Copulas in Financial Risk Management

Dr Jörn Rank
Department for Continuing Education
University of Oxford

A thesis submitted for the diploma in
Mathematical Finance

4 August 2000

Diploma in Mathematical Finance

Department for Continuing Education

OXFORD UNIVERSITY

Authenticity Statement

I confirm that the work in this project is my own work and I guarantee its authenticity.

Name Dr Jörn Rank

Address Mainstr. 7e
65520 Bad Camberg
Germany

Signed

Date

To Christiane and Maximilian

Acknowledgements

I would like to thank my employer Arthur Andersen Wirtschaftsprüfungsgesellschaft, Steuerberatungsgesellschaft mbH, and in particular Dr Hans-Peter Deutsch, head of the Financial and Commodity Risk Consulting Group, for offering me the opportunity to take part in the course "Diploma in Mathematical Finance" at the University of Oxford. In addition, I would like to thank Dr Jeff Dewynne, Dr Sam Howison and Dr Paul Wilmott for all their interesting lectures during the modules. Of course, many thanks also go to Ms Tiffany Fliss for organizing everything so well and reliably.

My special thank-you is dedicated to my beloved wife Christiane for all her patience during the many weekends that I had to spend on writing this thesis and on other diploma-related work.

Abstract

The Value-at-Risk (VaR) is of central importance in modern financial risk management. Of the various methods that exist to compute the VaR, the most popular are historical simulation, the variance-covariance method, and Monte Carlo (MC) simulation. While historical simulation is not based on particular assumptions as to the behaviour of the risk factors, both the variance-covariance method and MC simulation assume some kind of multinormal distribution of the (relative) risk factors. Therefore the dependence structure between different risk factors is shown by the covariance or correlation between these factors.

It is shown in [1, 2] that the concept of correlation entails several pitfalls. As this might result in an unreliable prediction of the VaR of a financial portfolio, new methods have to be found to avoid these pitfalls. The authors of [1, 2] therefore propose the use of *copulas* to quantify risk.

In [3], the method of copulas is introduced. These functions can be used to describe the dependence between two or more variables with arbitrary marginal distributions. In rough terms, a copula is a function $C : [0,1]^n \to [0,1]$ with certain special properties; thus the joint distribution can be written as

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = C\left(F_1(x_1), \ldots, F_n(x_n)\right) \ ,$$

where $F_1, \ldots, F_n$ denote the cumulated probability functions of the n variables under consideration. In general, a copula C depends on one or more parameters $p_1, \ldots, p_k$ that determine the dependence between the variables $x_1, \ldots, x_n$. In this sense, these parameters assume the role of the correlations.

In this thesis we investigate two market risk factors only, the FX rates USD vs. DEM and GBP vs. DEM, and use the *Gumbel-Hougaard copula* [4] to describe their dependence structure. We present algorithms to estimate the parameter of this copula and to generate pseudo random numbers due to a copula dependence.

Based on about 2000 items of historical data, we compute the VaR using a copula-modified MC algorithm. To see the advantage of this method, we compare these results with VaR results obtained from "traditional" MC simulations and from the variance-covariance method. On the basis of our backtesting results, we find that the "copula method" is more reliable than the other two methods.

Contents

List of Figures

Chapter 1

Introduction

In modern financial risk management, one fundamental quantity of interest is the *Value-at-Risk* (VaR). It denotes the amount by which a portfolio of financial assets might fall at the most with a given probability during a particular time horizon. For the sake of simplification we will only consider time horizons of one day's length in this thesis. This covers portfolios that include assets that can be sold in one day. The discussions would have to be modified slightly if one were interested in longer time horizons.

Typical choices of the probabilities that are connected with VaRs are *confidence levels* of 90%, 95%, and 99%, respectively. If a risk manager is free to choose[1], the particular value he might select will depend on weighing up the reliability that he expects from the VaR prediction against the amount a VaR should have. Using a high confidence level, e.g. 99%, will result in a higher absolute VaR compared with a 90% VaR. If the risk manager compares the predicted VaRs of the portfolio with the real losses of the following business days over a sufficiently long period of time, he will observe that about 1% of the real losses exceed the calculated 99% VaRs, assuming the method the VaR predictions are based on were reliable. However, if he had chosen a confidence level of 90%, backtesting of the data would deliver underestimated losses of 10%.

In statistics, the gap between the confidence level and 100% is called the *quantile* and usually denoted by the Greek letter α. Once the risk manager has chosen a confidence level $1-\alpha$, it is vital that he can trust the predicted VaRs, i.e. that he can be sure that the fraction of underestimated losses is in fact very close to α. Otherwise, the larger the difference is between the fraction of underestimated losses and α, the

[1] In some countries the confidence level to choose is prescribed by law. For example, in the Capital Adequacy Directive, a confidence level of 99% is required for internal models.

less useful the VaR is for managing the risk of the portfolio. Therefore, a reliable method is needed to extract the VaR.

In literature, several methods to calculate the VaR are known. The most popular are

- *historical simulation,*
- the *variance-covariance method*, and
- *Monte Carlo* (MC) *simulation.*

As it is not within the scope of this thesis, we will not explain the first two methods. An excellent description of these and of what we will term "traditional" Monte Carlo can be found in [5].

We next present general information on how to compute the VaR of a portfolio using Monte Carlo simulation. The value of the portfolio at present time t will be denoted by V_t. Let us assume that V_t depends on n risk factors which might be (relative) interest rates, foreign exchange (FX) rates, share prices, etc. Then a Monte Carlo computation of the VaR would consist of the following steps:

1. Choose the level of confidence $1-\alpha$ to which the VaR refers.
2. Simulate the evolution of the risk factors from time t to time $t+1$ by generating n-tupels of pseudo random numbers (PRNs) with an "appropriate" joint distribution that describes the behaviour of the risk factors. The number m of these n-tupels has to be large enough (typically $m = \mathcal{O}(1000)$) to obtain sufficient statistics in step 5.
3. Calculate the m different values of the portfolio at time $t+1$ using the values of the simulated n-tupels of the risk factors. Let us denote these values by $V_{t+1,1}, V_{t+1,2}, \ldots, V_{t+1,m}$.
4. Calculate the simulated profits and losses, i.e. the differences between the simulated future portfolio values and the present portfolio value, $\Delta V_i = V_{t+1,i} - V_t$ for $i = 1, \ldots, m$.
5. Ignore the fraction of the α worst changes ΔV_i. The minimum of the remaining ΔV_i's is then the VaR of the portfolio at time t. It will be denoted by $\mathrm{VaR}(\alpha, t, t+1)$.

As soon as the time evolves from t to $t+1$, the real value of the (unchanged) portfolio changes from V_t to V_{t+1}. With this data at hand, one can backtest $\mathrm{VaR}(\alpha, t, t+1)$ by comparing it with $\Delta V = V_{t+1} - V_t$.

It is obvious that the central point in the above algorithm is step two, the generation of the PRNs according to the "appropriate" joint distribution. Therefore the following questions have to be answered:

Question 1 *What is the "appropriate" joint distribution of the risk factors under consideration?*

Question 2 *Is it possible to simulate PRNs using this joint distribution?*

Finding answers to the questions above for two risk factors is the main objective of this work. For this purpose, this thesis is organized as follows:
In the following chapter, the "traditional" Monte Carlo method is presented. In chapter 3, *copulas* are defined. Examples of these functions are given, and the basic properties are summarized. Copulas can be used to describe the dependence structure between two or more variables. In this sense they replace the concept of covariance or correlation that is usually used to describe the dependence between variables. Chapter 4 is concerned with the simultaneous generation of pairs of PRNs whose dependence structure is determined by a copula.

After the theoretical overview we present in chapter 5 numerical results of a VaR Monte Carlo simulation based on copulas. The portfolio for which the VaR is calculated is affected by two risk factors only, the FX rates USD/DEM and GBP/DEM. To test the quality of this method, we backtest our data and compare the results with numerical data obtained from "traditional" Monte Carlo computations and with variance-covariance data. Finally, we give a summary and our conclusions in chapter 6.

Chapter 2

"Traditional" Monte Carlo Simulation of the Value-at-Risk

In this chapter we want to recall the joint simulation of risk factors which is used with most Monte Carlo based methods that are used for computing the VaR. In other words, we want to explain the second step on page 2 in more detail.

As the algorithm we will present next is the one that is most often used in practice, we will refer to it as the "traditional method". This is to distinguish it from the new method which will be presented in chapter 4 and which we will call the "copula method".

Let us briefly recapitulate the content of the second step mentioned above. It states that we have to generate n-tupels of PRNs with an "appropriate" joint distribution that describes the behaviour of the risk factors. If one only takes FX rates as risk factors into account, the "traditional" ansatz to achieve these PRNs consists of the following steps:

1. Collect historical data of the n FX rates, i.e. n time series spanning $N+1$ business days. We denote these data by $x_{i,0}, x_{i,1}, \ldots, x_{i,N}$ for $i = 1, \ldots, n$, where today's data are given by the $x_{i,N}$'s.
   A typical choice for $N+1$ is, for example, $N+1 = 250$.

