# Adaptive Estimation of Time Varying Copulae

Master Thesis submitted to


**Prof. Dr. Wolfgang Härdle**
**Prof. Dr. Vladimir Spokoiny**

Institute for Statistics and Econometrics

CASE - Center for Applied Statistics and Economics

**Humboldt-Universität zu Berlin**


by


**Ekaterina Ignatieva**

(184944)


in partial fulfillment of the requirements

for the degree of **Master of Sciences in Statistics**

Berlin, September 6, 2005

# Declaration of Authorship

I hereby confirm that I have authored this master thesis independently and without use of others than the indicated sources. All passages which are literally or in general matter taken out of publications or other sources are marked as such.

Berlin, September 6, 2005

Ekaterina Ignatieva

# Contents

# List of Figures

# List of Tables

# 1 Introduction

> *"...as well as being one of the most ubiquitous concepts in modern finance and insurance, correlation is also one of the most misunderstood concepts. Some of the confusion may arise from the literary use of the word to cover any notion of dependence. To a mathematician correlation is only one particular measure of stochastic dependence among many. It is the canonical measure in the world of multivariate normal distribution functions, and more generally for spherical and elliptical distributions. However, empirical research in finance and insurance shows that the distributions of the real world are seldom in this class."*
>
> Embrechts, Mc Neil and Straumann, 1999

In this chapter we will consider a concept of dependence. As we already know, the cumulative distribution function (cdf) of a 2-dimensional vector $(X_1, X_2)$ is given by

$$F(x_1, x_2) = P(X_1 \leq x_1, Y_1 \leq y_1) \tag{1.1}$$

For the case that $X_1$ and $X_2$ are independent, their joint cumulative distribution function $F(x_1, x_2)$ can be written as a product of their 1-dimensional marginals:

$$F(x_1, x_2) = F_{X_1}(x_1) F_{X_2}(x_2) = P(X_1 \leq x_1) P(X_2 \leq x_2) \tag{1.2}$$

But how can we represent $F(x_1, x_2)$ without having an information concerning the dependence of $X_1$ and $X_2$? How can we model the relationship between them? Most people would model this by means of linear correlation. Several authors (e.g. Embrechts, Mc Neil and Straumann, 1999) show therefore that the correlation is an appropriate measure of dependence only when the random variables have an eliptical or spherical distribution, which include the normal multivariate distribution. Although the terms "correlation" and "dependency" are often used interchangeably, correlation is actually a rather imperfect measure of dependency, and there are many circumstances where correlation should not be used. We therefore need an alternative dependency measure that is reliable when correlation is not - and the answer is to use copulae.

Copulae represent an elegant concept of connecting marginals with joint cummulative distribution functions. Copulae are functions that join or "couple" multivariate distribution functions to their 1-dimensional marginal distribution functions. Let us consider a d-dimensional vector $X = (X_1, ..., X_d)$. Using copulae, the marginal distribution functions $F_{X_i}(i = 1, ..., d)$ can be separately modelled from their dependence structure and then coupled together to form the multivariate distribution $F_X$.

Copula functions have a long history in probability theory and statistics since they date back to (Sklar, 1959). Their application in finance is very recent: the idea first appears in (Embrechts, Mc Neil and Straumann, 1999) in connection with correlation as a measure of dependence. Copulae constitute an essential part in quantitative finance (Härdle, Kleinow and Stahl, 2002) and are recognized as an important tool in *Value-at-Risk* (VaR) calculations. VaR is a measure that characterizes the riskiness of a portfolio; it keeps the probability of a negative outcome (portfolio losses) below some level $\lambda$.

**DEFINITION 1.1.** *For a financial position $X$, we define its VaR at level $\lambda$ as*

$$VaR_\lambda(X) = \inf\{m|P[X + m < 0] \leq \lambda\} \tag{1.3}$$

*In other terms, VaR is a quantile at level $\lambda$ of the distribution of X:*

$$VaR_\lambda(X) = F_\lambda^{-1}(X) \tag{1.4}$$

To be able to model the VaR of a portfolio, we have to know as well the distribution of losses associated with the risk factors as a dependence structure within the portfolio. Formerly, there were two main approaches for modelling dependence structure: multivariate normality or independence. In both cases, risk aggregation is straightforward but often too far away from realistic models, since risk factors are seldom independent and can even have a dependence structure. To overcome this problem, we separate the study of the univariate behaviour of each variable from the study of dependence structure. The dependence structure of the random variables is known as a *copula.* The combination of a copula with specific univariate distributions then allows to construct a multivariate distribution, so-called copula-based models, which allow to calculate VaR more efficiently.

In this study we perform the copulae estimations using adaptive and non-adaptive techniques. This thesis can be seen as a preliminary study to risk management with adaptive copulae because it is the accurate basis for VaR calculations.

The thesis is organized as follows: chapter 1 and 2 are devoted to the introduction of basic definitions, and later to the definition and basic properties of a copula. Chapter 2 refers to the modelling dependence with copulae. Some examples of bivariate copulae and their extension to the multivariate case are represented in chapter 4 and chapter 5, respectively. Chapter 6 and 7 are devoted to the time-varying copulae.

*Full maximum likelihood* (FML) estimation and *Inference for margins* (IFM) methods are performed in the 6th chapter. Adaptive copula estimation, based on *Local change point analysis*(LCPD) (Mercurio, Spokoiny, 2004) is represented in chapter 7. The change point test on the copula dependence parameter $\theta$ can improve substantially the dependence modelling. We illustrate the method considering some simulated examples (chapter 8) and then apply it to the 2-, 4- and 6-dimensional data sets from the Dax Index.

First let us concentrate on the 2-dimensional case, then we will extend this concept to the d-dimensional case, for a random variable in $\mathbb{R}^d$ with $d \geq 1$.

To be able to define a copula function, first we need to represent a concept of the *volume of a rectangle, a 2-increading function* and *a grounded function.*
Let $U_1$ and $U_2$ be two sets in $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ and consider the function $F : U_1 \times U_2 \longrightarrow \overline{\mathbb{R}}$.

**DEFINITION 1.2.** *The $F$-volume of a rectangle $B = [x_1, x_2] \times [y_1, y_2] \subset U_1 \times U_2$ is defined as:*

$$V_F(B) = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \tag{1.5}$$

**DEFINITION 1.3.** *$F$ is said to be a 2-increasing function if for every $B = [x_1, x_2] \times [y_1, y_2] \subset U_1 \times U_2$,*

$$V_F(B) \quad \geq \quad 0 \tag{1.6}$$

**REMARK 1.1.** *Note, that "to be 2-increasing function" neither implies nor is implied by "to be increasing in each argument".*

The following lemmas (Nelsen, 1999) will be very useful later for establishing the continuity of copulae.

**LEMMA 1.1.** *Let $U_1$ and $U_2$ be nonempty sets in $\overline{\mathbb{R}}$ and let $F : U_1 \times U_2 \longrightarrow \overline{\mathbb{R}}$ be a two-increasing function. Let $x_1$, $x_2$ be in $U_1$ with $x_1 \leq x_2$, and $y_1$, $y_2$ be in $U_2$ with $y_1 \leq y_2$. Then the function $t \mapsto F(t, y_2) - F(t, y_1)$ is nondecreasing on $U_1$ and the function $t \mapsto F(x_2, t) - F(x_1, t)$ is nondecreasing on $U_2$.*

**DEFINITION 1.4.** *If $U_1$ and $U_2$ have a smallest element* $\min U_1$ *and* $\min U_2$ *respectively, then we say, that a function $F : U_1 \times U_2 \longrightarrow \mathbb{R}$ is grounded if :*

$$\text{for all } x \in U_1 : \quad F(x, \min U_2) \quad = \quad 0 \text{ and} \tag{1.7}$$
$$\text{for all } y \in U_2 : \quad F(\min U_1, y) \quad = \quad 0 \tag{1.8}$$

In the following, we will refer to this definition of a distribution function :

**DEFINITION 1.5.** *A distribution function is a function from* $\overline{\mathbb{R}}^2 \mapsto [0,1]$ *which:*

- *is grounded*
- *is 2-increasing*
- *satisfies* $F(\infty, \infty) = 1$

**LEMMA 1.2.** *Let* $U_1$ *and* $U_2$ *be nonempty sets in* $\overline{\mathbb{R}}$ *and let* $F : U_1 \times U_2 \longrightarrow \overline{\mathbb{R}}$ *be a grounded two-increasing function. Then* $F$ *is nondecreasing in each argument.*

**Proof:**

Set $x_1 = \min U_1$ and $y_1 = \min U_2$ in Lemma 1.1. □

**DEFINITION 1.6.** *If* $U_1$ *and* $U_2$ *have a greatest element* $\max U_1$ *and* $\max U_2$ *respectively, then we say, that a function* $F : U_1 \times U_2 \longrightarrow \mathbb{R}$ *has margins and that the margins of* $F$ *are given by:*

$$F(x) = F(x, \max U_2) \text{ for all } x \in U_1 \tag{1.9}$$

$$F(y) = F(\max U_1, y) \text{ for all } y \in U_2 \tag{1.10}$$

**LEMMA 1.3.** *Let* $U_1$ *and* $U_2$ *be nonempty sets in* $\overline{\mathbb{R}}$ *and let* $F : U_1 \times U_2 \longrightarrow \overline{\mathbb{R}}$ *be a grounded two-increasing function which has margins. Let* $(x_1, y_1)$, $(x_2, y_2) \in S_1 \times S_2$. *Then*

$$|F(x_2, y_2) - F(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |F(y_2) - F(y_1)| \tag{1.11}$$

**Proof:**

From the triangle inequality we have:

$$|F(x_2, y_2) - F(x_1, y_1)| \leq |F(x_2, y_2) - F(x_1, y_2)| + |F(x_1, y_2) - F(x_1, y_1)| \tag{1.12}$$

Let us assume that $x_1 \leq x_2$. Since $F$ is grounded, 2-increasing and has margins, it follows from Lemma 1.1 and 1.2 that:

$$0 \leq F(x_2, y_2) - F(x_1, y_2) \leq F(x_2) - F(x_1) \tag{1.13}$$

A similar inequality holds in the case $x_2 \leq x_1$. Thus we have that for any $x_1$, $x_2 \in S_1$ holds:

$$|F(x_2, y_2) - F(x_1, y_2)| \leq |F(x_2) - F(x_1)| \tag{1.14}$$

An analogous inequality holds for any $y_1$, $y_2 \in S_1$:

$$|F(x_1, y_2) - F(x_1, y_1)| \leq |F(y_2) - F(y_1)| \tag{1.15}$$

which completes the proof. □

# 2 Copulae. Definition and some Properties

One can easily verify, that $F_1(X)$ and $F_2(Y)$, where $F_1$ and $F_2$ are the marginal distributions of $X$ and $Y$ respectively, are two uniform variables if $F_1$ and $F_2$ are continuous. Hence, if the marginals $F_1$ and $F_2$ of the bivariate distribution F are continuous, there exists a unique copula, which is a cumulative distribution function, with its marginals being uniform.

**DEFINITION 2.1.** *Formally a function $C : [0,1]^2 \to [0,1]$ such that*

$$F(x,y) = C\left\{F_1(x), F_2(y)\right\} \tag{2.1}$$

*is a **copula**. On the other hand, if $C(u,v)$ and $F_1$ and $F_2$ are given, then there exists an $F$ such that:*

$$F\left\{F_1^{-1}(u), F_2^{-1}(v)\right\} = C(u,v) \tag{2.2}$$

In fact, the above defenition is just an implication of the theorem below, known as Sklar's theorem.

**THEOREM 2.1.** *Sklar's Theorem:*
*If the marginals $F_1, F_2$ of a 2-dimensional distribution function $F$ are continuous, there exists a unique copula $C : [0,1]^2 \to [0,1]$ such that*

$$F(x_1,x_2) = C\left\{F_{X_1}(x_1), F_{X_2}(x_2)\right\} \tag{2.3}$$

*Conversely, if $C$ is a copula and $F_{X_1}, F_{X_2}$ are distribution functions then $F$ defined by 2.3 is a 2-dimensional distribution function with marginals $F_{X_1}, F_{X_2}$.*

If the density $f(.,.)$ of $F(.,.)$ exists, one can derive the relationship between the density $f$ of $F$ and $c$ of $C$:

$$f(x,y) = c\left\{F_1(x), F_2(y)\right\} f_1(x) f_2(y) \tag{2.4}$$

with a copula density

$$c(u_1,u_2) = \frac{\partial^2 C(u_1,u_2)}{\partial u_1 \partial u_2} \tag{2.5}$$

where $f_1(x)$ and $f_2(y)$ are the marginal densities of $F$.

## 2.1 Elementary properties

For every $0 \le u \le 1$ and $0 \le v \le 1$ holds:

$$C(u,1) = P(U \le u, V \le 1) = P(U \le u) = u \tag{2.6}$$

and similary

$$C(1,v) = P(U \le 1, V \le v) = P(V \le v) = v \tag{2.7}$$

Also

$$C(u,0) = P(U \le u, V \le 0) = 0 \tag{2.8}$$

$$C(0,v) = P(U \le 0, V \le v) = 0 \tag{2.9}$$

It follows from 2.8 and 2.9 that $C$ is grounded.

## 2.2 Rectangular inequality

Since $C(u,v)$ is a distribution function, it satisfies for all $0 \le u_1 < u_2 \le 1$ and $0 \le v_1 < v_2 \le 1$

$$\begin{aligned} & P(u_1 < U \le u_2, v_1 < V \le v_2) \\ &= C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) > 0 \end{aligned} \tag{2.10}$$

i.e. $C$ is 2-increasing. When $C(u,v)$ has a density $c(u,v)$, this inequality becomes $c(u,v) > 0$.

Given these two properties one can state the equivalent definition of copulae.

**DEFINITION 2.2.** *A two-dimensional copula is a function $C$ defined on the unit square $I^2 = I \times I$ with $I$ the unit interval ($I = [0,1]$) such that*

- *for every $u \in I$ holds: $C(u,0) = C(0,v) = 0$, i.e. $C$ is grounded.*
- *for every $u_1, u_2, v_1, v_2 \in I$ with $u_1 \le u_2$ and $v_1 \le v_2$ holds:*

  $$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0, \tag{2.11}$$

  *i.e. $C$ is 2-increasing.*
- *for every $u \in I$ holds $C(u,1) = u$ and $C(1,v) = v$.*

Informally, a copula is a joint distribution function defined on the unit square $[0,1]^2$ which has uniform marginals. That means that if $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$ are univariate distribution functions, then $C\{F_{X_1}(x_1), F_{X_2}(x_2)\}$ is a 2-dimensional distribution function with uniform marginals $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$.

## 2.3 Continuity

A copula is continuous in $u$ and $v$; actually it satisfies the stronger Lipschitz condition:

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1| \tag{2.12}$$

The inequation follows directly from Lemma 1.3 by setting $F(x_1) = u_1$, $F(x_2) = u_2$, $F(y_1) = v_1$, $F(y_2) = v_2$ in inequation 1.11. From 2.12 it follows that every copula $C$ is uniformly continuous on its domain.

Another important property of copulae concerns the partial derivatives of a copula with respect to its variables.

## 2.4 Differentiability

Since $C(u, v)$ is increasing and continuous in the two variables, it is differentiable almost everywhere and we see that

$$0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1 \tag{2.13}$$

and

$$0 \leq \frac{\partial}{\partial v} C(u, v) \leq 1 \tag{2.14}$$

Moreover, the functions

$$u \mapsto C_v(u) = \partial C(u, v)/\partial v \text{ and}$$

$$v \mapsto C_u(v) = \partial C(u, v)/\partial u$$

are defined and nonincreasing almost everywhere on $I = [0, 1]$.

## 2.5 Invariance under strictly monotone transformations

Copulae are invariant under strictly monotone transformations of the random variables.

**THEOREM 2.2.** *If $(X_1, X_2)$ have a copula $C$ and set $g_1$, $g_2$ two continuous increasing functions, then $\{g_1(X_1), g_2(X_2)\}$ have a copula $C$, too.*

**Proof:**

Let $F_1, F_2$ denote the distribution functions of $X_1, X_2,$ and $G_1, G_2$ denote the distribution functions of $g_1(X_1), g_2(X_2)$ respectively. Let $(X_1, X_2)$ have a copula $C$ and $\{g_1(X_1), g_2(X_2)\}$ have a copula $C_g$. Then

$$\begin{aligned} C_g\{G_1(x_1), G_2(x_2)\} &= P\{g_1(X_1) \leq x_1, g_2(X_2) \leq x_2\} \\ &= P\left\{X_1 \leq g_1^{-1}(x_1), X_2 \leq g_2^{-1}(x_2)\right\} \\ &= C\left[F_1\left\{g_1^{-1}(x_1)\right\}, F_2\left\{g_2^{-1}(x_2)\right\}\right] \\ &= C\{G_1(x_1), G_2(x_2)\} \end{aligned}$$

□

## 2.6 The survival function of copulae

Using 2.6 and 2.7 one can define the survival function $\overline{C}$ corresponding to $C(u, v)$:

$$\overline{C}(u, v) = 1 - u - v + C(u, v) \tag{2.15}$$

From 2.6 and 2.7 we have

$$\overline{C}(u, 1) = 0 \tag{2.16}$$

and considering 2.8 and 2.9 we obtain

$$\overline{C}(u, 0) = 1 - u \tag{2.17}$$

Starting from the pair (1-U,1-V) one can define another copula $C^{'}(u, v)$ whose survival function is connected with $C$. Namely, given

$$C^{'}(u, v) = P(1 - U \leq u, 1 - V \leq v) \tag{2.18}$$

$$= \overline{C}(1 - u, 1 - v) = u + v - 1 + C(1 - u, 1 - v)$$

we then obtain the survival function $\overline{C^{'}}$ corresponding to $C^{'}(u, v)$:

$$\overline{C^{'}}(u, v) = C(1 - u, 1 - v) \tag{2.19}$$

## 2.7 The Fréchet bounds for copulae

Each copula function is bounded by the minimum and maximum one

$$C^{-}(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = C^{+}(u, v) \tag{2.20}$$

The functions $\max(u + v - 1, 0)$ and $\min(u, v)$ can be easily checked to be copula functions. They are called the minimum and the maximum copula, respectively. The minimum and the maximum copulae are assumed to be an *upper Fréchet bound* and a *lower Fréchet bound* for copulae, respectively.

## 2.8 Convex combination of copulae

A linear convex combination of copulae is a copula. For example if $0 \leq \alpha \leq 1$, then the function

$$C(u, v) = \alpha C^{+}(u, v) + (1 - \alpha) C^{-}(u, v) \tag{2.21}$$

is a copula function.

# 3 Modelling dependence with copulae

Dependence relations between random variables are one of the most widely studied topics in probability theory and statistics. Copulae are a very important part in the study of dependence between two variables, since they allow to separate the effect of dependence from effects of the marginal distributions. As we have already seen, all copula properties are invariant under strictly increasing transformations of the underlying random variables.

## 3.1 The Bravais-Pearson correlation coefficient

The Bravais-Pearson correlation coefficient (or simple linear correlation) is the most frequently used measure of dependence. Let $X$, $Y$ be two random variables with $Var(X) < \infty$, $Var(X) \neq 0$ and $Var(Y) < \infty$, $Var(Y) \neq 0$. The Bravais-Pearson correlation coefficient for $X$, $Y$ is defined as

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} \tag{3.1}$$

The Bravais-Pearson correlation coefficient is a measure of linear dependence. In the case of perfect linear dependence we have $|\rho(X,Y)| = 1$, otherwise it holds $-1 < \rho(X,Y) < 1$. Linear correlation is invariant under strictly increasing *linear* transformations, i.e. for $a, b \in \mathbb{R}\backslash 0$ and $c, d \in \mathbb{R}$ it holds:

$$\rho(aX + c, bY + d) = sgn(ab)\rho(X,Y) \tag{3.2}$$

As it was already mentioned, linear correlation is a rather imperfect measure of dependence. In (Embrechts, Lindskog and McNeil, 2001) is shown, that linear correlation fails in the case if the underlying variables are not jointly elliptically distributed (such as multivariate normal or multivariate t-distribution).

