determines the speed of mean-reversion towards the stationary value $\mu/(1 - \alpha)$. Situations where current interest rates are high, imply a negative drift, until rates revert to the long-run value, and low current rates are associated with positive expected drift. The stationary variance of the process is given by

$$\text{Var}(Y_t) = \frac{\sigma^2}{1 - \alpha^2}.$$

To avoid negative interest rates, one usually models the logarithm of the rate rather than the original value. Figure 2.2 shows a simulated AR(1)-process, where $\alpha = 0.95$, $\mu = 0.00022$ and $\sigma = 0.013$. The dotted line represents the stationary level of the process.

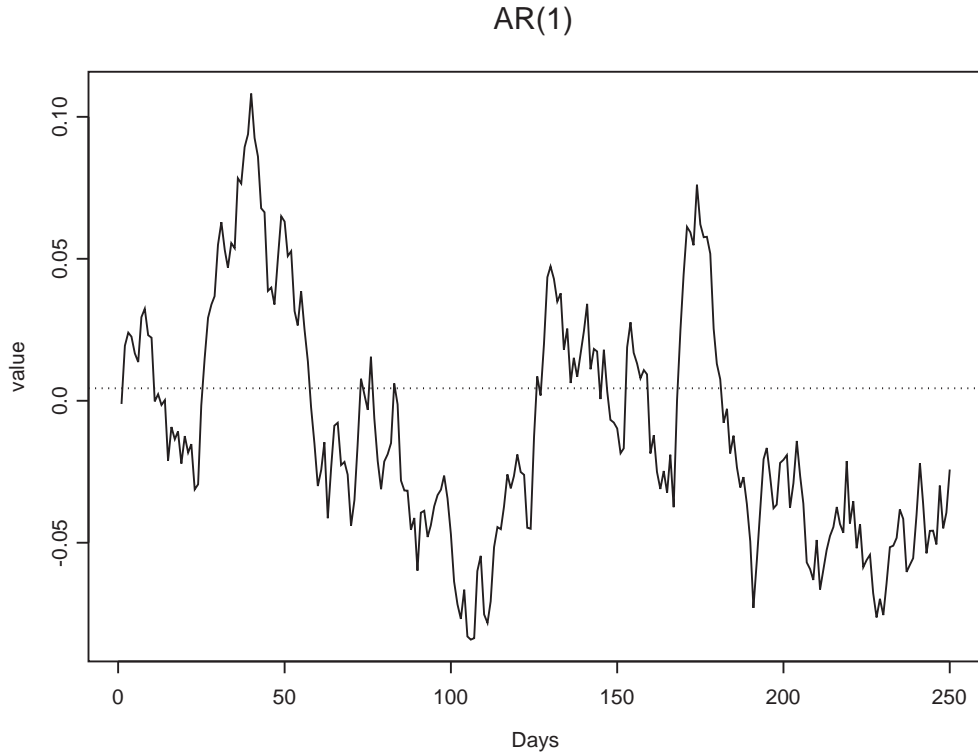


Figure 2.2: Simulated $AR(1)$ -process with $\alpha = 0.95$, $\mu = 0.00022$ and $\sigma = 0.013$.

As in the random walk model, it is possible to let the volatility σ depend on time. A very common assumption for interest rates is to parameterise the volatility as a function of interest rate level (see Chan et al. (1992)), i.e

$$\sigma_t = \kappa Y_{t-1}^\gamma,$$

where κ is a parameter. If one sets $\gamma = 0$, one is back to the $AR(1)$ -model, or the Ornstein-Uhlenbeck process (Vasicek, 1977) in the continuous case. Setting $\gamma = 1/2$ gives the well-known CIR-model (Cox et al., 1985). A different class of models to capture volatility dynamics is the family of GARCH-models that we will describe in Lesson 3.

2.3 Stationarity

A sequence of random variables $\{X_t\}$ is covariance stationary if there is no trend, and if the covariance does not change over time, that is

$$E[X_t] = \mu \text{ for all } t$$

and

$$\text{Cov}(X_t, X_{t-k}) = E[(X_t - \mu)(X_{t-k} - \mu)] = \gamma_k \text{ for all } t \text{ and any } k.$$

The increments $\epsilon_1, \epsilon_2, \dots$ of a random walk model and those of an $AR(1)$ -model are both stationary. Moreover, a process that follows an $AR(1)$ -model is itself stationary, while the random walk process is not, since its variance increases linearly with time.

It is possible to formally test whether a time series is stationary or not. Statistical tests of the null hypothesis that a time series is non-stationary against the alternative that it is stationary are called unit root tests. One such test is the Dickey-Fuller test (Alexander, 2001). In this test one rewrites the AR(1) model to

$$Y_t - Y_{t-1} = \mu + (\alpha - 1) Y_{t-1} + \epsilon_t,$$

and simply performs a regression of $Y_t - Y_{t-1}$ on a constant and Y_{t-1} and tests whether the coefficient of Y_{t-1} is significantly different from zero¹.

It should be noted that tests for stationarity are not very reliable, and we advise against trusting the results of such tests too much. Be aware that the tests are typically derived under the assumption of constant variance. As will be shown in Lesson 3 it is widely assumed that the volatility of financial time series follows some time-varying function. Hence, this may cause distortions in the performance of the conventional tests for unit root stationarity.

2.4 Autocorrelation function

Assume that we have a stationary time series $\{X_t\}$ with constant expectation and time independent covariance. The autocorrelation function (ACF) for this series is defined as

$$\begin{aligned} \rho_k &= \frac{\text{Cov}(X_t, X_{t-k})}{\sqrt{\text{Var}(X_t) \text{Var}(X_{t-k})}} \\ &= \frac{\gamma_k}{\gamma_0}, \end{aligned}$$

for $k \geq 0$ and

$$\rho_{-k} = \rho_k.$$

The value k denotes the lag. By plotting the autocorrelation function as a function of k , we can determine if the autocorrelation decreases as the lag gets larger, or if there is any particular lag for which the autocorrelation is large.

For a non-stationary time series, $\{Y_t\}$, the covariance is *not* independent of t , but given by

$$\text{Cov}(Y_t, Y_{t-k}) = (t - k) \sigma^2.$$

This means that the autocorrelation at time t and lag k is given by

$$\begin{aligned} \rho_{k,t} &= \frac{(t - k) \sigma^2}{\sqrt{t \sigma^2 (t - k) \sigma^2}} \\ &= \sqrt{\frac{t - k}{t}}. \end{aligned}$$

We see that if t is large relative to k , then $\rho_{k,t} \approx 1$.

The following general properties provide hints of the structure of the underlying process of a financial time series²:

¹It can be shown that standard t-ratios of $(\alpha - 1)$ to its estimated standard error does not have a Student's t distribution and that the appropriate critical values have to be increased by an amount that depends on the sample size.

²It should be noted that if the variance of the time series is time-varying, this will influence the autocorrelation function, and the “rules” given here are not necessarily applicable.

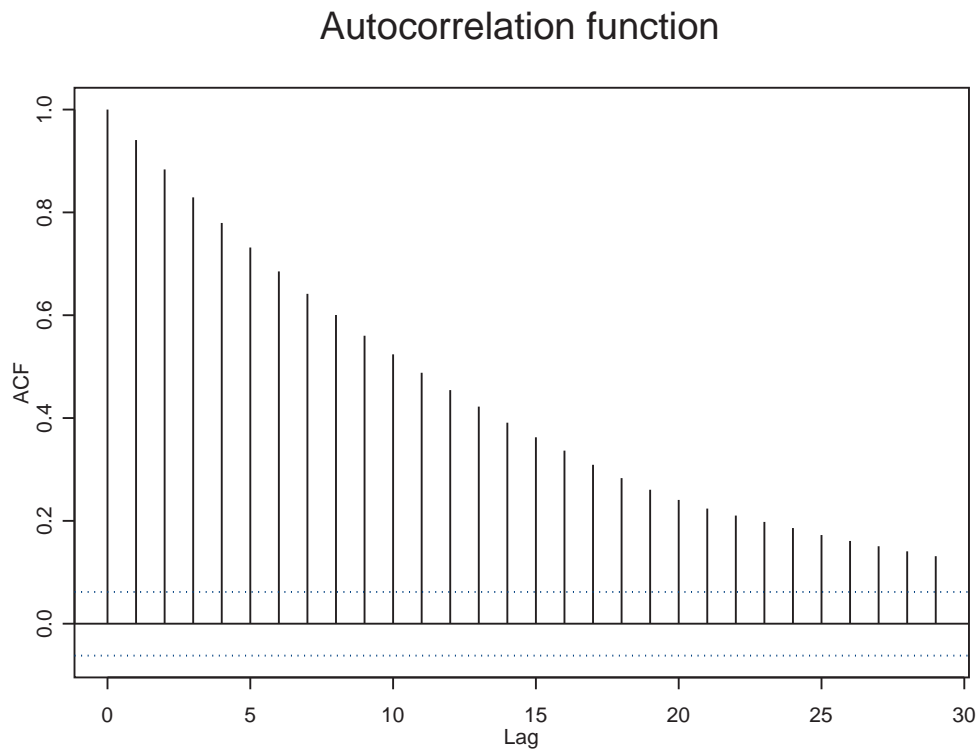


Figure 2.3: *The autocorrelation function of the AR(1)-process in Figure 2.2.*

- For both the random walk model and the AR(1)-model, the autocorrelation function for the increments $\epsilon_1, \epsilon_2, \dots$ should be zero for any lag $k > 0$.
- For the AR(1)-model, the autocorrelation function for Y_t should be equal to α^k for lag k (a geometric decline).
- For a random walk model, the autocorrelation function for Y_t is likely to be close to 1 for all lags.

