Copulas and Stochastic Processes

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines Doktors
der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Mathematiker

Volker Schmitz

aus Düren

Berichter: Universitätsprofessor Dr. H.-H. Bock
Universitätsprofessor Dr. B. Rauhut

Tag der mündlichen Prüfung: 21. Januar 2003

Diese Dissertation ist auch als Verlagsveröffentlichung
(Shaker Verlag, Aachen, ISBN 3-8322-1278-7) erhältlich.

Preface

This thesis was written during my time at the Institute of Statistics of Aachen University.

I would like to thank Professor Hans-Hermann Bock for giving me the opportunity and freedom to choose this topic, his valuable comments, continuous interest and support for this work. I am very grateful for having been given the opportunity to work in this highly interesting and motivating field.

Many thanks to Professor B. Rauhut for also taking the effort to referee this thesis.

A special thankyou goes to Dr. E. Pier-Ribbert who drew my attention to the interesting field of copulas during my stay at Dresdner Bank, Frankfurt.

I would also like to thank my colleagues at the Institute of Statistics of whom many contributed to this thesis in one way or another. In particular, I have to mention Stefan Merx who provided me with many useful hints and critical remarks and Eric Beutner for his proof-reading of some of the chapters.

Last, I want to thank my wife Fiona for her patience and perseverance in supporting me during the past years.

Aachen, February 2003 Volker Schmitz

Contents

1 Introduction

The modelling of dependence relations between random variables is one of the most widely studied subjects in probability theory and statistics. Thus, a great variety of concepts for dependence structures has emerged (see Jogdeo, 1982). However, without specific assumptions about the dependence, no meaningful stochastic model can be developed.

The dependence of random variables can basically be classified into *spatial* and *temporal* dependencies where the former means the dependence between a number of variables *at the same time* whereas the latter means the intertemporal dependence structure of a process.

The classic approach uses second order moments of the underlying random variables, known as the *covariance*. It is well known that only linear dependence relationships can be captured this way and that it is characterizing only for special classes of distributions, for example the normal distribution. The question arises if there is a possibility to capture the whole dependence structure without any disturbing effects coming from the marginal distributions. This is exactly where copulas provide a beneficial approach.

As the details will be given in the next chapter we will only say that copulas allow to separate the effect of dependence from the effects of the marginal distributions, i. e., if $X_1, \dots, X_n$ are real-valued random variables on some probability space $(\Omega, \mathcal{A}, \mathsf{P})$ with joint distribution function H, i. e., $H(x_1, \dots, x_n) := \mathsf{P}(X_1 \le x_1, \dots, X_n \le x_n)$, $x_i \in \mathbb{R}$, $1 \le i \le n$, a copula C of $(X_1, \dots, X_n)$ is given by the relation $H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$ for all $x_i \in \mathbb{R}$ where the F_i are the univariate marginal distributions of the X_i, i. e., $F_i(x) := \mathsf{P}(X_i \le x)$, $x \in \mathbb{R}$, $1 \le i \le n$.

Simply put, a copula is a multivariate distribution function on the n-dimensional unit cube with uniform marginal distributions.

The question of existence and uniqueness of such a copula was answered by Sklar (1959) who also introduced the term "copula" which is now the most widely used term. ("Copula" is derived from the Latin verb "copulare", meaning "to join together".) Also used are the terms "dependence function" (e. g., Galambos, 1978, Definition 5.2.1) in the context of multivariate extreme value theory, "uniform representation" (e. g., Kimeldorf and Sampson, 1975) and "standard form" (e. g., Cook and Johnson, 1981).

Although the concept is rather old, the first volume of the *Encyclopedia of Statistical Sciences* to contain the entry "copulas" was the *Update Volume* from 1997 (Fisher, 1997). This may be due to the fact that from 1959 to 1974 results concerning copulas were obtained in the course of the development of the theory of probabilistic metric spaces and distributions with given marginals (for a historic account see Schweizer, 1991). Nowadays, however, interest in copulas has grown especially in mathematical finance; spatial dependence between assets is tried to be captured by copulas instead of the classic models such as correlation structures. For an example of this approach see Embrechts et al. (1999) or Schmitz and Pier-Ribbert (2001).

Most of the current research in copulas is done for spatial dependence due to great interest in practice for new spatial dependence models, but the analysis of *temporal* dependence is also possible by the copula approach. The first paper dealing exclusively with copulas and stochastic processes was presented by Darsow et al. (1992), who established the connection between copulas and Markov processes. Only few papers have appeared since then, and most are working papers not yet published in a journal.

This thesis extends the relationship between copulas and stochastic processes into new directions and tries to show that the copula approach can provide new structural insights into the temporal dependence underlying the processes. It will become clear that many properties of stochastic processes can be analyzed by splitting conditions up into a "marginal part" and a "dependence part", thus providing a better insight into the laws governing the temporal behaviour of the process.

Many known theorems can be given alternative and in most cases simpler proofs via copulas.

Although the focus is laid upon univariate stochastic processes, we also will—in a small excursion—deal with two special and interesting aspects of copulas for spatial dependence (Chapter 4).

The structure of this thesis is as follows:

After this general **Introduction**, **Chapter 2** (Copulas and Their Properties) will give a basic introduction to the definition and theory of copulas. It will be a basic reference for all other chapters to follow. A special focus should be laid upon Section 2.2 (The Role of Partial Derivatives) in which the connection between conditional distribution functions and partial derivatives of copulas will be developed in a mathematical rigorous way. Although many of the results are well known, it is difficult to find mathematically satisfying proofs for them.

As the family of Archimedean copulas is the most prominent of all currently used models, we give a short introduction to them in Section 2.3.

As a preparation for the analysis of symmetric stochastic processes we need to introduce the survival copula and some symmetry concepts for multivariate random variables in

Section 2.4. However, this section is of its own interest. The most important result is that radial symmetry of a n-variate distribution function can be equivalently established by radial symmetry of the underlying copula and (univariate) symmetry of the marginals (Corollary 2.41).

As we learned above, copulas provide a means of dividing the multivariate distribution function into separate models for the marginals and the dependence between the variables. It is thus necessary to introduce some of the most widely used dependence concepts and their relation to copulas in **Chapter 3**. The results (most of which are known) will be used in the subsequent chapters. Especially the concepts of tail dependence and positive likelihood ratio dependence will provide us with an elegant tool to compare Kendall's τ and Spearman's ρ between the minimum and maximum of n independent and identically distributed (iid) random variables in Section 4.1.

Chapter 4 (Two Examples: Analysis of Spatial Dependence With Copulas) consists of two independent sections. In the first one we will take a closer look at the dependence structure between the minimum and maximum of n iid random variables. By deriving their copula we are able to give an alternative proof for the asymptotic independence which is essentially shorter than the known ones and uses only elementary calculus.

We will further determine Kendall's τ and Spearman's ρ and show the relation $3\tau \geq \rho \geq \tau > 0$ for them with the help of tail dependence and positive likelihood ratio dependence.

In the second section we will derive the copula of Brownian motion and its supremum process. The rather astonishing result is that this copula is independent of time.

Chapter 5 deals with the main topic of this thesis: the connection between univariate stochastic processes and copulas. As we will look at a rather wide field of different concepts, an introduction to definitions used in the sequel is given in Section 5.1.

Section 5.2 introduces basic relations between copulas and stochastic processes. Concepts such as strict stationarity, equivalence of stochastic processes and symmetry of processes are characterized by copulas.

Section 5.3 will specialize on processes in continuous time, e. g., Brownian motion. We develop sufficient conditions to the univariate marginals and the copula structure of a process to ensure stochastic continuity (Section 5.3.1). These conditions are then used to give an elegant proof of the findings of Cambanis (1991) for EFGM-processes. He showed that a stochastic process based on multivariate EFGM-distributions cannot be stochastically continuous. However, he established this result by a number of inequalities whereas we are able to give a direct proof with the help of Theorem 5.30 in Theorem 5.35.