## 3.2 Perfect dependence with copulae

As we have already seen, each copula function is bounded by the minimum and maximum one

$$C^-(u,v) = \max(u+v-1,0) \le C(u,v) \le \min(u,v) = C^+(u,v) \qquad (3.3)$$

and the functions $C^-(u,v)$ and $C^+(u,v)$, known as Fréchet bounds, are also copula functions. In (Embrechts, Lindskog and McNeil, 2001) is shown, that in the 2-dimensional case $C^-(u,v)$ and $C^+(u,v)$ are the bivariate distributions functions of the random vectors $(U,1-U)^T$ and $(U,U)^T$ respectively, where $U \sim U[0,1]$ (i.e. uniformly distributed on the unit interval $[0,1]$). In this case one says that $C^-(u,v)$ describes perfect negative dependence and $C^+(u,v)$ describes perfect positive dependence.

## 3.3 Lower and upper tail dependence

The concept of lower and upper tail dependence refers to the study of dependence between extreme values. In the case of copulae, where $(U,V)$ is a pair of uniform variables on the unit square $[0,1]^2$, the upper tail dependence is defined as

$$\delta = \lim_{u \to 1-} P(U > u | V > v) \qquad (3.4)$$

or equivalent definition, using survival function of a copula:

$$\delta = \lim_{u \to 1-} \frac{\overline{C}(u,u)}{1-u} > 0 \qquad (3.5)$$

If the coefficient of upper tail dependence $\delta \in (0,1]$, $U$ and $V$ are said to be asymptotically dependent in the upper tail; if $\delta = 0$, $U$ and $V$ are said to be asymptotically independent in the upper tail. Similary, lower tail dependence holds if $\frac{C(u,u)}{u}$ has a limit different from 0 when $u$ tends to 0:

$$\gamma = \lim_{u \to 0+} \frac{C(u,u)}{u} > 0 \qquad (3.6)$$

or equivalent

$$\gamma = \lim_{u \to 0+} P(U \le u | V \le v) \qquad (3.7)$$

If $\gamma = 0$, $U$ and $V$ are said to be asymptotically independent in the lower tail. There is a connection between the tail dependence of $C$ and of the copula $C^{'}(u,v) = \overline{C}(1-u,1-v)$. The lower tail dependence of $C$ is the upper tail dependence of $C^{'}$

and vice-versa:

$$\lim_{u\to 1-}\frac{\overline{C}(u,u)}{1-u} = \lim_{u\to 0}\frac{\overline{C}(1-u,1-u)}{u} = \lim_{u\to 0}\frac{C^{'}(u,u)}{u} \tag{3.8}$$

The examples of the lower and upper tail dependence are discussed in the next chapter.

# 4 Examples

**EXAMPLE 4.1.** *The independent or product copula is a copula associated with a pair $(U, V)$ of independent variables.*

$$C^0(u,v) = uv \tag{4.1}$$

**EXAMPLE 4.2.** *Let $(X, Y)$ be a pair of bivariate normal distributed random varibles: $(X, Y) \sim N((0,0), \Sigma)$ with $\Sigma = \binom{1\rho}{\rho 1}$. Let $\phi$ and $\Phi$ be the density and the cdf, respectively, of the $N(0,1)$ and $\phi_\rho$ the density of the Bivariate Normal Distribution (BVN) distribution. Let $U = \Phi(X)$ and $U = \Phi(Y)$ be two uniform variables with densities $c$ and cumulative distribution function $C$. Then*

$$C_\rho^{Gauss}(u,v) = \Phi_\rho \left\{\Phi^{-1}(u), \Phi^{-1}(v)\right\} \tag{4.2}$$

*is a copula associated with the Bivariate Normal Distribution or, in another term, a Gaussian copula. The density of the Gaussian copula is given by:*

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v} = \frac{\phi_\rho \left\{\Phi^{-1}(u), \Phi^{-1}(v)\right\}}{\phi\left\{\Phi^{-1}(u)\right\}\phi\left\{\Phi^{-1}(v)\right\}} \tag{4.3}$$

$$= (1/\sqrt{1-\rho^2}) \cdot \exp\left[-\frac{\rho^2}{2(1-\rho^2)}\left\{\Phi^{-1}(u)\right\}^2 + \left\{\Phi^{-1}(v)\right\}^2 + \rho\Phi^{-1}(u)\Phi^{-1}(v)\right]$$

*The copula dependence parameter $\theta$ is the collection of all the unknown correlation coefficients in $\Sigma$. If $\theta = 0$ ($\rho = 0$), i.e. vanishing correlation, the Gaussian copula reduces to the product copula:*

$$\begin{aligned} C_0^{Gauss}(u,v) &= \int_{-\infty}^{\phi_1^{-1}(u)} f(x_1)dx_1 \int_{-\infty}^{\phi_2^{-1}(v)} f(x_2)dx_2 \\ &= uv \\ &= \Pi(u,v) \text{ if } \rho = 0 \end{aligned} \tag{4.4}$$

*If $\theta \neq 0$, then the corresponding Gaussian copula generates joint symmetric dependence, but no tail dependence (i.e., there are no joint extreme events). In the figure 4.1 and figure 4.2 the density function of a Gaussian copula and the density of $F(x_1, x_2)$ with $\rho = 0$ are represented.*

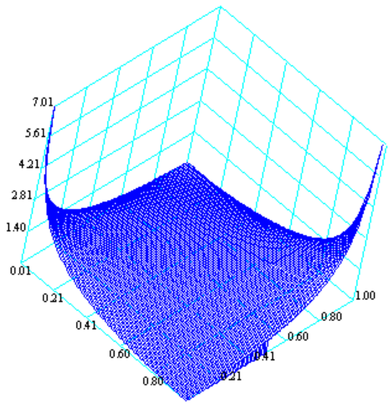


Figure 4.1. Density of a Gaussian copula with $\rho = 0.5$.

ga.xpl

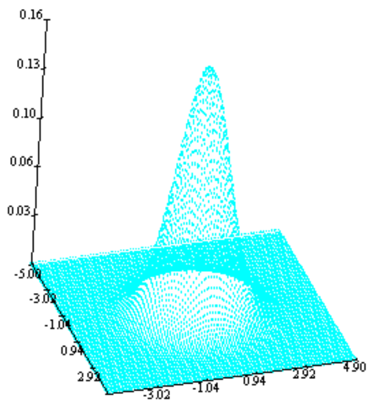


Figure 4.2. Density function of $F(x_1, x_2)$ with $\rho = 0$.

gapdf.xpl

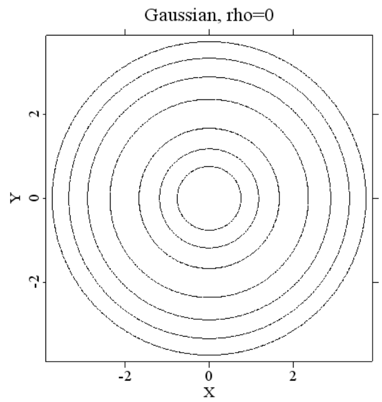


Figure 4.3. Contour plot of the Gaussian copula, $\rho = 0$.
gauscont.xpl

A simple, but useful way to represent the shape of a copula is the contour diagram, that is, graphs of its level sets - the sets in $I^2 = [0,1]^2$ given by $C(u,v) = const$ , for selected constants in $I = [0,1]$. In the figure 4.3 we present the countour diagrams of the Gaussian copula in the case of independence, $\rho = 0$.

**EXAMPLE 4.3.** *One can easily verify that the function*

$$C_\theta^{GH}(u,v) = \exp\left[-\left\{(-\ln u)^\theta + (-\ln v)^\theta\right\}^{1/\theta}\right] \tag{4.5}$$

*is a copula function, a so called Gumbel-Hougaard copula. The parameter $\theta$ may take values in the interval $[1,\infty)$. This copula allows to generate upper tail dependence. Indeed,*

$$\begin{aligned} \delta &= \lim_{u\to 1-}\frac{\overline{C}(u,u)}{1-u} = \lim_{u\to 1-}\frac{1-2u+C(u,u)}{1-u} \\ &= \lim_{u\to 1-}\frac{1-2u+\exp(2^{1/\theta}\ln u)}{1-u} = \lim_{u\to 1-}\frac{1-2u+u^{2^{1/\theta}}}{1-u} \\ &= 2-\lim_{u\to 1-}2^{1/\theta}u^{2^{1/\theta}-1} = 2-2^{1/\theta} \end{aligned} \tag{4.6}$$

*Thus, for $\theta > 1$, the Gumbel-Hougaard copula has an upper tail dependence. For $\theta = 1$, the Gumbel-Hougaard copula reduces to the product copula, i.e.*

$$C_1(u,v) = \Pi(u,v) = uv \tag{4.7}$$

*For $\theta \to \infty$, one finds for the Gumbel-Hougaard copula:*

$$C_\theta(u,v) \longrightarrow \min(u,v) = C^+(u,v) \tag{4.8}$$

*Figure 4.4 represents a density of a Gumbel-Hougaard copula with $\theta = 2.5$.*

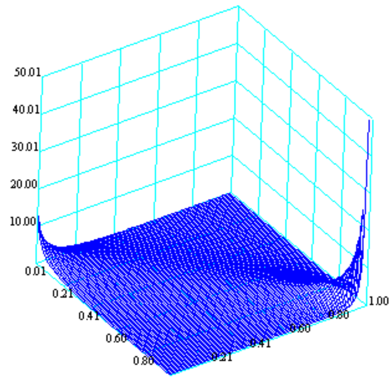


Figure 4.4. Density of a Gumbel-Hougaard copula with $\theta = 2.5$.

gh.xpl

**EXAMPLE 4.4.** *The copula*

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \tag{4.9}$$

*with parameter $\theta > 0$ is assumed to be the Clayton copula. The density function of the Clayton copula is given by*

$$c_\theta(u, v) = (1 + \theta) u^{-(\theta+1)} v^{-(\theta+1)} (u^{-\theta} + v^{-\theta} - 1)^{-(1/\theta+2)} \tag{4.10}$$

*As the parameter $\theta$ tends to infinity, dependence becomes maximal and as $\theta$ tends to zero, the pair $(U, V)$ becomes independent. As $\theta \to 1$, the distribution tends to the lower Fréchet bound. Unlike the Gaussian copula, the Clayton copula can generate asymmetric dependence and lower tail dependence, but no upper tail dependence that is the limit $\delta = \lim_{u \to 1-} \frac{\overline{C}(u,u)}{1-u} = 0$. The density of a Clayton copula for the case of $\theta = 1.5$ is represented by the figure 4.5.*

Figure 4.6 represents the countour diagrams of the Gumbel-Hougaard and Clayton copulae with parameter $\theta = 1.5$.

Recall that the Gumbel-Hougaard copula generates an upper tail dependence and the Clayton copula is able to model a lower tail dependence. Figure 4.7 represents an example of Gumbel-Hougaard and Clayton copulae sampling with fixed marginal parameters $\sigma_1 = 1$, $\sigma_2 = 1$ and $\theta = 1.5$.

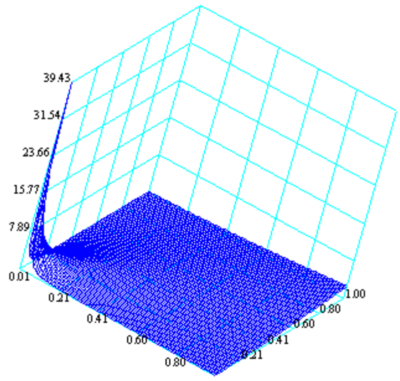


Figure 4.5. Density of a Clayton copula with $\theta = 1.5$

clayton.xpl

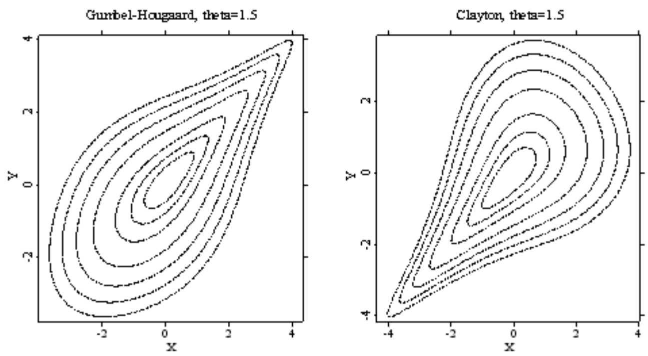


Figure 4.6. Contour plots of the Gumbel-Hougaard and Clayton copulae, $\theta = 1.5$

ghccont.xpl

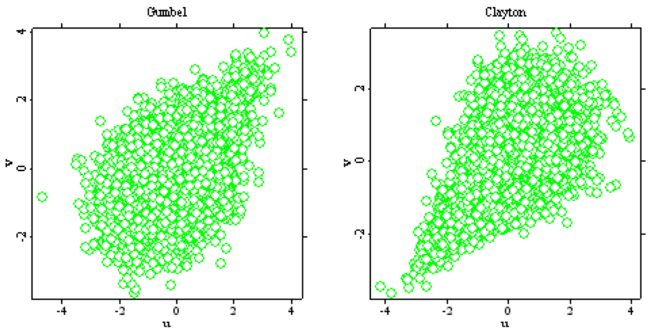


Figure 4.7. 10000-sample output with fixed $\sigma_1 = 1$, $\sigma_2 = 1$ of the Gumbel-Hougaard and Clayton copulae, $\theta = 1.5$

ghcsampleout.xpl

**EXAMPLE 4.5.** *The function*

$$C_\theta(u, v) = \frac{uv}{1 - \theta(1-u)(1-v)} \tag{4.11}$$

*represents the Ali-Mikhail-Haq family of copulae, $|\theta| \leq 1$. If $\theta = 0$, then we have an independence: $C_\theta(u, v) = C^0$. This family does not contain the Fréchet bounds. The density of an AMH copula with $\theta = 0.8$ is represented by figure 4.8.*

**EXAMPLE 4.6.** *The Frank copula with dependence parameter $0 < \theta \leq \infty$ is represented by the function:*

$$C_\theta(u, v) = -\frac{1}{\theta} \log \left\{ 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right\} \tag{4.12}$$

*The dependence becomes maximal when $\theta$ tends to infinity and independence is achieved when $\theta = 0$. The density of a Frank copula for the case of $\theta = 5.0$ is given by figure 4.9.*

Figure 4.10 represents the countour diagrams of the Ali-Mikhail-Haq copula with parameter $\theta = 0.8$ and of the Frank copula with parameter $\theta = 5.0$.

Figure 4.11 represents an example of Ali-Mikhail-Haq and Frank copulae sampling with fixed marginal parameters $\sigma_1 = 1$, $\sigma_2 = 1$ and copulae dependence parameters of $\theta = 0.8$ and $\theta = 5.0$ respectively.

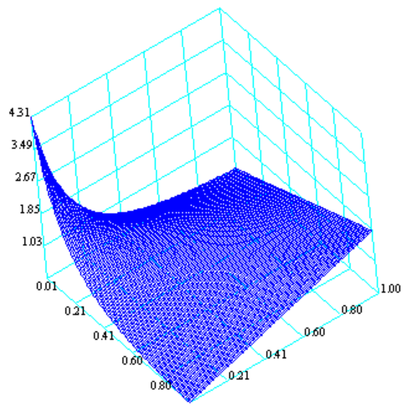


Figure 4.8. Density of a AMH copula with $\theta = 0.8$

amh.xpl

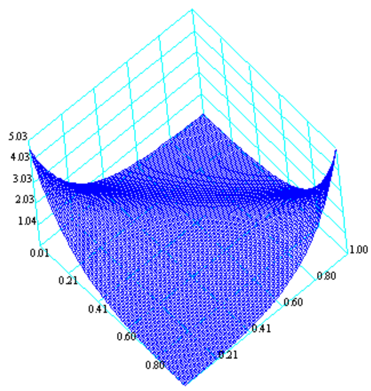


Figure 4.9. Density of a Frank copula with $\theta = 5.0$

frank.xpl

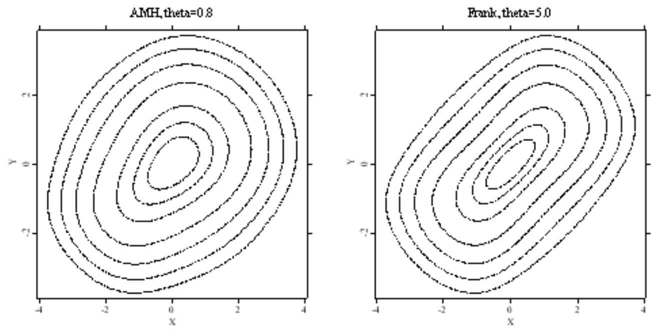


Figure 4.10. Contour plots of the AMH copula with $\theta = 0.8$ and Frank copula with $\theta = 5.0$

amhfcont.xpl

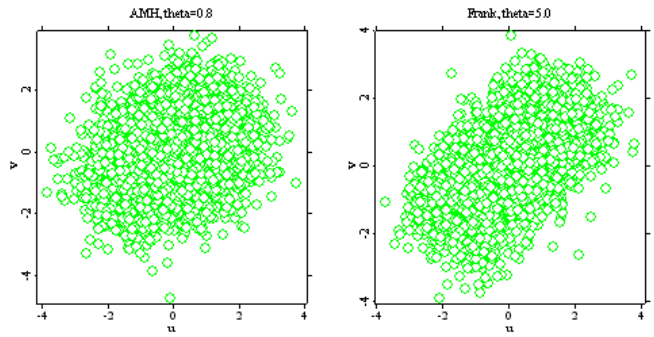


Figure 4.11. 10000-sample output with fixed $\sigma_1 = 1$, $\sigma_2 = 1$ of the AMH copula, $\theta = 0.8$ and Frank copula, $\theta = 5.0$

amhfsampleout.xpl

Until now, we have considered copulae only in a 2-dimensional setting. Let us now extend this concept to the d-dimensional case, for a random variable in $\mathbb{R}^d$ with $d \geq 1$.

# 5 Multivariate copulae

Let $U_1, U_2, ..., U_d$ be nonempty sets in $\overline{\mathbb{R}}$ and consider the function $F : U_1 \times U_2 \times ... \times U_d \longrightarrow \overline{\mathbb{R}}$. For $a = (a_1, a_2, ..., a_d)$ and $b = (b_1, b_2, ..., b_d)$ with $a \leq b$ (i.e. $a_k \leq b_k$ for all $k$) let $B = [a, b] = [a_1, b_1] \times [a_2, b_2] \times ... \times [a_n, b_n]$ be the d-box with vertices $c = (c_1, c_2, ..., c_d)$. It is obvious, that each $c_k$ is either equal to $a_k$ or to $b_k$.