Figure 2.3 shows the first 30 lags of the autocorrelation function for the simulated AR(1)-process in Figure 2.2.

Volatility

Returns from financial market variables measured over short time intervals (i.e. intra-daily, daily, or weekly) are uncorrelated, but not independent. In particular, it has been observed that although the signs of successive price movements seem to be independent, their magnitude, as represented by the absolute value or square of the price increments, is correlated in time. This phenomena is denoted *volatility clustering*, and indicates that the volatility of the series is time varying. Small changes in the price tend to be followed by small changes, and large changes by large ones. A typical example is shown in Figure 3.1, where the Norwegian stock market index (TOTX) during the period January 4th, 1983 to August, 26th, 2002 is plotted together with the geometric returns of this series.

Since volatility clustering implies a strong autocorrelation in the absolute values of returns, a simple method for detecting volatility clustering is calculating the autocorrelation function (defined in Section 2.4) of the absolute values of returns. If the volatility is clustered, the autocorrelation function will have positive values for a relatively large number of lags. As can be seen from Figure 3.2, this is the case for the returns in Figure 3.1.

The issue of modelling returns accounting for time-varying volatility has been widely analysed in financial econometrics literature. Two main types of techniques have been used: Generalised Autoregressive Conditional Heteroscedasticity (GARCH)-models (Bollerslev, 1986) and stochastic volatility models (Aquilari and West, 2000; Kim et al., 1998). The success of the GARCH-models at capturing volatility clustering in financial markets is extensively documented in the literature. Recent surveys are given in Ghysels et al. (1996) and Shepard (1996). Stochastic volatility models are more sophisticated than the GARCH-models, and from a theoretical point of view they might be more appropriate to represent the behaviour of the returns in financial markets. The main drawback of the stochastic volatility models is, however, that estimating them is a statistically and computationally demanding task. This has prevented their wide-spread use in empirical applications and they are little used in the industry compared to the GARCH-models. In this course we will therefore concentrate on the GARCH-models.

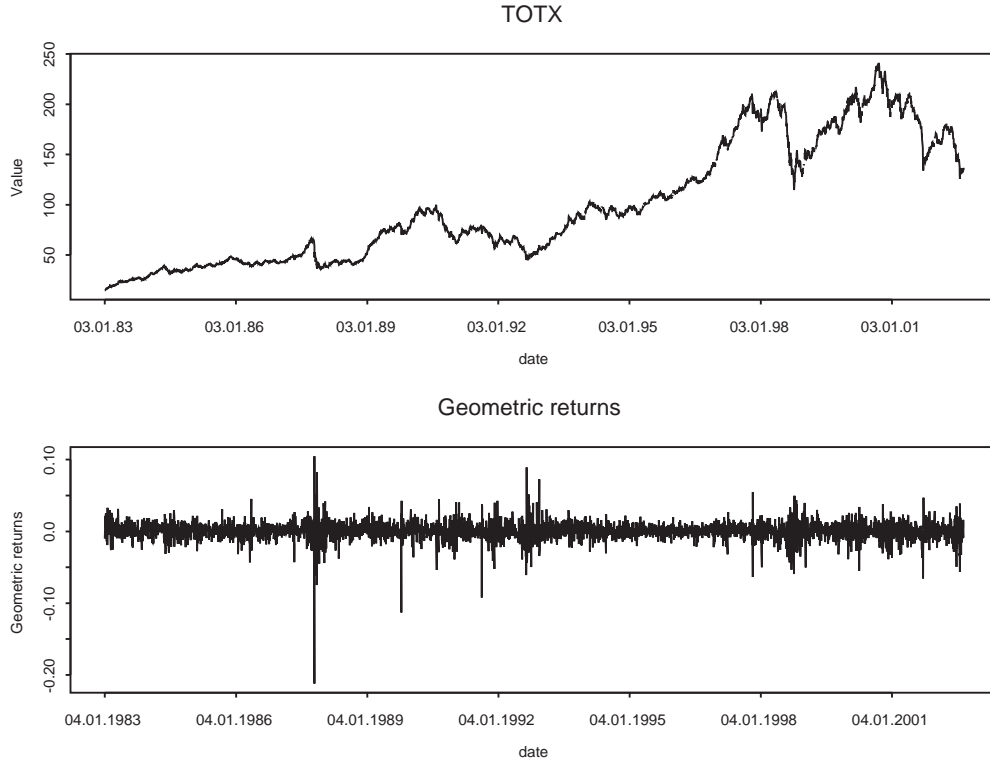


Figure 3.1: The Norwegian stock market index (TOTX) during the period January 4th, 1983 to August, 26th, 2002. Upper panel: Original price series. Lower panel: Geometric returns.

3.1 GARCH-models

Let

$$Y_t = \mu + \alpha Y_{t-1} + \epsilon_t$$

$$\epsilon_t \sim N(0, \sigma_t^2)$$

where $\epsilon_t, t = 1, \dots$, are serially independent. If $\alpha = 1$ then $\{Y_t\}$ follows a random walk model and an AR(1)-model otherwise. The fundamental idea of the GARCH(1,1)-model (Bollerslev, 1986) is to describe the evolution of the variance σ_t^2 as

$$\sigma_t^2 = a_0 + a \epsilon_{t-1}^2 + b \sigma_{t-1}^2. \quad (3.2)$$

The parameters satisfy $0 \leq a \leq 1$, $0 \leq b \leq 1$, and $a + b \leq 1$. The variance process is stationary if $a + b < 1$, and the stationary variance is given by $a_0/(1 - a - b)$.

The parameter $\eta = a + b$ is known as persistence and defines how slowly a shock in the market is forgotten. To see this, consider the expected size of the variance σ_k^2 k time units ahead, given the present value σ_0^2 . It turns out to be

$$E(\sigma_k^2 | \sigma_0^2) = a_0 \frac{1 - \eta^k}{1 - \eta} + \eta^k \sigma_0^2.$$

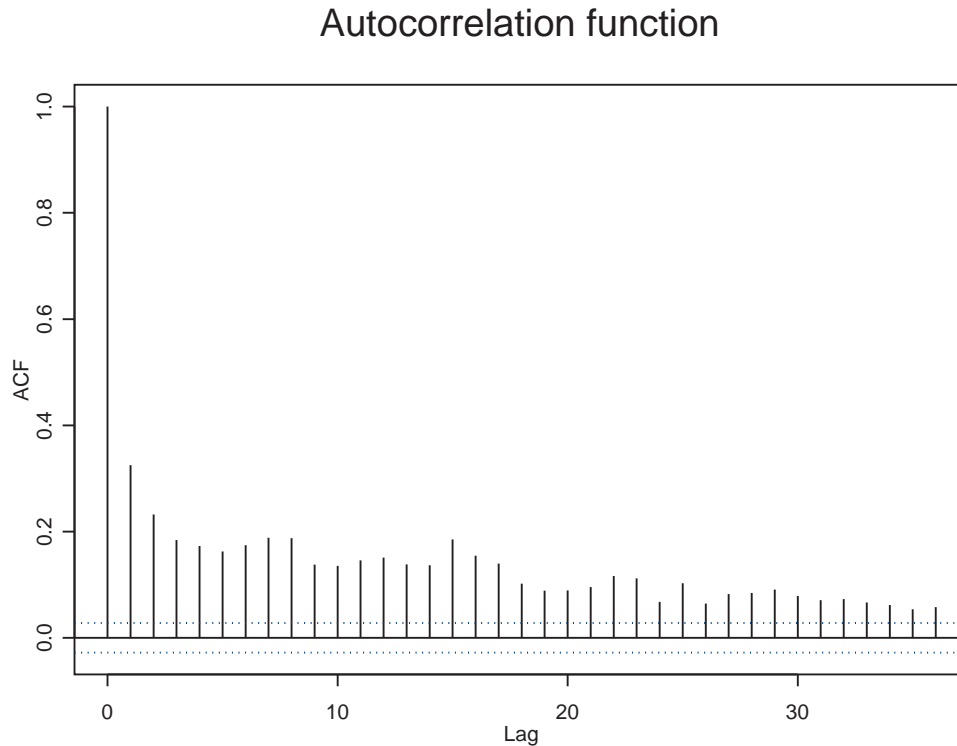


Figure 3.2: *The autocorrelation function of the absolute values of the geometric returns in Figure 3.1.*

As $k \rightarrow \infty$, the left-hand side approaches the stationary variance $a_0/(1 - \eta)$. The speed of convergence depends on the size of η . If η is close to one, it takes very long before the stationary level is reached.

A special case of the GARCH(1,1)-model arises when $a + b = 1$ and $a_0 = 0$. In this case it is common to use the symbol λ for b , and Equation 3.2 takes the simpler form

$$\sigma_t^2 = (1 - \lambda) \epsilon_{t-1}^2 + \lambda \sigma_{t-1}^2 = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i \epsilon_{t-i}^2,$$

which is known as the IGARCH-model Nelson (1990). The variance can in this case be interpreted as a weighted average of all past squared returns with the weights declining exponentially as we go further back. For this reason, this model is known as the Exponentially Weighted Moving Average (EWMA) model, and is the standard RiskMetrics model (Morgan, 1996). The problem with this model is that even though it is strictly stationary, the variance in the stationary distribution does not exist (Nelson, 1990).

Equation 3.2 is the simplest GARCH model, having only one lagged term for both ϵ and σ , and a Gaussian error distribution. More general models can be obtained by considering longer lag polynomials in ϵ and σ , and using non-normal distributions. We will discuss alternative distributions in Lesson 4.