Section 5.4 (Markov Processes) introduces the results of Darsow, Nguyen, and Olsen (1992) and extends the findings to a characterization of symmetric Markov processes by radial symmetry of the bivariate copulas (Section 5.4.2). Interestingly, within the class of Archimedean copulas, symmetric Markov processes can be characterized by the so-called Frank copula.

Finally, the copula structure of continuous local martingales is derived in Section 5.5. As continuous local martingales are not defined by their finite-dimensional distributions, a different technique has to be used to determine the underlying copulas. We show that the copula of Brownian motion plays a central role as each continuous local martingale is equivalent to a time change of Brownian motion, a result due to Dambis (1965) and Dubins and Schwarz (1965). It is then possible, for example, to derive the copula of the Ornstein-Uhlenbeck process and the Brownian bridge. These can be used to construct and/or simulate processes which have the same intertemporal dependence structure but with other marginal distributions.

In the **Appendix** we will provide some additional information about basic tools used within this thesis which are not worth being dealt with in the main part. A special focus is laid on the simulation of random variables and processes from copulas, since in practice simulation from copulas is one of the main tools to analyze the underlying models.

The **List of Symbols and Abbreviations** tries to give the reader a central point of access for all symbols used within the text. Where appropriate, the page number is indicated where the corresponding symbol is defined or introduced.
I tried to used notations and symbols which seem to be standard in the field concerned, the only speciality is the symbol "■" which marks the end of an example or remark.

Concerning the **Bibliography**, two things should be mentioned. Firstly, the numbers at the end of each entry refer to the page on which the work is cited in this thesis. Secondly, it is clear that the average reader does not necessarily have access to every source in the references. However, I decided to refer to the literature wherein the *exact* formulation of the needed theorems can be found, but it is possible to read the text with a single reference, e. g., Rogers and Williams (2000a,b).

Finally, the **Index** should certainly help the reader find what he or she is looking for. If page numbers behind entries are emphasized (such as *4711* in contrast to 4711), they refer to the most important pages concerning these entries. It usually means that the definition can be found there.

2

Copulas and Their Properties

In this chapter we give a basic introduction to copulas and their properties. As there seems to be a lack of literature connecting copulas and conditional probabilities in a mathematical rigorous way, we also shed some light on this aspect of copulas which turns out to be a very powerful tool when dealing with the modelling of stochastic processes with copulas.

2.1 Basic Properties

The basic idea underlying the concept of copulas can be described as follows: Assume a pair of real-valued[1] random variables X and Y on a common probability space $(\Omega, \mathcal{F}, \mathsf{P})$ with distribution functions $F(x) := \mathsf{P}(X \leq x)$, $x \in \mathbb{R}$, and $G(y) := \mathsf{P}(Y \leq y)$, $y \in \mathbb{R}$, respectively, and a joint distribution function $H(x, y) := \mathsf{P}(X \leq x, Y \leq y)$, $x, y \in \mathbb{R}$. If we associate the three numbers $F(x)$, $G(y)$ and $z := H(x, y)$ to every pair of real numbers (x, y), we observe that each of them lies in the unit interval $I := [0, 1]$, i. e., each pair (x, y) leads to a point $\big(F(x), G(y)\big)$ in the unit plane I^2 and this ordered pair in turn corresponds to the number $z \in I$ by the relation $z = H(x, y) = C\big(F(x), G(y)\big)$ for some function $C : I^2 \to I$. The last correspondence turns out to be characterizing for copulas.

Let us define some notions first. We denote the extended real line by $\overline{\mathbb{R}}$, i. e., $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$, analogously, $\overline{\mathbb{R}}^n$ denotes the extended n-dimensional real space.

For two vectors $\boldsymbol{x} = (x_1, \ldots, x_n), \boldsymbol{y} = (y_1, \ldots, y_n) \in \overline{\mathbb{R}}^n$, we will write $\boldsymbol{x} \leq \boldsymbol{y}$ if $x_i \leq y_i$, $1 \leq i \leq n$, and $\boldsymbol{x} < \boldsymbol{y}$ if $x_i < y_i$ for all $1 \leq i \leq n$.

[1] All random variables in this thesis will be assumed to be real-valued if not stated otherwise.

2.1 Definition
A (halfopen) rectangle *or* interval *in* $\overline{\mathbb{R}}^n$ *is the Cartesian product of* n *one-dimensional intervals of the form*

$$R_{\boldsymbol{x}}^{\boldsymbol{y}} := \bigtimes_{i=1}^{n} (x_i, y_i] = (x_1, y_1] \times \cdots \times (x_n, y_n], \quad \boldsymbol{x} \le \boldsymbol{y},$$

where $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n) \in \overline{\mathbb{R}}^n$.
The set of all rectangles in $\overline{\mathbb{R}}^n$ *will be denoted by* $\mathcal{R}^n$.

The vertices *of a rectangle* $R_{\boldsymbol{x}}^{\boldsymbol{y}} \in \mathcal{R}^n$ *(of which we clearly have* 2^n*) are the points* $\operatorname{vert}(R_{\boldsymbol{x}}^{\boldsymbol{y}}) := \{(u_1, \ldots, u_n) : u_i \in \{x_i, y_i\},\ 1 \le i \le n\}$.

The unit cube I^n *is the product* $I \times \cdots \times I$ *(n times) where* $I = [0, 1]$ *is the unit interval.*

The domain *and* range *of a function* H *will be denoted by* $\operatorname{dom}(H)$ *and* $\operatorname{ran}(H)$*, respectively.*

As we will use n-dimensional Lebesgue-(Stieltjes-)integration, we need a n-dimensional analogue of an increasing[2] function (i. e., some prototype of a measure generating function). This can be established via the n-dimensional volume of a rectangle as we will see later.

2.2 Definition (*H*-volume)
Let $S_1, \ldots, S_n \subset \overline{\mathbb{R}}$ *be nonempty sets and* $H : \overline{\mathbb{R}}^n \to \mathbb{R}$ *a function such that* $\operatorname{dom}(H) \subset S_1 \times \cdots \times S_n$. *Let* $R_{\boldsymbol{a}}^{\boldsymbol{b}} = \bigtimes_{i=1}^{n} (a_i, b_i] \in \mathcal{R}^n$, $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$, *be a rectangle with* $R_{\boldsymbol{a}}^{\boldsymbol{b}} \subset \operatorname{dom}(H)$. *Then the* H-volume V_H *of* $R_{\boldsymbol{a}}^{\boldsymbol{b}}$ *is given by*

$$V_H(R_{\boldsymbol{a}}^{\boldsymbol{b}}) := \Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H(x_1, \ldots, x_n) \tag{2.1}$$

with

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H(x_1, \ldots, x_n) := \Delta_{(a_n, b_n)}^{(n)} \Delta_{(a_{n-1}, b_{n-1})}^{(n-1)} \cdots \Delta_{(a_1, b_1)}^{(1)} H(x_1, \ldots, x_n) \tag{2.2}$$

where

$$\Delta_{(a_i, b_i)}^{(i)} H(x_1, \ldots, x_n) := H(x_1, \ldots, x_{i-1}, b_i, x_{i+1}, \ldots, x_n) - H(x_1, \ldots, x_{i-1}, a_i, x_{i+1}, \ldots, x_n) \tag{2.3}$$

is the first order difference operator. $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H(x_1, \ldots, x_n)$ *is therefore the n-th order difference of* H *with regard to* $\boldsymbol{a}$ *and* $\boldsymbol{b}$.

2.3 Remark
The above notation is used to clarify on which coordinate the operator Δ operates. As this is notationally inconvenient, we will use the notation

$$\Delta_{a_i}^{b_i} \equiv \Delta_{(a_i, b_i)}^{(i)}$$

where the coordinate is clear from the context. Where confusion could arise, we will use the full notation. ■

[2] Note that the terms “increasing” and “decreasing” mean $f(x) \le f(y)$, $x \le y$, for a function f and $f(x) \ge f(y)$, $x \le y$, respectively. If a strict inequality holds, we will use the term “strictly in-/decreasing”.