**DEFINITION 5.1.** *The F-volume of a d-box $B = [a, b] = [a_1, b_1] \times [a_2, b_2] \times ... \times [a_d, b_d] \subset U_1 \times U_2 \times ... \times U_d$ is defined as follows:*

$$V_F(B) = \sum_{k=1}^{d} sgn(c_k)F(c_k) \tag{5.1}$$

*where $sgn(c_k) = 1$, if $c_k = a_k$ for even $k$ and $sgn(c_k) = -1$, if $c_k = a_k$ for odd $k$.*

**EXAMPLE 5.1.** *For the case of $d = 3$, the F-volume of a 3-box $B = [a, b] = [x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$ is defined as:*

$$\begin{aligned} V_F(B) = & F(x_2, y_2, z_2) - F(x_2, y_2, z_1) - F(x_2, y_1, z_2) - F(x_1, y_2, z_2) \\ & + F(x_2, y_1, z_1) + F(x_1, y_2, z_1) + F(x_1, y_1, z_2) - F(x_1, y_1, z_1) \end{aligned}$$

**DEFINITION 5.2.** *F is said to be a d-increasing function if for all d-boxes B with vertices in $U_1 \times U_2 \times ... \times U_d$ holds:*

$$V_F(B) \geq 0 \tag{5.2}$$

**DEFINITION 5.3.** *If $U_1, U_2, ..., U_d$ have a smallest element* $\min U_1$, $\min U_2$,..., $\min U_d$, *respectively, then we say, that a function $F : U_1 \times U_2 \times ... \times U_d \longrightarrow \overline{\mathbb{R}}$ is grounded if :*

$$F(x) = 0 \text{ for all } x \in U_1 \times U_2 \times ... \times U_d \tag{5.3}$$

*such that $x_k = \min U_k$ for at least one $k$.*

While the lemmas, which we have presented for the 2-dimensional case, have analogous multivariate versions (Nelsen, 1999), we will leave them out here.

**DEFINITION 5.4.** *A d-dimensional copula (or d-copula) is a function $C$ defined on the unit d-cube $I^d = I \times I \times ... \times I$ such that*

- *for every $u \in I^d$ holds: $C(u) = 0$, if at least one coordinate of $u$ is equal to 0; i.e. $C$ is grounded.*
- *for every $a, b \in I^d$ with $a \leq b$ holds:*

$$V_C([a,b]) \geq 0; \tag{5.4}$$

  *i.e. $C$ is d-increasing.*
- *for every $u \in I^d$ holds: $C(u) = u_k$, if all coordinates of $u$ are 1 except $u_k$.*

Analogously to the 2-dimensional setting, let us state Sklar's theorem for the d-dimensional case.

**THEOREM 5.1.** *Sklar's Theorem:*
*If the marginals $F_1, ...F_d$ of the multiveriate distribution function $F$ are continuous, there exists a unique copula $C : [0,1]^d \rightarrow [0,1]$ such that*

$$F(x_1, ..., x_d) = C\left\{F_{X_1}(x_1), ..., F_{X_d}(x_d)\right\} \tag{5.5}$$

*Conversely, if $C$ is a copula and $F_{X_1}, F_{X_2}, ..., F_{X_d}$ are distribution functions then $F$ defined by 5.5 is a d-dimensional distribution function with marginals $F_{X_1}, F_{X_2}, ..., F_{X_d}$.*

If the density $f$ of $F$ exists, one can derive the relationship between the density $f$ of $F$ and $c$ of $C$:

$$f(x_1, ..., x_d) = c\left\{F_{X_1}(x_1), ..., F_{X_d}(x_d)\right\} \prod_{j=1}^{d} f_j(x_j) \tag{5.6}$$

where the copula density is

$$c(u_1, ..., u_d) = \frac{\partial^d C(u_1, ..., u_d)}{\partial u_1 ... \partial u_d} \tag{5.7}$$

and $u_j = F_{X_j}(x_j)$, $f_j(x_j) = F^{'}_{X_j}(x_j)$.

In order to illustrate the d-copulae we present the following examples:

**EXAMPLE 5.2.** *Let $\Phi$ denote the univariate standard normal distribution function and $\Phi_{\Sigma,d}$ the d-dimensional standard normal distribution function with correlation matrix $\Sigma$. Then a function*

$$\begin{aligned} C^{Gauss}(u_1, ..., u_d, \Sigma) &= \Phi_{\Sigma,d}\left\{\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right\} \\ &= \int_{-\infty}^{\phi_1^{-1}(u_d)} ... \int_{-\infty}^{\phi_2^{-1}(u_1)} f_\Sigma(x_1, ..., x_n) dx_1 ... dx_d \end{aligned} \tag{5.8}$$

*is a d-dim Gaussian copula. The density of the d-dim Gaussian copula is given by*

$$f_(u_1, ..., u_d, \Sigma) = \frac{\partial^d C(u1, ..., u_d)}{\partial u_1, ..., \partial u_d} = \frac{1}{\sqrt{\det(\Sigma)}} \quad \times \tag{5.9}$$

$$\times \quad \exp\left\{-\frac{(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d))'(\Sigma^{-1} - \mathcal{I}_d)(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d))}{2}\right\}$$

**EXAMPLE 5.3.** *A d-dim Gumbel-Hougaard copula with dependence parameter $\theta$ from the interval $[1, \infty)$ is a function:*

$$C_\theta^{GH}(u_1, ..., u_d) = \exp\left[-\left\{\sum_{j=1}^{d}(-\log u_j)^\theta\right\}^{1/\theta}\right] \tag{5.10}$$

*If $\theta = 1$, Gumbel-Hougaard copula reduces to the d-dimensional product copula, i.e.*

$$C_1(u_1, ..., u_d) = \prod_{j=1}^{d} u_j = \Pi^d(u) \tag{5.11}$$

**EXAMPLE 5.4.** *A d-dim Frank copula with parameter $\theta > 0$ is given by*

$$C_\theta(u_1, ..., u_d) = -(1/\theta)\log\left\{1 + \frac{\prod_{i=1}^{d}(e^{-\theta u_i} - 1)}{(e^{-\theta} - 1}) ^{d-1}\right\} \tag{5.12}$$

**EXAMPLE 5.5.** *A d-dim Ali-Mikhail-Haq copula with $-1 \leq \theta < 1$ is defined by*

$$C_\theta(u_1, ..., u_d) = \frac{\prod_{i=1}^{d} u_i}{1 - \theta\left\{\prod_{i=1}^{d}(1 - u_i)\right\}} \tag{5.13}$$

**EXAMPLE 5.6.** *A d-dim Clayton copula with copula dependence parameter $\theta > 0$ is given by*

$$C_\theta(u_1, ..., u_d) = (u_1^{-\theta} + ... + u_d^{-\theta} - d + 1)^{-1/\theta} \tag{5.14}$$

*and its density function is*

$$c_\theta(u_1, ..., u_d) = \prod_{j=1}^{d}\{1 + (j-1)\theta\}\prod_{j=1}^{d} u_j^{-(\theta+1)}\left\{\sum_{j=1}^{d} u_j^{-\theta} - d + 1\right\}^{-(1/\theta+d)} \tag{5.15}$$

# 6 Copulae estimation

Consider a vector of random variables: $X = (X_1, ..., X_d)^T$. Let $F_{X_1}(x_1, \delta_1)$,..., $F_{X_d}(x_d, \delta_d)$ denote the distribution functions of $X_1, ..., X_d$. Recall from Sklar's theorem, that we can write the following copula-based model for the distribution of the vector X:

$$F(x_1, ..., x_d; \delta_1, ..., \delta_d; \theta) = C\left\{F_{X_1}(x_1, \delta_1), ..., F_{X_2}(x_2, \delta_2); \theta\right\} \tag{6.1}$$

Suppose we want to estimate this model, i.e. we want to estimate the copula dependence parameter $\theta$ and the parameters from the marginals $\delta_1, ..., \delta_d$.
Assume that the density of the copula $C$ is given by $c$:

$$c(u_1, ..., u_d; \theta) = \frac{\partial^d C(u_1, ..., u_d; \theta)}{\partial u_1 ... \partial u_d} \tag{6.2}$$

If the density $f$ of $F$ exists, one can derive the following relationship between the density $f$ of $F$ and $c$ of $C$:

$$f(x_1, ..., x_d; \delta_1, ..., \delta_d; \theta) = c\left\{F_{X_1}(x_1, \delta_1), ..., F_{X_d}(x_d, \delta_d); \theta\right\} \prod_{j=1}^{d} f_j(x_j; \delta_j) \tag{6.3}$$

Suppose that we have $n$ i.i.d. $d$-dimensional vectors of observations $(x_1, ..., x_n)$, i.e. $\{x_i\}_{i=1}^n$ with $(x_{1,i}, ..., x_{d,i})^T$. The "full" likelihood function for $\{x_i\}_{i=1}^n$ is then given by

$$L(x_1, ..., x_d; \delta_1, ..., \delta_d; \theta) = \prod_{i=1}^{n} f(x_{1,i}, ..., x_{d,i}; \delta_1, ..., \delta_d; \theta) \tag{6.4}$$

and the corresponding "full" log-likelihood function is given by

$$l(x_1, ..., x_d; \delta_1, ..., \delta_d; \theta)$$

$$= \sum_{i=1}^{n} \left[\log c\left\{F_{X_1}(x_{1,i}, \delta_1), ..., F_{X_d}(x_{d,i}, \delta_d); \theta\right\} + \sum_{j=1}^{d} \log f_j(x_{j,i}; \delta_j)\right] \tag{6.5}$$

$$= \sum_{i=1}^{n} \log c\left\{F_{X_1}(x_{1,i}, \delta_1), ..., F_{X_d}(x_{d,i}, \delta_d); \theta\right\} + \sum_{i=1}^{n} \sum_{j=1}^{d} \log f_j(x_{j,i}; \delta_j)$$

Our objective is to maximize this log-likelihood numerically.

## 6.1 Full maximum likelihood estimation

In the full maximum likelihood (FML) procedure the marginal parameters $\delta_1, ..., \delta_d$ and the copula dependence parameter $\theta$ are estimated simultaneously. The estimates vector $\left(\hat{\delta}_1, ..., \hat{\delta}_d, \hat{\theta}\right)^T$ is a solution of

$$\left(\frac{\partial l}{\partial \delta_1}, ..., \frac{\partial l}{\partial \delta_d}, \frac{\partial l}{\partial \theta}\right) = 0 \tag{6.6}$$

The drawback of the FLM estimation is that with an increasing scale of the problem, the algorithm becomes too burdensome computationally.

## 6.2 The inference for margins method

In the inference for margins (IFM) method, the marginal parameters $\delta_1, ..., \delta_d$ are estimated in the first step. The estimator of the copula dependence parameter is obtained in the second step by substituting $\hat{\delta}_1, ..., \hat{\delta}_d$ in the "full" log-likelihood function and by numerical maximisation with respect to $\theta$.

1. The estimates vector $(\hat{\delta}_1, ..., \hat{\delta}_d)$ is obtained by maximising the log-likelihood function for each marginal:

$$l_j(\delta_j) = \sum_{i=1}^{n} f_j(x_{j,i}; \delta_j) \text{ for } j = 1, ..., d, \tag{6.7}$$

   i.e. for $j = 1, ..., d$ the estimates are given by

$$\hat{\delta}_j = \arg\max_{\delta} l_j(\delta_j) \tag{6.8}$$

2. Substitute these marginal estimates in the "full" log-likelihood and maximize

$$l(\hat{\delta}_1, ..., \hat{\delta}_d; \theta) = \sum_{i=1}^{n} \ln c\left\{F_{X_1}(x_1, \hat{\delta}_1), ..., F_{X_d}(x_d, \hat{\delta}_d); \theta\right\} \tag{6.9}$$

   with respect to $\theta$, i.e. the estimator of the copula dependence parameter $\theta$ is given by

$$\hat{\theta} = \arg\max_{\theta} l(\hat{\delta}_1, ..., \hat{\delta}_d; \theta) \tag{6.10}$$

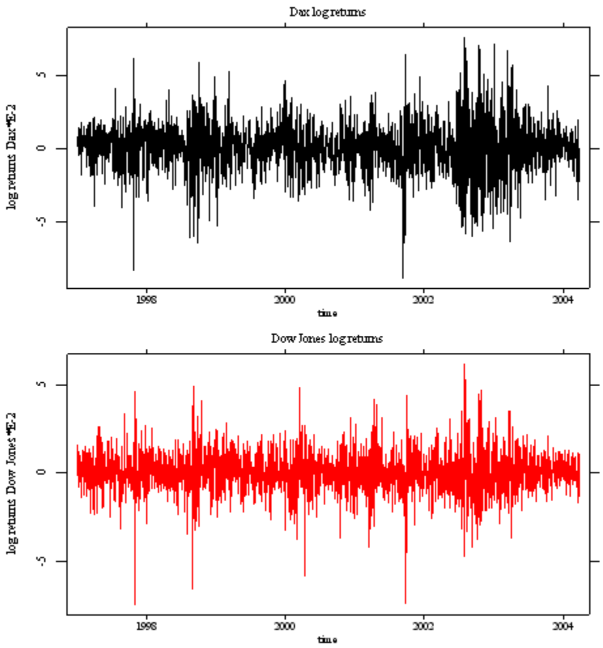


Figure 6.1. Dax and Dow Jones log-returns for the period January 1, 1997 to March 20, 2004 (1573 observations).

daxdowlogreturns.xpl

## 6.3 Application of the IFM procedure

The IFM procedure is applied to daily observations of a portfolio consisting of the Dax and the Dow Jones for the period January 1, 1997 to March 20, 2004 (1823 observations). The log-returns are shown the picture 6.1. We consider a moving time window with $n = 250$ observations, i.e. the parameter estimates at each point are obtained from the last 250 observations. The univariate margins (log-returns) are assumed to be normally distibuted $X_{j,i} \sim N(0, \sigma_j^2)$, $j = 1, 2$ with parameters $\delta_j = \sigma_j^2$ estimated from the data. Figure 6.2 and 6.3 represent the density of the Dax and Dow Jones respectively, estimated nonparametrically using Quartic kernel with $\widehat{h} = 1.06\widehat{\sigma}n^{-\frac{1}{5}}$.

To make procedure running, we have to specify the copula function. Selected copula belongs to the Gumbel-Hougaard family of copulae:

$$C_\theta^{GH}(u, v) = \exp\left[-\left\{(-\log u)^\theta + (-\log v)^\theta\right\}^{1/\theta}\right] \tag{6.11}$$

Recall, that for this copula parameter $\theta$ may take values in the interval $[1, \infty)$. Independence is achieved if $\theta = 1$.
According to the IFM procedure, we estimate at first parameters from the marginal distributions. Figure 6.4 represents estimates of parameters $\delta_1$ from the Dax and

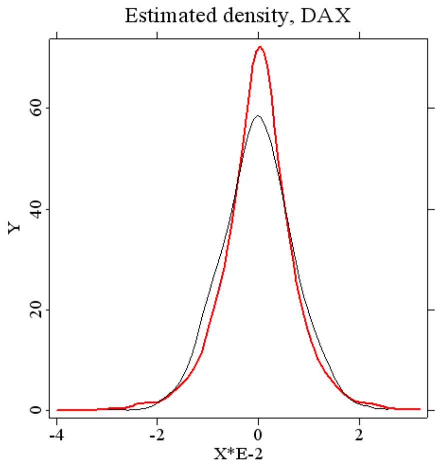


Figure 6.2. Density of the Dax log-returns (red) and normal density (black), estimated nonparametrically using Quartic kernel with $\widehat{h} = 1.06\widehat{\sigma}n^{-\frac{1}{5}}$ .

kernelest.xpl

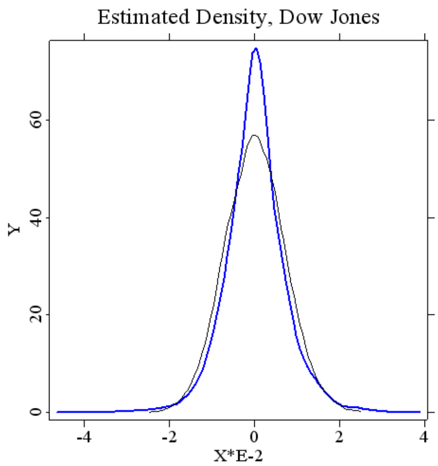


Figure 6.3. Density of the Dow Jones log-returns (blue) and normal density (black), estimated nonparametrically using Quartic kernel with $\widehat{h} = 1.06\widehat{\sigma}n^{-\frac{1}{5}}$ .

kernelest.xpl

| | min | max | mean | median | std error |
|---|---|---|---|---|---|
| $\hat{\delta}_1$ | 0.013345 | 0.029246 | 0.018396 | 0.017827 | 0.0042519 |
| $\hat{\delta}_2$ | 0.008681 | 0.016996 | 0.012817 | 0.012586 | 0.0017591 |
| $\hat{\theta}$ | 1 | 1.12866 | 1.033 | 0.0195815 | 0.034162 |

Table 6.1. Descriptive statistics for $\hat{\delta}_1$, $\hat{\delta}_2$, $\hat{\theta}$

$\delta_2$ from the Dow Jones log-returns (upper respectively middle panel) and copula dependence parameter $\theta$ between Dax and Dow Jones log-returns (lower panel). Descriptive statistics for $\hat{\delta}_1$, $\hat{\delta}_2$ and $\hat{\theta}$ are given by the table 6.1.

Estimates of the copula dependence parameter $\theta$ reaches its minimum equal to 1, which indicates independence for the Gumbel-Hougaard copula, at the time point corresponding to November 30, 2000. The maximum is reached at the time point corresponding to March 11, 2004. Figures 6.5 and 6.6 represent the last 250 realizations of the Dax and Dow Jones log returns, corresponding to minimal (November 30, 2000) and maximal (March 11, 2004) dependence.