2. Assuming $x_{i,j} \neq 0$, compute the relative changes in the data:

$$r_{i,j} = \frac{x_{i,j} - x_{i,j-1}}{x_{i,j-1}} \quad \text{for} \quad i = 1, \ldots n \quad \text{and} \quad j = 1, \ldots, N\,. \tag{2.1}$$

   For each $i = 1, \ldots, n$, the $r_{i,1}, \ldots, r_{i,N}$ will be considered as elements belonging to a random variable r_i.

3. The next step is to make an assumption as to the marginal distributions $f_1, \ldots, f_n$ of the random variables $r_1, \ldots, r_n$.
   As the $r_1, \ldots, r_n$ originally come from FX rates, the assumption usually made in finance at this point is that the f_i's are given by normal distributions $N(\mu_i, \sigma_i^2)$,

$$f_i(r_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{-\frac{(r_i - \mu_i)^2}{2\sigma_i^2}\right\} \tag{2.2}$$

   for $i = 1, \ldots, n$. In (2.2), μ_i and σ_i^2 denote the mean and the variance of r_i,

$$\mu_i = \mathbb{E}(r_i)\,, \quad \sigma_i^2 = \mathbb{E}((r_i - \mu_i)^2)\,. \tag{2.3}$$

4. Once the marginal distributions have been chosen, one has to determine the parameters of the distributions.
   In the case of normal distributions, this task is easy.[1] It is reduced to calculate the estimators for the means and the variances,

$$\hat{\mu}_i = \frac{1}{N}\sum_{j=1}^{N} r_{i,j} \quad \text{and} \tag{2.4}$$

$$\hat{\sigma}_i^2 = \frac{1}{N-1}\sum_{j=1}^{N} (r_{i,j} - \hat{\mu}_i)^2 \tag{2.5}$$

   for $i = 1, \ldots, n$.

5. As stated above, we want to generate n-tupels of PRNs according to a joint distribution. Therefore we have to make a further assumption, i.e. an assumption that describes the dependence structure of the random variables. As the marginal distributions have already been chosen, one is, of course, not completely free to choose the joint distribution. If $f(\vec{r})$ denotes the joint distribution, one has to ensure that the following condition is true for each $i = 1, \ldots, n$:

$$f_i(r_i) = \int_{-\infty}^{\infty} \mathrm{d}r_1 \ldots \int_{-\infty}^{\infty} \mathrm{d}r_{i-1} \int_{-\infty}^{\infty} \mathrm{d}r_{i+1} \ldots \int_{-\infty}^{\infty} \mathrm{d}r_n\, f(\vec{r})\,. \tag{2.6}$$

   Let us come back to our "traditional" Monte Carlo simulation. Here one assumes at this point a multinormal distribution,

$$f(\vec{r}) = \frac{1}{\sqrt{(2\pi)^n \det \mathcal{C}}} \exp\left\{-\frac{1}{2}(\vec{r} - \vec{\mu})^t\, \mathcal{C}^{-1}\, (\vec{r} - \vec{\mu})\right\} \tag{2.7}$$

[1] "Easy" in the sense that one does not have to fit the data to obtain the parameters.

where

$$\vec{r} = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} , \quad \vec{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} , \tag{2.8}$$

and the normalization condition

$$\int_{-\infty}^{\infty} \mathrm{d}r_1 \ldots \int_{-\infty}^{\infty} \mathrm{d}r_n \, f(\vec{r}) = 1 \, . \tag{2.9}$$

In (2.7), $\mathcal{C}$ denotes the covariance matrix,

$$\mathcal{C} = \begin{pmatrix} \sigma_1^2 & c_{1,2} & c_{1,3} & \cdots & c_{1,n} \\ c_{1,2} & \sigma_2^2 & c_{2,3} & \cdots & c_{2,n} \\ c_{1,3} & c_{2,3} & \sigma_3^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & c_{n-1,n} \\ c_{1,n} & c_{2,n} & \cdots & c_{n-1,n} & \sigma_n^2 \end{pmatrix} , \tag{2.10}$$

and $c_{i,j}$ are the covariance between r_i and r_j,

$$c_{i,j} = \mathbb{E}\left((r_i - \mu_i)(r_j - \mu_j)\right) \, . \tag{2.11}$$

For the sake of completeness, we introduce the term of correlation at this point. It is given by

$$\rho_{i,j} \equiv \frac{c_{i,j}}{\sigma_i \sigma_j} \quad \text{if } \sigma_i, \sigma_j \neq 0 \, . \tag{2.12}$$

One can easily show that the function f given by (2.7) fulfils condition (2.6). However, we would like to emphasize that the normal marginal distributions do *not* enforce the assumption of a joint distribution of the multinormal kind. This point is essential for understanding the copula ansatz. Furthermore, it was pointed out in [1, 2] that a dependence structure, which is, like the multinormal distribution, described by correlation, entails several pitfalls.

6. This step of the algorithm consists of a determination of the covariances (2.11). In a similar way as the means and variances in step 4 they can be estimated as follows:

$$\hat{c}_{i,j} = \frac{1}{N-1} \sum_{k=1}^{N} \left(r_{i,k} - \hat{\mu}_i\right)\left(r_{j,k} - \hat{\mu}_j\right) \, . \tag{2.13}$$

7. As we have chosen a joint distribution (step 5) and have estimated its parameters (steps 4 and 6) we now have all tools at hand to generate the desired n-tupels of PRNs. Because this is a standard technical procedure, we present it in detail in appendix A.3. The relevant steps required are given on page 31, steps (a)

and (b) for n dimensions, and steps ($\tilde{\text{a}}$) and ($\tilde{\text{b}}$) for two dimensions only. Let us therefore assume at this point that we have generated n-tupels of PRNs. We denote these numbers by

$$\vec{r}^{\,k} = \begin{pmatrix} r_1^k \\ \vdots \\ r_n^k \end{pmatrix} \tag{2.14}$$

where $k = 1, \ldots, m$, and m is the number of Monte Carlo iterations.

8. The previous step provided us with m independent n-tupels of PRNs. However, these numbers are related to the relative changes in the data, see step 2. Therefore one finally has to compute the absolute values of the simulated risk factors. From (2.1) and (2.14) the m simulated values of the n risk factors at time step $N + 1$ are given by

$$x_{i,N+1}^k = x_{i,N} \, (1 + r_i^k) \tag{2.15}$$

where $i = 1, \ldots, n$ and $k = 1, \ldots, m$.

So far, we have presented the "traditional method" of generating PRNs, assuming that the joint distribution of the relative risk factors is given by a binormal distribution, see step 5 and especially equation (2.7). As already mentioned, the assumption of the binormal distribution is not conclusive. At this point, the ansatz for the "copula method" comes in. To be able to understand it, we present the basic features of copulas in the following chapter. Chapter 4 then shows how to replace the steps 5–7 from the algorithm above.

Chapter 3

Basic Features of Copulas

In this chapter we summarize the basic definitions that are necessary to understand the concept of copulas. We then present the most important properties of copulas that are needed to understand the usage of copulas in finance.

We will follow the notation used in [3]. As it would exceed the scope of this thesis, we do not present the proofs of the theorems cited. At this point we again refer to the excellent textbook by Nelson [3]. Furthermore, we will restrict ourselves to copulas of two dimensions only. The generalisation to n dimensions is straightforward and can also be found in [3].

3.1 Definition of a Copula

Definition 1 A *two-dimensional copula* is a function $C : [0,1] \times [0,1] \to [0,1]$ with the following properties:

1. For every $u, v \in [0,1]$:
$$C(u,0) = C(0,v) = 0\,. \tag{3.1}$$

2. For every $u, v \in [0,1]$:
$$C(u,1) = u \quad \text{and} \quad C(1,v) = v\,. \tag{3.2}$$

3. For every $u_1, u_2, v_1, v_2 \in [0,1]$ with $u_1 \leq u_2$ and $v_1 \leq v_2$:
$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \geq 0\,. \tag{3.3}$$

Below we will use the shorthand notation *copula* for C. As the usage of the name copula for the function C is not obvious from definition 1, it will be explained at the end of section 3.3.

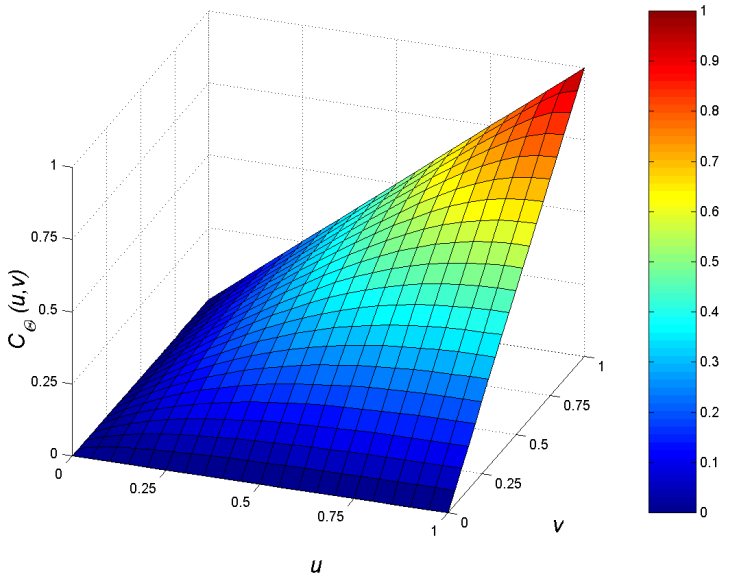

Figure 3.1: The Gumbel-Hougaard copula $C_\theta(u, v)$ for $\theta = 2.0$.