Note that $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H$ can also be written as

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H(\boldsymbol{x}) = \sum_{(\varepsilon_1,\dots,\varepsilon_n)\in\{0,1\}^n} (-1)^{\sum_{i=1}^n \varepsilon_i} \cdot H\bigl(\varepsilon_1 a_1 + (1-\varepsilon_1)b_1, \dots, \varepsilon_n a_n + (1-\varepsilon_n)b_n\bigr). \quad (2.4)$$

The proof of this is a straightforward induction and omitted.

The following definition is the n-dimensional analogue[3] of the one-dimensional monotonicity:

2.4 Definition (n-increasing)
Let H be a real-valued function with $\mathrm{dom}(H) \subset \overline{\mathbb{R}}^n$. *$H$ is said to be* n-increasing *if* $V_H(R) \geq 0$ *for all rectangles $R \in \mathcal{R}^n$ with* $\mathrm{vert}(R) \subset \mathrm{dom}(H)$.

Some authors refer to n-increasing functions as *quasi-monotone* or Δ-monotone (e. g., Behnen and Neuhaus, 1995).

The next definition is intuitive in the sense that the notion of *groundedness* is a natural requirement for a multivariate distribution function on $\mathbb{R}^n$.

2.5 Definition (grounded)
Let H be a real-valued function with $\mathrm{dom}\, H = S_1 \times \cdots \times S_n$, *$S_i \subset \overline{\mathbb{R}}$, and let each S_i have a smallest element a_i. H is called* grounded *if for all $i = 1, \dots, n$,*

$$H(x_1, \dots, x_{i-1}, a_i, x_{i+1}, \dots, x_n) = 0$$

for all $x_j \in S_j$, $1 \leq j \leq n$, $j \neq i$, $a_i \in S_i$.

We can now introduce the definition of a copula:

2.6 Definition (copula)
A copula *(or n-*copula*) is a function $C : I^n \to I$ with the following properties:*

a) C is grounded, i. e., for every $\boldsymbol{u} = (u_1, \dots, u_n) \in I^n$,

$$C(\boldsymbol{u}) = 0 \text{ if there exists an } i \in \{1, \dots, n\} \text{ with } u_i = 0, \quad (2.5)$$

and:

$$\text{If all coordinates of } \boldsymbol{u} \in I^n \text{ are 1 except } u_k, \text{ then } C(\boldsymbol{u}) = u_k. \quad (2.6)$$

b) C is n-increasing, i. e., for all $\boldsymbol{a}, \boldsymbol{b} \in I^n$ such that $\boldsymbol{a} \leq \boldsymbol{b}$,

$$V_C(R_{\boldsymbol{a}}^{\boldsymbol{b}}) \geq 0. \quad (2.7)$$

The set of all n-copulas will be denoted by $\mathcal{C}_n$.

[3] In view of the property of being a measure inducing function

We easily see that $C(u_1,\dots,u_n) = \Delta_{\mathbf{0}}^{\mathbf{u}} C(x_1,\dots,x_n) = V_C((0,u_1]\times\cdots\times(0,u_n])$ due to (2.1)–(2.3) and (2.5), so one can think of $C(u_1,\dots,u_n)$ as an assignment of a number in I to the rectangle $(0,u_1]\times\cdots\times(0,u_n]$.

We will now present some simple but basic properties of copulas. **All these properties hold for the general case $\boldsymbol{n \geq 2}$, although formulated and proved for the case $\boldsymbol{n = 2}$.** The general case can be proved analogously, but a rigorous treatment would be notationally complex.

2.7 Lemma
Let $C \in \mathcal{C}_2$ and $0 \leq x_1 \leq x_2 \leq 1$, then

$$y \mapsto C(x_2, y) - C(x_1, y) \tag{2.8}$$

is increasing on I.
Similarly, for $0 \leq y_1 \leq y_2 \leq 1$, the function

$$x \mapsto C(x, y_2) - C(x, y_1) \tag{2.9}$$

is increasing on I.

Proof. We only consider the former case as the latter one follows analogously. Let $0 \leq y_1 \leq y_2 \leq 1$ and $0 \leq x_1 \leq x_2 \leq 1$. As C is 2-increasing (see Def. 2.4), we have (with $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{y} = (y_1, y_2)$)

$$\begin{aligned} V_C(R_{\boldsymbol{x}}^{\boldsymbol{y}}) &= C(x_2,y_2) - C(x_2,y_1) - C(x_1,y_2) + C(x_1,y_1) \geq 0 \\ \iff\ & C(x_2,y_2) - C(x_1,y_2) \geq C(x_2,y_1) - C(x_1,y_1), \end{aligned} \tag{2.8*}$$

which is just the claimed monotonicity property. □

2.8 Corollary
Copulas are Lipschitz-continuous (hence continuous), i. e., for any $C \in \mathcal{C}_2$ and $0 \leq x_1 \leq x_2 \leq 1$, $0 \leq y_1 \leq y_2 \leq 1$, and $x, y \in I$, it holds that

$$0 \leq C(x_2,y) - C(x_1,y) \leq x_2 - x_1 \text{ and} \tag{2.10 a}$$
$$0 \leq C(x,y_2) - C(x,y_1) \leq y_2 - y_1 \text{ so that} \tag{2.10 b}$$
$$|C(x_2,y_2) - C(x_1,y_1)| \leq |x_2 - x_1| + |y_2 - y_1|. \tag{2.10}$$

Proof. Set $y = 0$, $y = 1$ and $x = 0$, $x = 1$, respectively, in Lemma 2.7, (2.8) and (2.9). □

2.9 Corollary
Let $C \in \mathcal{C}_2$, then $y \mapsto C(x,y)$ is increasing for all $x \in I$ and analogously, $x \mapsto C(x,y)$ is increasing for all $y \in I$.

Proof. Take $x_1 = 0$ and $y_1 = 0$, respectively, in Lemma 2.7. □

2.10 Remark
The set $\mathcal{C}_n$ of all n-copulas is a compact and convex subset of the space of all continuous real valued functions defined on I^n under the topology of uniform convergence. It follows that—in $\mathcal{C}_n$—pointwise convergence implies uniform convergence.

The convexity follows directly by the definiton of a copula. The compactness follows by noticing that the set of all continuous real-valued functions on I^n is a compact metric space and the subset of all n-copulas is closed. ■

Two functions, of which only one is a copula for all dimensions (see Remark 2.12), are special in that they provide bounds for all copulas:

2.11 Theorem (Fréchet bounds)
Let C be a 2-copula. Then for $u, v \in I$,

$$W(u,v) := \max(u+v-1,0) \le C(u,v) \le \min(u,v) =: M(u,v). \tag{2.11}$$

The functions W and M are called lower *and* upper Fréchet bounds.

Proof. Due to monotonicity we have $C(u,v) \le C(u,1) = u$ and $C(u,v) \le C(1,v) = v$, so $C(u,v) \le M(u,v)$. $C(u,v) \ge 0$ is trivially true and $V_C\big((u,1] \times (v,1]\big) \ge 0$ yields $C(1,1) + C(u,v) - C(u,1) - C(1,v) \ge 0$, so $C(u,v) \ge u+v-1$ from which the theorem follows. □

2.12 Remark
The multivariate version of Theorem 2.11 is given by

$$W(\boldsymbol{u}) := \max\left(\sum_{i=1}^n u_i - n + 1, 0\right) \le C(\boldsymbol{u}) \le \min_{1\le i\le n}(u_i) =: M(\boldsymbol{u}) \tag{2.12}$$

for $\boldsymbol{u} = (u_1, \dots, u_n) \in I^n$, $n \ge 2$, $C \in \mathcal{C}_n$.