While GARCH models for equities and exchange rates usually turn out to be stationary, one often gets $a + b > 1$ when fitting GARCH models to short-term interest rates. For instance,

Engle et al. (1990) estimates the sum to be 1.01 for U.S Treasury securities and Gray (1996) find the sum of the coefficients to be 1.03 for one-month T-bills. Several authors (Bali, 2003; Rodrigues and Rubia, 2003) claim that this behaviour is due to model misspecification, and that the specification in Equation 3.2 can not be used for modelling the volatility of short-term interest rates, because it fails to capture the relationship between interest rate levels and volatility described in Section 2.2. In their opinion, both the level and the GARCH effects have to be incorporated into the volatility process. During the last few years, several such models have been presented in the literature (Bali, 2003; Brenner et al., 1996; Koedijk et al., 1997; Longstaff and Schwartz, 1992). The presentation of such models is however outside the scope of this course.

3.2 Estimation of GARCH-models

The most common GARCH models are available as procedures in econometric packages such as SAS, S-PLUS and Matlab, and practitioners do not need to code up any program for estimating the parameters of a GARCH model. Interested readers may for instance see Alexander (2001) or Zivot and Wang (2003) for an overview of the most used estimation approaches. What the practitioner has to consider, however, is how one should choose the data period and time resolution to be used when estimating the parameters. We will focus on these issues here.

When estimating a GARCH-model, it will be a trade-off between having enough data for the parameter estimates to be stable, and too much data so that the forecast do not reflect the current market conditions. The parameter estimates of a GARCH-model depend on the historic data used to fit the model. A very long time series with several outliers is unlikely to be suitable, because extreme moves in the past can have great influence on the long-term volatility forecasts made today. Hence, in choosing the time period of historical data used for estimating a GARCH model, the first consideration is whether major market events from several years ago should be influencing forecasts today.

On the other hand, it is important to understand that a certain minimum amount of data is necessary to ensure proper convergence of the model, and to get robust parameter estimates. If the data period is too small, the estimates of the GARCH model might lack stability as the data window is slightly altered.

We have fitted the GARCH-model in Equation 3.2 to the geometric returns in Figure 3.1. The a and b parameters were estimated to be 0.19 and 0.77, respectively, and the stationary standard deviation to 0.0130. The estimated volatility is shown in Figure 3.3. The largest peak corresponds to Black Monday (October 1987). We also fitted the GARCH-model to only the last period of the geometric returns, from November, 26th, 1998 to August, 26th, 2002. The estimated a and b were then 0.18 and 0.78, respectively, and the asymptotic standard deviation 0.0135. Hence, for this data set, the estimated parameters does not vary much with the period selected.

As far as time resolution is concerned, it might be difficult to estimate GARCH models for low-frequency (i.e. monthly and yearly) data, even if this is the desired resolution, because the historical data material is limited. However, temporal aggregation may be used to estimate a low frequency model with high frequency data. Drost and Nijman (1993) give analytical expressions for the parameters in temporally aggregated GARCH models. They show that the volatility is asymptotically constant, meaning that the GARCH effects eventually disappear. Other authors warn that temporal aggregation must be used with care (see e.g. Meddahi and Renault (2004)), especially in the case of a near non-stationary high frequency process and a

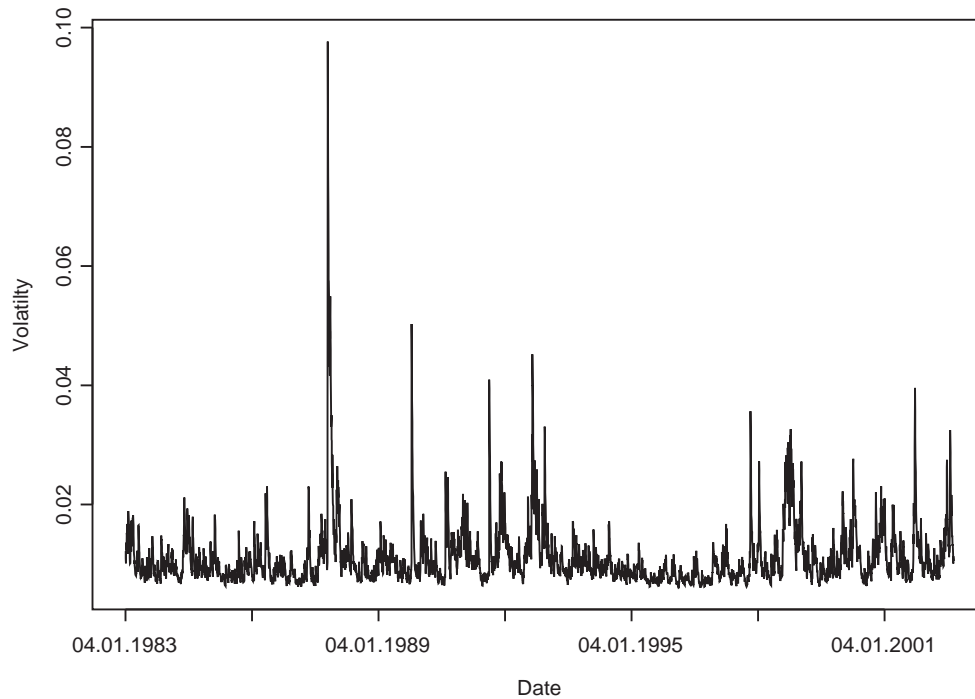


Figure 3.3: *The estimated volatility obtained when fitting the GARCH-model in Equation 3.2 to the geometric returns in Figure 3.1.*

large aggregation level.

3.3 Goodness-of-fit of a GARCH model

The goodness-of-fit of a GARCH-model is evaluated by checking the significance of the parameter estimates and measuring how well it models the volatility of the process. If a GARCH-model adequately captures volatility clustering, the absolute values of the *standardised returns*, given by

$$\epsilon_t^* = \hat{\epsilon}_t / \hat{\sigma}_t,$$

where $\hat{\sigma}_t$ is the estimate of the volatility, should have no autocorrelation. Such tests may be performed by graphical inspection of the autocorrelation function, or by more formal statistical tests, like the Ljung-Box statistic (Zivot and Wang, 2003). Figure 3.4 shows the autocorrelation function of the absolute values of the standardised returns obtained when fitting the GARCH-model in Equation 3.2 to the geometric returns in Figure 3.1. As can be seen from the figure, there is not much autocorrelation left when conditioning on the estimated volatility.

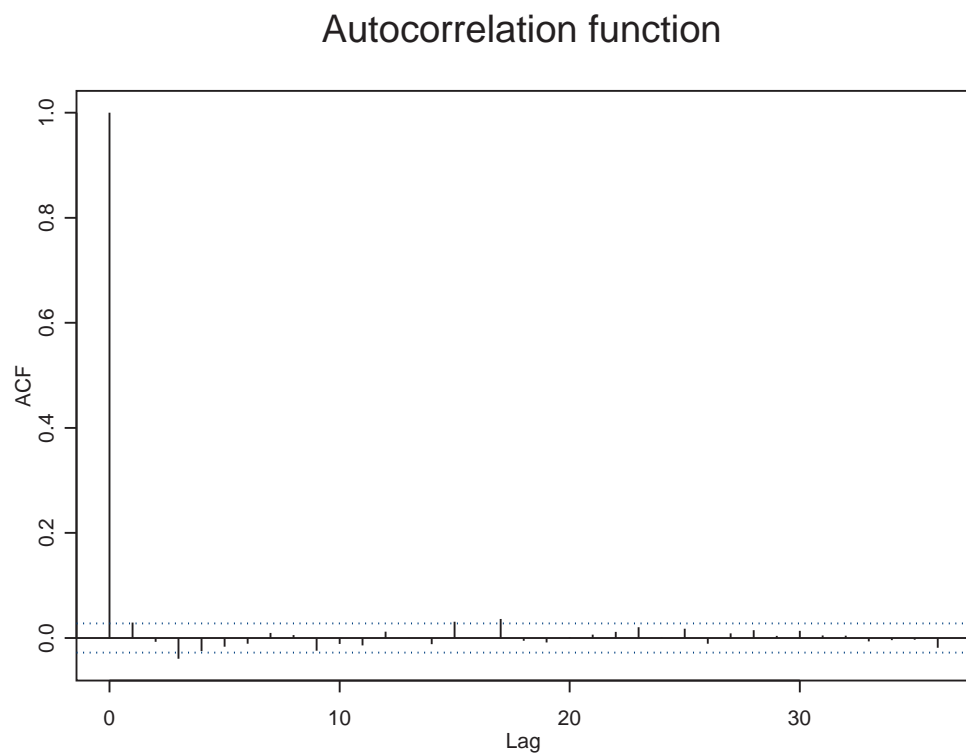


Figure 3.4: *The autocorrelation function of the absolute values of the standardised returns obtained when fitting the GARCH-model in Equation 3.2 to the geometric returns in Figure 3.1.*

Return distributions

4.1 Marginal and conditional return distribution

In risk management there are two types of return distributions to be considered, the *marginal* (or stationary) distribution of the returns, and the *conditional* return distribution. If we have the model

$$\begin{aligned} Y_t &= \mu + \alpha Y_{t-1} + \epsilon_t \\ \epsilon_t &\sim N(0, \sigma_t^2) \end{aligned} \tag{4.1}$$

where $\epsilon_t, t = 1, \dots$, are serially independent, the marginal distribution is the distribution of the errors ϵ_t , while the conditional distribution is the distribution of ϵ_t/σ_t .