A function that fulfils property 1 is also said to be *grounded*. Property 2 is the two-dimensional analogue of a nondecreasing one-dimensional function. A function with this feature is therefore called *2-increasing*.

3.2 Examples of Copulas

3.2.1 Gumbel-Hougaard Copulas

The first example we want to consider is the *Gumbel-Hougaard* family of copulas [4]. It is given by the function

$$C_\theta(u, v) \equiv \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right\} . \tag{3.4}$$

The parameter θ may take all values in the interval $[1, \infty[$. In figure 3.1 we present the graph of a Gumbel-Hougaard copula for $\theta = 2.0$.

A discussion in [3] shows that C_θ is suited to describe bivariate extreme value distributions. We will come back to this point later in this thesis on page 25. For $\theta = 1$, expression (3.4) reduces to $C_1(u, v) = \Pi(u, v)$, the *product copula*,

$$\Pi(u, v) \equiv u\, v . \tag{3.5}$$

We will discuss Π in the context of random variables in more detail in theorem 4 in section 3.4.

For $\theta \to \infty$ one finds for the Gumbel-Hougaard copula $C_\theta(u,v) \stackrel{\theta\to\infty}{\longrightarrow} \min(u,v) \equiv M(u,v)$. It can be shown that M is also a copula. Furthermore, for any given copula C one has $C(u,v) \leq M(u,v)$, and M is called the *Fréchet-Hoeffding upper bound.*[1]

3.2.2 Gaussian Copulas

The second important copula that we want to investigate is the *Gaussian* or *normal copula* [2],

$$C_{\text{Gauss}}(u,v) \equiv \int_{-\infty}^{\Phi_1^{-1}(u)} \mathrm{d}r_1 \int_{-\infty}^{\Phi_2^{-1}(v)} \mathrm{d}r_2 \; f(r_1,r_2) \; . \tag{3.6}$$

In (3.6), f denotes the bivariate normal density function, i.e. equation (2.7) for $n=2$. The function Φ_1 in (3.6) refers to the one-dimensional, cumulated normal density function,

$$\Phi_1(r_1) = \int_{-\infty}^{r_1} \mathrm{d}r_1' \; f_1(r_1') \; , \tag{3.7}$$

where f_1 is given by equation (2.2). For Φ_2 an analogical expression holds.

If the covariance between r_1 and r_2 is zero, the Gaussian copula becomes

$$\begin{aligned} C_{\text{Gauss}}(u,v) &= \int_{-\infty}^{\Phi_1^{-1}(u)} \mathrm{d}r_1 \; f_1(r_1) \int_{-\infty}^{\Phi_2^{-1}(v)} \mathrm{d}r_2 \; f_2(r_2) \\ &= u\,v \\ &= \Pi(u,v) \quad \text{if} \quad c_{1,2} = 0 \; . \end{aligned} \tag{3.8}$$

As we will see in section 3.4, result (3.8) is a direct consequence of theorem 4.

As $\Phi_1(r_1), \Phi_2(r_2) \in [0,1]$, one can replace u, v in (3.6) by $\Phi_1(r_1), \Phi_2(r_2)$. If one considers r_1, r_2 in a probabilistic sense, i.e. r_1 and r_2 being values of two random variables R_1 and R_2, one obtains from (3.6)

$$C_{\text{Gauss}}(\Phi_1(r_1), \Phi_2(r_2)) = \mathbb{P}(R_1 \leq r_1, R_2 \leq r_2) \; . \tag{3.9}$$

In other words: $C_{\text{Gauss}}(\Phi_1(r_1), \Phi_2(r_2))$ is the binormal cumulated probability function.

3.3 Sklar's Theorem and Further Important Properties of Copulas

In this section we focus on the properties of copulas. The theorem we will present next establishes the continuity of copulas via a Lipschitz condition on $[0,1] \times [0,1]$:

[1] In addition, the two-dimensional function $W(u,v) \equiv \max(u+v-1, 0)$ defines a copula with $W(u,v) \leq C(u,v)$ for any other copula C. W is called the *Fréchet-Hoeffding lower bound.*

Theorem 1 Let C be a copula. Then for every $u_1, u_2, v_1, v_2 \in [0,1]$:

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1| \,. \tag{3.10}$$

From (3.10) it follows that every copula C is uniformly continuous on its domain.

A further important property of copulas concerns the partial derivatives of a copula with respect to its variables:

Theorem 2 Let C be a copula. For every $u \in [0,1]$, the partial derivative $\partial C/\partial v$ exists for almost all[2] $v \in [0,1]$. For such u and v one has

$$0 \leq \frac{\partial}{\partial v} C(u,v) \leq 1 \,. \tag{3.11}$$

The analogous statement is true for the partial derivative $\partial C/\partial u$.
In addition, the functions $u \to C_v(u) \equiv \partial C(u,v)/\partial v$ and $v \to C_u(v) \equiv \partial C(u,v)/\partial u$ are defined and nondecreasing almost everywhere on [0,1].

To give an example of this theorem, we consider the partial derivative of the Gumbel-Hougaard copula (3.4) with respect to u,

$$\begin{aligned} C_{\theta,u}(v) = \frac{\partial}{\partial u} C_\theta(u,v) \quad &= \quad \exp\left\{ - \left[(-\ln u)^\theta + (-\ln v)^\theta \right]^{1/\theta} \right\} \times \\ & \qquad \left[(-\ln u)^\theta + (-\ln v)^\theta \right]^{-\frac{\theta-1}{\theta}} \; \frac{(-\ln u)^{\theta-1}}{u} \,. \end{aligned} \tag{3.12}$$

In figure 3.2, $C_{\theta,u}$ is shown for $u = 0.1, 0.2, \ldots, 0.9$ and $\theta = 2.0$. Note that for $u \in]0,1[$ and for all $\theta \in \mathbb{R}$ where $\theta > 1$, $C_{\theta,u}$ is a strictly increasing function of v. Therefore the inverse function $C_{\theta,u}^{-1}$ is well defined. However, as one might guess from (3.12), $C_{\theta,u}^{-1}$ can not be calculated analytically so that some kind of numerical algorithm has to be used for this task. As C_θ is symmetric in u and v, the partial derivative of C_θ with respect to v shows an identical behaviour for the same set of parameters.

We now turn to the central theorem in the theory of copulas, *Sklar's theorem* [6]. To be able to understand it, we recall some terms known from statistics.

Definition 2 A *distribution function* is a function $F : \overline{\mathbb{R}} \to [0,1]$ with the following properties:

1. F is nondecreasing.

2. $F(-\infty) = 0$ and $F(+\infty) = 1$.

[2]The expression "almost all" is used in the sense of the Lebesgue measure.

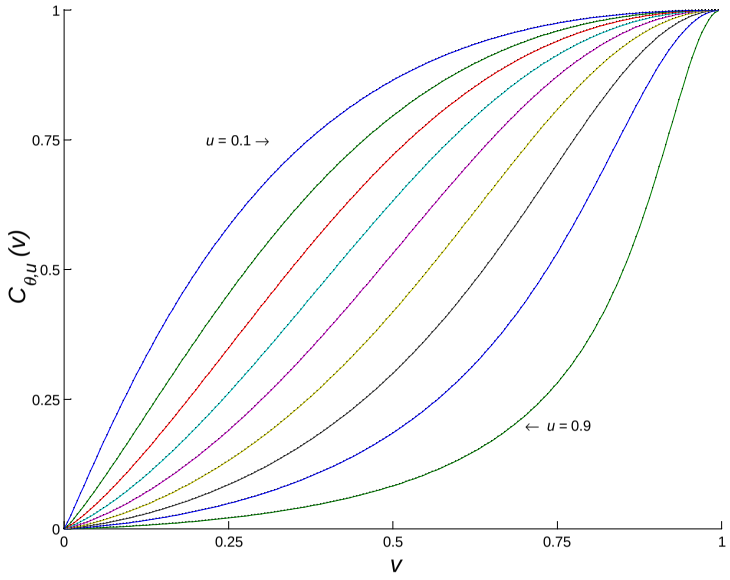

Figure 3.2: The partial derivative $C_{\theta,u}(v)$ for $u = 0.1, 0.2, \ldots, 0.9$ and $\theta = 2.0$.