Although the lower bound W is never a copula[4] for $n \ge 3$, the left-hand inequality in (2.12) is "best-possible" in the sense that for any $n \ge 3$ and any $\boldsymbol{u} \in I^n$, there is a n-copula C such that $C(\boldsymbol{u}) = W(\boldsymbol{u})$ (see Nelsen, 1999, Theorem 2.10.12). ■

By $\Pi(\boldsymbol{u}) := \prod_{i=1}^n u_i$, $\boldsymbol{u} \in I^n$, we denote the so-called *independence copula.*

For a stochastic interpretation of the functions M, W and Π we need the following classic result of Sklar (1959, 1973) which shows the relationship between copulas and random variables and their distribution functions, respectively. This will explain the interest in the study of copulas which also underlies this thesis.

[4]This can be seen from the fact that $V_W(\boldsymbol{(1/2,1]}) = 1 - n/2$ with $\boldsymbol{(1/2,1]} := (1/2,1] \times \cdots \times (1/2,1]$ so that W is not grounded.

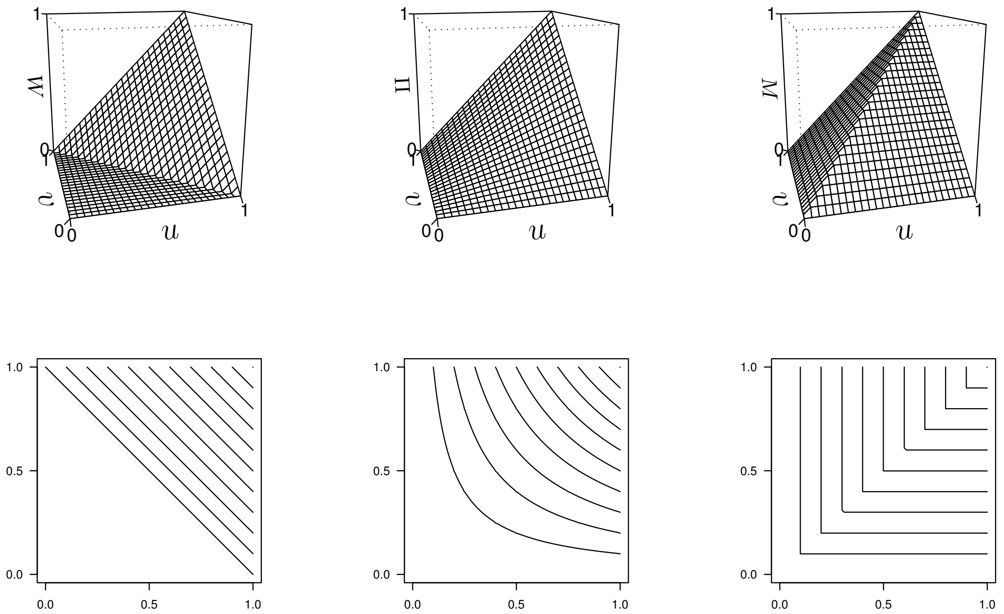

FIGURE 2.1. Lower Fréchet bound, independence copula and upper Fréchet bound; perspective and related contour plots for the case $n = 2$.

2.13 Theorem (Sklar, 1959)
Let H be a n-dimensional distribution function with one-dimensional marginals $F_1, \ldots, F_n$. Then there exists a n-copula C such that for $(x_1, \ldots, x_n) \in \overline{\mathbb{R}}^n$,

$$H(x_1, \ldots, x_n) = C\big(F_1(x_1), \ldots, F_n(x_n)\big). \tag{2.13}$$

If $F_1, \ldots, F_n$ are all continuous, then C is unique; otherwise, C is uniquely determined on $\operatorname{ran} F_1 \times \cdots \times \operatorname{ran} F_n$.

Conversely, if C is a n-copula and $F_1, \ldots, F_n$ are distribution functions on $\mathbb{R}$, then the function H defined by (2.13) is a n-dimensional distribution function with marginals $F_1, \ldots, F_n$.

For a proof, see Nelsen (1999, Theorem 2.10.9) and the references therein.

In Figure 2.1, plots of the Fréchet bounds and the independence copula are given. The geometric interpretation of Theorem 2.11 is that the surface of any copula of two random variables must lie between the surfaces of the lower and the upper Fréchet bound. Every kind of dependence between two random variables has a corresponding copula with a surface which can be identified in the unit cube.

2.14 Remark (cf. Nelsen, 1999, Section 2.5)
With the theorem of Sklar, the copulas W, Π and M have the following stochastic interpretations:

Two random variables U, V distributed uniformly on $(0, 1)$ have the joint distribution function

- M iff (if and only if) U is almost surely (a. s.) an increasing function of V:

$$U = f(V) \text{ for some monotone, increasing function } f : \mathbb{R} \to \mathbb{R},$$

- W iff U is a. s. a decreasing function of V:

$$U = g(V) \text{ for some monotone, decreasing function } g : \mathbb{R} \to \mathbb{R},$$

- Π iff U and V are independent.

These properties were firstly observed by Hoeffding (1940) and Fréchet (1957). ■

We may now ask: Which are the copulas of *monotone transformations* of the underlying random variables? The answer is given in the following theorem (see Nelsen, 1999, Theorems 2.4.2 and 2.4.3):

2.15 Theorem
Let X and Y be continuous random variables with copula $C_{X,Y}$, and let f and g be real-valued functions on $\operatorname{ran}(X)$ *and* $\operatorname{ran}(Y)$*, respectively.*

a) *If f and g are strictly increasing, then $C_{f(X),g(Y)}(u,v) = C_{X,Y}(u,v)$ for all $u,v \in I$.*

b) *If f is strictly increasing and g is strictly decreasing, then $C_{f(X),g(Y)}(u,v) = u - C_{X,Y}(u,1-v)$ for all $u,v \in I$.*

c) *If f is strictly decreasing and g is strictly increasing, then $C_{f(X),g(Y)}(u,v) = v - C_{X,Y}(1-u,v)$ for all $u,v \in I$.*

d) *If f and g are both strictly decreasing, then $C_{f(X),g(Y)}(u,v) = u + v - 1 + C_{X,Y}(1-u,1-v)$ for all $u,v \in I$.*

This theorem will be needed later to determine the copula of continuous local martingales.

We will now see that property (2.7) guarantees that C induces a Lebesgue-Stieltjes-measure on $(I^n, \mathcal{B}^n \cap I^n)$, where $\mathcal{B}^n$ denotes the Borel-σ-algebra.

2.16 Theorem
Let C be a n-copula. Then C induces a unique probability measure P_C on $(I^n, \mathcal{B}^n \cap I^n)$.

Proof. As C is a distribution function on I^n, the fact follows easily from Shiryaev (1996, Theorem II.3.2, p. 160) or Billingsley (1995, Example 1.1, p. 9). □

However, it is *not* true that every probability measure P on I^n is induced by a copula C.[5] For a measure to be induced by a copula, it must spread mass in a manner consistent with the boundary conditions of a copula. Therefore, any probability measure on I^n where at least one one-dimensional margin is not uniformly distributed cannot be induced by a copula.

In detail, we have:

[5] This means that for a rectangle $R_{\mathbf{0}}^{\boldsymbol{u}}$ being spanned by $\mathbf{0}$ and $\boldsymbol{u}$ (both in I^n), there is always a n-copula C such that $\mathsf{P}(R_{\mathbf{0}}^{\boldsymbol{u}}) = C(\boldsymbol{u})$ holds for all $\boldsymbol{u} \in I^n$.

2.17 Corollary
A probability measure P *on* $(I^n, \mathcal{B}^n \cap I^n)$ *is induced by a copula* $C \in \mathcal{C}_n$ *if and only if*

$$\mathsf{P}\big((\mathbf{0}, \boldsymbol{x_i}]\big) = x_i, \quad \textit{for all } x_i \in I,\ 1 \le i \le n,$$

with $\mathbf{0} := (0, \ldots, 0)$, $\boldsymbol{x_i} := (1, \ldots, 1, x_i, 1, \ldots, 1)$.