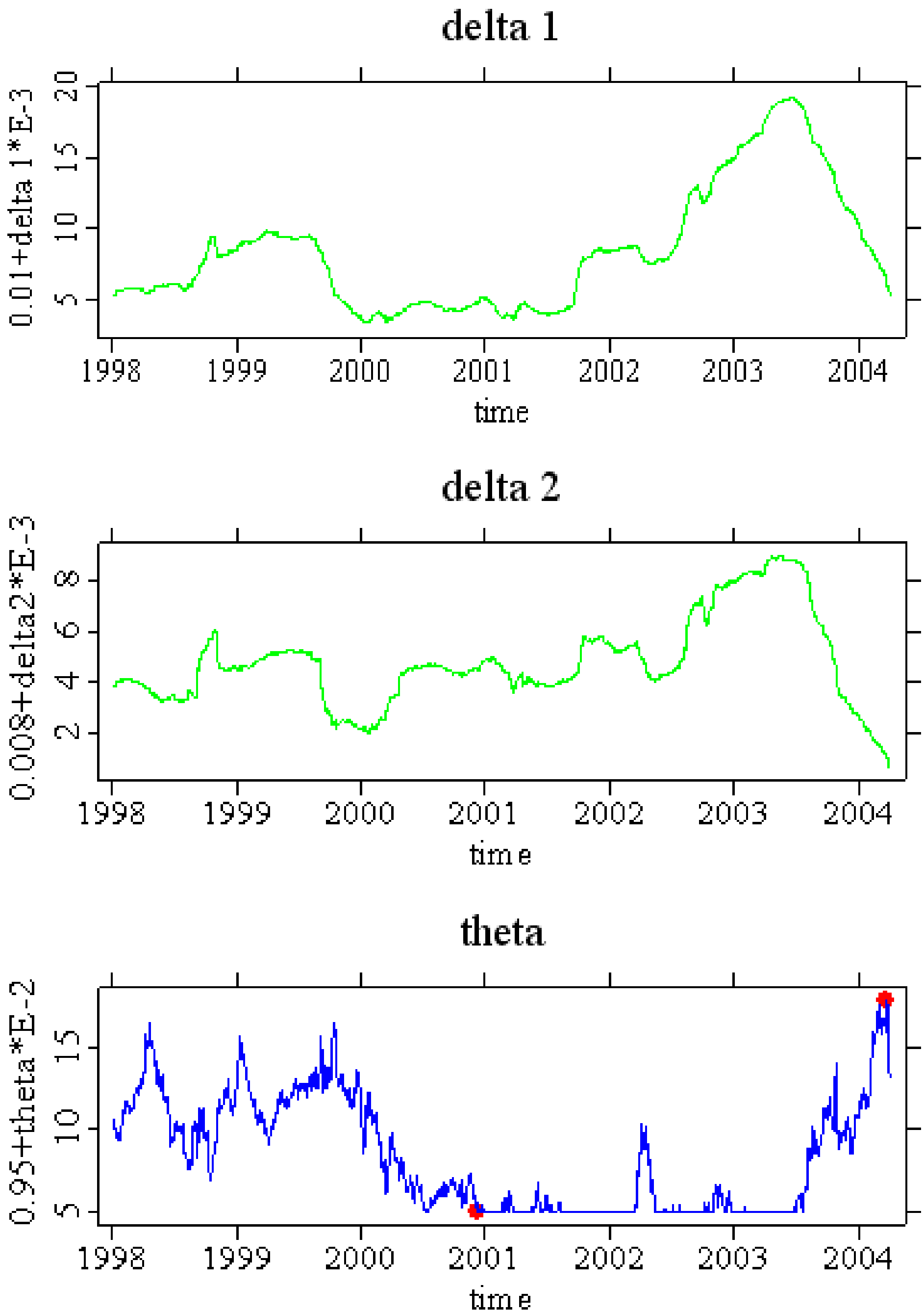


Figure 6.4. Parameters $\hat{\delta}_1$ and $\hat{\delta}_2$ estimated from normal marginals of the Dax and Dow Jones log-returns (upper respectively middle panel) and estimated copula dependence parameter $\hat{\theta}$ (lower panel). The resultes are obtained using IFM method with moving time window $n = 250$.

delta12thetaIFM.xpl

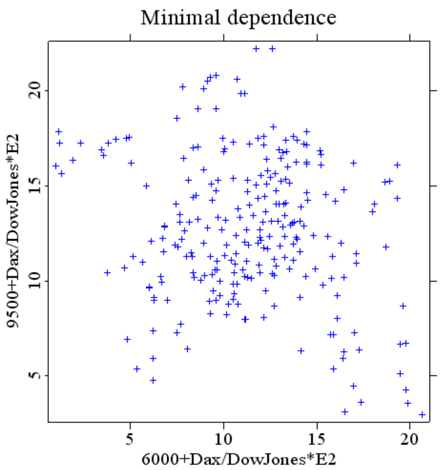


Figure 6.5. Dax and Dow Jones log-returns at minimal dependence (November 30, 2000).

maxmindep.xpl

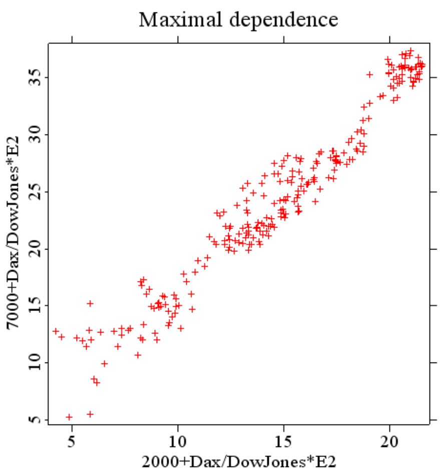


Figure 6.6. Dax and Dow Jones log-returns at maximal dependence (March 11, 2004).

maxmindep.xpl

# 7 Adaptive copulae estimation

In this part we discuss a new procedure for estimating the copula dependence parameter $\theta$ using adaptive techniques. The approached, called *local change point analysis*(LCPD) (Mercurio, Spokoiny, 2004), is based on the assumption of local time homogeneity, i.e. for every moment $n$ there is exists a historic time interval $[n-m, n[$, in which the copula parameter $\theta$ is nearly constant. Our main objective is therefore to describe the interval of homogeneity and to estimate copula dependence parameter $\theta_n$ from this interval.

## 7.1 Choice of the interval of homogeneity

The approach is based on the adaptive choice of the interval of homogeneity for the endpoint $n$. This choice is done by using *local change point detection* (LCPD) algorithm. One defines a family of intervals of the form $\mathcal{I} = \{I_k, k = 0, 1, ...\}$ such that $I_k = [n - m_k, n]$ with $m_k$: $m_0 < m_1 < m_2 < ... \leq n$. Thus, the intervals $I_k$ are ordered by their length $m_k$.

The *LCPD procedure* is made up by the following parts:

1. start from the smallest interval $I_0$
2. test the hypothesis of homogeneity within $I_0$
3. if the hypothesis is not rejected, take the next larger interval
4. continue the procedure until change point $\widehat{\nu}$ is detected or the largest possible interval $[0, n[$ is reached
5. if the hypothesis of homogeneity within some $I_k$ is rejected, the estimated interval of homogeneity is given by $\widehat{I} = [\widehat{\nu}, n[$, otherwise we take $\widehat{I} = [0, n[$
6. estimate the copula dependence parameter $\theta$ from observation $S_t$ for $t \in \widehat{I}$, assuming the homogeneous model within $\widehat{I}$, i.e. define $\widehat{\theta}_n = \widetilde{\theta}_{\widehat{I}}$

To make the procedure running, we have to perform the homogeneity test.

## 7.2 Test of homogeneity against a change point alternative

Let $I = [n-m, n[$ be an interval candidate. The null hypothesis means that the observations $S_t$ for $t \in I$ follow the model with dependence parameter $\theta$. The alternative hypothesis claims that the parameter $\theta$ changes spontaneously in some internal point $\tau$ of the interval $I$. Let $\mathcal{T}_I$ be a set of internal points within $I$. Therefore we test:

$H_0$: $\theta = const\ \forall \tau \in \mathcal{T}_I$ against
$H_1$: $\exists \tau \in \mathcal{T}_I$: $\theta_t = \theta$ for $t \in J = [\tau, n[$ and $\theta_t = \theta^{'}(\neq \theta)$ for $t \in J^c = I \backslash J = [n-m, \tau[$

$H_0$ corresponds to the log-likelihood function $l_I(\theta)$ and $H_1$ leads to the following log-likelihood: $l_J(\theta) + l_{J^c}(\theta^{'})$. The likelihood ratio test statistics for the change point alternative with the change point location at the point $\tau$ is then can be written as:

$$\begin{aligned} T_{I,\tau} &= \max_{\theta,\theta^{'}} \left\{ l_J(\theta) + l_{J^{'}}(\theta^{'}) \right\} - \max_{\theta} l_I(\theta) \\ &= l_J(\widetilde{\theta}_J) + l_{J^c}(\widetilde{\theta}_{J^c}) - l_I(\widetilde{\theta}_I) = \widehat{l}_J + \widehat{l}_{J^c} - \widehat{l}_I \end{aligned} \tag{7.1}$$

The change point test for the interval $I$ is defined as a maximum of such a test statistics over the internal points $\tau \in \mathcal{T}_I$:

$$T_I = \max_{\tau \in \mathcal{T}_I} T_{I,\tau} \tag{7.2}$$

The change point test compares this test statistics with a critical value $\lambda_I$ which may depend on the interval $I$ and the nominal first kind error probability $\alpha$. The way of choosing the critical value is described below. One rejects the hypothesis of homogeneity if $T_I > \lambda_I$. The estimator of the change point is then defined as $\widehat{\nu} = \arg\max_{\tau \in \mathcal{T}_\mathcal{I}} T_{I,\tau}$.

## 7.3 Parameters of the LCPD procedure

To start the procedure running, we have to specify some parameters. This includes:

- Selection of interval candidates $\mathcal{I}$ and setting the internal points $\mathcal{T}_I$ for each of this intervals.
- Choice of the critical values $\lambda_I$, which may depend on the interval $I$ and the nominal first kind error probability $\alpha$.

### 7.3.1 Selection of $\mathcal{I}$

It is usefull to take the set $\mathcal{I}$ of interval candidates in form of an arithmetic or a geometric grid. We fix the length of the first interval to $m_0$, that is, for the end point $n$ we define the first interval as $I_0 = [n - m_0, n[$. At every iteration this length is increased by adding or by multiplying with some fixed step $c > 0$. For our simulated examples, which are discussed below, we set $\mathcal{I}$ in form of a geometric grid, the family of tested intervals has then the following form:

$$I_k = [n - m_k, n[ \text{ where } m_k = [m_0 c^k] \text{ for } \ k = 0, 1, 2, ... \tag{7.3}$$

Here $[x]$ means the integer part of $x$.

### 7.3.2 Setting of $\mathcal{T}_I$

For every interval $I \in \mathcal{I}$, $I = [n - m, n[$, one defines $\mathcal{T}_I$ as the set of all internal points of $I$ separated away form the end point. That is, we fix some parameters $\rho_1 \leq 1/3$ and $\rho_2 \leq 1/3$ and set $\mathcal{T}_I = \{t : n - m + \rho_1 m \leq t \leq n - \rho_2 m\}$. A reasonable choice of $\rho_2$ is one third of the interval length and of $\rho_1$ the decimal place of $c$, i.e. if we set $c = 1.1$, than $\rho = 0.1$. The right choice of the parameters $\rho_1$ and $\rho_2$ is very important. On the one side, the behavior of the log-likelihood test statistics $T_{I,\tau}$ becomes quite irregular when $\tau$ approaches the end-points of the interval $I$. This is the case if $\rho_1$ or $\rho_2$ is too small. On the other side, if we set $\rho_1$ or $\rho_2$ large, the change points which are close to the end-points of the interval could not be detected. In order to illustrate this problem, we consider the following example.
A set of 200 observations was simulated from the Gumbel-Hougaard copula with parameter $\theta = 1$ for the first 100 observations and $\theta = 3$ for the second half. The test statistics $T_{I,\tau}$ corresponding to this case (length of the interval $I$ is fixed and equal to 200) is given by the upper panel of the figure 7.1. The maximum of the test statistics $T_{I,\tau}$ is equal to 40.902 and reached at $\tau = 101$. Since the values of $\rho_1$ and $\rho_2$ will never exceed one third of the interval length, there are no obstacles in detecting the change point in the middle of the interval, that is, one rejects $H_0$ very significantly.
Let us now consider another example. We simulate again 200 observations from the Gumbel-Hougaard copula. Now however, we set $\theta = 1$ for the first 30 observations and $\theta = 3$ for the last 170 observations. The test statistics $T_{I,\tau}$ corresponding to this case is given by the lower panel of the figure 7.1. The maximum of the test statistics $T_{I,\tau}$ is equal to 15.744 and is reached at $\tau = 28$. In this case a choice of $\rho_1$ constitutes a problem: by setting $m\rho_1 = 40$ (that is $\rho_1 = 0.2$), which corresponds to one fifth of the interval length, a change point at $\tau = 30$ will either not be detected at all, or will be detected at a wrong place (dependent of the critical value $\lambda_I$), since all the points $\tau < 40$ will not be taken into account by searching the maximum of $T_{I,\tau}$ over all $\tau \in \mathcal{T}_I$.

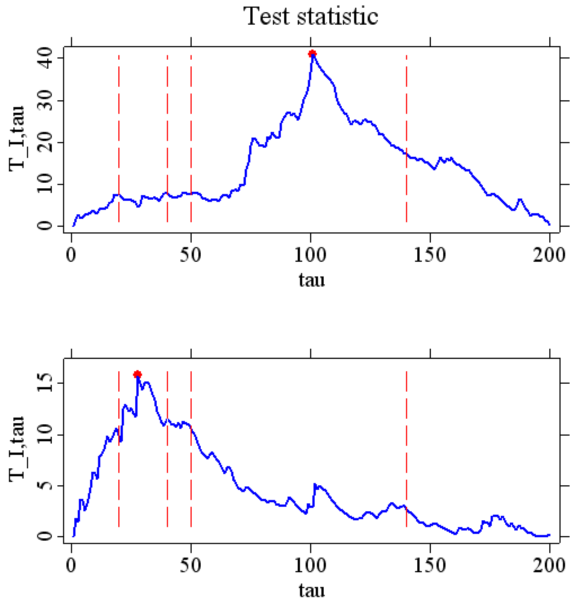


Figure 7.1. Test statistics $T_{I,\tau}$ for one fixed interval $I$ plotted against $\tau$. Jump at the middle point ($\tau = 100$) and not at the middle point ($\tau = 30$) are represented by the upper and lower panel, respectively.

testChangePoint.xpl

### 7.3.3 Choice of the critical values $\lambda_I$

The standard approach for choosing the critical values is to provide a prescribed first kind error probability $\alpha$ (e.g. $\alpha = 0.05$ or $\alpha = 0.1$), that is the number of rejections of the test statistics, obtained from Monte Carlo simulations for the homogeneous case should not exceed the given level $\alpha$. We define for every $I$ a value $\beta_I$ using Bonferonni method: we set $\beta_I$ in such a way that $\sum_{I \in \mathcal{I}} \beta_I = \alpha$. A reasonable proposal is to set

$$\beta_{I_k} = \alpha m_k^{-1} \left( \sum_{l=0}^{\infty} m_l^{-1} \right)^{-1} \approx \frac{\alpha(1 - c^{-1})}{c^k} \tag{7.4}$$

and accordingly the value $\alpha_{I_k}$:

$$\alpha_{I_k} \approx \alpha(1 - c^{-(k+1)}) \tag{7.5}$$

Then the critical values $\lambda_{I_k}$ are selected by Monte Carlo in order to provide a prescribed first kind error probability $\alpha_{I_k}$ for every interval $I_k$, that is, it holds:

$$P\left( max_{k' \leq k} T_{I'_k} > \lambda_{I'_k} \right) = \alpha_{I_k} \tag{7.6}$$

# 8 Some simulated examples

**EXAMPLE 8.1.** *In this example a set of 160 observations was simulated from a bivariate Clayton copula*

$$C_\theta(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$$

*with parameter:*

$$\theta_t = \begin{cases} 1 & \textit{if} \quad 1 \le t \le 40 \\ 5 & \textit{if} \quad 41 \le t \le 80 \\ 10 & \textit{if} \quad 81 \le t \le 120 \\ 3 & \textit{if} \quad 121 \le t \le 160 \end{cases}$$

Figure 8.1 represents one simulation of the copula parameter $\theta$, estimated by LCPD algorithm using parameters $\alpha = 0.05$, $m_0 = 20$, $c = 1.1$ and $\rho_1 = 0.1$, $\rho_2 = 0.3$.

**EXAMPLE 8.2.** *In this example a set of 240 observations was simulated from a bivariate Gumbel-Hougaard copula*

$$C_\theta(u,v) = \exp\left[-\left\{(-\ln u)^\theta + (-\ln v)^\theta\right\}^{1/\theta}\right]$$

*with parameter:*

$$\theta_t = \begin{cases} 1 & & \textit{if} \quad 1 \le t \le 80 \\ 3 \quad \textit{or} \quad 2 & & \textit{if} \quad 81 \le t \le 160 \\ 1 & & \textit{if} \quad 161 \le t \le 240 \end{cases}$$

We assume at first that all the data comes from the homogeneous model following a Gumbel-Hougaarg copula. We assume the marginals beeing normal with variance 1. Based on the assumption of homogeneity, that is, setting a constant parameter $\theta$ (in our simulated example we use $\theta$=1) for all 240 data points, we simulate by Monte-Carlo sequentially growing series of observations. The series length corresponds to the length of the interval $I_k$. The number of simulations performed in our example

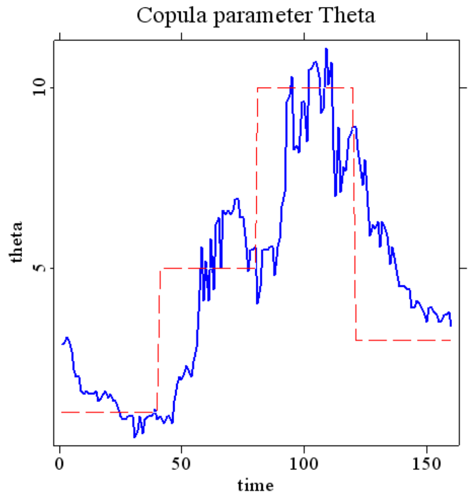


Figure 8.1. One simulation of the copula parameter $\theta$: real (red) and estimated (blue). The results are obtained with parameters $\alpha = 0.05$, $m_0 = 20$, $c = 1.1$, $\rho_1 = 0.1$, $\rho_2 = 0.3$

claytononesim.xpl

is 1000. Using this simulated data, we calculate 1000 test statistics for each interval $I_k$. Starting from the smallest interval $I_0$, we define a critical value $\lambda_{I_0}$ for $I_0$ as a $(1-\alpha_{I_0})$-quantile of the test statistics $T_{I_0}$. A setting of $\alpha_{I_k}$ (prescribed first kind error probability for each interval $I_k$) was discussed in chapter 7.3.3. After computing $\lambda_{I_{k'}}$ for all $k' < k$, every following $\lambda_{I_k}$ for the interval $I_k$ is computed in a way that the interval $I_k$ is accepted with a frequency $1 - \alpha_{I_k}$. Critical values, computed in such a way, are in fact empirical quantiles of the test statistics under the null hypothesis of homogeneity. Figure 8.2 plots critical values, computed by 1000 Monte-Carlo simulations using the following parameters of the procedure: initial length $m_0 = 20$; parameter $\rho_2 = 0.3$; $\alpha = 0.05$ (solid line) and $\alpha = 0.1$ (dashed line); $c = 1.1$ and $\rho_1 = 0.1$ (black line); $c = 1.2$, $\rho_1 = 0.2$ (red line); $c = 1.25$, $\rho_1 = 0.25$ (blue line).

A choice of $\alpha$ controls a trade off between type-1-error and type-2-error. Parameter $c$ is responsible for the speed with which the length of the intervals $I_k$ grows: a large $c$ results in a rapid growth of the interval length. The problem can occur if at some iteration a tested interval $I_k$ contains more than one change point: the change point detection procedure may simply break down. To overcome this problem, one can set $c$ sufficiently small. However, with decreasing $c$, critical values $\lambda_{I_k}$ become larger, which results in less sensitivity of the procedure.
For our simulated example we compare the results for different settings of parameters: first, we set $c = 1.1$, then $c = 1.2$ and $c = 1.25$, which corresponds to the choice of $\rho_1 = 0.1$, $\rho_1 = 0.2$ and $\rho_1 = 0.25$, respectively. We also vary a jump size: the left panel of figure 8.3 and figure 8.4 represent a pointwise mean and a pointwise

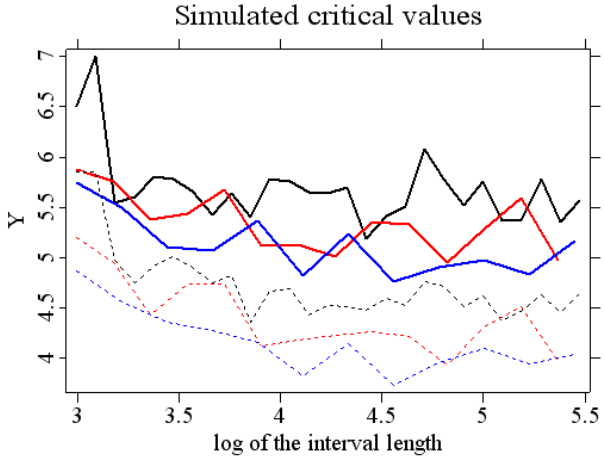


Figure 8.2. Critical values for $\alpha = 0.05$ (solid line) and $\alpha = 0.1$ (dashed line), computed by simulations using Gumbel-Hougaard copula with parameters $m_0 = 20$, $\rho_2 = 0.3$ and $c = 1.1$, $\rho_1 = 0.1$ (black line), $c = 1.2$, $\rho_1 = 0.2$ (red line), $c = 1.25$, $\rho_1 = 0.25$ (blue line).

critplot.xpl

median of the estimated parameter $\theta$ respectively, based on 200 simulations of the data, simulated with a jump size of 2. The right panel of these figures describes the same for the data, simulated with a jump size of 1.