The marginal distribution is known to be more heavy-tailed than the Gaussian distribution. If the variance of the time-series is assumed to follow a GARCH model, this implies that the marginal distribution for the returns has fatter tails than the Gaussian distribution, even though the conditional return distribution is Gaussian. According McNeil and Frey (2000) and Bradley and Taqqu (2003), the reason for this is that the GARCH(1,1)-process is a mixture of normal distributions with different variances. It can be shown analytically (Bradley and Taqqu, 2003) that the kurtosis implied by a GARCH-process with Gaussian conditional return distribution under certain conditions¹ is $6\alpha^2/(1 - \beta^2 - 2\alpha\beta - 3\alpha^2)$, which for most financial return series is greater than 0, the kurtosis of a normal distribution. Still, it is generally well recognised that GARCH-models coupled with conditionally normally distributed errors is unable to fully account for the tails of the marginal distributions of daily returns. Several alternative conditional distributions have therefore been proposed in the GARCH literature. In this section we will review some of the most important ones, but before that we will show how to test for normality.

¹The model must be stationary and $\beta^2 + 2\alpha\beta + 3\alpha^2 < 1$.

4.2 Testing for normality

It is crucial to recognise that not everything can be modelled or approximated by the normal distribution. The skillful practitioner will make it a habit to always examine if the data at hand is indeed normal or if other distributions are more suitable. There are several statistical methods that can be used to see if a data set is from the Gaussian distribution. The most common is the qq-plot, a scatter-plot of the empirical quantiles of the data against the quantiles of a standard normal random variable. If the data is normally distributed, then the quantiles will lie on a straight line. Figure 4.1 shows the qq-plot for the standardised returns obtained when fitting the GARCH-model in Equation 3.2 to the geometric returns in Figure 3.1. As can be seen from the figure, there is a significant deviation from the straight line in the tails, especially in the lower tail, showing that the distribution of the standardised returns is more heavy-tailed than the normal distribution.

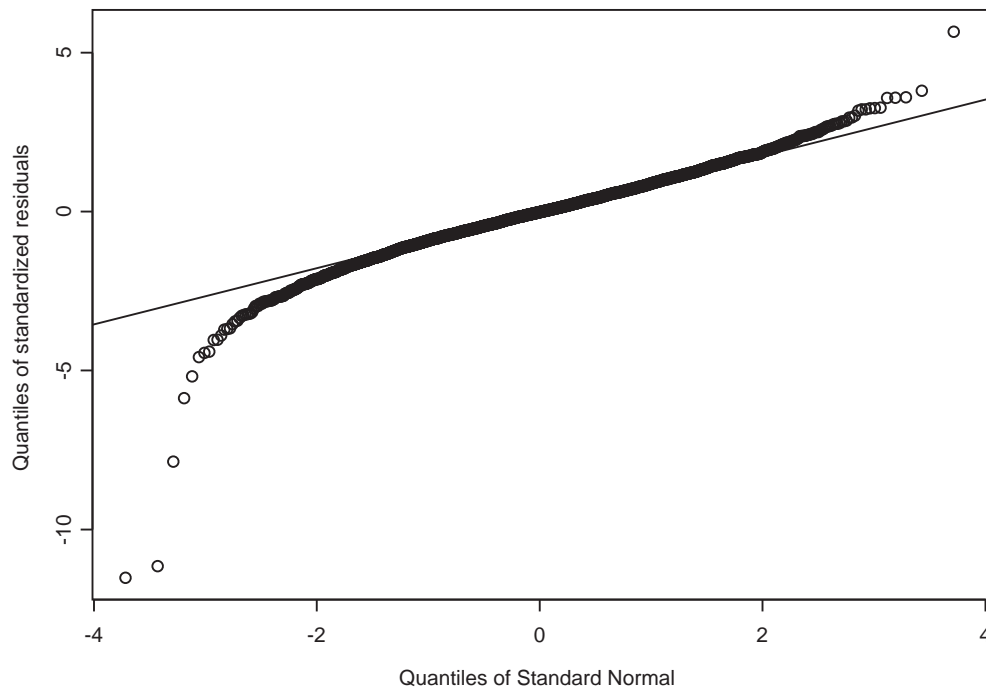


Figure 4.1: The qq-plot for the standardised returns obtained when fitting the GARCH-model in Equation 3.2 to the geometric returns of the Norwegian stock market index in Figure 3.1.

The qq-plot is an informal graphical diagnostic. In addition, most statistical software supply several formal statistical tests for normality, like the Shapiro-Wilk test (c.f. Zivot and Wang (2003)). The reliability of such tests is generally small. Even though they provide numerical measures of goodness of fit, which is seemingly more precise than inspecting qq-plots, it is our opinion that they are in fact no better.

4.3 The scaled Students-t distribution

The density function of the scaled Students-t-distribution is given by

$$f_x(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sigma \sqrt{\pi(\nu-2)}} \left[1 + \frac{(x-\mu)^2}{(\nu-2)\sigma^2}\right]^{-\left(\frac{\nu+1}{2}\right)}.$$

Here, μ and σ represent a location and a dispersion parameter respectively, and ν is the degrees of freedom parameter, controlling the heaviness of the tails. The mean and variance of this distribution are given by

$$E(X) = \mu \quad \text{and} \quad \text{Var}(X) = \sigma^2.$$

This distribution is also denoted a Pearson's type VII distribution and has been used as the conditional distribution for GARCH by Bollerslev (1987), among others.

The conditional distribution of financial returns is usually heavy-tailed, but it is also skewed, with the left tail heavier than the right. The scaled Students-t distribution allows for heavy tails, but is symmetric around the expectation. A better choice would be a distribution that allows for skewness in addition to heavy tails, and in Section 4.4 we present one such distribution.

4.4 The Normal Inverse Gaussian (NIG) distribution

The normal inverse Gaussian (NIG) distribution was introduced as the conditional distribution for returns from financial markets by Barndorff-Nielsen (1997). Since then, applications in finance have been reported in several papers, both for the conditional distribution (Andersson, 2001; Forsberg and Bollerslev, 2002; Jensen and Lunde, 2001; Venter and de Jongh, 2002) and the marginal distribution (Bølviken and Benth, 2000; Eberlein and Keller, 1995; Lillestøl, 2000; Prause, 1997; Rydberg, 1997).

The univariate NIG distribution can be parameterised in several ways. We follow Karlis (2002) and Venter and de Jongh (2002) and let

$$f_x(x) = \frac{\delta \alpha}{\pi q(x)} \exp(p(x)) K_1(\alpha q(x)),$$

where

$$p(x) = \delta \sqrt{\alpha^2 - \beta^2} + \beta(x - \mu),$$

$$q(x) = \sqrt{\delta^2 + (x - \mu)^2},$$

$\delta > 0$ and $0 < |\beta| < \alpha$. K_1 is the modified Bessel function of the second kind of order 1 (Abramowitz and Stegun, 1972). The parameters μ and δ determine the location and scale respectively, while α and β control the shape of the density. In particular $\beta = 0$ corresponds to a symmetric distribution. The parameter given by

$$\xi = \left(1 + \delta \sqrt{\alpha^2 - \beta^2}\right)^{-1/2}, \quad (4.2)$$

determines the heaviness of the tails. The closer ξ is to 1, the heavier are the tails (the limit is the Cauchy distribution), and when $\xi \rightarrow 0$, the distribution approaches the normal distribution. The parameter

$$\chi = \beta\xi/\alpha,$$

determines the skewness of the distribution. For $\chi < 0$, the left tail is heavier than the right, for $\chi = 0$ the distribution is symmetric, and for $\chi > 0$ the right tail is the heaviest. Due to the restrictions on δ , β , and α , we have that $0 \leq |\chi| < \xi < 1$. The expectation and variance of the NIG distribution are given by

$$E(X) = \mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} \quad \text{and} \quad \text{Var}(X) = \frac{\delta\alpha^2}{(\alpha^2 - \beta^2)^{3/2}}.$$

We fitted the NIG-distribution to the standardised returns (see Section 3.3) obtained when fitting the GARCH-model in Equation 3.2 to the geometric returns in Figure 3.1. The result is shown in Figure 4.2. The Gaussian distribution is superimposed. As can be seen from the figure, the NIG-distribution is more peaked and heavy-tailed than the Gaussian distribution (the estimated NIG-parameters were $\mu = 0.11$, $\alpha = 1.43$, $\beta = -0.15$, $\delta = 1.38$).

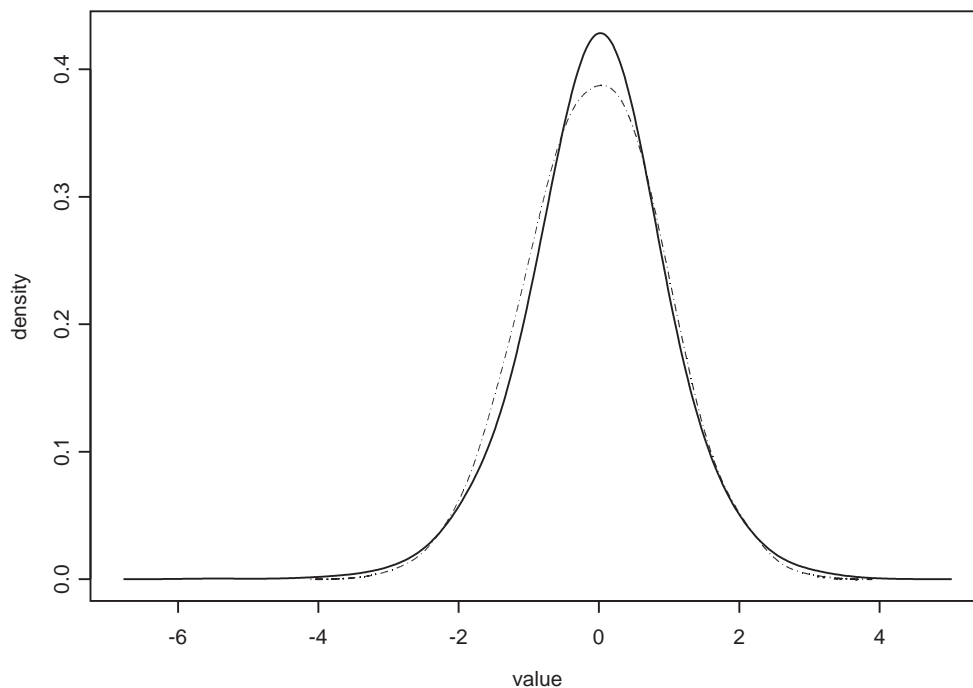


Figure 4.2: *The density of the NIG-distribution fitted to the standardised returns obtained when fitting the GARCH-model in Equation 3.2 to the geometric returns of the Norwegian stock market index in Figure 3.1 (solid line), with the Gaussian distribution superimposed (dotted line).*