Proof. If P is induced by a copula C we must have $\mathsf{P}\big((\mathbf{0}, \boldsymbol{x_i}]\big) = C(1, \ldots, 1, x_i, 1, \ldots, 1) = x_i$ due to (2.6) for all $x_i \in I$. On the other hand, if we have $\mathsf{P}\big((\mathbf{0}, \boldsymbol{x_i}]\big) = x_i$, we can define the copula C by $C(\boldsymbol{u}) := \mathsf{P}((\mathbf{0}, \boldsymbol{u}])$, $\boldsymbol{u} \in I^n$, and easily check the definition. □

The following theorem concerning the denseness of the subset of absolutely continuous copulas in the space of all copulas was firstly stated for the case $n = 2$ in Darsow et al. (1992). For a proof, the authors refer to an oral presentation given at their institute. To make things more rigorous, we state the theorem for the general case and give a detailed proof.

As copulas are uniformly continuous on I^n, the topology of uniform convergence and its induced supremum metric $d_{\sup}$ provide a natural framework for measuring distances between copulas.

For notational convenience, λ^n will be used for the usual Lebesgue measure on $(\mathbb{R}^n, \mathcal{B}^n)$ and the Lebesgue measure restricted on I^n, i. e., $(I^n, \mathcal{B}^n \cap I^n)$. The meaning depends on the context; where confusion could arise, additional remarks will be made.

2.18 Theorem
Let $\mathcal{C}_n$ *denote the set of all* n*-copulas and* $\mathcal{C}_{\lambda^n} \subset \mathcal{C}_n$ *the subset of all copulas which are absolutely continuous with respect to (w. r. t.) the Lebesgue measure* λ^n *restricted on* I^n*. Then* $\mathcal{C}_{\lambda^n}$ *is dense in* $\mathcal{C}_n$ *w. r. t. the supremum metric* $d_{\sup}$*, i. e., the metric generating the uniform topology on* I^n*.*

Proof. Let $d_{\sup}$ denote the supremum metric on $\mathcal{C}_n$, i. e., for $C_1, C_2 \in \mathcal{C}_n$,

$$d_{\sup}(C_1, C_2) := \|C_1 - C_2\| := \sup_{\boldsymbol{x} \in I^n} |C_1(\boldsymbol{x}) - C_2(\boldsymbol{x})|,$$

where $\|\cdot\|$ denotes the corresponding supremum norm on I^n.

As we know that copulas are continuous on the compact set I^n, we could replace “sup” by “max”.

Now, let $C^* \in \mathcal{C}_n$ be a copula and $\varepsilon > 0$ arbitrary. We need to construct a copula $C \in \mathcal{C}_n$ with $\|C - C^*\| < \varepsilon$ whose induced measure P_C (see Theorem 2.16) is absolutely continuous with respect to the Lebesgue measure $\lambda^n\,|_{I^n}$. Let P_{C^*} denote the unique probability measure on $(I^n, \mathcal{B}^n \cap I^n)$ induced by C^*. For some $N \in \mathbb{N}$ which we will choose later, let

$$S_{i_1,\ldots,i_n} := \left(\frac{i_1 - 1}{N}, \frac{i_1}{N}\right] \times \cdots \times \left(\frac{i_n - 1}{N}, \frac{i_n}{N}\right], \quad 1 \le i_j \le N,\ i_j \in \mathbb{N},\ 1 \le j \le n, \quad (2.14)$$

be n-dimensional cubes (of which there are N^n) so that

$$(0,1]^n = \sum_{1 \le i_1,\dots,i_n \le N} S_{i_1,\dots,i_n}, \tag{2.15}$$

and let

$$R^N_{i_1,\dots,i_n} := \sum_{\substack{1 \le j_k \le i_k \\ 1 \le k \le n}} S_{j_1,\dots,j_n} \tag{2.16}$$

be the n-dimensional rectangle defined by the points $(0,\dots,0)$ and $(i_1/N,\dots,i_n/N)$.

N is chosen such that $N > n/\varepsilon$.

Now, for $1 \le i_1,\dots,i_n \le N$, let C in a first step be defined on the points $(i_1/N,\dots,i_n/N)$ by

$$C\left(\frac{i_1}{N},\dots,\frac{i_n}{N}\right) := \mathsf{P}_{C^*}(R^N_{i_1,\dots,i_n}) = C^*\left(\frac{i_1}{N},\dots,\frac{i_n}{N}\right). \tag{2.17}$$

Note that the C-volume of any rectangle of the form $R = (0, i_1/N] \times \dots \times (0, i_n/N]$ can be determined by (2.16) from these values, i.e.,

$$V_C(R^N_{i_1,\dots,i_n}) = \sum_{\substack{1 \le j_k \le i_k \\ 1 \le k \le n}} V_C(S_{j_1,\dots,j_n}).$$

Recall that using the notation for the n-th order difference operator $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}}$ for $\boldsymbol{a}, \boldsymbol{b} \in \overline{\mathbb{R}}^n$ we have

$$V_C(S_{i_1,\dots,i_n}) = \Delta^{(i_1/N,\dots,i_n/N)}_{((i_1-1)/N,\dots,(i_n-1)/N)} C(u_1,\dots,u_n),$$

and this is well defined by (2.17).

In a second step, let us now define C by a λ^n-density in the following way which guarantees that the mass of C is spread uniformly on the cubes $S_{i_1,\dots,i_n}$:

$$\frac{dC}{d\lambda^n}(x_1,\dots,x_n) \equiv c(x_1,\dots,x_n) := N^n \sum_{1 \le i_1,\dots,i_n \le N} \mathbb{1}_{S_{i_1,\dots,i_n}}(\boldsymbol{x}) \cdot V_C(S_{i_1,\dots,i_n}), \tag{2.18}$$

where $\boldsymbol{x} := (x_1,\dots,x_n) \in I^n$. Obviously, $c(\boldsymbol{x}) \ge 0$ for all $\boldsymbol{x} \in I^n$ and

$$\begin{aligned}\int_{I^n} \mathrm{d}C = \int_{I^n} \frac{\mathrm{d}C}{\mathrm{d}\lambda^n}\,\mathrm{d}\lambda^n &= \sum_{1 \le i_1,\dots,i_n \le N} N^n V_C(S_{i_1,\dots,i_n}) \underbrace{\int_{I^n} \mathbb{1}_{S_{i_1,\dots,i_n}}(\boldsymbol{x})\,\mathrm{d}\lambda^n(\boldsymbol{x})}_{=N^{-n}} \\ &= \sum_{1 \le i_1,\dots,i_n \le N} V_C(S_{i_1,\dots,i_n}) = V_C(I^n) \overset{(2.17)}{=} V_{C^*}(I^n) = 1.\end{aligned}$$

Then C is absolutely continuous w. r. t. λ^n by definition.

As a third step, we still need to show that C indeed is a copula and $\|C - C^*\| < \varepsilon$.

That C is a copula can be seen as follows:

- C is defined on I^n, if $x_i = 0$ for some $i \in \{1, \dots, n\}$, then, by (2.18) and x_i not belonging to any of the $S_{i_1,\dots,i_n}$, we get $C(x_1, \dots, x_n) = 0$. Therefore, C is grounded.
- Let $u_i = k/N$ for some $1 \leq k \leq N$, then

$$C(1, \dots, 1, u_i, 1, \dots, 1) \overset{(2.17)}{=} C^*(1, \dots, 1, u_i, 1, \dots, 1) = u_i$$

as C^* is a copula by assumption. Let $u_i \neq k/N$ for all $0 < k \leq N$. Then

$$\boldsymbol{u_i} := (1, \dots, 1, u_i, 1, \dots, 1) \in S_{N,\dots,N,j_i,N,\dots,N} \quad \text{for some } 1 \leq j_i \leq N,$$