Small jumps seem to be more difficult to recognize, that is, the detection speed decreases with a decreasing jump size. In order to prove this, we have a look on the descriptive statistics for detection speeds to sudden jumps of the parameter $\theta$. Table 8.1 and table 8.2 introduce a percentage rule which tells after how many steps a sudden jump in parameter $\theta$ was detected at 40%, 50% and 60% level of a jump size of 2 and a jump size of 1, respectively. For example, for the data simulated from the Gumbel-Hougaard copula with jump size of 2, the first jump (occurring at the 81st point), was detected on average after 9.6 steps at the 40% level. The number increases to 10.2 for the 60% rule. The detection speed for downward jumps is faster, that is, for example only 6.4, respectively 7.8, time steps are on average necessary to detect the second jump (occurring at the 161st point) at the 40% respectively at the 60% level of a jump size. Additionally, we consider the data simulated from the Gumbel-Hougaard copula with a jump size of 1. The detection speeds decrease, that is, the average number of time points necessary to detect the jump at the 40%, 50% or 60% level of a jump size of 1 is larger than those for the jump size of 2.

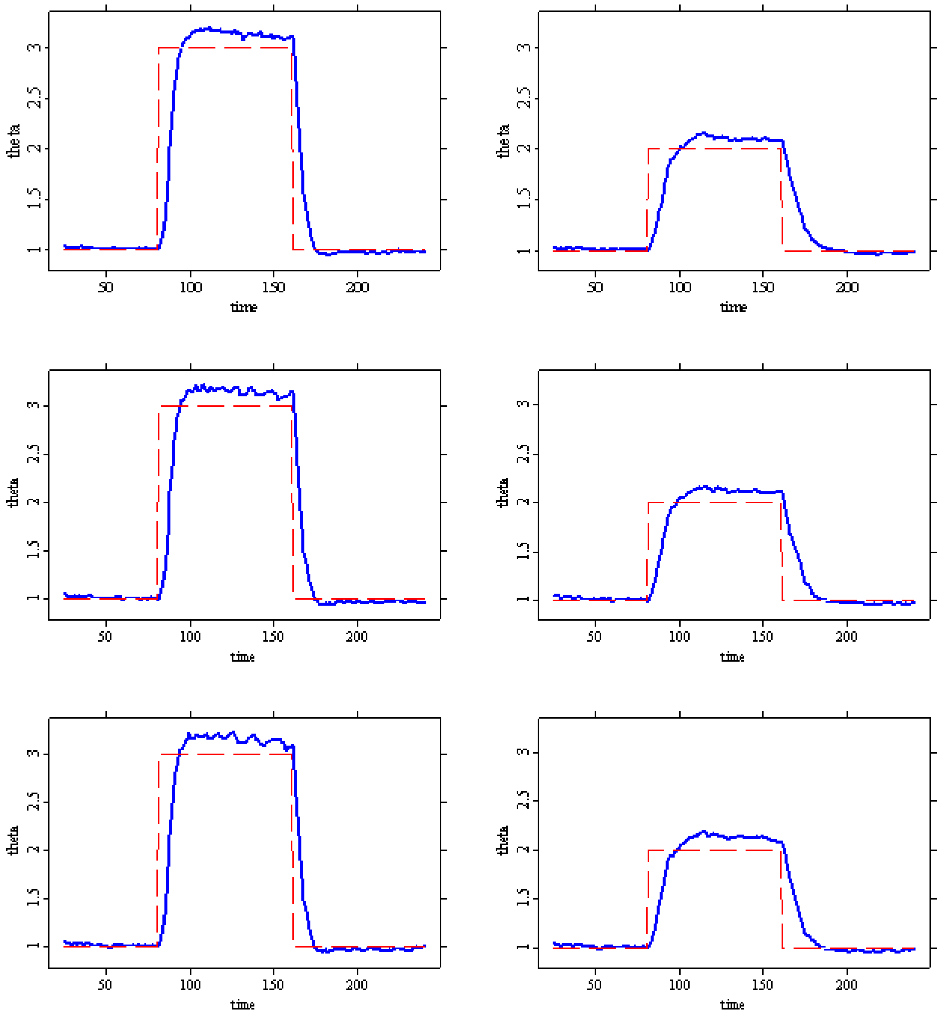


Figure 8.3. Pointwise mean (blue line) based on 200 simulations of the data, simulated from the Gumbel-Hougaard copula and real parameter (dashed lines). The left panel represents a jump size equal to 2, the right panel represents a jump size equal to 1. The results are obtained with parameters $m_0 = 20$, $\rho_2 = 0.3$ and $c = 1.1$, $\rho_1 = 0.1$ (upper panel); $c = 1.2$, $\rho_1 = 0.2$ (middle panel); $c = 1.25$, $\rho_1 = 0.25$ (lower panel).

thetaplotmean.xpl

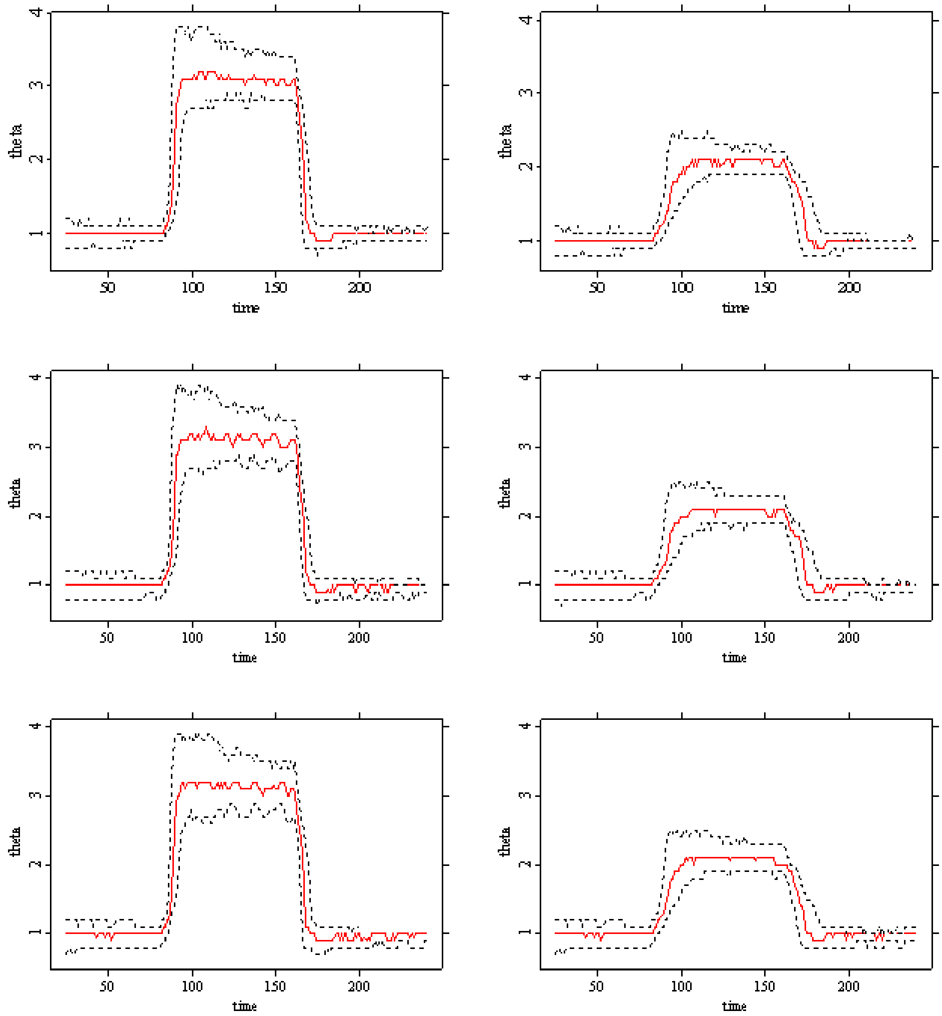


Figure 8.4. Pointwise median (red line) and quartiles (dashed lines) for the estimates of the copula parameter $\theta_t$, based on 200 simulations of the data, simulated from the Gumbel-Hougaard copula. The left panel represents a jump size equal to 2, the right panel represents a jump size equal to 1. The results are obtained with parameters $m_0 = 20$, $\rho_2 = 0.3$ and $c = 1.1$, $\rho_1 = 0.1$ (upper panel); $c = 1.2$, $\rho_1 = 0.2$ (middle panel); $c = 1.25$, $\rho_1 = 0.25$ (lower panel).

thetaplot.xpl

Detection decays with parameters $m_0 = 20$, $c = 1.1$, $\rho_1 = 0.1$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 9.595 | 3.722 | 24 | 1 |
| 50% rule | 9.930 | 3.903 | 25 | 1 |
| 60% rule | 10.165 | 4.043 | 26 | 2 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 6.390 | 3.410 | 20 | 1 |
| 50% rule | 7.080 | 3.418 | 20 | 1 |
| 60% rule | 7.795 | 3.774 | 21 | 2 |

Detection decays with parameters $m_0 = 20$, $c = 1.2$, $\rho_1 = 0.2$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 9.465 | 3.567 | 22 | 1 |
| 50% rule | 9.780 | 3.774 | 24 | 1 |
| 60% rule | 10.010 | 3.832 | 28 | 2 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 6.140 | 3.365 | 20 | 1 |
| 50% rule | 6.965 | 3.400 | 20 | 1 |
| 60% rule | 7.650 | 3.644 | 21 | 2 |

Detection decays with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 9.215 | 3.596 | 21 | 1 |
| 50% rule | 9.475 | 3.697 | 21 | 2 |
| 60% rule | 9.740 | 3.860 | 24 | 2 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 6.175 | 3.158 | 15 | 1 |
| 50% rule | 6.890 | 3.216 | 17 | 1 |
| 60% rule | 7.605 | 3.634 | 21 | 2 |

Table 8.1. Descriptive statistics for the detection speeds to sudden jumps of the Gumbel-Hougaard copula dependence parameter with a jump size of 2. The results are obtained with parameters $\alpha = 0.05$, $m_0 = 20$, $\rho_2 = 0.3$ and different parameters $c$ and $\rho_1$. Statistics are based on 200 simulations.

Detection decays with parameters $m_0 = 20$, $c = 1.1$, $\rho_1 = 0.1$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 11.115 | 6.256 | 30 | 1 |
| 50% rule | 12.400 | 6.552 | 31 | 1 |
| 60% rule | 13.890 | 7.324 | 37 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 9.835 | 6.396 | 28 | 1 |
| 50% rule | 11.395 | 6.887 | 41 | 1 |
| 60% rule | 12.700 | 7.903 | 53 | 1 |

Detection decays with parameters $m_0 = 20$, $c = 1.2$, $\rho_1 = 0.2$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 10.585 | 5.997 | 30 | 1 |
| 50% rule | 11.905 | 6.175 | 31 | 1 |
| 60% rule | 13.345 | 6.678 | 37 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 9.085 | 6.114 | 32 | 1 |
| 50% rule | 10.510 | 6.324 | 41 | 1 |
| 60% rule | 12.385 | 7.681 | 60 | 1 |

Detection decays with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 10.210 | 5.829 | 28 | 1 |
| 50% rule | 11.535 | 6.150 | 29 | 1 |
| 60% rule | 13.030 | 6.801 | 37 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 8.885 | 5.428 | 28 | 1 |
| 50% rule | 10.075 | 5.719 | 32 | 1 |
| 60% rule | 11.800 | 7.880 | 61 | 1 |

Table 8.2. Descriptive statistics for the detection speeds to sudden jumps of the Gumbel-Hougaard copula dependence parameter with a jump size of 1. The results are obtained with parameters $\alpha = 0.05$, $m_0 = 20$, $\rho_2 = 0.3$ and different parameters $c$ and $\rho_1$. Statistics are based on 200 simulations.

**EXAMPLE 8.3.** *In this example a set of 240 observations was simulated from a bivariate Clayton copula*

$$C_\theta(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$$

*with parameter:*

$$\theta_t = \begin{cases} 1 & \text{if} \quad 1 \le t \le 80 \\ 3 \quad or \quad 2 & \text{if} \quad 81 \le t \le 160 \\ 1 & \text{if} \quad 161 \le t \le 240 \end{cases}$$

Figure 8.5 represents critical values, computed by simulations using Clayton copula. Computation is done in the same way as for the data simulated from the Gumbel-Hougaard copula. Pointwise mean and pointwise median are represented by figure 8.6 and 8.7, respectively. Descriptive statistics for the detection speeds are given by table 8.3 for a the jump size of 2 and table 8.4 for a jump size of 1.
As well as in the case of a Gumbel-Hougaard copula, a decreasing jump size results in a decrease of the detection speed. Although, the detection speed for downward jumps is faster, some cases have occured, when the second jump at the 161st point was not detected at all. For the data simulated with a jump size of 2, the number of such a failure amounts to 1 for 40%, 50% or 60% rule, i.e. for one simulation the jump in $\theta$ was not detected already at the 40% level of a jump size. For the data simulated with a jump size of 1, this number amounts to 2 failures for the 40% rule and increases to 6 for the 60% rule, i.e. for 2 out of 200 simulations the jump at $\theta$ was not detected already at the 40% level and for 6 out of 200 simulations at the 60% level.

In order to compare the results obtained from the data simulated with a Gumbel-Hougaard and with a Clayton copula, we introduce a concept of the *Kullback-Leibler divergence*. In general, the Kullback-Leibler divergence (or Kullback-Leibler information number) for two probability densities $p(y,\theta)$ and $p(y,\theta^{'})$ is defined as

$$K(\theta,\theta^{'}) = \boldsymbol{E}_\theta \log\left(p(y,\theta)/p(y,\theta^{'})\right) = \int \log\frac{p(y,\theta)}{p(y,\theta^{'})} p(y,\theta)dy \qquad (8.1)$$

The Kullback-Leibler divergence number aims to measure the similarity between two probability densities. It fulfills

$$K(\theta,\theta^{'}) \ge 0 \qquad (8.2)$$

and

$$K(\theta,\theta^{'}) = 0 \Leftrightarrow p(y,\theta^{'}) = p(y,\theta^{'}) \qquad (8.3)$$

In the case of a copula, we are interested in the change of the information connected with a jump of a copula parameter $\theta$; first for the upward jump, and then for the downward jump. Thus, for varying parameter $\theta_2$ and fixed $\theta_1 = 1$, we denote

Detection decays with parameters $m_0 = 20$, $c = 1.1$, $\rho_1 = 0.1$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 11.725 | 6.966 | 34 | 1 |
| 50% rule | 13.585 | 7.916 | 38 | 1 |
| 60% rule | 15.125 | 9.052 | 48 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 10.874 | 7.992 | 54 | 1 |
| 50% rule | 11.945 | 8.336 | 54 | 1 |
| 60% rule | 12.975 | 8.554 | 54 | 1 |

Detection decays with parameters $m_0 = 20$, $c = 1.2$, $\rho_1 = 0.2$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 11.065 | 6.234 | 33 | 1 |
| 50% rule | 12.730 | 7.322 | 37 | 1 |
| 60% rule | 13.655 | 7.654 | 41 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 9.814 | 6.996 | 36 | 1 |
| 50% rule | 10.919 | 7.913 | 54 | 1 |
| 60% rule | 11.925 | 8.230 | 54 | 1 |

Detection decays with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 10.940 | 6.519 | 32 | 1 |
| 50% rule | 12.440 | 7.326 | 39 | 1 |
| 60% rule | 13.410 | 7.708 | 39 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 9.688 | 6.716 | 36 | 1 |
| 50% rule | 10.764 | 7.416 | 53 | 1 |
| 60% rule | 12.352 | 8.109 | 53 | 1 |

Table 8.3. Descriptive statistics for the detection speeds to sudden jumps of the Clayton copula dependence parameter with a jump size of 2. The results are obtained with parameters $\alpha = 0.05$, $m_0 = 20$, $\rho_2 = 0.3$ and different parameters $c$ and $\rho_1$. Statistics are based on 200 simulations.

Detection decays with parameters $m_0 = 20$, $c = 1.1$, $\rho_1 = 0.1$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 10.650 | 10.183 | 52 | 1 |
| 50% rule | 13.215 | 11.766 | 53 | 1 |
| 60% rule | 15.545 | 12.697 | 62 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 14.364 | 14.128 | 78 | 1 |
| 50% rule | 16.700 | 15.204 | 78 | 1 |
| 60% rule | 18.660 | 16.238 | 78 | 1 |

Detection decays with parameters $m_0 = 20$, $c = 1.2$, $\rho_1 = 0.2$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 10.100 | 9.755 | 48 | 1 |
| 50% rule | 12.045 | 10.576 | 51 | 1 |
| 60% rule | 14.030 | 11.454 | 60 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 17.565 | 15.578 | 79 | 1 |
| 50% rule | 20.000 | 17.192 | 79 | 1 |
| 60% rule | 21.772 | 17.682 | 79 | 1 |

Detection decays with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$

| to the first jump at $t = 80$ | mean | standard deviation | maximum | minimum |
|---|---|---|---|---|
| 40% rule | 10.100 | 10.051 | 60 | 1 |
| 50% rule | 11.745 | 10.918 | 60 | 1 |
| 60% rule | 13.870 | 11.912 | 60 | 1 |
| to the second jump at $t = 160$ | mean | standard deviation | maximum | minimum |
| 40% rule | 16.626 | 14.837 | 74 | 1 |
| 50% rule | 19.843 | 17.547 | 75 | 1 |
| 60% rule | 21.727 | 18.064 | 79 | 1 |

Table 8.4. Descriptive statistics for the detection speeds to sudden jumps of the Clayton copula dependence parameter with a jump size of 1. The results are obtained with parameters $\alpha = 0.05$, $m_0 = 20$, $\rho_2 = 0.3$ and different parameters $c$ and $\rho_1$. Statistics are based on 200 simulations.

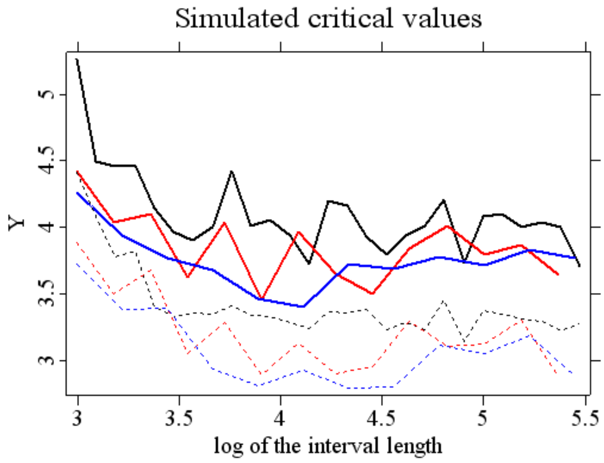


Figure 8.5. Critical values for $\alpha = 0.05$ (solid line) and $\alpha = 0.1$ (dashed line), computed by simulations using Clayton copula with parameters $m_0 = 20$, $\rho_2 = 0.3$ and $c = 1.1$, $\rho_1 = 0.1$ (black line), $c = 1.2$, $\rho_1 = 0.2$ (red line), $c = 1.25$, $\rho_1 = 0.25$ (blue line).

critplot.xpl

| | $K(1,2)$ | $K(1,3)$ | $K(2,1)$ | $K(3,1)$ |
|---|---|---|---|---|
| Gumbel-Hougaard | 123.01 | 339.36 | 79.377 | 158.08 |
| Clayton | 24.938 | 84.439 | 18.096 | 48.803 |

Table 8.5. Kullback-Leibler information number $K(\theta_1, \theta_2)$ and $K(\theta_2, \theta_1)$ for fixed $\theta_1 = 1$ and parameter $\theta_2 = 2.0, 3.0$; for the Gumbel-Hougaard and the Clayton copula.

by $K(\theta_1, \theta_2)$ the Kullback-Leibler information number for the upward jump, from $\theta_1 = 1.0$ to $\theta_2 = 1.0, 1.1, 1.2, ..., 10.9$. $K(\theta_2, \theta_1)$ denotes then the Kullback-Leibler information number for the downward jump, from $\theta_2 = 1.0, 1.1, 1.2, ..., 10.9$ to $\theta_1 = 1.0$. The upper and lower panel of figure 8.8 plot $K(\theta_1, \theta_2)$ and $K(\theta_2, \theta_1)$ against $\theta_2$ for fixed $\theta_1 = 1.0$. The blue line refers to the Gumbel-Hougaard copula and the red line to the Clayton copula. Table 8.5 represents the Kullback-Leibler information number for $\theta_2 = 2.0, 3.0$ that corresponds to the jump size in our simulated examples.