For example, if f is a probability density function, then the cumulated probability function

$$F(x) = \int_{-\infty}^{x} \mathrm{d}t\, f(t) \tag{3.13}$$

is a distribution function.

Definition 3 A *joint distribution function* is a function $H : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to [0, 1]$ with the following properties:

1. H is 2-increasing.

2. For every $x_1, x_2 \in \overline{\mathbb{R}}$: $H(x_1, -\infty) = H(-\infty, x_2) = 0$ and $H(+\infty, +\infty) = 1$.

It is shown in [3] that H has margins F_1 and F_2 that are given by $F_1(x_1) \equiv H(x_1, +\infty)$ and $F_2(x_2) \equiv H(+\infty, x_2)$, respectively. Furthermore, F_1 and F_2 themselves are distribution functions.

In section 3.2.2 we have already given an example of a two-dimensional function satisfying the properties of definition 3. If one considers C_{Gauss} as given in (3.9) as a function of r_1 and r_2, it can be seen as the joint distribution function belonging to the binormal distribution.

We now have all the definitions that we need for

Theorem 3 (Sklar's theorem) Let H be a joint distribution function with margins F_1 and F_2. Then there exists a copula C with

$$H(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \tag{3.14}$$

for every $x_1, x_2 \in \overline{\mathbb{R}}$. If F_1 and F_2 are continuous, then C is unique. Otherwise, C is uniquely determined on Range $F_1 \times$ Range F_2. On the other hand, if C is a copula and F_1 and F_2 are distribution functions, then the function H defined by (3.14) is a joint distribution function with margins F_1 and F_2.

With Sklar's theorem in mind the use of the name "copula" becomes obvious. It was chosen by Sklar to describe [7] "a function that links a multidimensional distribution to its one-dimensional margins" and appeared in mathematical literature for the first time in [6]. Usually, the term "copula" is used in grammar to describe an expression that links a subject and a predicate. The origin of the word is Latin.

3.4 Copulas and Random Variables

So far we have considered functions of one or two variables. We now turn our attention to random variables. As briefly mentioned at the end of section 3.2.2, we will denote random variables by capital letters and their values by lower case letters. In the definition of a random variable and its distribution function we again follow [3]:

Definition 4 A *random variable* R is a quantity whose values are described by a probability function $f(r)$, where r takes all possible values of R.

Note that it is not stated whether or not the probability function f is known.

Definition 5 The *distribution function* of a random variable R is a function F that assigns all $r \in \overline{\mathbb{R}}$ a probability $F(r) = \mathbb{P}(R \leq r)$. In addition, the *joint distribution function* of two random variables R_1, R_2 is a function H that assigns all $r_1, r_2 \in \overline{\mathbb{R}}$ a probability $H(r_1, r_2) = \mathbb{P}(R_1 \leq r_1, R_2 \leq r_2)$.

We want to emphasize that (joint) distribution functions in the sense of definition 5 are also (joint) distribution functions in the sense of definitions 2 and 3. Therefore equation (3.14) from Sklar's theorem can be written as

$$H(r_1, r_2) = C(F_1(r_1), F_2(r_2)) = \mathbb{P}(R_1 \leq r_1, R_2 \leq r_2) \tag{3.15}$$

Before we can present the next important property of copulas we first have to introduce the term of independent variables:

Definition 6 Two random variables R_1 and R_2 are *independent* if and only if the product of their distribution functions F_1 and F_2 equals their joint distribution function H,

$$H(r_1, r_2) = F_1(r_1)\, F_2(r_2) \quad \text{for all} \quad r_1, r_2 \in \overline{\mathbb{R}}\,. \tag{3.16}$$

In section 3.2 we have already mentioned the product copula Π, see equation (3.5). The next theorem clarifies the importance of this copula.

Theorem 4 Let R_1 and R_2 be random variables with continuous distribution functions F_1 and F_2 and joint distribution function H. From Sklar's theorem we know that there exists a unique copula $C_{R_1R_2}$ with

$$\mathbb{P}(R_1 \leq r_1, R_2 \leq r_2) = H(r_1, r_2) = C_{R_1R_2}(F_1(r_1), F_2(r_2))\,. \tag{3.17}$$

Then R_1 and R_2 are independent if and only if $C_{R_1R_2} = \Pi$.

We will end this section with a statement on the behaviour of copulas under strictly monotone transformations of random variables.

Theorem 5 Let R_1 and R_2 be random variables with continuous distribution functions and with copula $C_{R_1R_2}$. If α_1 and α_2 are strictly increasing functions on Range R_1 and Range R_2, then $C_{\alpha_1(R_1)\,\alpha_2(R_2)} = C_{R_1R_2}$. In other words: $C_{R_1R_2}$ is invariant under strictly increasing transformations of R_1 and R_2.

Chapter 4

Generation of Pseudo Random Numbers Using Copulas

In the last chapter we presented the definition of copulas and their most important properties, especially Sklar's theorem. We are thus ready to return to the Monte Carlo simulation of the Value-at-Risk. The strategy, as already outlined at the end of chapter 2, is to give a general guide on how to generate pairs of PRNs based on random variables whose dependence structure is defined by a copula. This guide will follow very closely the instructions given in [3]. As the numerical results presented in the following chapter assume a Gumbel-Hougaard copula, we will discuss this copula in more detail.

4.1 The General Method

Let us assume that the copula C we are interested in is known, i.e. all parameters of the copula are known. The task is then to generate pairs (u, v) of observations of in $[0, 1]$ uniformly distributed random variables U and V whose joint distribution function is C. To reach this goal we will use the method of conditional distributions. Let c_u denote the conditional distribution function for the random variable V at a given value u of U,

$$c_u(v) \equiv \mathbb{P}(V \leq v, U = u) \; . \tag{4.1}$$

From (3.15) we have[1]

$$c_u(v) = \lim_{\Delta u \to 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} = \frac{\partial}{\partial u} C(u, v) = C_u(v) \; , \tag{4.2}$$

[1]Note that if F_{uniform} denotes the distribution function of a random variable U which is uniformly distributed in $[0, 1]$, one has $F_{\text{uniform}}(u) = u$.

where C_u is the partial derivative of the copula. From theorem 2 we know that $c_u(v)$ is nondecreasing and exists for almost all $v \in [0,1]$.

For the sake of simplicity, we assume from now on that c_u is strictly increasing and exists for all $v \in [0,1]$. As discussed in section 3.3, this is at least true for the Gumbel-Hougaard family of copulas. If these conditions are not fulfilled, one has to replace the term "inverse" in the remaining part of this section by "quasi-inverse" (see [3] again for details).

With result (4.2) at hand we can now use the method of variable transformation, which is described in appendix A.1, to generate the desired pair (u, v) of PRNs. The algorithm consists of the following two steps:

- Generate two independent uniform PRNs $u, w \in [0,1]$. u is already the first number we are looking for.
- Compute the inverse function of c_u. In general, it will depend on the parameters of the copula and on u, which can be seen, in this context, as an additional parameter of c_u. Set $v = c_u^{-1}(w)$ to obtain the second PRN.

It may happen that the inverse function cannot be calculated analytically. In this case one has to use a numerical algorithm to determine v. For example, this situation occurs when Gumbel-Hougaard copulas are used, as already mentioned in section 3.3.

4.2 Pseudo Random Numbers According to the Gumbel-Hougaard Copula

In this section we present a detailed example of the application of the "copula method" in the calculation of the Value-at-Risk. We refer to the algorithm presented in chapter 2 and will use the same notation.

As stated at the end of chapter 2, we want to replace steps 5–7 of the "traditional" algorithm. On a more abstract level, these steps have the following content:

Step 5: Choose a joint distribution function that describes the dependence structure of the problem under consideration.

Step 6: Determine the parameter(s) of the joint distribution function from historical data.

Step 7: Generate n-tupels (or simply pairs) of PRNs according to this distribution.

Let us first turn to step 5 as mentioned on page 5. There we have chosen the multinormal distribution function to describe the dependence structure. As the numerical results that we will present in the next chapter are based on two risk factors only, we will only consider the case $n = 2$ below. We thus have to replace a binormal distribution function by an alternative dependence structure.

With Sklar's theorem in mind, we now make the following assumption:

Step 5 ("copula method"): The joint distribution function of the two risk factors is given by a Gumbel-Hougaard copula,

$$C_\theta(\Phi_1(r_1), \Phi_2(r_2)) = \mathbb{P}(R_1 \leq r_1, R_2 \leq r_2) \ , \tag{4.3}$$

where Φ_i is the cumulated normal density function as given by equation (3.7).

By comparing (4.3) with expression (3.9) it becomes obvious that what we call the "copula method" actually involves replacing the Gaussian copula in the "traditional method" by the Gumbel-Hougaard copula.