and we have

$$\begin{aligned}
C(\boldsymbol{u_i}) &= C\left(\frac{N}{N}, \dots, \frac{N}{N}, u_i, \frac{N}{N}, \dots, \frac{N}{N}\right) \\
&\overset{(2.18)}{=} \int_0^1 \dots \int_0^{u_i} \dots \int_0^1 \sum_{1 \leq i_1,\dots,i_n \leq N} N^n \mathbb{1}_{S_{i_1,\dots,i_n}}(\boldsymbol{x}) \cdot V_C(S_{i_1,\dots,i_n}) \, \mathrm{d}\lambda^n(\boldsymbol{x}) \\
&= \sum_{1 \leq i_1,\dots,i_n \leq N} N^n V_C(S_{i_1,\dots,i_n}) \Bigg[\int_0^1 \dots \int_0^{\frac{j_i-1}{N}} \dots \int_0^1 \mathbb{1}_{S_{i_1,\dots,i_n}}(\boldsymbol{x}) \, \mathrm{d}\lambda^n(\boldsymbol{x}) \\
&\qquad + \int_0^1 \dots \int_{\frac{j_i-1}{N}}^{u_i} \dots \int_0^1 \mathbb{1}_{S_{i_1,\dots,i_n}}(\boldsymbol{x}) \, \mathrm{d}\lambda^n(\boldsymbol{x}) \Bigg] \\
&= \sum_{\substack{1 \leq k_1,\dots,k_n \leq N \\ k_i \leq j_i - 1}} N^n V_C(S_{k_1,\dots,k_n}) \cdot \left(\frac{1}{N}\right)^n \\
&\qquad + \sum_{\substack{1 \leq k_1,\dots,k_n \leq N \\ k_i = j_i}} N^n V_C(S_{k_1,\dots,k_n}) \left(u_i - \frac{j_i - 1}{N}\right) \left(\frac{1}{N}\right)^{n-1} \\
&= \sum_{\substack{1 \leq k_1,\dots,k_n \leq N \\ k_i \leq j_i - 1}} V_C(S_{k_1,\dots,k_n}) + N \left(u_i - \frac{j_i - 1}{N}\right) \sum_{\substack{1 \leq k_1,\dots,k_n \leq N \\ k_i = j_i}} V_C(S_{k_1,\dots,k_n}) \\
&\overset{(2.17)}{=} C^*(1, \dots, 1, \frac{j_i - 1}{N}, 1, \dots, 1) + N \left(u_i - \frac{j_i - 1}{N}\right) \\
&\qquad \cdot \left[C^*(1, \dots, 1, \frac{j_i}{N}, 1, \dots, 1) - C^*(1, \dots, 1, \frac{j_i - 1}{N}, 1, \dots, 1) \right] \\
&\overset{C^* \text{ is copula}}{=} \frac{j_i - 1}{N} + N \left(u_i - \frac{j_i - 1}{N}\right) \left(\frac{j_i}{N} - \frac{j_i - 1}{N}\right) = u_i.
\end{aligned}$$

- Trivially, the C-volume of any rectangle is non-negative by (2.18).

Therefore, C is a copula.

As a last step, let $\boldsymbol{x} = (x_1, \dots, x_n) \in I^n$ with $x_i \neq 0$ for all $1 \leq i \leq n$. (Remember that if $x_i = 0$ for some $i \in \{1, \dots, n\}$, $C^*(\boldsymbol{x}) = C(\boldsymbol{x}) = 0$.) Then $\boldsymbol{x} \in S_{i_1,\dots,i_n}$ for some $(i_1, \dots, i_n) \in \{1, \dots, N\}^n$. For $\boldsymbol{v}^* := \operatorname{argmin}\{\|\boldsymbol{x} - \boldsymbol{v}\|_1 : \boldsymbol{v} \in \operatorname{vert}(S_{i_1,\dots,i_n})\}$, we get

$$\|\boldsymbol{x} - \boldsymbol{v}^*\|_1 := \sum_{i=1}^{n} \underbrace{|x_i - v_i^*|}_{\leq (2N)^{-1}} \leq \frac{n}{2N} < \frac{\varepsilon}{2}. \tag{2.19}$$

Therefore,

$$\begin{aligned} |C(\boldsymbol{x}) - C^*(\boldsymbol{x})| &= |C(\boldsymbol{x}) - C(\boldsymbol{v}^*) + C(\boldsymbol{v}^*) - C^*(\boldsymbol{v}^*) + C^*(\boldsymbol{v}^*) - C^*(\boldsymbol{x})| \\ &\leq |C(\boldsymbol{x}) - C(\boldsymbol{v}^*)| + \underbrace{|C(\boldsymbol{v}^*) - C^*(\boldsymbol{v}^*)|}_{= 0 \text{ by construction}} + |C^*(\boldsymbol{v}^*) - C^*(\boldsymbol{x})| \\ &\overset{(2.10)}{\leq} 2\|\boldsymbol{x} - \boldsymbol{v}^*\|_1 \overset{(2.19)}{<} \varepsilon, \end{aligned}$$

thus

$$\|C - C^*\| < \varepsilon,$$

which completes the proof. □

2.19 Remark
The same construction idea is used for the bivariate case by Li et al. (1998) and Kulpa (1999), who call it a “checkerboard approximation”. ■

2.2 The Role of Partial Derivatives

In this section we will take a closer look at the role played by the partial derivatives of a copula C. We will see that there is a very strong connection between the partial derivatives and conditional distributions related to C. To give a heuristic motivation of what may be the interpretation of the partial derivatives, consider $\mathsf{P}(X \le x \mid Y = y)$ for two random variables X and Y with distribution functions F_X and F_Y where F_Y is continuous and strictly increasing and with copula C and joint cumulative distribution function F_{XY}.[6] If $\mathsf{P}(X \le x \mid Y = y)$ is assumed to be continuous from the right w. r. t. y, we have

$$\begin{aligned}
\mathsf{P}(X \le x \mid Y = y) &= \lim_{h \searrow 0} \mathsf{P}(X \le x \mid y \le Y \le y + h) \\
&= \lim_{h \searrow 0} \frac{F_{XY}(x, y+h) - F_{XY}(x,y)}{F_Y(y+h) - F_Y(y)} \\
&= \lim_{h \searrow 0} \frac{C\big(F_X(x), F_Y(y+h)\big) - C\big(F_X(x), F_Y(y)\big)}{F_Y(y+h) - F_Y(y)} \\
&= \lim_{h \searrow 0} \frac{C\big(F_X(x), F_Y(y) + \Delta(h)\big) - C\big(F_X(x), F_Y(y)\big)}{\Delta(h)} \\
&= \frac{\partial}{\partial v} C(u,v)\Big|_{\left(F_X(x), F_Y(y)\right)}
\end{aligned}$$

with $\Delta(h) := F_Y(y+h) - F_Y(y)$ wherever the derivative exists. This is the basic idea behind the interpretation of partial derivatives which will be put in a more rigorous framework hereafter.

For a n-copula $C \in \mathcal{C}_n$, denote the n first partial derivatives of C by

$$D_k C(u_1, \dots, u_n) := \frac{\partial}{\partial u_k} C(u_1, \dots, u_n), \quad 1 \le k \le n,\ u_i \in (0,1),\ 1 \le i \le n. \tag{2.20}$$

Since monotone functions are differentiable almost everywhere, it follows that for any $k \in \{1, \dots, n\}$, $u_i \in I$, $i \ne k$, $D_k C(u_1, \dots, u_n)$ exists for almost all $u_k \in (0,1)$, and Corollaries 2.8 and 2.9 (for n dimensions) yield

$$0 \le D_k C(u_1, \dots, u_n) \le 1 \quad \text{almost surely.} \tag{2.21}$$

Analogously, by Lemma 2.7 and Corollary 2.8, the functions in Lemma 2.7 have derivatives almost everywhere, and

$$D_k C(u_1, \dots, u_n)\big|_{u_k = y} - D_k C(u_1, \dots, u_n)\big|_{u_k = x} \ge 0 \tag{2.22}$$

if $0 \le x \le y \le 1$, $u_i \in (0,1)$.

[6] We recall the definition $\mathsf{P}(X \le x \mid Y) := \mathsf{E}(\mathbb{1}_{\{X \le x\}} \mid Y)$.