Comparing the results for the Gumbel-Hougaard and the Clayton copula, it is easy to see, that a decreasing Kullback-Leibler information number results in increasing detection decays.

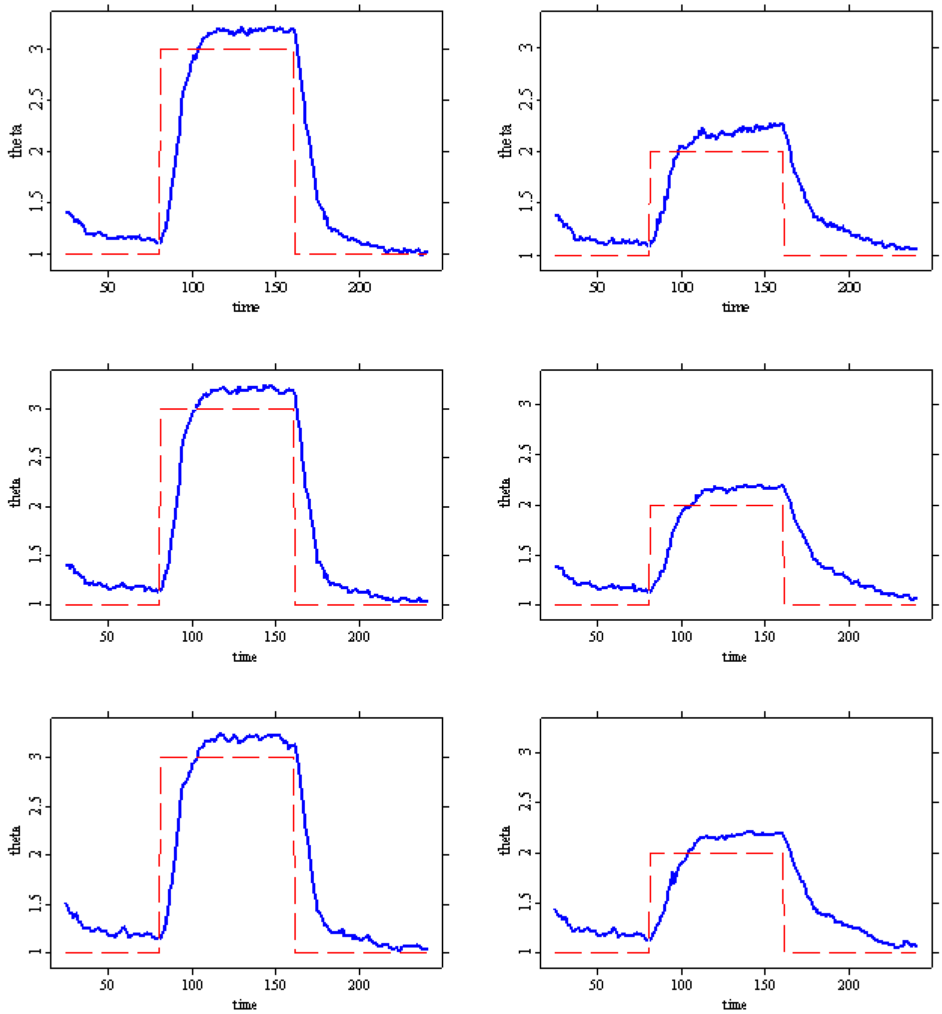


Figure 8.6. Pointwise mean (blue line) based on 200 simulations of the data, simulated from the Clayton copula and real parameter (dashed line). The left panel represents a jump size equal to 2, the right panel represents a jump size equal to 1. The results are obtained with parameters $m_0 = 20$, $\rho_2 = 0.3$ and $c = 1.1$, $\rho_1 = 0.1$ (upper panel); $c = 1.2$, $\rho_1 = 0.2$ (middle panel); $c = 1.25$, $\rho_1 = 0.25$ (lower panel).

thetaplotmean.xpl

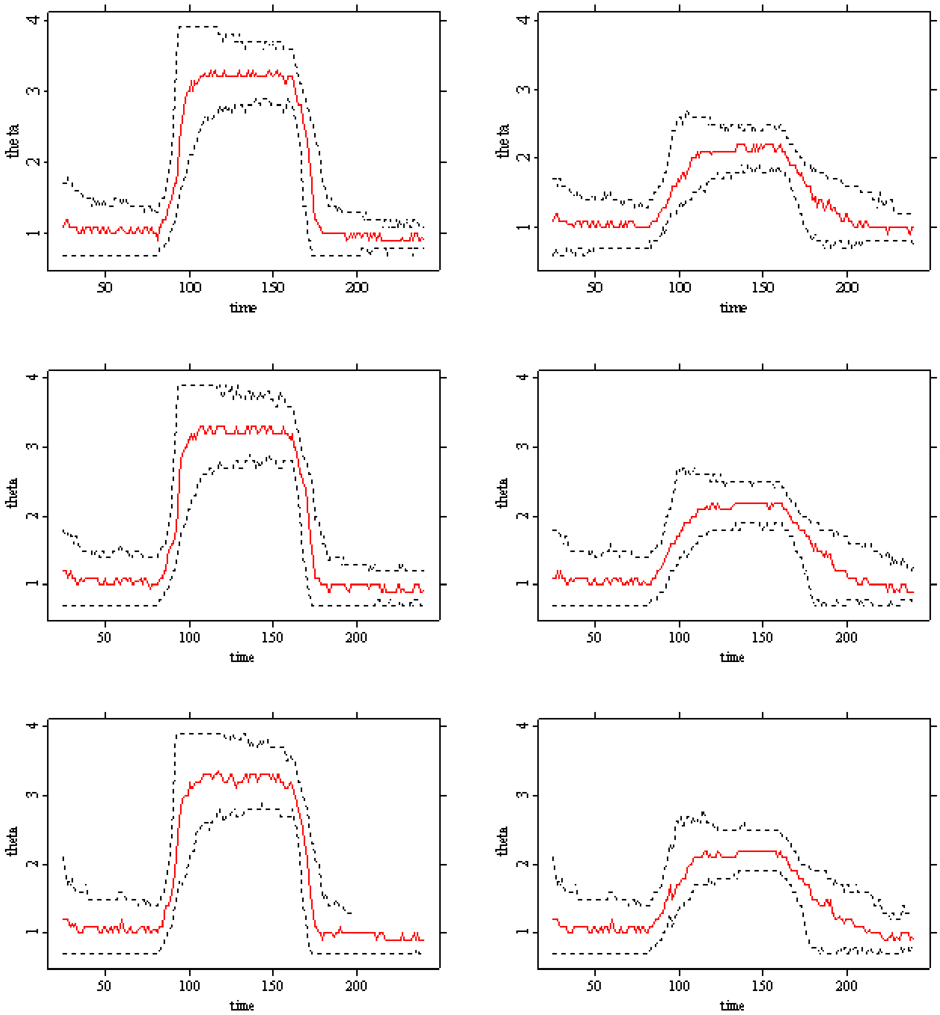


Figure 8.7. Pointwise median (red line) and quartiles (dashed lines) for the estimates of the copula parameter $\theta_t$, based on 200 simulations of the data, simulated from the Clayton copula. The left panel represents a jump size equal to 2, the right panel represents a jump size equal to 1. The results are obtained with parameters $m_0 = 20$, $\rho_2 = 0.3$ and $c = 1.1$, $\rho_1 = 0.1$ (upper panel); $c = 1.2$, $\rho_1 = 0.2$ (middle panel); $c = 1.25$, $\rho_1 = 0.25$ (lower panel).

thetaplot.xpl

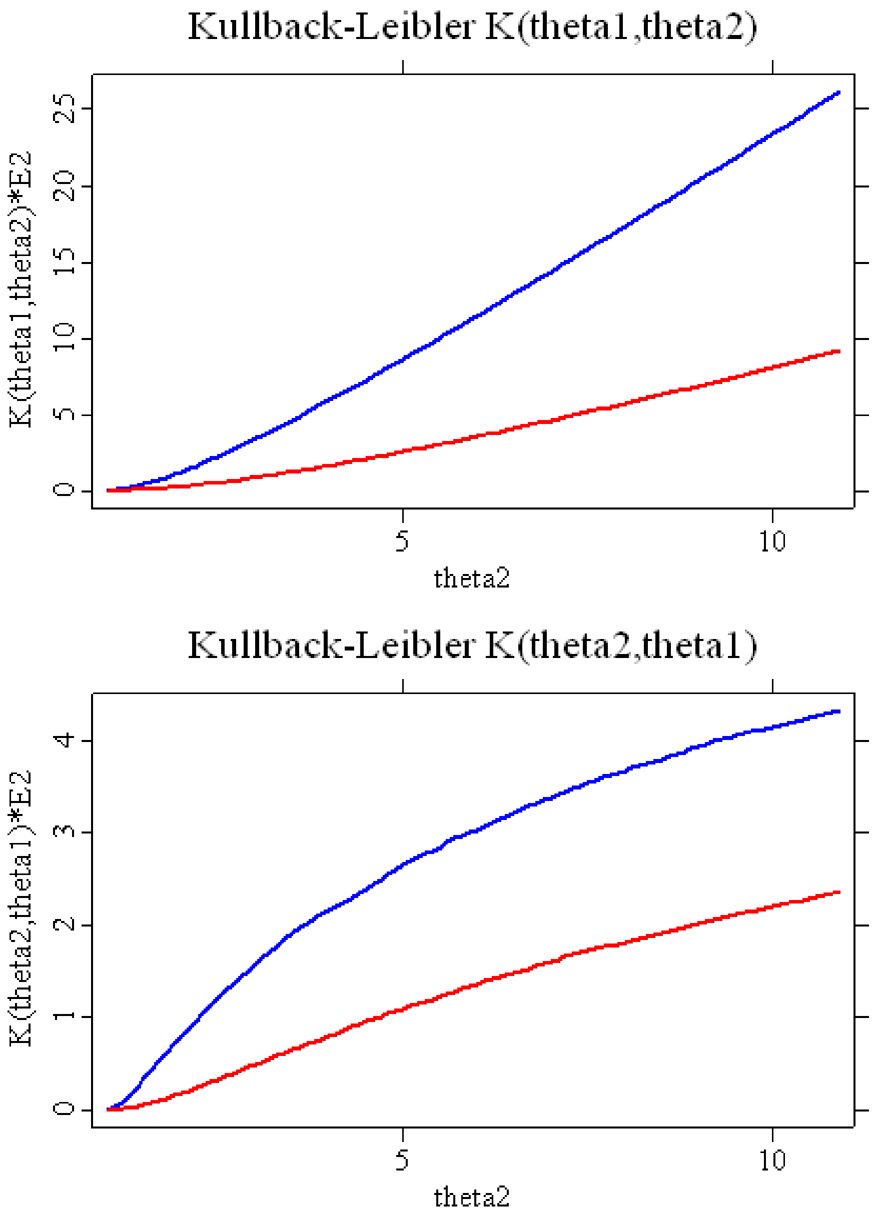


Figure 8.8. Kullback-Leibler information number $K(\theta_1, \theta_2)$ (upper panel) and $K(\theta_2, \theta_1)$ (lower panel) plotted against $\theta_2$ for fixed $\theta_1 = 1$. The blue line refers to the Gumbel-Hougaard copula and the red line to the Clayton copula.

KullbackLeibler.xpl

# 9 Applications to real data

## 9.1 Bivariate case

In this chapter the performance of the LCPD procedure is illustarted by means of the different data sets from the Dax Index (the data is available under `http://sfb649.wiwi.hu-berlin.de/fedc`). The data sets represent closing prices of the following stocks: DaimlerChrysler (DCX) and Volkswagen (VW) in the first example; Allianz (ALV) and Münchener Rückversicherung (MUV2) in the second example; Bayer (BAY) and BASF (BAS) in the third example. The period under consideration goes from January 1st, 2000 to December 31rd, 2004. Each time series consists of 1270 observations.
Figures 9.1, 9.4 and 9.7 represent parameters estimated from the marginals (log returns) for DCX and VW; ALV and MUV2; BAY and BAS. The estimation is done by using exponential smoothing techniques. At every time point $t$, marginal parameters $\hat{\delta}_{t,j} = \hat{\sigma}^2_{t,j}$, $j = 1, 2$ are estimated from observations of the previous period using the exponential weighting scheme:

$$\hat{\delta}_{t,j} = \hat{\sigma}^2_{t,j} = (e^{\lambda} - 1) \sum_{s<t} e^{-\lambda(t-s)} X^2_{s,j} \tag{9.1}$$

where $X_{s,j}$ denotes log returns of DCX and VW; ALV and MUV2; or BAY and BAS at time $s$ and $0 \leq \lambda \leq 1$ is a smoothing parameter (we set $\lambda = 1/20$).
In the next step, we standardize the data by the estimated parameters from the marginals. A new data is then used for the estimation of the copula dependence parameter $\theta$. The chosen copula belongs to the Gumbel-Hougaard family of copulae:

$$C_{\theta}(u, v) = \exp\left[-\left\{(-\ln u)^{\theta} + (-\ln v)^{\theta}\right\}^{1/\theta}\right]$$

Racall that for the Gumbel-Hougaard $\theta = 1$ indicates independence.
To start the LCPD procedure running, we specify the parameters of the procedure. We set the nominal first kind error probability $\alpha = 0.05$; initial interval length $m_0 = 20$; the parameter controlling the growth rate of the tested intervals $c = 1.25$ and correspondingly $\rho_1 = 0.25$; parameter $\rho_2 = 0.3$.
Figures 9.2, 9.5 and 9.8 display a stock price process (upper panel), log returns (middle panel) and esdimated dependence parameter $\theta$ (lower panel) between DCX and VW; ALV and MUV2; or BAY and BAS. Figures 9.3, 9.6 and 9.9 represent

as well parameter $\theta$ with its mean (upper panel) as estimated intervals of time homogeneity $\widehat{I}$ (lower panel).

## 9.2 Multivariate case

In this section we apply the LCPD procedure to the same data, but in the multivariate setting. In the first example, we consider 4-dimensional data: DaimlerChrysler, Volkswagen, Bayer and BASF; in the second example 6-dimesional data: Daimler-Chrysler, Volkswagen, Bayer, BASF, Allianz and Münchener Rückversicherung. As well as in the bivariate case, the period under consideration is from January 1st, 2000 to December 31rd, 2004 (1270 observations). A Clayton copula is chosen:

$$C_\theta(u_1, ..., u_d) = (u_1^{-\theta} + ... + u_d^{-\theta} - d + 1)^{-1/\theta} \tag{9.2}$$

with copula density

$$c_\theta(u_1, ..., u_d) = \prod_{j=1}^{d} \{1 + (j-1)\theta\} \prod_{j=1}^{d} u_j^{-(\theta+1)} \left\{ \sum_{j=1}^{d} u_j^{-\theta} - d + 1 \right\}^{-(1/\theta + d)} \tag{9.3}$$

Recall that $\theta \to 0$ indicates independence for the Clayton copula. We estimate parameters from the marginals in the same way as before, using exponential smoothing:

$$\hat{\delta}_{t,j} = \hat{\sigma}^2_{t,j} = (e^\lambda - 1) \sum_{s<t} e^{-\lambda(t-s)} X^2_{s,j} \tag{9.4}$$

Here $X_{s,j}$, $j = 1, ..., 6$ denotes log returns of DCX, VW, ALV, MUV, BAY and BAS at time point $s$. The data standardized by the estimated parameters from the marginals is used for the estimation of parameter $\theta$. Parameters of the LCPD procedure are chosen as before: $\alpha = 0.05$, $m_0 = 20$, $c = 1.25$ and $\rho_1 = 0.25$, $\rho_2 = 0.3$. For the 4-dimensional case (DCX, VW, BAY, BAS) we represent parameter $\theta$ and its mean in the upper panel of figure 9.10 and estimated intervals of time homogeneity in the lower panel. Equivalently, figure 9.11 dislays $\theta$, its mean and the intervals of time homogeneity for the 6-dimensional data (DCX, VW, BAY, BAS, ALV, MUV2). As expected, the length of the intervals of time homogeneity decreases with increasing dimensions of the data. Since the esimates of the copula dependence parameter $\theta$ is now obtained from shorter interval lengths, it is becoming more volatile.