4.5 Extreme value theory

Traditional parametric and non-parametric methods for estimating distributions and densities work well when there are many observations, but they give poor fit when the data is sparse, such as in the extreme tails of the distribution. This result is particularly troubling because risk management often requires estimating quantiles and tail probabilities beyond those observed in the data. The methods of extreme value theory (Embrechts et al., 1997) focus on modelling the tail behaviour using extreme values beyond a certain threshold rather than all the data. These methods have two features, which make them attractive for tail estimation; they are based on sound statistical theory, and they offer a parametric form for the tail of a distribution.

The idea is to model tails and the central part of the empirical distribution by different kinds of (parametric) distributions. One introduces a threshold u , and consider the probability distribution of the returns, given that they exceed u . It can be shown (c.f. Embrechts et al. (1997)) that when u is large enough, this distribution is well approximated by the generalised Pareto distribution (GPD) with density function

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

where $\beta > 0$. The distribution is defined for $x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ when $\xi < 0$. The case $\xi > 0$ corresponds to heavy-tailed distributions whose tails decay like power functions, such as the Pareto, Student's t and the Cauchy distribution. The case $\xi = 0$ corresponds to distributions like the normal, exponential, gamma and lognormal, whose tails essentially decay exponentially. Finally, the case $\xi < 0$ gives short-tailed distributions with a finite right endpoint, such as the uniform and beta distribution.

The choice of the threshold u is crucial for the success of these methods. The appropriate value is typically chosen via ad hoc rules of thumb, and the best choice is obtained by a trade-off between bias and variance. A smaller value of u reduces the variance because more data are used, but increases the bias because GPD is only valid as u approaches infinity.

Multivariate time series

Most of the papers analysing financial returns have concentrated on the univariate case, not so many are concerned with their multivariate extensions. Appropriate modelling of time-varying dependencies is fundamental for quantifying different kinds of financial risk, for instance for assessing the risk of a portfolio of financial assets. Indeed, financial volatilities co-vary over time across assets and markets. The challenge is to design a multivariate distribution that both models the empirical return distribution in an appropriate way, and at the same time is sufficiently simple and robust to be used in simulation-based models for practical risk management.

5.1 Vector Autoregression models

The natural extension of the autoregressive model in Section 2.2 to dynamic multivariate time series is the Vector Autoregression Model of order 1 (VAR(1)-model):

$$\begin{aligned} \mathbf{Y}_t &= \boldsymbol{\mu} + \boldsymbol{\Omega} \mathbf{Y}_{t-1} + \boldsymbol{\epsilon}_t \\ \mathbb{E}[\boldsymbol{\epsilon}_t] &= \mathbf{0} \\ \text{Cov}[\boldsymbol{\epsilon}_t] &= \boldsymbol{\Sigma}_t, \end{aligned} \tag{5.1}$$

where $\boldsymbol{\epsilon}_t, t = 1, \dots$, are serially independent vectors. In the most simple setting, $\boldsymbol{\Omega}$ a diagonal matrix, with the mean-reversion parameter of each time series along the diagonal. It may however be more complicated, specifying for instance the causal impact of an unexpected shock to one variable on other variables in the model. An example is the stock market return in US, influencing the returns in Europe and Asia the following day. A multivariate random walk model is easily obtained from Equation 5.1 if $\boldsymbol{\Omega}$ is a diagonal matrix with all the diagonal elements equal to 1.

The covariance matrix $\boldsymbol{\Sigma}_t$ may either be constant or time varying. One way to make it time-varying, is to extend the the GARCH-model framework to the multivariate case, and in Section 5.2 we review the most common variants of the multivariate GARCH(1,1) model.

5.2 Multivariate GARCH models

The different variants of the multivariate GARCH(1,1)-model differ in the model for Σ_t . In this section we describe the two most commonly used in practice; the Diagonal-Vec and the Constant Conditional Correlation model. For a review of other models, see e.g. Alexander (2001) or Zivot and Wang (2003).

5.2.1 The Diagonal-Vec model

The first multivariate GARCH model was specified by Bollerslev et al. (1998). It is the so-called Diagonal-Vec model, where the conditional covariance matrix is given by

$$\Sigma_t = \mathbf{C} + \mathbf{A} \otimes \epsilon_{t-1} \epsilon_{t-1}' + \mathbf{B} \otimes \Sigma_{t-1}.$$

The symbol \otimes stands for the Hadamard product, that is, element-by-element multiplication. Let d be the dimension of the multivariate time series. Then, \mathbf{A} , \mathbf{B} , and \mathbf{C} are $d \times d$ matrices and ϵ_t is a $d \times 1$ column vector. The matrices \mathbf{A} , \mathbf{B} and \mathbf{C} are restricted to be symmetric, to ensure that the covariance matrix Σ_t is symmetric.

The Diagonal-Vec model appears to provide good model fits in a number of applications, but it has two large deficiencies: (a) it does not ensure positive semi-definiteness of the conditional covariance matrix Σ_t , and (b) it has a large number of parameters. If N is the number of assets, it requires the estimation of $3N(N+1)/2$ parameters. According to Ledoit et al. (2003), the method is not computationally feasible for problems with dimension larger than 5. For these reasons, several other multivariate GARCH(1,1)-models that avoid these difficulties by imposing additional structure on the problem have been proposed.

5.2.2 The Constant Conditional Correlation model

The Constant Conditional Correlation (CCC) model was introduced by Bollerslev (1990). In this model, the conditional correlation matrix \mathbf{R} is assumed to be constant across time. That is,

$$\begin{aligned} \Sigma_t &= \mathbf{D}_t \mathbf{R} \mathbf{D}_t \\ \mathbf{D}_t &= \begin{bmatrix} \sigma_{1,t} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \sigma_{d,t} \end{bmatrix} \\ \sigma_{i,t}^2 &= a_{i,0} + a_i \epsilon_{i,t-1}^2 + b_i \sigma_{i,t-1}^2. \end{aligned} \tag{5.2}$$

Thus, Σ_t is obtained by a simple transformation involving diagonal matrices, whose diagonal elements are conditional standard deviations which may be obtained from a univariate GARCH-model. The CCC model is much simpler to estimate than most other multivariate models in the GARCH family, contributing to its popularity. With N assets, only $N(3+N)$ parameters has to be estimated.

5.3 Time-varying correlation

An observation often reported by market professionals is that “during major market events, correlations change dramatically”. The possible existence of changes in the correlation, or

more precisely, of changes in the dependency structure between assets, has obvious implications in risk assessment and portfolio management. For instance, if all stocks tend to fall together as the market falls, the value of diversification may be overstated by those not taking the increase in downside correlations into account.

There are two ways of understanding “changes of correlations” that are not necessarily mutually exclusive (Malevergne and Sornette, 2002): There might be genuine changes in correlation with time. Longin and Solnik (2001) and Ang and Chen (2000) and many others claim to have produced evidence that stock returns exhibit greater dependence during market downturns than during market upturns. In contrast, the other explanation is that the many reported changes of correlation structure may not be real, but instead be attributed to a change in volatility, or a change in the market trend, or both (Forbes and Rigobon, 2002; Malevergne and Sornette, 2002; Sheedy, 1997). Malevergne and Sornette (2002) and Boyer et al. (1997) show by explicit analytical calculations that correlations conditioned on one of the variables being smaller than a limit, may deviate significantly from the unconditional correlation. Hence, the conditional correlations can increase when conditioning on periods characterised by larger fluctuations, even though the unconditional correlation coefficient is not altered.

Multivariate GARCH-models, like those described in Section 5.2, may be used to estimate time-varying correlations. However, different models give often very different estimates of the same correlation parameter, making it difficult to decide which one should trust. Moreover, when comparing different multivariate GARCH-models of increasing flexibility, using traditional statistical model selection criteria (i.e. the Akaike (AIC) and Bayesian (BIC) (c.f. Zivot and Wang (2003))), the Constant Conditional Correlation model often provides the best fit (see e.g. Aas (2003)). This might indicate that what at first step seem to be time-varying changes in correlation, can be attributed to changes in volatility instead. It might also be, however, that the correlation really is time-varying, and that the poor performance the time-varying correlation GARCH-models is a result of model misspecification in terms of the conditional covariances.

To summarise, all correlation estimates will change over time, whether the model is based on constant or time-varying correlation. In a constant correlation model, this variation in the estimates with time is only due to sampling error, but in a time-varying parameter model this variation in time is also ascribed to changes in the true value of the parameter. This makes it very difficult to verify whether there really is time-varying correlation between financial assets or not.

5.4 Multivariate return distributions

As in the univariate case, there are two types of return distributions to be considered, the marginal distribution of the returns, and the conditional return distribution. Both are known to be skewed and heavy-tailed, but until recently, most multivariate parametric approaches have been based on the multinormal assumption due to computational limitations. In this section we will describe the multivariate Gaussian distribution, and a very interesting alternative, the multivariate normal inverse Gaussian distribution.