Next we consider step 6 of the original algorithm. For two risk factors only, it states how to estimate the covariance (or the correlation, see (2.12)) between these factors. Note that this is the only parameter that results from the assumption that the joint distribution is of the binormal kind. From (4.3) it is obvious that the role of the correlation $\rho \equiv \rho_{1,2}$ in the "traditional method" will be assumed by θ in the "copula method". To estimate θ we use the *maximum likelihood method* [8]:

Step 6 ("copula method"): From the cumulated probability (4.3) we can define a probability density by calculating the partial derivative of C_θ with respect to both parameters,

$$f_\theta(r_1, r_2) \equiv \frac{\partial^2}{\partial r_1 \, \partial r_2} C_\theta(\Phi_1(r_1), \Phi_2(r_2)) \ . \tag{4.4}$$

With this result at hand the *likelihood function* can be computed:

$$L(\theta) \equiv \prod_{j=1}^{N} f_\theta(r_{1,j}, r_{2,j}) \ . \tag{4.5}$$

The N pairs $(r_{1,j}, r_{2,j})$ are again the relative changes of the historical data, see equation (2.1) in chapter 2.

The maximum likelihood method states that the particular $\theta = \theta_{\max}$, at which L reaches its maximum, is a good estimator of the "real" θ. In other words, one

has to find $\theta_{\max}$ so that $L(\theta) < L(\theta_{\max})$ for all $\theta \in [1, \infty[$ with $\theta \neq \theta_{\max}$. Here we assume that the likelihood function contains one unique maximum only. As this is, of course, not true in general, we will come back to this point later in this section.

In practice, one does not maximize L directly but the logarithm of L,

$$l(\theta) \equiv \ln L(\theta) = \sum_{j=1}^{N} \ln f_\theta(r_{1,j}, r_{2,j}) \ . \tag{4.6}$$

From (3.7) and (4.4) one has

$$l(\theta) = \sum_{j=1}^{N} \ln \left(\frac{\partial^2}{\partial u \, \partial v} C_\theta(u, v) \Big|_{u=\Phi_1(r_{1,j}), v=\Phi_2(r_{2,j})} f_1(r_{1,j}) \, f_2(r_{2,j}) \right) \ . \tag{4.7}$$

Further simplification gives

$$\begin{aligned} l(\theta) \ &= \ \sum_{j=1}^{N} \ln \left(\frac{\partial^2}{\partial u \, \partial v} C_\theta(u, v) \Big|_{u=\Phi_1(r_{1,j}), v=\Phi_2(r_{2,j})} \right) \ + \\ &\quad \sum_{j=1}^{N} \left(\ln f_1(r_{1,j}) + \ln f_2(r_{2,j}) \right) \ . \end{aligned} \tag{4.8}$$

The second sum in (4.8) does not depend on theta. Therefore it is sufficient to maximize the function

$$\hat{l}(\theta) = \sum_{j=1}^{N} \ln \left(\frac{\partial^2}{\partial u \, \partial v} C_\theta(u, v) \Big|_{u=\Phi_1(r_{1,j}), v=\Phi_2(r_{2,j})} \right) \tag{4.9}$$

with respect to θ. While the partial derivative in (4.9) can be calculated analytically from (3.12), there is no exact analytic expression to compute the values of the cumulated normal density functions $\Phi_i(r_{i,j})$ for $i = 1, 2$ and $j = 1, \ldots, N$. For this task we used the approximation method presented in [9]. As stated above, the maximization of $\hat{l}(\theta)$ finally yields the desired value of θ.

In general, the maximization of $\hat{l}(\theta)$ has to be done numerically. For an arbitrary historical date[2] we present in figure 4.1 a typical curve of $\hat{l}(\theta)$. As one can see, only one global maximum occurs in that figure. In particular, no local maxima can be seen. This behaviour was observed by us at all historical dates that we have tested. Therefore we conclude that the maximization algorithm that we have used to find θ is numerically stable and reliable.

[2] We have chosen 01/06/1995 as the relevant date plus the preceding 249 business days.

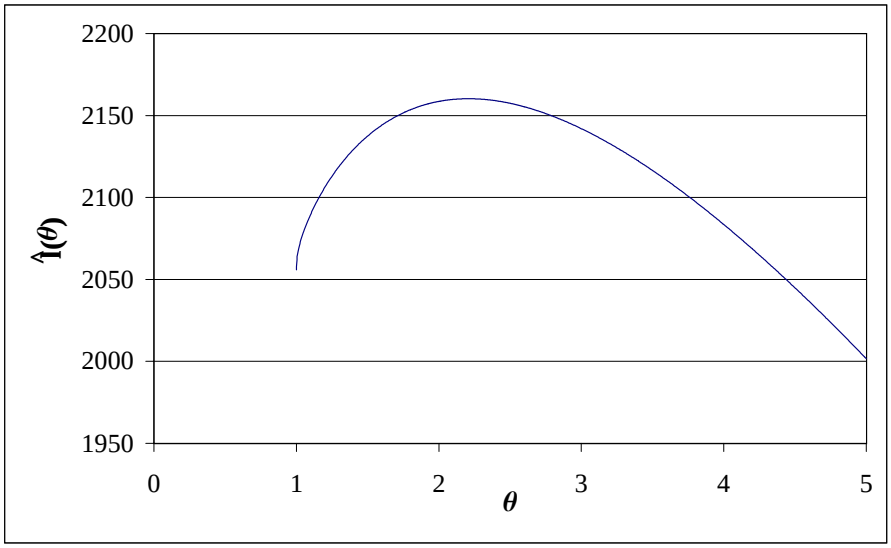

Figure 4.1: The modified likelihood function $\hat{l}(\theta)$ vs. θ.

Let us now come to step 7 of the original algorithm. Instead of creating pairs of correlated PRNs that belong to binormal distributed random variables we now have to generate PRNs that obey the copula dependence structure that was chosen in (4.3) with the parameter $\theta_{\max}$. Using the results from section 4.1, step 7 can now be reformulated:

Step 7 ("copula method"): First, generate two independent, in $[0,1]$ uniformly distributed PRNs u, w. Compute $v = C^{-1}_{\theta_{\max},u}(w)$, where $C_{\theta_{\max},u}$ is given by equation (3.12). Finally determine $r_1 = \Phi_1^{-1}(u)$ and $r_2 = \Phi_2^{-1}(v)$ to obtain one pair (r_1, r_2) of PRN with the desired copula dependence structure.

Let us end this chapter with two remarks on the last step. As already discussed in section 3.3, the determination of $C^{-1}_{\theta_{\max},u}(w)$ is no problem from a numerical point of view. Secondly, we also realized the computation of $\Phi_1^{-1}(u)$ and $\Phi_2^{-1}(v)$ numerically to obtain the results which are presented in following chapter.

Chapter 5

Numerical Results

As already mentioned in the previous chapters, we considered just two risk factors in our numerical simulations. These factors are the FX rates USD/DEM and GBP/DEM, respectively. To see how the "copula method" works and to demonstrates its benefits it is not necessary to use a complicated portfolio. The VaR results that we will discuss below are therefore based on a simple linear portfolio. Its value is given by

$$V = 1 \times \text{USD} - 1 \times \text{GBP} \ . \tag{5.1}$$

As we are interested in the value of V given in DEM, we can say that the portfolio contains the risk factors directly.

The data on which our investigation is based are the historical FX rates ranging from 2 Jan. 1991 until 9 Mar. 2000[1]. Due to the introduction of the euro in the EU, we used the fixed rate EUR/DEM = 1.95583 to convert the EUR based FX rates from 4 Jan. 1999 until 9 Mar. 2000 into DEM-based FX rates.

The time interval on which our VaR computations are based are 250 business days. Using the notation given in step 1 on page 4, we used $N + 1 = 250$ to obtain our results. To get an impression of the analysed data we show in figure 5.1 the means of the relative FX rates (see expression (2.1)) over these 250 business days.

Let us now start our comparison of the "traditional method" vs. the "copula method" of calculating the Value-at-Risk. First, we want to take a look at the two different parameters of the underlying joint distributions, i.e. the correlation $\rho = \rho_{1,2}$ of the binormal distribution and the parameter $\theta_{\max}$ of the Gumbel-Hougaard copula. To simplify the notation, we will skip the suffix "max" in the remainder of this thesis. The behaviour of these parameters is presented in figure 5.2. As already mentioned

[1]The latter date simply is the date when we started our investigation.