2.20 Notation
As we will need the operator D_k quite often, we will use the notation

$$D_k C(u_1, \ldots, u_{k-1}, x, u_{k+1}, \ldots, u_n) := \frac{\partial}{\partial u_k} C(u_1, \ldots, u_n)\Big|_{u_k = x}. \tag{2.23}$$

It follows that the functions $x \to D_k C(u_1, \ldots, u_{k-1}, x, u_{k+1}, \ldots, u_n)$ are defined and increasing almost everywhere.

Before formulating a lemma which relates the partial derivatives to conditional expectations, we need the following result from Elstrodt (1996) which—in connection with (2.21) and the above remarks—gives us the integrability of the partial derivatives.

2.21 Theorem (Elstrodt, 1996, Satz 6.1, p. 150)
Let $K \in \{\mathbb{R}, \mathbb{C}\}$ and $f : [\boldsymbol{a}, \boldsymbol{b}] \to K$ ($\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^p$, $\boldsymbol{a} < \boldsymbol{b}$, $p \in \mathbb{N}$) be a bounded function. Then f is Riemann-integrable if and only if the set of its points of discontinuity has λ^p-measure 0. Then, the Riemann-integral and the Lebesgue-integral are identical.

We can now formulate the basic result connecting partial derivatives of copulas with conditional distribution functions.

2.22 Lemma
Let X and Y be two real-valued random variables on the same probability space $(\Omega, \mathcal{A}, \mathsf{P})$ with corresponding copula C and continuous marginals F_X, F_Y. Define the conditional probabilities

$$\mathsf{P}(X \le x \mid Y) := \mathsf{E}(\mathbb{1}_{\{X \le x\}} \mid Y) \quad \textit{and} \quad \mathsf{P}(Y \le y \mid X) := \mathsf{E}(\mathbb{1}_{\{Y \le y\}} \mid X) \tag{2.24}$$

with $x, y \in \mathbb{R}$ as usual[7]. Then,

$$\begin{aligned} &D_2 C\big(F_X(x), F_Y(Y)\big) \quad \textit{is a version of} \quad \mathsf{P}(X \le x \mid Y) \quad \textit{and} \\ &D_1 C\big(F_X(X), F_Y(y)\big) \quad \textit{is a version of} \quad \mathsf{P}(Y \le y \mid X). \end{aligned} \tag{2.25}$$

Proof. (Cf. Darsow et al., 1992) We will only consider the first case, as the second one is completely analogous. Let $\sigma(Y)$ denote the σ-Algebra generated by Y and consider the setting $\big(\Omega, \sigma(Y)\big) \xrightarrow{Y} (\mathbb{R}^1, \mathcal{B}^1)$. We have to show

$$\int_A \mathsf{P}(X \le x \mid Y) \, \mathrm{d}\mathsf{P} = \int_A \mathbb{1}_{\{X \le x\}} \, \mathrm{d}\mathsf{P} \stackrel{!}{=} \int_A D_2 C\big(F_X(x), F_Y(Y)\big) \, \mathrm{d}\mathsf{P} \tag{2.26}$$

for all $A \in \sigma(Y)$, $x \in \mathbb{R}$. With the definitions $\mathsf{Q}_x(A) := \int_A \mathsf{P}(X \le x \mid Y) \, \mathrm{d}\mathsf{P}$ and $\tilde{\mathsf{Q}}_x(A) := \int_A D_2 C\big(F_X(x), F_Y(Y)\big) \, \mathrm{d}\mathsf{P}$ for $A \in \sigma(Y)$ we therefore must show

$$\mathsf{Q}_x(A) = \tilde{\mathsf{Q}}_x(A) \quad \text{for all } A \in \sigma(Y), x \in \mathbb{R}. \tag{2.27}$$

[7] See Billingsley, 1995, Sections 34 and 35, or Rogers and Williams, 2000a, II.40–42.

Note that Q_x is well defined due to the definition of conditional expectation and $\tilde{\mathsf{Q}}_x$ is well defined due to (2.21) and Theorem 2.21. It is elementary to show that Q_x and $\tilde{\mathsf{Q}}_x$ are σ-finite measures on $\sigma(Y)$.

It is known that $\mathcal{E} := \{(-\infty, a] : a \in \mathbb{R}\}$ is a σ-finite generator of $\mathcal{B}^1$ w.r.t. the Lebesgue-measure λ^1 which is closed under intersections. From Bauer (1992, Satz 5.4) or Rogers and Williams (2000a, Corollary II.4.7) it follows that we only need to show (2.27) for a σ-finite generator of $\sigma(Y)$ which is closed under intersections. We now have

$$\sigma(Y) \overset{(*)}{=} Y^{-1}(\mathcal{B}^1) = Y^{-1}\big(\sigma(\mathcal{E})\big) \overset{(**)}{=} \sigma\big(Y^{-1}(\mathcal{E})\big),$$

where $(*)$ is valid due to Billingsley (1995, Theorem 20.1, p. 255) and $(**)$ due to Elstrodt (1996, Satz 4.4, p. 19).

So we see that $\mathcal{E}' := \{Y^{-1}(E) : E \in \mathcal{E}\}$ is a generator of $\sigma(Y)$ which is closed under intersections (as $E_1' \cap E_2' = Y^{-1}(E_1) \cap Y^{-1}(E_2) = Y^{-1}(E_1 \cap E_2) \in \mathcal{E}'$ for some $E_1, E_2 \in \mathcal{E}$).

Therefore, we can assume the set A to be of the form $A = Y^{-1}\big((-\infty, a]\big)$ for some $a \in \mathbb{R}$. With Theorem 2.21, we have

$$\begin{aligned}
\int_A D_2C\big(F_X(x), F_Y(Y)\big)\,\mathrm{d}\mathsf{P} &= \int_{(-\infty,a]} D_2C\big(F_X(x), F_Y(\xi)\big) \underbrace{\mathrm{d}\mathsf{P}^Y(\xi)}_{\mathrm{d}F_Y(\xi) \text{ in Lebesgue-Stieltjes sense}} \\
&\overset{(*)}{=} \int_{(0,F_Y(a)]} D_2C\big(F_X(x), \eta\big)\,\mathrm{d}\eta \\
&= C\big(F_X(x), F_Y(a)\big) - \underbrace{C\big(F_X(x), 0\big)}_{=0} = \int_\Omega \mathbb{1}_{\{X \le x\}} \cdot \mathbb{1}_{\{Y \le a\}}\,\mathrm{d}\mathsf{P} = \int_A \mathbb{1}_{\{X \le x\}}\,\mathrm{d}\mathsf{P}.
\end{aligned}$$

$(*)$ is due to Kamke (1960, Satz 1, p. 164) with $\eta = F_Y(\xi)$ and the continuity of F_Y in combination with Lemma A.2.

This is exactly (2.27) from which the assertion follows. □

2.23 Remark

In view of Lemma A.3 we see that $F_X(X), F_Y(Y) \sim \mathsf{U}(0,1)$. Thus, (2.25) is equivalent to:

$$\begin{aligned}
&D_2C\big(F_X(x), V\big) \quad \text{is a version of} \quad \mathsf{P}(X \le x \mid Y) \quad \text{and} \\
&D_1C\big(U, F_Y(y)\big) \quad \text{is a version of} \quad \mathsf{P}(Y \le y \mid X),
\end{aligned} \tag{2.28}$$

where $U, V \sim \mathsf{U}(0,1)$ have copula C and are defined on the same probability space as X and Y. This means that for $\mathsf{P}(X \le x \mid Y)$, say, we only need the dependence information via C and the marginal F_X . ■

A statement analogous to Lemma 2.22 holds for the general case of n random variables and can be formulated as follows:

2.24 Corollary

Let $X_1, \ldots, X_n$ be real-valued random variables on the same probability space $(\Omega, \mathcal{A}, \mathsf{P})$ with corresponding copula C and continuous marginals $F_1, \ldots, F_n$.[8] *Then, for any $k \in \{1, \ldots, n\}$,*

$$D_k C\big(F_1(x_1), \ldots, F_{k-1}(x_{k-1}), F_k(X_k), F_{k+1}(x_{k+1}), \ldots, F_n(x_n)\big)$$

is a version of (2.29)

$$\mathsf{P}\big(X_i \le x_i, 1 \le i \le n, i \ne k \mid X_k\big)$$

for all $x_1, \ldots, x_n \in \mathbb{R}$.