5.4.1 The multivariate Gaussian distribution

A d -dimensional column vector \mathbf{X} is said to be multivariate Gaussian distributed if the probability density function can be written as

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \cdot \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\},$$

where \mathbf{x} and $\boldsymbol{\mu}$ are d -dimensional column vectors and Σ is a symmetric $d \times d$ matrix of covariances. The mean vector and covariance matrix of the multivariate Gaussian distribution are given by

$$\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu}, \quad \text{and} \quad \text{Cov}(\mathbf{X}) = \Sigma.$$

5.4.2 The multivariate normal inverse Gaussian (NIG) distribution

A d -dimensional column vector \mathbf{X} is said to be MNIG distributed if the probability density function can be written as (Øigård and Hanssen, 2002)

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\delta}{2^{\frac{d-1}{2}}} \left[\frac{\alpha}{\pi q(\mathbf{x})} \right]^{\frac{d+1}{2}} \exp(p(\mathbf{x})) K_{\frac{d+1}{2}}(\alpha q(\mathbf{x})),$$

where

$$p(\mathbf{x}) = \delta \sqrt{\alpha^2 - \boldsymbol{\beta}^T \boldsymbol{\Gamma} \boldsymbol{\beta}} + \boldsymbol{\beta}^T (\mathbf{x} - \boldsymbol{\mu}),$$

$$q(\mathbf{x}) = \sqrt{\delta^2 + (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Gamma}^{-1} (\mathbf{x} - \boldsymbol{\mu})},$$

$\delta > 0$ and $\alpha^2 > \boldsymbol{\beta}^T \boldsymbol{\Gamma} \boldsymbol{\beta}$. $K_i(x)$ is the modified Bessel function of the second kind of order i (Abramowitz and Stegun, 1972). As for the univariate density, the scalar δ and the d -dimensional vector $\boldsymbol{\mu}$ determine the scale and location, respectively, while the scalar α and the d -dimensional vector $\boldsymbol{\beta}$ control the shape of the density. Finally, $\boldsymbol{\Gamma}$ is a symmetric positive semidefinite $d \times d$ matrix with determinant equal to 1. It should be noticed that choosing $\boldsymbol{\Gamma} = \mathbf{I}$ is not sufficient to produce a diagonal covariance matrix. In addition, the parameter vector $\boldsymbol{\beta}$ has to be the zero-vector. The mean vector of the MNIG distribution is given by

$$\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu} + \frac{\delta \boldsymbol{\Gamma} \boldsymbol{\beta}}{\sqrt{\alpha^2 - \boldsymbol{\beta}^T \boldsymbol{\Gamma} \boldsymbol{\beta}}},$$

and the covariance matrix is given by

$$\text{Cov}(\mathbf{X}) = \delta (\alpha^2 - \boldsymbol{\beta}^T \boldsymbol{\Gamma} \boldsymbol{\beta})^{-1/2} [\boldsymbol{\Gamma} + (\alpha^2 - \boldsymbol{\beta}^T \boldsymbol{\Gamma} \boldsymbol{\beta})^{-1} \boldsymbol{\Gamma} \boldsymbol{\beta} \boldsymbol{\beta}^T \boldsymbol{\Gamma}^T].$$

We have recently done a study (Aas et al., 2003), in which we show that using this distribution as the conditional distribution for a multivariate GARCH-model outperforms the same GARCH-model with a multivariate Gaussian conditional distribution for a portfolio of American, European and Japanese equities.

5.5 Copulas

Unless asset returns are well represented by a multivariate normal distribution, correlation might be an unsatisfactory measure of dependence, see for instance Embrechts et al. (1999). If the multivariate normality assumption does not hold, the dependence structure of financial assets may be modelled through a *copula*. The application of copula theory to the analysis of economic problems is a relatively new (copula functions have been used in financial applications since Embrechts et al. (1999)) and fast-growing field. From a practical point of view, the advantage of the copula-based approach to modelling is that appropriate marginal distributions for the components of a multivariate system can be selected freely, and then linked through a copula suitably chosen to represent the dependence between the components. That is, copula functions allow one to model the dependence structure independently of the marginal distributions.

As an example of how copulas may be successfully used, consider the modelling of the joint distribution of a stock market index and an exchange rate. The Student's t distribution has been found to provide a reasonable fit to the conditional univariate distributions of stock market as well as exchange rate daily returns. A natural starting point in the modelling of the joint distribution might then be a bivariate Student's t distribution. However, the standard bivariate Student's t distribution has the restriction that both marginal distributions must have the same tail heaviness, while studies have shown that stock market and exchange rate daily returns have very different tail heaviness. Decomposing the multivariate distribution into the marginal distributions and the copula allows for the construction of better models of the individual variables than would be possible if only explicit multivariate distributions were considered.

In what follows, we first give a formal definition of a copula and then we review some of the commonly used parametric copulas.

5.5.1 Definition

The definition of a d -dimensional copula is a multivariate distribution, C , with uniformly distributed marginals $U(0, 1)$ on $[0, 1]$. It is obvious from this definition, that if F_1, F_2, \dots, F_n are univariate distribution functions, $C(F_1(x_1), F_2(x_2), \dots, F_n(x_d))$ is a copula, because $U_i = F_i(x_i)$ is a uniform random variable.

5.5.2 Parametric copulas

In what follows, we present the functional forms of three different copulas, the Gaussian copula, the Student- t copula and the Kimeldorf and Sampson copula. For the sake of simplicity we only consider the bivariate versions of the copulas.

Gaussian copula

The Gaussian copula is given by

$$C_R(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \exp\left\{-\frac{s^2 - 2R_{12}st + t^2}{2(1-R_{12}^2)}\right\} ds dt,$$

where R_{12} is the linear correlation between the two random variables and $\Phi^{-1}(\cdot)$ is the inverse of the standard univariate Gaussian distribution function.

Student's t copula

The Student's t copula generalizes the Gaussian copula to allow for joint fat tails and an increased probability of joint extreme events. This copula with ν degrees of freedom can be written as

$$C_{R,\nu}(u_1, u_2) = \int_{-\infty}^{t_{\nu}^{-1}(u_1)} \int_{-\infty}^{t_{\nu}^{-1}(u_2)} \frac{1}{2\pi(1-R_{12}^2)^{1/2}} \left\{ 1 + \frac{s^2 - 2R_{12}st + t^2}{\nu(1-R_{12}^2)} \right\}^{-(\nu+2)/2} ds dt,$$

where t_{ν}^{-1} is the inverse of the standard univariate student-t distribution with ν degrees of freedom, expectation 0 and variance $\frac{\nu}{\nu-2}$.

Kimeldorf and Sampson copula

The Student's t copula allows for joint extreme events but it does not allow for asymmetries. If one believes in the asymmetries in equity return dependence structures that have been reported by for instance Longin and Solnik (2001) and Ang and Chen (2000), the Student's t copula may also be too restrictive to provide a reasonable fit to equity data. Then, the Kimeldorf and Sampson copula, which is an asymmetric copula, exhibiting greater dependence in the negative tail than in the positive, might be a better choice. This copula, with parameter δ is given by

$$C_{\delta}(u_1, u_2) = \begin{cases} (u_1^{-\delta} + u_2^{-\delta} - 1)^{-1/\delta} & \text{if } \delta > 0 \\ u_1 \times u_2 & \text{if } \delta = 0 \end{cases}$$

The upper left corner of Figure 5.1 shows a plot of Norwegian vs. Nordic geometric returns. As can be seen from the figure, the correlation between the two indices is high (the coefficient is 0.63). We have fitted Gaussian, Kimeldorf and Sampson and Student's t-6 copulas to this data, respectively. The remaining three panels of Figure 5.1 show simulations from the estimated copulas. The marginals are Student's t-6 distributed in all three panels. It is up to the reader to decide which copula that best fit the data !

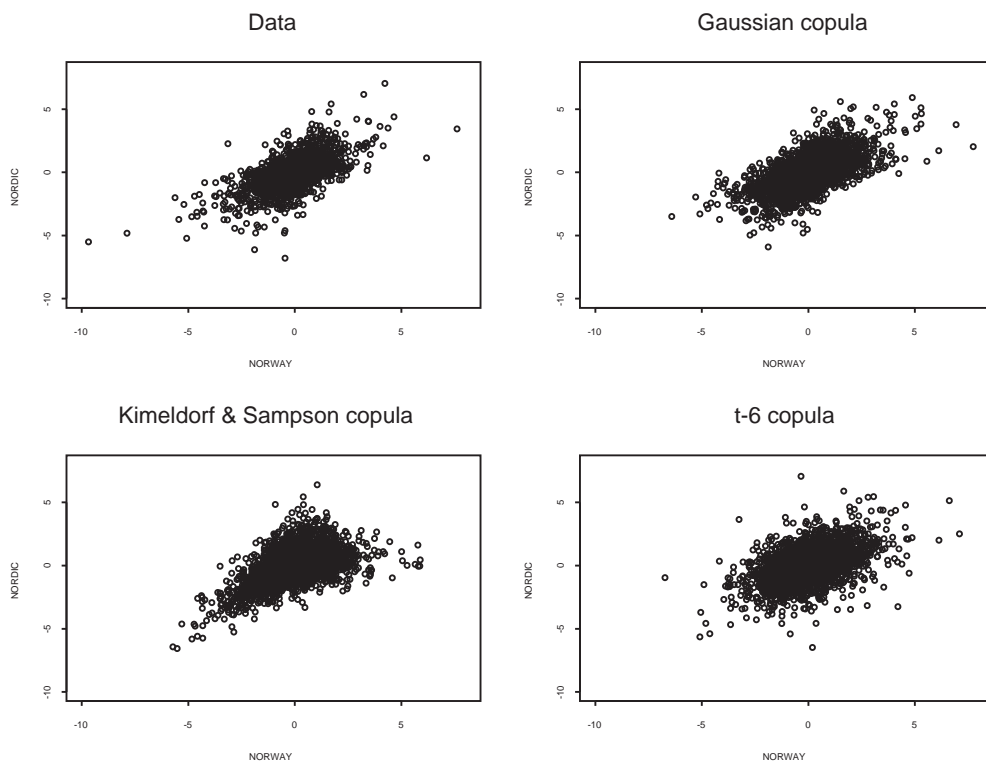


Figure 5.1: *Upper left corner: Norwegian vs. Nordic geometric returns. Upper right corner: Gaussian copula. Lower left corner: Kimeldorf and Sampson copula. Lower right corner: Student's t-6 copula. The marginals are Student's t-6 distributed in all three panels.*

With an understanding of how prices behave we are able to make better decisions. In this section we explore practical applications of the price models. First, in Section 6.1 we show how time series' models may be used for forecasting and then in Section 6.2 we describe an application to risk management.