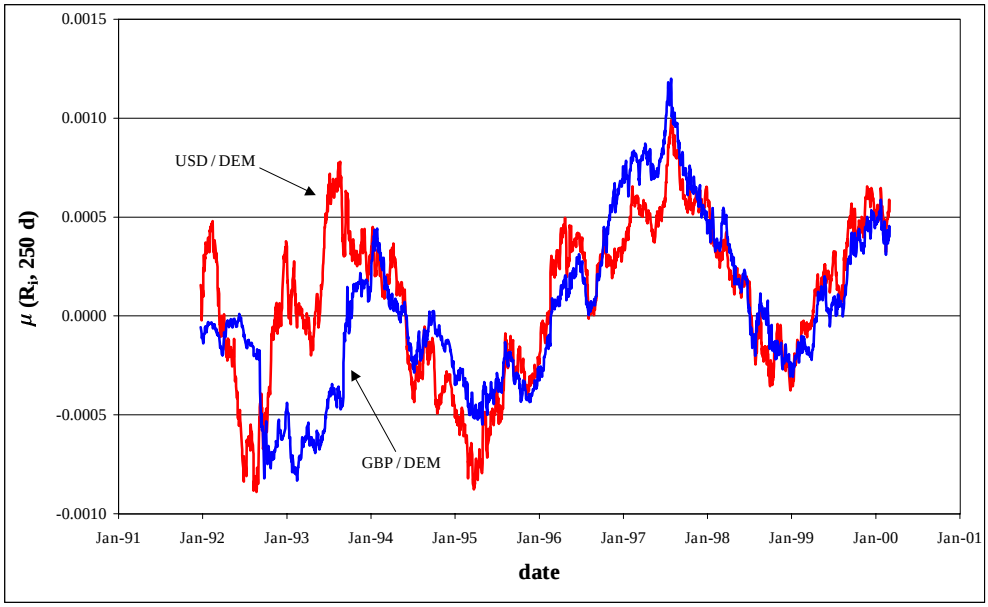

Figure 5.1: Means over 250 business days of the relative USD/DEM and GBP/DEM rates.

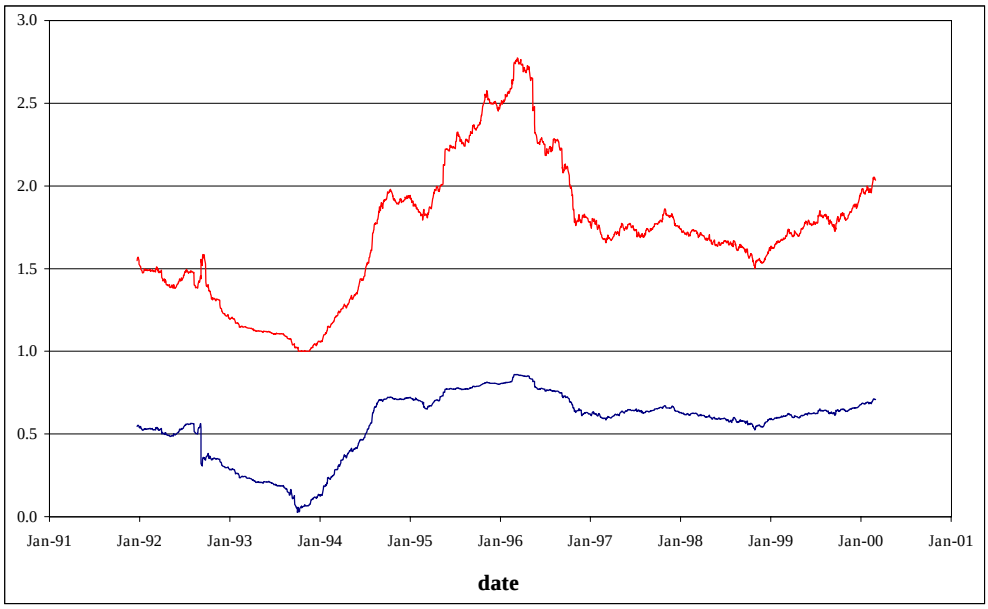

Figure 5.2: Linear correlation ρ (lower line) and Gumbel-Hougaard parameter θ (upper line) of USD/DEM and GBP/DEM rates.

in the previous chapters, the theoretical values that these parameters may take are $\rho \in [-1, +1]$ and $\theta \in [1, \infty[$, respectively.

First, a comparison of the two parameters shows that both curves behave in a very similar manner on a qualitative level. However, even if it is most often the case, an increase or decrease in one of the parameters does *not* necessarily result in an increase or decrease in the other.

Our second point is that both curves tend in the same region to significant minima, $\rho \to 0$ and $\theta \to 1$. Significant in this context means that, as stated on page 9 in section 3.2.1 and in theorem 4, for the Gumbel-Hougaard copula, $\theta = 1$ is equivalent to independence of the two risk factors. On the other hand, for the "traditional method", $\rho = 0$ is equivalent to independence[2]. This shows that both methods interpret the data correctly in qualitative terms.

However, in finance we are interested in quantitatively correct statements. Especially in the case of risk management, we want to compute a Value-at-Risk that best reflects reality. The message of this thesis is that for the example under consideration, which is defined by portfolio (5.1), the "copula method" provides more reliable results than the results obtained from the "traditional method". To confirm this statement we calculated the VaR in the time interval from 2 Jan. 1992[3] to 8 Mar. 2000[4]. For each business day in this interval we calculated three different VaRs, corresponding to confidence levels (or quantiles) of 90% (or $\alpha_1 = 10\%$), 95% (or $\alpha_2 = 5\%$), and 99% (or $\alpha_3 = 1\%$), respectively. Each of these $\mathrm{VaR}(\alpha_i, t, t+1)$ was compared with the change in the value of the portfolio from time t to time $t+1$, i.e. with $\Delta V = V_{t+1} - V_t$. The numbers of underestimated losses in relation to the total number of backtested values gave three values that could be compared with the three quantiles α_1, α_2, and α_3. To test the dependence of these numerical results on the number of Monte Carlo steps, we did this investigation for various numbers of MC steps ranging from 100 to 1500. In table 5.1 and figure 5.3 we present our results of the underestimated losses.

First, we wish to emphasize that, once a number of MC steps of about 900 is reached, the data do not change significantly under any additional increase of the number of MC steps. This behaviour can be observed for both algorithms and each confidence level, as one can see from table 5.1. Therefore the numerical results for

[2] In general, a correlation of zero does not imply independence. However, for the multinormal distribution $c_{i,j} = 0$ for all i, j implies independence of the variables as the distribution function (2.7) factorizes.

[3] On page 20 we give 2 Jan. 1991 as the start date of our investigations. However, to calculate a VaR one has to collect historical data (see step 1 on page 4) and our calculations are based on 250 business days.

[4] We use the final day (9 Mar. 2000) of our investigation for backtesting only.

	Percentage of underestimated losses					
	Traditional Monte Carlo			Copula Monte Carlo		
# MC-steps	$\alpha_1 = 10\%$	$\alpha_2 = 5\%$	$\alpha_3 = 1\%$	$\alpha_1 = 10\%$	$\alpha_2 = 5\%$	$\alpha_3 = 1\%$
100	8.44%	4.61%	1.60%	8.69%	5.24%	1.94%
200	8.05%	4.42%	1.50%	8.54%	5.00%	1.50%
300	8.25%	4.42%	1.41%	8.35%	4.80%	1.50%
400	8.05%	4.37%	1.31%	8.10%	4.46%	1.26%
500	7.81%	4.32%	1.36%	8.05%	4.61%	1.26%
600	7.76%	4.37%	1.26%	8.15%	4.66%	1.26%
700	7.71%	4.32%	1.36%	8.20%	4.66%	1.31%
800	7.76%	4.37%	1.31%	8.25%	4.71%	1.26%
900	7.76%	4.37%	1.41%	8.25%	4.90%	1.21%
1000	7.86%	4.46%	1.41%	8.25%	4.85%	1.31%
1100	7.81%	4.42%	1.31%	8.30%	4.85%	1.21%
1200	7.81%	4.42%	1.31%	8.25%	4.95%	1.26%
1300	7.81%	4.32%	1.26%	8.10%	4.95%	1.26%
1400	7.81%	4.37%	1.31%	8.15%	4.85%	1.21%
1500	7.76%	4.42%	1.31%	8.15%	4.80%	1.26%

Table 5.1: Backtesting results of the "traditional method" and the "copula method" for quantiles $\alpha_1 = 10\%$, $\alpha_2 = 5\%$, and $\alpha_3 = 1\%$, and various numbers of MC steps.

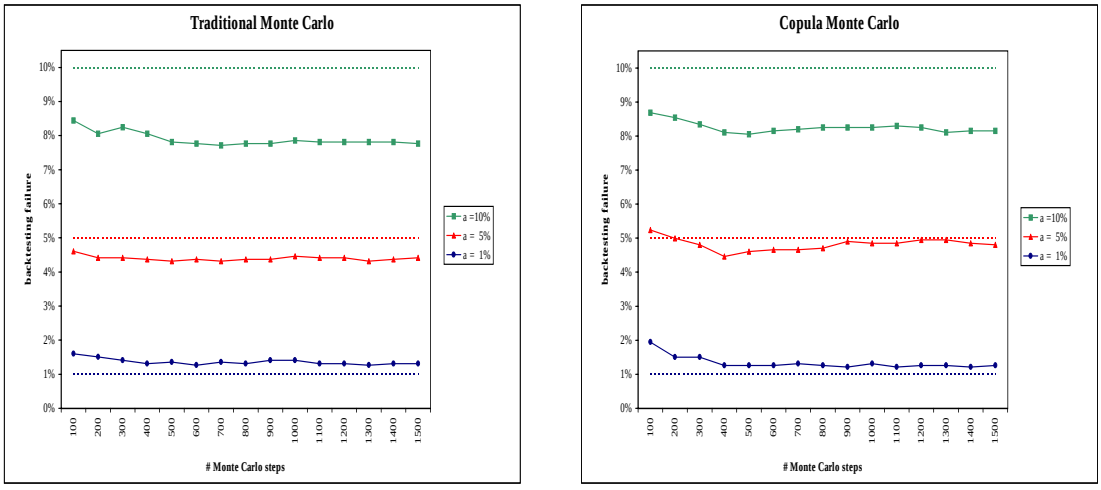

Figure 5.3: Backtesting results of the "traditional method" (left) and the "copula method" (right) for quantiles $\alpha_1 = 10\%$ (green squares), $\alpha_2 = 5\%$ (red triangles), and $\alpha_3 = 1\%$ (blue circles), and various numbers of MC steps.

numbers of MC steps of $\mathcal{O}(1000)$ seem to be reliable and are therefore suited as a basis for a discussion and comparison of the two methods.