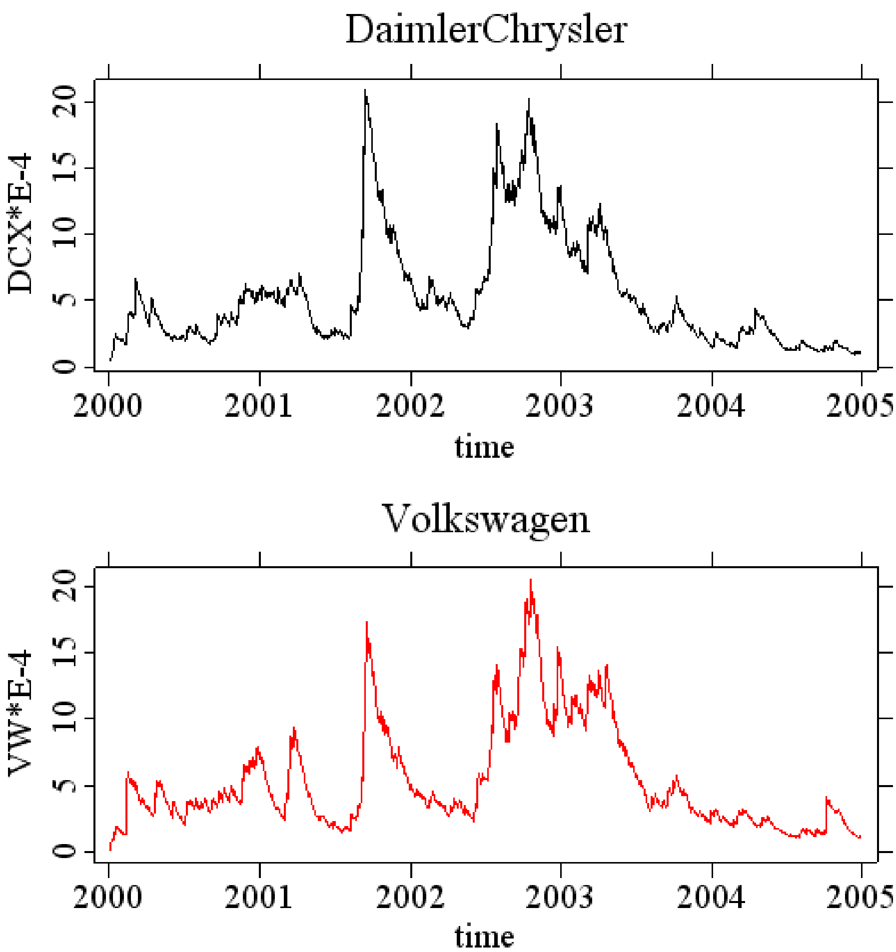


Figure 9.1. Marginal parameters for DaimlerChrysler (upper panel) and Volkswagen (lower panel) estimated by exponential smoothing with parameter $\lambda = 1/20$.

testrealdata.xpl

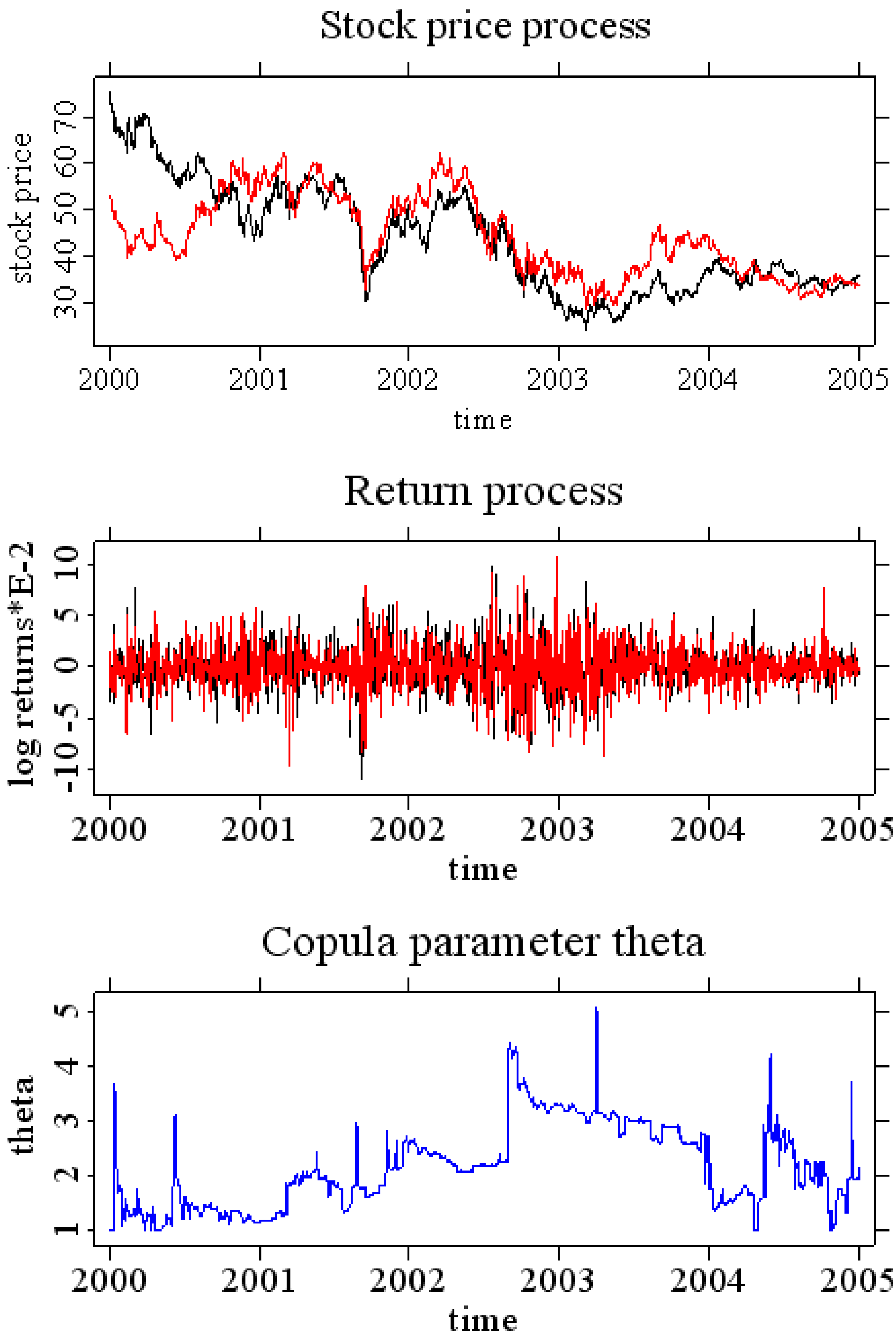


Figure 9.2. Stock price process (upper panel), log returns (middle panel) and copula dependence parameter $\theta$ (lower panel) for DaimlerChrysler (black line) and Volkswagen (red line). The estimates of $\theta$ are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

plot.xpl

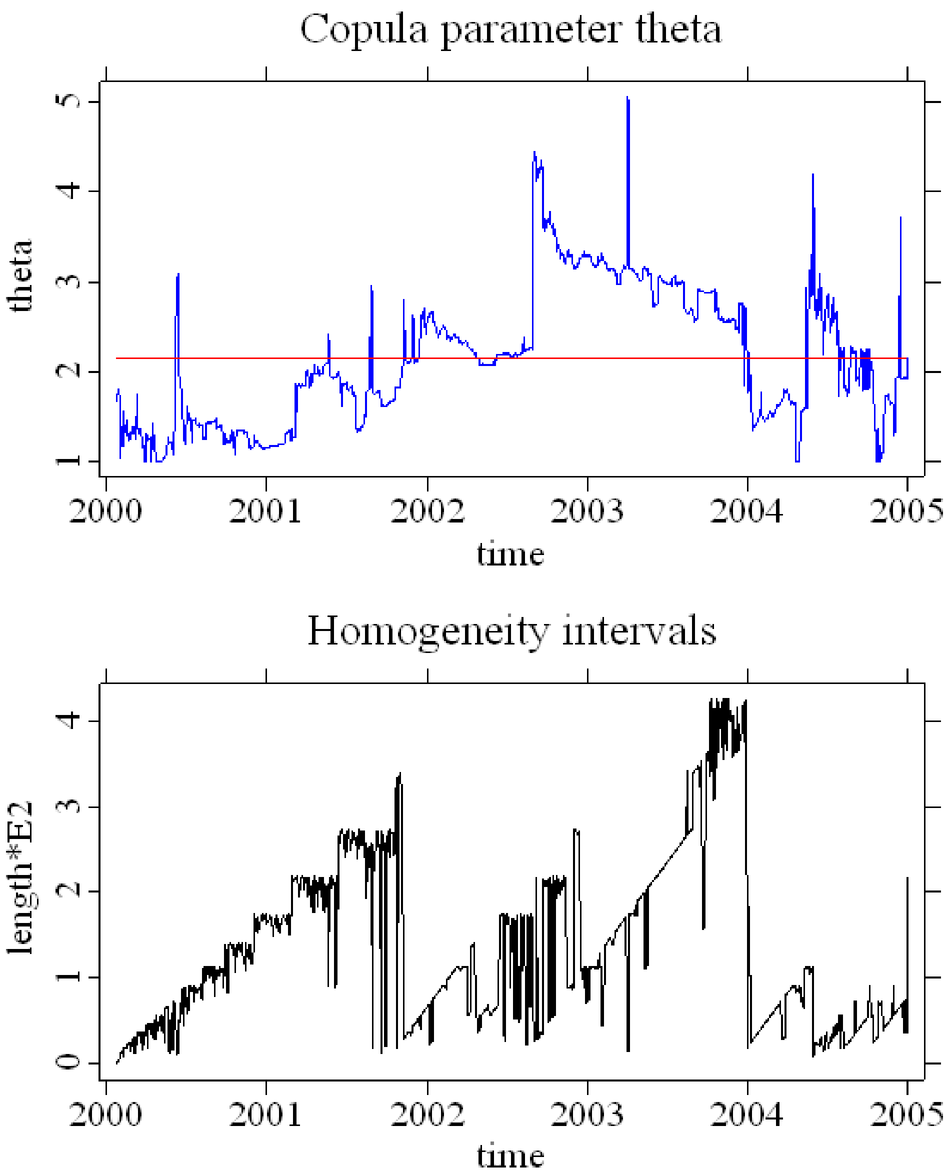


Figure 9.3. Upper panel: estimated copula dependence parameter $\theta$ for Daimler-Chrysler and Volkswagen (blue line) and its mean (red line). Lower panel: estimated intervals of time homogeneity. The results are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

realthetahomlength.xpl

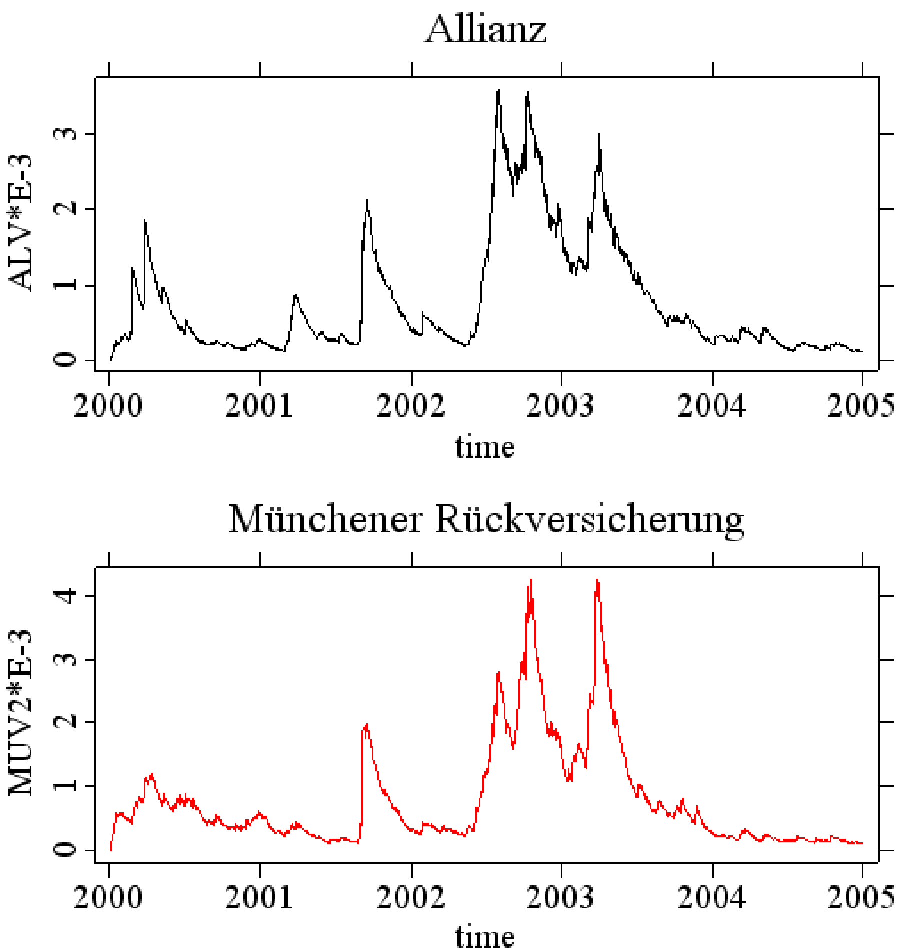


Figure 9.4. Marginal parameters for Allianz (upper panel) and Münchener Rückversicherung (lower panel) estimated by exponential smoothing with parameter $\lambda = 1/20$.

testrealdata.xpl

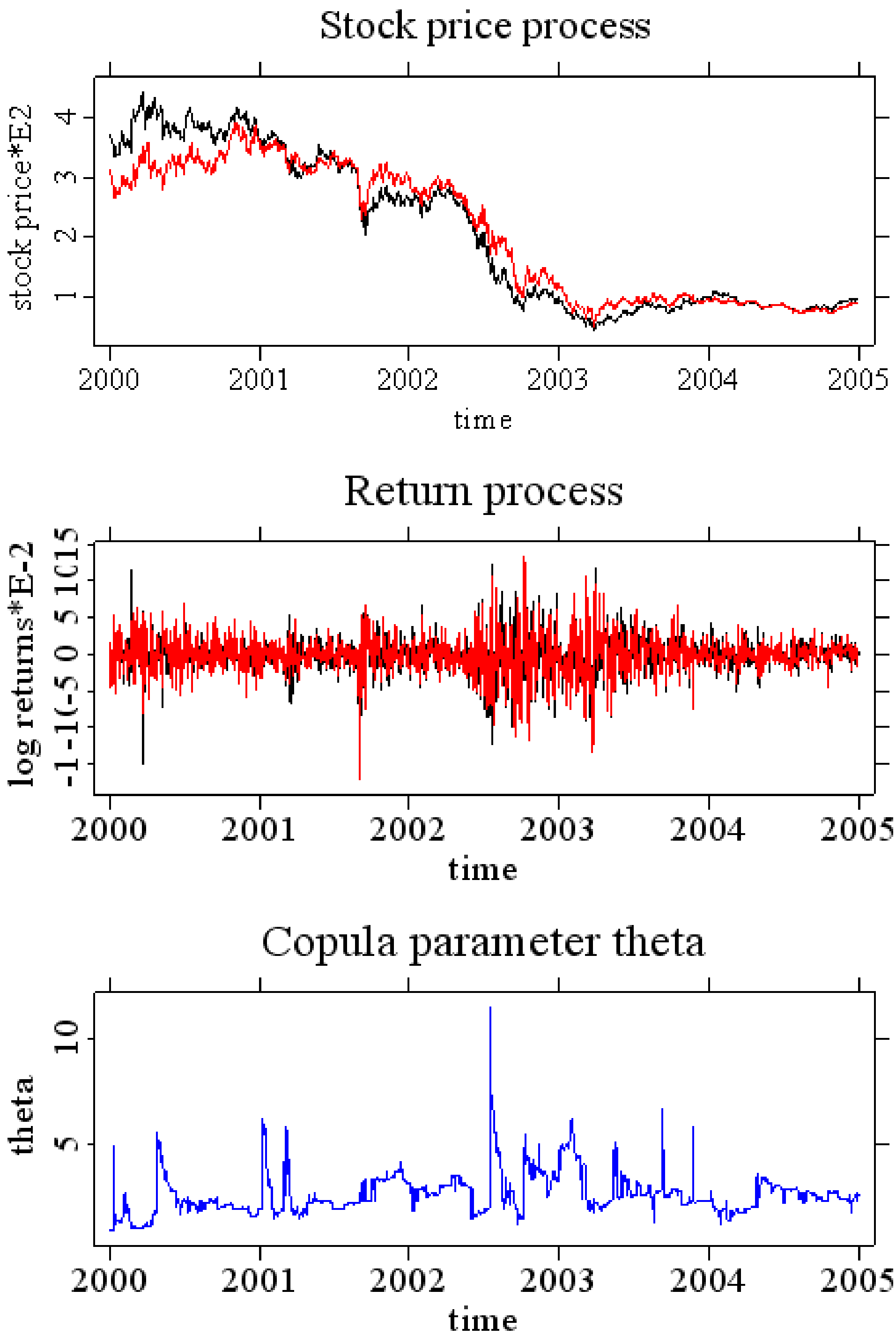


Figure 9.5. Stock price process (upper panel), log returns (middle panel) and copula dependence parameter $\theta$ (lower panel) for Allianz (black line) and Münchener Rückversicherung (red line). The estimates of $\theta$ are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

plot.xpl

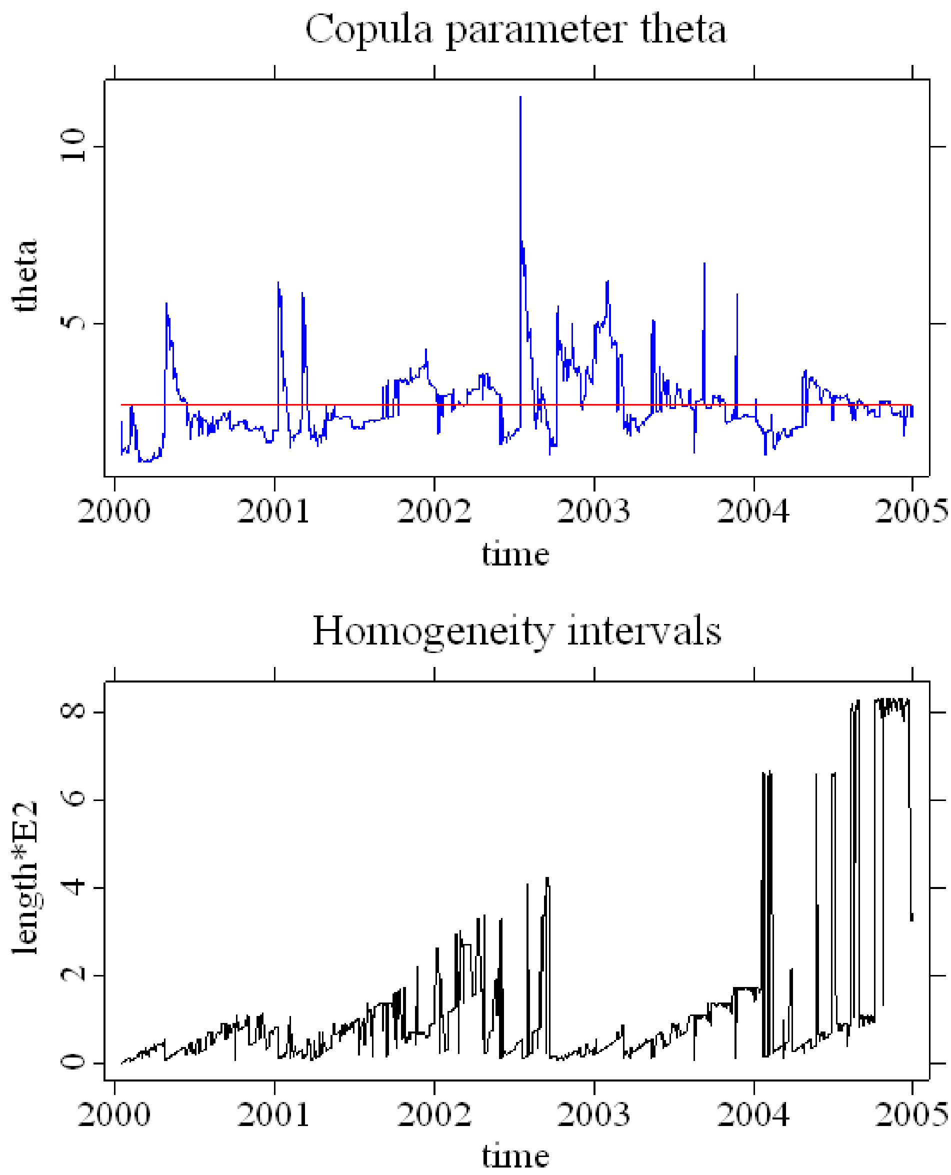


Figure 9.6. Upper panel: estimated copula dependence parameter $\theta$ for Allianz and Münchener Rückversicherung (blue line) and its mean (red line). Lower panel: estimated intervals of time homogeneity. The results are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

realthetahomlength.xpl

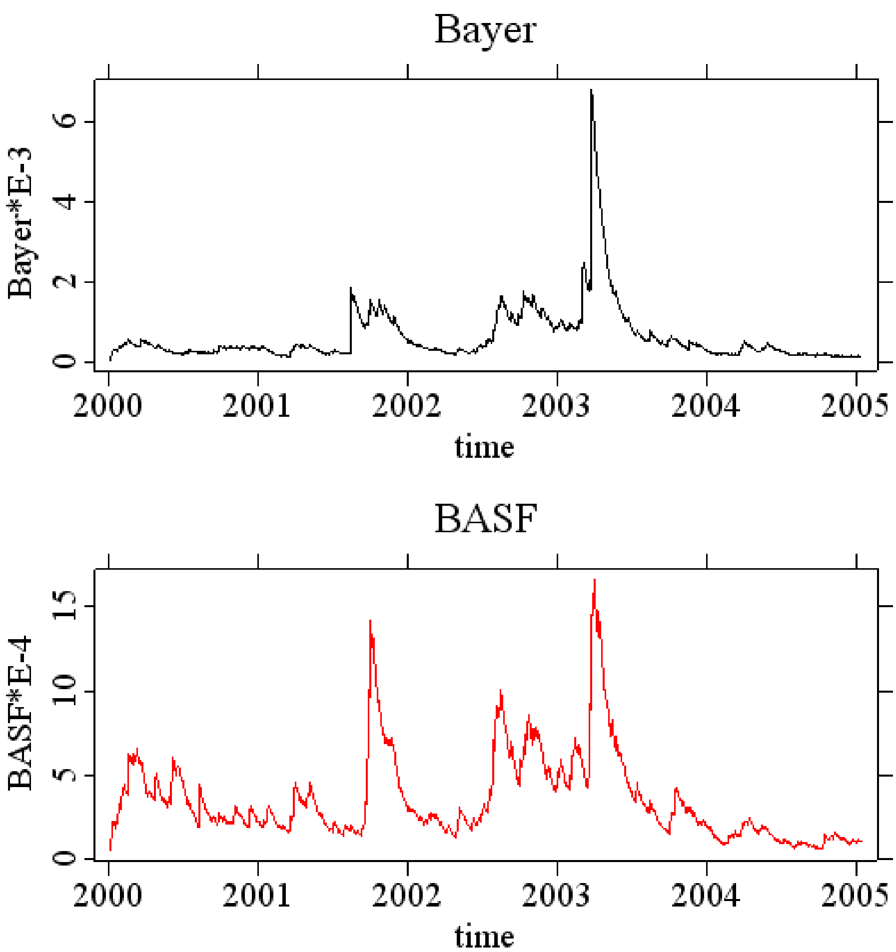


Figure 9.7. Marginal parameters for Bayer (upper panel) and BASF (lower panel) estimated by exponential smoothing with parameter $\lambda = 1/20$.