6.1 Prediction

A financial time series model is a useful tool to generate forecasts for both the future value and the volatility of the time series. Moreover, it is important to have knowledge of the uncertainty of such forecasts. In this section we will give examples of forecasting with an AR(1)-model and a GARCH(1,1)-model.

6.1.1 Forecasting with an AR(1)-model

In an AR(1)-model the 1-step forecast of Y_{T+1} at time T given $\{Y_T, Y_{T-1}, \dots\}$ is

$$E[Y_{T+1}|Y_T] = \mu + \alpha Y_T$$

The 1-step forecast can be iterated to get the k -step forecast ($k \geq 2$)

$$E[Y_{T+k}|Y_T] = \mu \sum_{i=0}^{k-1} \alpha^i + \alpha^k Y_T$$

As $k \rightarrow \infty$, the forecast approaches the stationary value $\mu/(1-\alpha)$. The forecast error is given by

$$E[Y_{T+k}|Y_T] - Y_{T+k} = \sum_{i=0}^{k-1} \alpha^{k-1-i} \epsilon_{T+i+1}.$$

The forecast error is therefore a linear combination of the unobservable future shocks entering the system after time T . The variance of the k -step forecast error is given by

$$\text{Var}(E[Y_{T+k}|Y_T] - Y_{T+k}) = \sigma^2 \sum_{i=0}^{k-1} \alpha^i = \sigma^2 \frac{1 - \alpha^{2k}}{1 - \alpha^2}.$$

Thus as $k \rightarrow \infty$ the variance increases to the stationary level, $\sigma^2/(1 - \alpha^2)$.

Figure 6.1 shows forecasts 100 days ahead for the AR(1)-process in Figure 2.2. The dashed curves are approximate pointwise 95% confidence intervals. The horizontal line represents the stationary level of the process. As can be seen from the figure, as we go further into the future, the forecast approaches the stationary level, while the confidence interval widens.

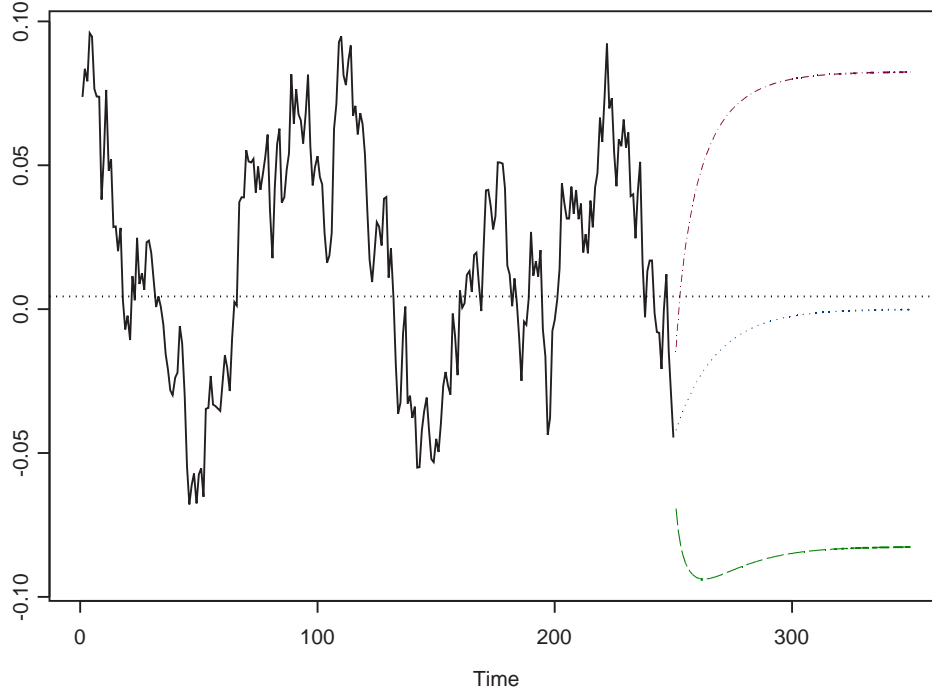


Figure 6.1: The AR(1)-process in Figure 2.2 and forecasts 100 days ahead. The dashed curves are approximate pointwise 95% confidence intervals. The horizontal line represents the stationary level of the process.

6.1.2 Forecasting with a GARCH(1,1)-model

Consider the basic GARCH(1,1)-model from Equation 3.2 which is fitted to a data set for the time period $t = 1, \dots, T$. The 1-step forecast of the variance, given the information at time T is given by

$$E[\sigma_{T+1}^2 | \sigma_T^2] = a_0 + a \epsilon_T^2 + b \sigma_T^2,$$

where ϵ_T^2 and σ_T^2 are the fitted values from the estimation process. The above derivation can be iterated to get the k -step forecast ($k \geq 2$)

$$E[\sigma_{T+k}^2 | \sigma_T^2] = a_0 \sum_{i=1}^{k-2} (a+b)^i + (a+b)^{k-1} (a_0 + a \epsilon_T^2 + b \sigma_T^2).$$

As $k \rightarrow \infty$, the variance forecast approaches the stationary variance $a_0/(1 - a - b)$ (if the GARCH process is stationary). The smaller this quantity $\eta = a + b$, the more rapid is the

convergence to the long-term volatility estimate. Convergence is typically rather slow in foreign exchange markets, while equity markets often have more rapidly convergent volatility.

The forecasted volatility can be used (like in the previous section) to generate confidence intervals of the forecasted series values. Moreover, in many cases the forecasted volatility is itself of central interest, for instance when valuing path-dependent options or volatility options. The variance of the forecasted volatility can be obtained. Explicit formulas for this variance is only known for some special models. Instead simulation techniques, like the one indicated in the following section can be used.

6.2 Risk management

Financial risks often relate to negative developments in financial markets. Movements in financial variables such as interest rates and exchange rates create risks for most corporations. Generally, financial risks are classified into the broad categories of market risks, credit risks, liquidity risks, and operational risks. The focus of this course is on market risk, which arises from changes in the prices of financial assets. Value-at-Risk (VaR) is considered the *de facto* standard risk measure used today by regulators and investment banks. Formally, VaR represents the worst expected loss over a given time interval under normal market conditions at a given confidence level. It may however be easier to consider it as a quantile in the distribution of asset values/returns. If this distribution can be assumed to be normal, the VaR value can be derived directly from the standard deviation of the returns, using a multiplicative factor that depends on the confidence level. As demonstrated previously in this note, however, the asset return distribution is seldom normal and a Monte Carlo simulation approach must be used instead. In brief, this method consists of two steps. First, the risk manager specifies a stochastic process for the financial variables and the parameters of this process. The latter are usually determined from historical data, but they may also be set manually. Second, fictitious price paths are simulated for all variables of interest. Each of these realizations is then used to compile a distribution of returns, from which the VaR value can be measured. The first step has been covered previously in this note. This section contains a short description of the latter.

6.2.1 Simulations

The basic concept behind Monte Carlo simulations is to simulate repetitions of a random process for the financial variable of interest, covering a wide range of possible simulations. In this course, we concentrate on one asset. We assume that the log-increments of the price process can be represented by a random walk model in which the time variation in the variances is simulated using a GARCH(1,1)-model. The class of GARCH-models have become increasingly popular for computing Value-at-Risk, due to two main reasons. First, it allows for volatility clustering. Second, since volatility is allowed to be volatile, the unconditional distribution will typically be not be thin-tailed, even when the conditional distribution is normal.

Assume that the GARCH-model is estimated over a time period that ends in time $t = T$, and that ϵ_T^2 and σ_T^2 are the fitted values from the estimation process. Moreover, let the starting value of the asset be Y_T . The following procedure generates the distribution of the value of the asset, Y_{T+k} , k time steps from today:

1. Start with σ_T , ϵ_T , and Y_T .

2. Repeat for $t = T + 1 \dots T + k$
 - Compute $\sigma_t^2 = a_0 + a \epsilon_{t-1}^2 + b \sigma_{t-1}^2$.
 - Draw $\epsilon_t \sim N(0, \sigma_t^2)$
 - Compute $\log(Y_t) = \mu + \log(Y_{t-1}) + \epsilon_t$.
 - Compute $Y_t = \exp(\log(Y_t))$

Steps 1-2 generates *one* price path for the value of the asset. This procedure is then repeated as many times as necessary, say N times, each time with the same starting values, obtaining a distribution of values, $Y_{T+k}^1, \dots, Y_{T+k}^N$. From this distribution the 1% VaR can be reported as the 1% lowest value.