Secondly, and this is the most important point that we wish to emphasize, the backtesting results show that the "copula method" works better then the "traditional method" for each confidence level. As mentioned previously, "better" in this context means that the absolute deviation of the percentage of the backtesting failure from the quantiles α_i is smaller in the case of the "copula method" than in the case of the "traditional method" for each $i = 1, 2, 3$.

To quantify the last statement, we consider below the results that we obtained from the simulations using 1500 Monte Carlo steps. For the 10% quantile, we obtained percentages of backtesting failures of 7.76% in the case of the "traditional method" and of 8.15% for the "copula method". In other words, the deviation from the theoretical result is about 21% higher in the "traditional" case. For the 5% and 1% quantiles we also found that the results obtained from the "copula method" are more reliable than those obtained from the "traditional method". The numerical values that we obtained from our simulations show deviations of 0.58% (traditional) and 0.20% (copula) from the 5% quantile, and 0.31% (traditional) and 0.26% (copula) from the 1% quantile. A comparison of these results shows that the deviations from the theoretical results are about 3 times higher for the "traditional method" than for the "copula method" in the case of the 5% quantile and about 19% higher in the case of the 1% quantile.

Even though we do not explain the variance-covariance method for calculating the VaR in this thesis (see, for example, [5]), we still want to compare our numerical results with the analytical results obtained using this method. For the three quantiles, we measured percentages of backtesting failures of 7.76%, 4.37%, and 1.36%. These results are also less reliable than the results from the "copula method". In particular, the variance-covariance results are very close to the "traditional" results. This can be seen as a kind of consistence check of the algorithm of the "traditional method" as, in the case of a portfolio which is a linear combination of the risk factors (see equation (5.1)), the variance-covariance VaR is the limit of the "traditional" Monte Carlo VaR for an infinite number of MC steps.

We now focus on the interpretation of our results. As discussed in detail in the previous chapters, the only difference between both methods is due to the different dependence structures that underlie the methods. As a consequence, the pseudo random numbers that are generated in the Monte Carlo simulations (see step 7 of each algorithm, pages 6 and 19) show a different behaviour for each of the methods,

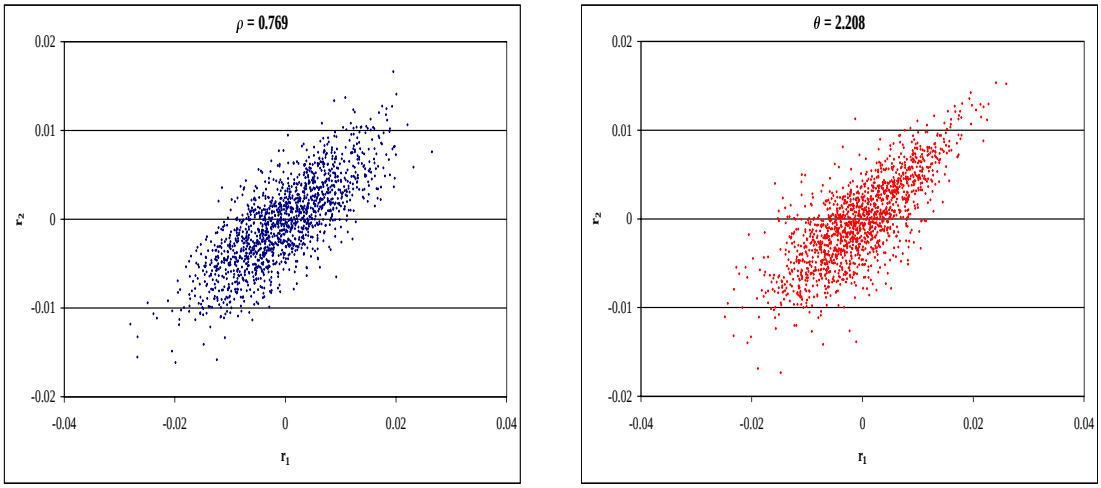

Figure 5.4: 1500 pairs of pseudo random numbers, generated by using a binormal distribution with $\rho = 0.769$ (left) and a Gumbel-Hougaard Copula dependence with $\theta = 2.208$ (right).

even if the historical data the simulations are based on are the same. Inspired by [1], we present in figure 5.4 1500 pairs of PRNs that have been generated at an arbitrary historical date (see footnote 2 on page 18) using to both methods. The PRNs that have been generated by using the binormal distribution yield a figure that resembles an ellipsoid. This behaviour is well known in statistics and can therefore be used as additional evidence that our numerical Monte Carlo algorithm works fine in the "traditional" case.

Let us now turn to the right part of figure 5.4. It shows that the use of the Gumbel-Hougaard copula results in more extreme values in the sense that, if one PRN is already in the tail of its marginal distribution, the probability that the corresponding second PRN will also be in its tail is higher than in the "traditional method". In other words, "there is less "diversification" of large risk in the Gumbel dependence model" [2]. On the other hand it is clear (see step 5 on page 2) that the accuracy of a Monte Carlo VaR is dominated by the α most extreme simulated losses of the portfolio under investigation. However, in our case of a linear portfolio, these losses are directly related to the values of the simulated pairs of PRNs. From this and from the fact that the backtesting results obtained from the "copula method" are more reliable than those obtained using the "traditional method" we conclude that the dependence structure given by the Gumbel-Hougaard copula better describes the dependence

between the USD/DEM and GBP/DEM rates than a binormal distribution based on the concept of correlation.

Chapter 6

Summary and Conclusions

In this thesis we have demonstrated an alternative method to compute the Value-at-Risk of a financial portfolio. For this purpose, we have introduced the concept of copulas and presented the most important properties of these functions. For the sake of simplicity, we have restricted ourselves to two risk factors and therefore to two-dimensional copulas only. We have discussed one representative of the copula family in detail, the Gumbel-Hougaard copula. Using this tool, we have modified one basic assumption underlying the common method of computing the VaR using Monte Carlo techniques: usually, the dependence structure between two risk factors is assumed to be described by the correlation between these factors. Especially, the joint distribution is assumed to be of the binormal kind. We have replaced this dependence structure by a dependence structure that is defined by the Gumbel-Hougaard copula. In doing so, we have replaced the role of the correlation by the parameter that determines the behaviour of a Gumbel-Hougaard copula.

To test whether the copula ansatz improves the computation of the VaR, we have considered a simple linear portfolio which consists of the two FX rates USD/DEM and GBP/DEM, respectively. Assuming confidence levels of 90%, 95%, and 99%, we have calculated the VaRs for this portfolio on the basis of 250 banking days, using three different methods:

- "traditional" Monte Carlo simulation
- the variance-covariance method
- "copula" Monte Carlo simulation

The computations and the corresponding backtesting of the results have been performed on the basis of historical FX rates ranging over more than nine years. This means that our backtesting statistics are based on approximately 2000 measurements.

The results that we have obtained in this thesis show that, at least for the portfolio under consideration, the "copula method" is more reliable than both the "traditional" Monte Carlo method and the variance-covariance method. In other words, for each confidence level the absolute deviations of the percentages of the backtesting failures from the quantiles are smaller in the case of the "copula method" than both in the case of the "traditional method" and in the case of the variance-covariance method.

With these results at hand we now want to formulate the answers to the two questions that were asked in the introduction. Question 1 asked for the "appropriate" joint distribution of the risk factors. Based on our numerical data the answer has to be that a binormal distribution describes the data well, but that a dependence structure defined by the Gumbel-Hougaard copula describes the data better. The content of question 2 was of a more technical nature. It concerned a possibility of simulating pseudo random numbers using this joint distribution. We have given an answer to this question in chapter 4. There we have discussed in general how to generate PRNs using a copula dependence structure. Furthermore, we have given detailed information on what an algorithm looks like in the case of the Gumbel-Hougaard copula.