testrealdata.xpl

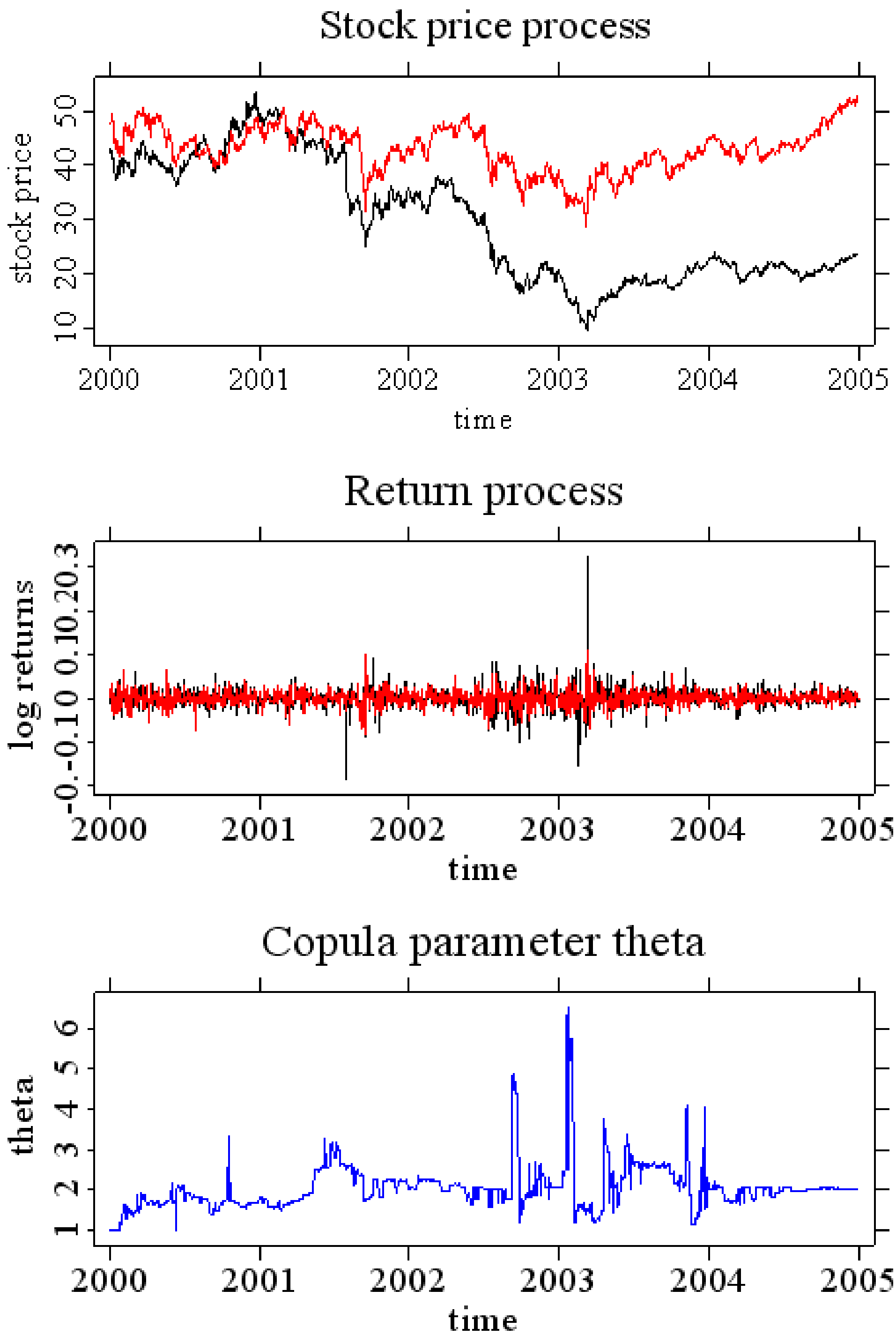


Figure 9.8. Stock price process (upper panel), log returns (middle panel) and copula dependence parameter $\theta$ (lower panel) for Bayer (black line) and BASF (red line). The estimates of $\theta$ are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

plot.xpl

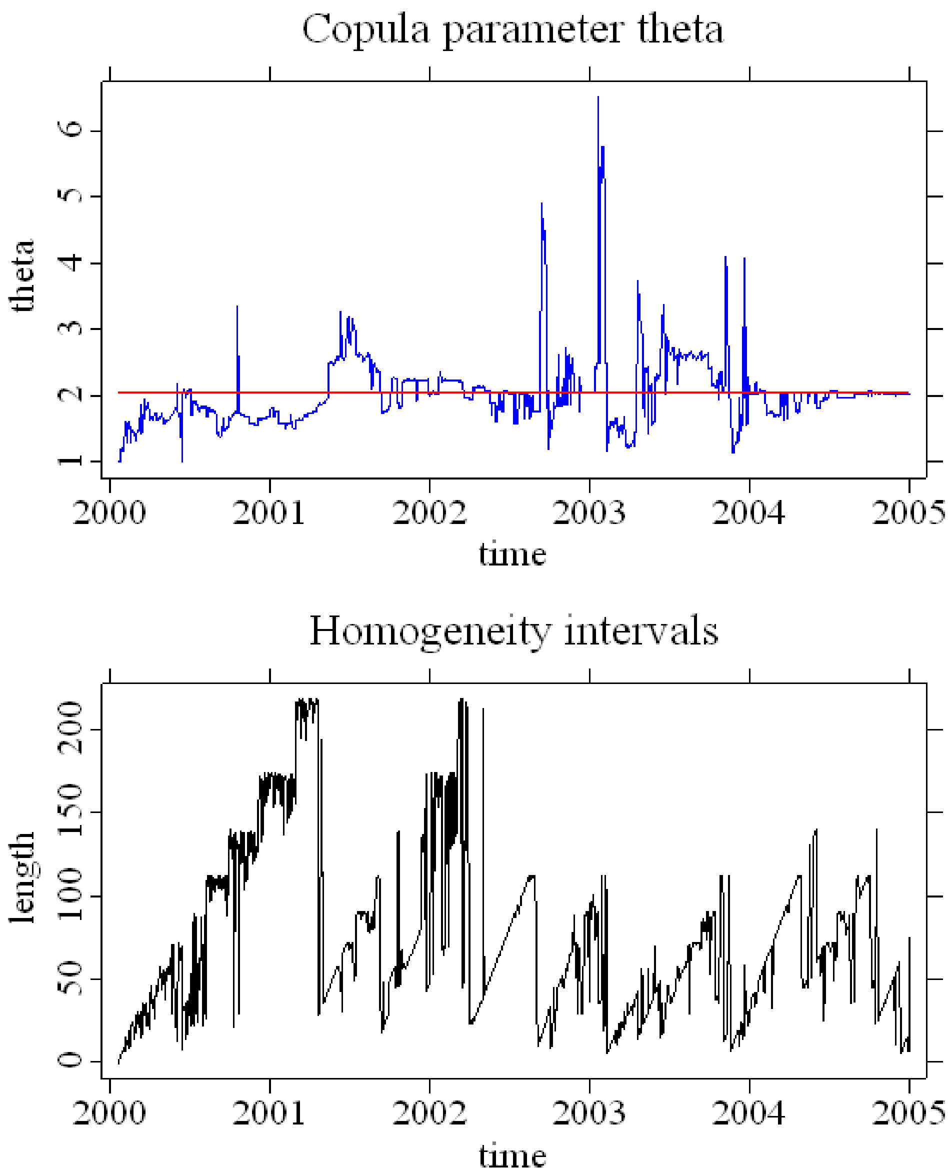


Figure 9.9. Upper panel: estimated copula dependence parameter $\theta$ for Bayer and BASF (blue line) and its mean (red line). Lower panel: estimated intervals of time homogeneity. The results are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

realthetahomlength.xpl

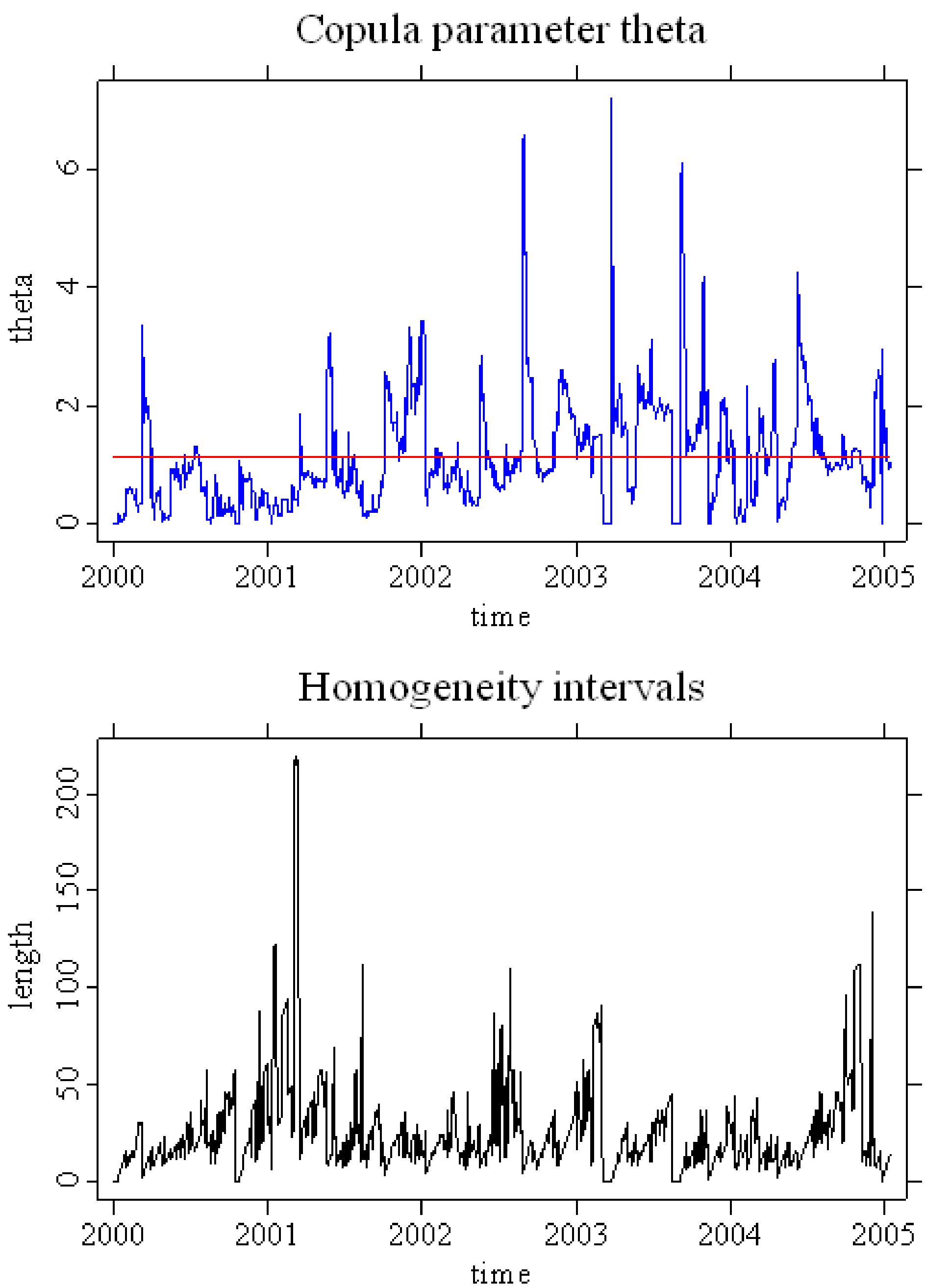


Figure 9.10. Upper panel: estimated copula dependence parameter $\theta$ for 4-dim data: DaimlerChrysler, Volkswagen, Bayer and BASF (blue line) and its mean (red line). Lower panel: estimated intervals of time homogeneity. The results are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

`realthetahomlength.xpl`

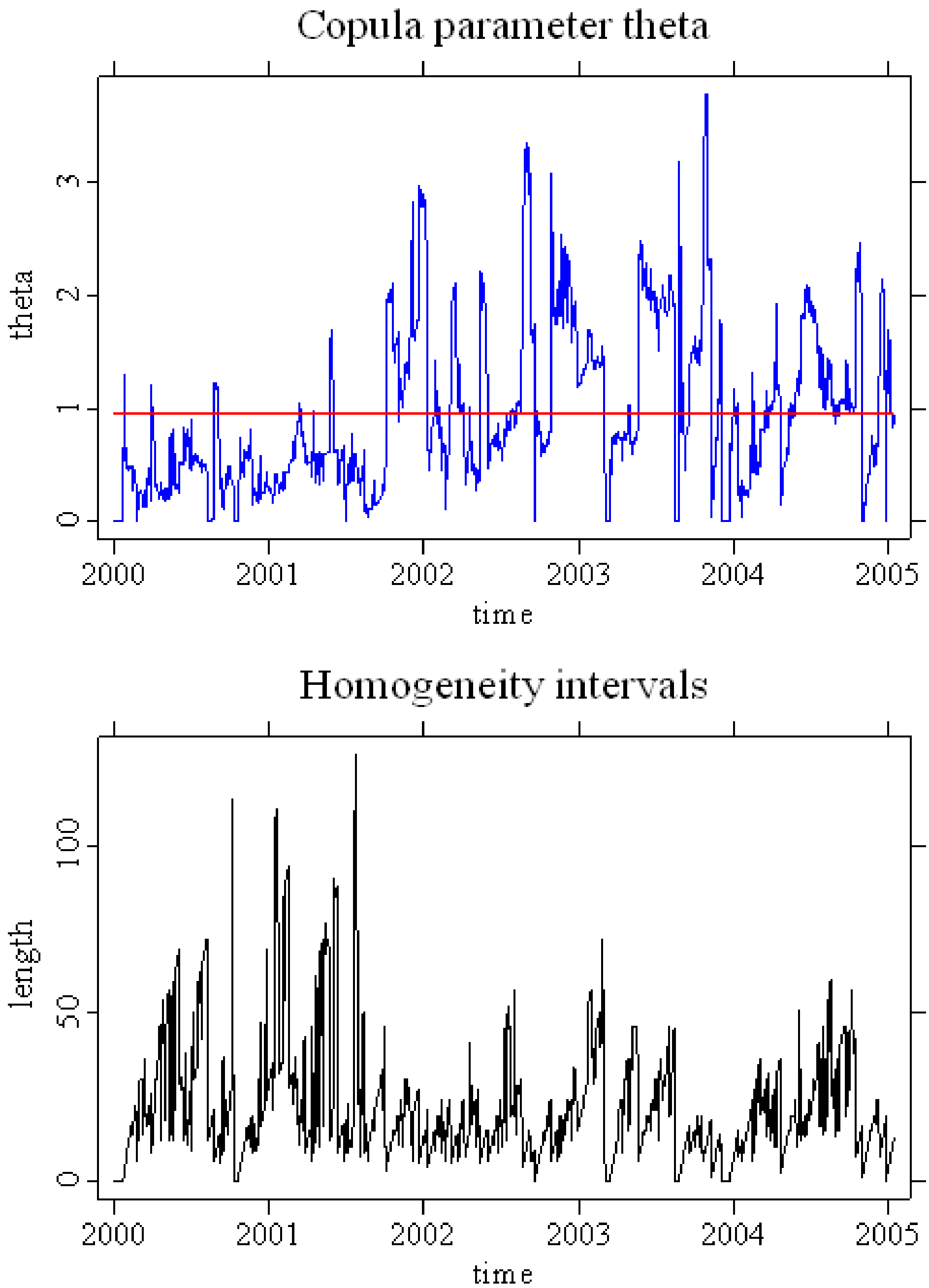


Figure 9.11. Upper panel: estimated copula dependence parameter $\theta$ for 6-dim data: DaimlerChrysler, Volkswagen, Bayer, BASF, Allianz and Münchener Rückversicherung (blue line) and its mean (red line). Lower panel: estimated intervals of time homogeneity. The results are obtained with parameters $m_0 = 20$, $c = 1.25$, $\rho_1 = 0.25$, $\rho_2 = 0.3$ and $\alpha = 0.05$.

realthetahomlength.xpl

# 10 References


Chen, X. and Fan, Y. (2004). Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models Under Copula Misspecification, *working paper.*

Chen, X. and Fan, Y. (2004). Estimation of Copula-Based Semiparametric Time Series Models, *working paper.*

Chen, X., Fan, Y. and Tsyrennikov, V. (2004). Efficient Estimation of Semiparametric Multivariate Copula Models, *working paper.*

Dias, A. and Embrechts, P. (2004) Dynamic Copula Models for Multivariate High-Frequency Data in Finance, *working paper.*

Dias, A. (2004) Copula Inference for Finance and Insurance, *dissertation, ETHZ.*

Embrechts, P., Lindskog, F. and McNeil, A. (2001). Modelling Dependence with Copulas and Application to Risk Management, *working paper.*

Embrechts, P., McNeil, A. and Straumann, D. (2001). Correlation and Dependence in Risk Management: Properties and Pitfalls, *working paper.*

Franke, J., Härdle, W. and Hafner, C. (2004). *Statistics of Financial Markets*, Springer-Verlag, Heidelberg.

Giacomini, E., Härdle, W. (2005). Value-at-Risk Calculations with Time Varying Copulae, *Bulletin of the International Statistical Institute, Vol.55.*

Härdle, W., Kleinow, T. and Stahl, G. (2002). *Applied Quantitative Finance*, Springer-Verlag, Heidelberg.

Mercurio, D. and Spokoiny, V. (2000). Statistical inference for time-inhomogeneous volatility models, *working paper*.

Mercurio, D. and Spokoiny, V. (2004). Estimation of time dependent volatility via local change point analysis with applications to Value-at-Risk, *working paper*.

Nelsen, R. (1998). *An Introduction to Copulas*, Springer-Verlag, New York.

Spokoiny, V. (2006). *Local parametric methods in nonparametric estimation*, Springer-Verlag, Berlin, Heidelberg, NY.