In Figure 6.2 we have shown the distribution (dotted line) obtained by using this procedure with $Y_T = 135$, $k = 10$, $N = 10,000$ and the GARCH-model fitted to the geometric returns in Figure 3.1. In the same figure, we have also shown the distribution (solid line) obtained by replacing the Gaussian conditional distribution with the NIG-distribution in Figure 4.2. The 1 % VaR for the two choices of conditional distribution are shown as a dotted and solid vertical line, respectively. As can be seen from the figure, using the NIG-distribution gives a more conservative estimate of VaR than that obtained using the Gaussian distribution.

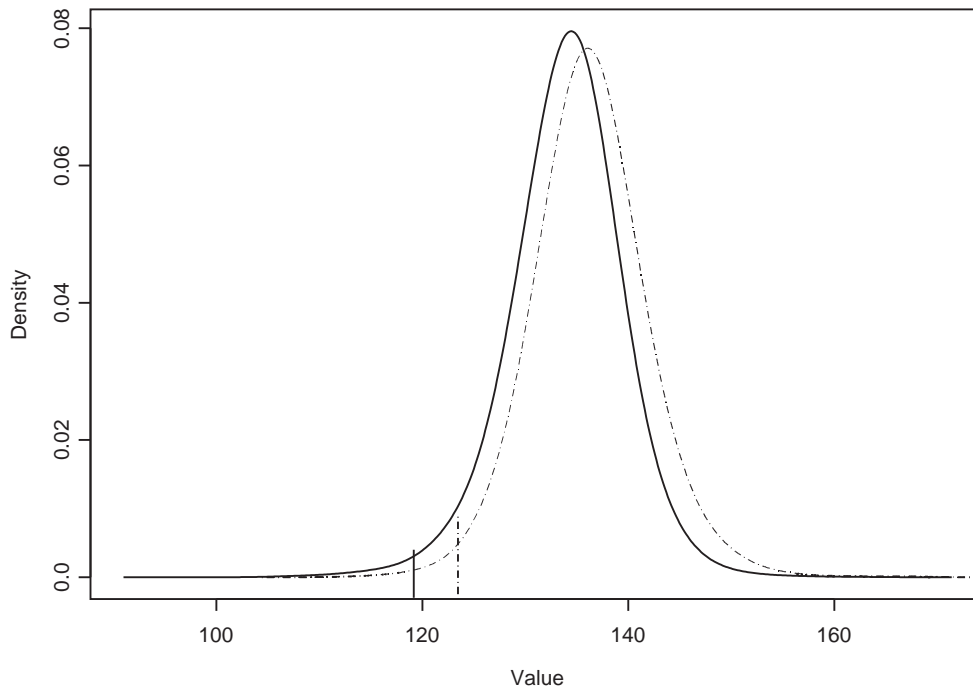


Figure 6.2: The distribution of the value of an asset 10 steps from today if it starts at 135. Solid line: NIG conditional distribution. Dotted line: Gaussian conditional distribution.

References

- K. Aas. Return modelling in financial markets: A survey, 2003. NR note, SAMBA/22/03, November 2003.
- K. Aas, X. K. Dimakos, and Ingrid H. Haff. Risk estimation using the multivariate normal inverse gaussian distribution, 2003. Submitted to *Journal of Business & Economic Statistics*.
- M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Function*. Dover, New York, 1972.
- C. Alexander. *Market models*. Wiley, New York, 2001.
- J. Andersson. On the normal inverse Gaussian stochastic volatility model. *Journal of Business and Economic Statistics*, 19:44–54, 2001.
- A. Ang and J. Chen. Asymmetric correlations of equity portfolios, 2000. Working Paper, Stanford, September 2000.
- O. Aquilar and M. West. Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics*, 18(3), 2000.
- T. G. Bali. Modeling the stochastic behavior of short-term interest rates: Pricing implications for discount bounds. *Journal of Banking & Finance*, 27:201–228, 2003.
- O. E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24:1–13, 1997.
- F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:167–179, 1973.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.
- T. Bollerslev. A conditionally heteroscedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, 69:542–547, 1987.
- T. Bollerslev. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics*, 72:498–505, 1990.

- T. Bollerslev, R. F. Engle, and J. M. Wooldridge. A capital asset pricing model with time varying covariances. *Journal of Political Economy*, 96:116–131, 1998.
- E. Bølviken and F. E. Benth. Quantification of risk in Norwegian stocks via the normal inverse Gaussian distribution. In *Proceedings of the AFIR 2000 Colloquium, Tromsø, Norway*, pages 87–98, 2000.
- B. H. Boyer, M. S. Gibson, and M. Loretan. Pitfalls in tests for changes in correlation, 1997. International Finance Discussion Paper, 597, Board of the Governors of the Federal Reserve System.
- B. Bradley and M. Taqqu. Financial risk and heavy tails. In S.T.Rachev, editor, *Handbook of Heavy Tailed Distributions in Finance*. Elsevier, North-Holland, 2003.
- R. J. Brenner, R. Harjes, and K. Kroner. Another look at models of the short-term interest rate. *Journal of Financial and Quantitative Analysis*, 31:85–107, 1996.
- K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders. An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, 47:1209–1227, 1992.
- J.C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–408, 1985.
- F. C. Drost and T. E. Nijman. Temporal aggregation of GARCH processes. *Econometrica*, 61:909–927, 1993.
- E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1(3):281–299, 1995.
- P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*. Springer-Verlag, Berlin Heidelberg, 1997.
- P. Embrechts, A. J. McNeil, and D. Straumann. Correlation: Pitfalls and alternatives. *Risk*, 12:69–71., 1999.
- R. F. Engle, V. Ng, and M. Rothschild. Asset pricing with a factor arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, 45:213–237, 1990.
- K. J. Forbes and R. Rigobon. No contagion, only interdependence: Measuring stock market co-movements. *Journal of Finance*, 57(5):2223–2261, 2002.
- L. Forsberg and T. Bollerslev. Bridging the gap between the distribution of realized (ecu) volatility and ARCH modeling (of the euro): The GARCH-NIG model. *Journal of Applied Econometrics*, 17(5):535–548, 2002.
- E. Ghysels, A. C. Harvey, and E. Renault. Stochastic volatility. In C. R. Rao and G. S. Maddala, editors, *Statistical Methods in Finance*. North-Holland, Amsterdam, 1996.
- S. F. Gray. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42:27–62, 1996.
- M. Berg Jensen and A. Lunde. The NIG-S and ARCH model: A fat-tailed stochastic, and autoregressive conditional heteroscedastic volatility model, 2001. Working paper series No. 83, University of Aarhus.

- P. Jorion. *Value at Risk*. McGraw-Hill, 1997.
- D. Karlis. An EM type algorithm for maximum likelihood estimation of the normal-inverse Gaussian distribution. *Statistics & Probability Letters*, 57:43–52, 2002.
- S. Kim, N. Shepard, and S. Chib. Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, 65:361–293, 1998.
- K. G. Koedijk, F. G. J. A. Nissen, P. C. Schotman, and C. C. P. Wolff. The dynamics of short-term interest rate volatility reconsidered. *European Finance Review*, 1:105–130, 1997.
- O. Ledoit, P. Santa-Clara, and M. Wolf. Flexible multivariate garch modeling with an application to international stock markets. *Review of Economics and Statistics*, 85(3), 2003.
- J. Lillo. Risk analysis and the NIG distribution. *Journal of Risk*, 2(4):41–56, 2000.
- F. Longin and B. Solnik. Extreme correlation of international equity markets. *The Journal of Finance*, 56(2):649–676, 2001.
- F. A. Longstaff and E. S. Schwartz. Interest rate volatility and the term structure: A two-factor general equilibrium model. *The Journal of Finance*, 47:1259–1282, 1992.
- Y. Malevergne and D. Sornette. Investigating extreme dependences: Concepts and tools, 2002. SSRN Electronic Paper Collection.
- H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.
- A. J. McNeil and R. Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7:271–300, 2000.
- N. Meddahi and E. Renault. Temporal aggregation of volatility models, 2004. To appear in *Journal of Econometrics*.
- JP Morgan. Riskmetrics - technical document, 1996. New York.
- D. Nelson. ARCH models as diffusion approximations. *Journal of Econometrics*, 45:7–38., 1990.
- T. A. Øigård and A. Hanssen. The multivariate normal inverse Gaussian heavy-tailed distribution: Simulation and estimation, 2002. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Orlando Florida, May 2002.
- K. Prause. Modelling financial data using generalized hyperbolic distributions, 1997. FDM preprint 48, University of Freiburg.
- P. H. M. Rodrigues and A. Rubia. Testing non-stationarity in short-term interest rates, 2003. 14th EC2 Conference: Endogeneity, Instruments and Identification in Economics, London, 12-13 December.
- T. H. Rydberg. The normal inverse Gaussian Levy process: Simulation and approximation. *Commun. Statist.-Stochastic Models*, 34:887–910, 1997.
- E. Sheedy. Correlation in international equity and currency markets: A risk adjusted perspective, 1997. CMBF Papers, No. 17, Macquarie University, Sydney.

N. Shepard. Statistical aspects of ARCH and stochastic volatility. In D.V. Lindley D. R. Cox and O. E. Barndorff-Nielsen, editors, *Time series Models in Econometrics, Finance, and Other Fields*. Chapman-Hall, London, 1996.

O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

J. H. Venter and P. J. de Jongh. Risk estimation using the normal inverse Gaussian distribution. *Journal of Risk*, 4(2):1–24, 2002.

E. Zivot and J. Wang. *Modelling Financial Time Series with S-PLUS*. Insightful Corporation, 2003.