We finally want to consider the application of copulas in financial risk management. In this thesis it has been shown that the use of copulas may improve the computation of the VaR of a financial portfolio. As we considered two risk factors only, a natural extension of studies in this field would have to include three or more risk factors. Furthermore, there is a large number of other copulas [3] some of which also may be suited for use in a VaR computation. Finally one could make use of the fact that the dependence structure between the risk factors can be described by a copula. As a consequence, one is completely free in choosing the marginal distributions of the risk factors. Especially, different marginal distributions may be chosen for each risk factor. The replacement of the normal marginal distributions that also have been used in this thesis by a suitable selection of alternative marginal distributions may provide an additional improvement of the VaR computation. By this replacement, the typical problem of the heavy tails [5] of the historical data, which are not described in a satisfactory way by a normal distribution, may be weakened.

Appendix A

Generation of Pseudo Random Numbers

In this appendix we deal with the generation of pseudo random numbers that are distributed according to normal and multivariate normal distributions.

As almost every computer language provides independent PRNs that are uniformly distributed in the $[0,1]$ interval we do not describe how to generate these numbers. The interested reader may find several algorithms describing this topic in any good text book on numerical statistics (see, for example, [8]).

A.1 Transformation of Variables

A general method to generate PRNs according to a given distribution uses a transformation of variables to obtain these numbers from uniformly distributed PRNs.

Starting with a uniformly distributed PRN $u \in [0,1]$ we want to generate a PRN $r = r(u)$ with a given distribution $f(r)$,

$$\int_a^b \mathrm{d}r\, f(r) = 1\,. \tag{A.1}$$

First one equalizes the integrated probability density to u,

$$\mathbb{P}(r) = \int_a^r \mathrm{d}r'\, f(r') = u\,. \tag{A.2}$$

If possible, one next has to invert the function $\mathbb{P}$. Then

$$r(u) = \mathbb{P}^{-1}(u) \tag{A.3}$$

is the PRN with the desired distribution. However, it is not always possible to compute (A.3) analytically. Especially normally distributed PRNs cannot be computed directly by the method described above.

A.2 The Box-Muller Method

To generate normally distributed PRNs one cannot directly apply the method described in the previous section. Below, we shortly summarize what is known in literature as the *Box-Muller method*. This refers to an algorithm which is used to generate simultaneously two independent normally distributed PRNs from two independent uniformly distributed PRNs.

Assume that we want to generate r_1 from a $N(\mu_1, \sigma_1^2)$ distribution and r_2 from a $N(\mu_2, \sigma_2^2)$ distribution[1]. If we denote the two uniformly distributed PRNs by u_1 and u_2, the introduction of polar coordinates, followed by a variable transformation as in A.1, shows (see [8] for details) that r_1 and r_2 are given by

$$r_1 = \sqrt{-2\sigma_1^2 \ln u_1}\, \sin(2\pi\, u_2) + \mu_1 \quad \text{and} \tag{A.4}$$

$$r_2 = \sqrt{-2\sigma_2^2 \ln u_1}\, \cos(2\pi\, u_2) + \mu_2 \ . \tag{A.5}$$

A.3 Generation of Multivariate Normally Distributed Pseudo Random Numbers

Below we will discuss how to generate a multivariate normal distribution of n correlated PRNs represented by a vector $\vec{r}$. The density function we have to consider is presented in equation (2.7).

The basic idea behind generating the desired PRNs is to apply what is called *Cholesky decomposition* [10] to the inverse covariance matrix. The ansatz of this decomposition is

$$\mathcal{C}^{-1} = \mathcal{A}^t \mathcal{A} \ , \tag{A.6}$$

where $\mathcal{C}$ is given by (2.10). The matrix $\mathcal{A}$ in (A.6) is a lower triangular matrix,

$$\mathcal{A} = \begin{pmatrix} \alpha_{1,1} & 0 & \cdots & 0 \\ \alpha_{2,1} & \alpha_{2,2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{pmatrix} \ . \tag{A.7}$$

Using (A.6) one now performs a linear variable transformation in (2.7),

$$\vec{r} \to \vec{s} = \mathcal{A}\,(\,\vec{r} - \vec{\mu}) \ . \tag{A.8}$$

[1] See equation (2.2) on page 5 for the density function $N(\mu_i, \sigma_i^2)$ of a normal distribution.

One can show that the corresponding determinant of the Jacobian matrix is given by

$$\frac{\partial(r_1,\ldots,r_n)}{\partial(s_1,\ldots,s_n)} = \sqrt{\det \mathcal{C}}\ . \tag{A.9}$$

Therefore, the transformed density function takes the form

$$f(\vec{s}) = \frac{1}{\sqrt{(2\pi)^n}}\ \exp\left\{-\frac{1}{2}\vec{s}^{\,t}\ \vec{s}\right\} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\ \exp\left\{-\frac{1}{2}s_i^2\right\}\ . \tag{A.10}$$

Equation (A.10) is the density function of n uncorrelated, standard normally distributed variables. Therefore the simultaneous generation of n correlated, normally distributed PRNs is achieved in two steps:

(a) Generate n uncorrelated, standard normally distributed PRNs $s_1,\ldots,s_n$ using the Box-Muller method presented in section A.2.

(b) Compute the desired PRNs $r_1,\ldots,r_n$ by the reverse transformation of (A.8),

$$\vec{s} \rightarrow\ \vec{r} = \mathcal{A}^{-1}\vec{s} + \vec{\mu}\ . \tag{A.11}$$

As we only need the two-dimensional case in this thesis, we will now focus on the generation of bivariate normally distributed PRNs. The n-dimensional case is a simple generalisation and can be found, for example, in [5].

The first step is (see equations (A.4) and (A.5)):

(ã) Generate two uncorrelated, standard normally distributed PRNs

$$s_1 = \sqrt{-2\,\ln u_1}\,\sin(2\pi\,u_2) \quad \text{and} \tag{A.12}$$

$$s_2 = \sqrt{-2\,\ln u_1}\,\cos(2\pi\,u_2) \tag{A.13}$$

from two uncorrelated PRNs u_1 and u_2 that are uniformly distributed in $[0,1]$.

From (2.10), (A.7) and condition (A.6) one can show that $\mathcal{A}$ can be written as

$$\mathcal{A} = \frac{1}{\sqrt{\det \mathcal{C}}} \begin{pmatrix} \sigma_2\sqrt{1-\rho^2} & 0 \\ -\rho\,\sigma_2 & \sigma_1 \end{pmatrix} \tag{A.14}$$

where $\det\mathcal{C} = \sigma_1^2\sigma_2^2\,(1-\rho^2)$ and ρ denotes the correlation coefficient (see (2.12)), $\rho = \rho_{1,2} = c_{1,2}/(\sigma_1\sigma_2)$. The correlation is normalized so it can take all values in the interval $[-1,1]$. For $\rho = \pm 1$ expression (A.14) is not defined. However, in this section we can neglect these cases as $\rho = \pm 1$ means that the variables r_1 and r_2 are completely positively or negatively correlated. Therefore one only has to simulate one PRN using to the method described in section A.2. The second PRN will be deterministic.

From (A.11) and (A.14) one finds that the second step mentioned above becomes:

($\tilde{\mathbf{b}}$) Perform the variable transformation $s_1, s_2 \rightarrow r_1, r_2$ as follows:

$$r_1 = \sigma_1 \, s_1 + \mu_1 \quad \text{and} \tag{A.15}$$

$$r_2 = \sigma_2 \, (\rho \, s_1 + \sqrt{1 - \rho^2} \, s_2) + \mu_2 \; . \tag{A.16}$$

r_1 and r_2 are the two desired, correlated PRNs that obey a bivariate normal distribution.

Bibliography

[1] P. Embrechts, A. McNeil, D. Straumann, May 1999 *Correlation: Pitfalls and Alternatives*. RISK, pages 69-71

[2] P. Embrechts, A. McNeil, D. Straumann, July 1999 *Correlation and Dependence in Risk Management: Properties and Pitfalls*. Preprint ETH Zürich

[3] R. B. Nelsen, 1999 *An Introduction to Copulas*. Springer, New York

[4] T. P. Hutchinson, C. D. Lai, 1990 *Continuous Bivariate Distributions, Emphasising Applications*. Rumsby Scientific Publishing, Adelaide.

[5] H.-P. Deutsch, R. Eller, 1999 *Derivatives and Internal Models*. Macmillan Press.

[6] A. Sklar, 1959 *Fonctions de répartition à n dimensions et leurs marges*. Publ. Inst. Statist. Univ. Paris 8, pages 229-231.

[7] A. Sklar, 1996 *Random Variables, Distribution Functions, and Copulas – a Personal Look Backward and Forward* published in *Distributions with Fixed Marginals and Related Topics*, edited by L. Rüschendorf, B. Schweizer, and M.D. Taylor, Institute of Mathematical Statistics, Hayward, CA, pages 1-14.

[8] S. Brandt, 1970 *Statistical and Computational Methods in Data Analysis*. North-Holland.

[9] J.C. Hull, 2000 *Options, Futures, and Other Derivates*. 4^{th} edition. Prentice Hall.

[10] J. Stoer, 1989 *Numerische Mathematik 1*. 5^{th} edition. Springer.