We will omit the proof as it corresponds to the one in Lemma 2.22.

The next goal should be to connect partial derivatives of n-copulas to expressions such as $\mathsf{P}(X_n \le x_n \mid X_1, \ldots, X_{n-1})$. This would be particularly helpful when dealing with modelling conditionally specified time series by copulas. Indeed we can show more than Lemma 2.22: If the higher partial derivatives of a n-copula exist and are continuous, they are related to conditional probabilities in an analogous way as exemplified in Lemma 2.22. Before stating the corresponding lemma let us introduce a notation for higher-order partial derivatives of n-copulas as a generalization of (2.20):

2.25 Notation

Let C be a n-copula. For $k \in \{1, \ldots, n\}$, we define

$$D_{1,\ldots,k} C(x_1, \ldots, x_n) := \frac{\partial^k}{\partial u_1 \cdots \partial u_k} C(u_1, \ldots, u_n)\Big|_{(x_1,\ldots,x_n)} \tag{2.30}$$

for all $(x_1, \ldots, x_n), (u_1, \ldots, u_n) \in (0,1)^n$ where the derivative exists. For $k = n$, we write

$$c_{1,\ldots,n}(x_1, \ldots, x_n) := D_{1,\ldots,n} C(x_1, \ldots, x_n)$$

as this is the probability density function of $(U_1, \ldots, U_n)$ where the $U_i \sim \mathsf{U}(0,1)$ are connected by the copula C.

More generally, let $J \subset \{1, \ldots, n\}$ be such that $J = \{j_1, \ldots, j_k\}$, $j_i \ne j_l$ for $i \ne l$. Then,

$$D_J\, C(x_1, \ldots, x_n) := \frac{\partial^k}{\partial u_{j_1} \partial u_{j_2} \cdots \partial u_{j_k}} C(u_1, \ldots, u_n)\Big|_{(x_1,\ldots,x_n)}. \tag{2.31}$$

2.26 Remark

Due to our convention, expressions such as $D_{1,2} C(0.2, 0.2, 0.4)$ make sense and are well defined. ■

[8] Note that for a specific k we only need the continuity of the marginal F_k for the statement to hold.

We can formulate the following theorem for uniformly distributed random variables. The general expression will follow as a corollary.

2.27 Theorem
Let $U = (U_1, \dots, U_n)$ be n real-valued random variables with uniform marginal distributions $\mathsf{U}(0,1)$, $n \in \mathbb{N}$, and let $C_{1,\dots,n} = C\langle U\rangle$, $C_{1,\dots,n-1} = C\langle U_1, \dots, U_{n-1}\rangle$ where $C\langle\,\cdot\,\rangle$ denotes the corresponding copula. Assume that $C_{1,\dots,n}$ is $(n-1)$ times partially differentiable w. r. t. the first $(n-1)$ arguments and the partial derivatives are continuous in these arguments. Then,

$$\frac{D_{1,\dots,n-1}C_{1,\dots,n}(U_1, \dots, U_{n-1}, u_n)}{D_{1,\dots,n-1}C_{1,\dots,n-1}(U_1, \dots, U_{n-1})} \quad \text{is a version of} \quad \mathsf{P}(U_n \le u_n \mid U_1, \dots, U_{n-1}) \tag{2.32}$$

for all $u_n \in (0,1)$.

Proof. The proof is essentially analogous to the one of Lemma 2.22. However, care has to be taken of the fact that, e. g., $D_{1,\dots,n-2}C_{1,\dots,n}(0, u_2, \dots, u_n)$ is not necessarily equal to 0.[9] As we will see, we do not need this either. We will follow the outline of the proof of Lemma 2.22.

Let $\sigma(U_1, \dots, U_{n-1})$ denote the σ-algebra generated by $U_1, \dots, U_{n-1}$. We need to show that

$$\int_A \mathsf{P}(U_n \le u_n \mid U_1, \dots, U_{n-1})\,\mathrm{d}\mathsf{P} = \int_A \frac{D_{1,\dots,n-1}C_{1,\dots,n}(U_1, \dots, U_{n-1}, u_n)}{D_{1,\dots,n-1}C_{1,\dots,n-1}(U_1, \dots, U_{n-1})}\,\mathrm{d}\mathsf{P}$$

for all $u_n \in (0,1)$, $A \in \sigma(U_1, \dots, U_{n-1})$. As $\mathcal{E}_{n-1} := \{(\mathbf{0}, \boldsymbol{a}] : \boldsymbol{a} = (a_1, \dots, a_{n-1}) \in (0,1)^{n-1}\}$ is a σ-finite generator of $\mathcal{B}^{n-1} \cap (0,1)^{n-1}$ which is closed under intersections, we can assume A to be of the form

$$A = U_1^{-1}\big((0, a_1]\big) \cap \dots \cap U_{n-1}^{-1}\big((0, a_{n-1}]\big)$$

(the same theorems as in the previous proof apply). Then,

[9] This can be easily seen by $\Pi(u_1, \dots, u_n) = \prod_{i=1}^n u_i$: We have $D_{1,\dots,n-2}\Pi(0, u_2, \dots, u_n) = u_n u_{n-1}$ which is greater than 0 for $u_1, u_2 \in (0,1)$.

$$
\begin{aligned}
&\int_A \frac{D_{1,\dots,n-1}C_{1,\dots,n}(U_1,\dots,U_{n-1},u_n)}{D_{1,\dots,n-1}C_{1,\dots,n-1}(U_1,\dots,U_{n-1})}\,\mathrm{dP} \\
&\quad= \int_{(\mathbf{0},\boldsymbol{a}]} \frac{D_{1,\dots,n-1}C_{1,\dots,n}(u_1,\dots,u_{n-1},u_n)}{D_{1,\dots,n-1}C_{1,\dots,n-1}(u_1,\dots,u_{n-1})}\,\mathrm{d}C_{1,\dots,n-1}(u_1,\dots,u_{n-1}) \\
&\quad\overset{(*)}{=} \int_{(\mathbf{0},\boldsymbol{a}]} D_{1,\dots,n-1}C_{1,\dots,n}(u_1,\dots,u_{n-1},u_n)\,\mathrm{d}u_1\cdots\mathrm{d}u_{n-1} \\
&\quad\overset{\text{Fubini}}{=} \int_{(0,a_1]\times\cdots\times(0,a_{n-2}]} \big[D_{1,\dots,n-2}C_{1,\dots,n}(u_1,\dots,u_{n-2},a_{n-1},u_n) \\
&\qquad\qquad - D_{1,\dots,n-2}C_{1,\dots,n}(u_1,\dots,u_{n-2},0,u_n)\big]\,\mathrm{d}u_{n-2}\cdots\mathrm{d}u_1 \\
&\quad= \int_{(0,a_1]\times\cdots\times(0,a_{n-3}]} \big[D_{1,\dots,n-3}C_{1,\dots,n}(u_1,\dots,u_{n-3},a_{n-2},a_{n-1},u_n) \\
&\qquad\qquad - D_{1,\dots,n-3}C_{1,\dots,n}(u_1,\dots,u_{n-3},0,a_{n-1},u_n) \\
&\qquad\qquad - D_{1,\dots,n-3}C_{1,\dots,n}(u_1,\dots,u_{n-3},a_{n-2},0,u_n) \\
&\qquad\qquad + D_{1,\dots,n-3}C_{1,\dots,n}(u_1,\dots,u_{n-3},0,0,u_n)\big]\,\mathrm{d}u_{n-3}\cdots\mathrm{d}u_1 \\
&\quad= \cdots = \sum_{\substack{b_i\in\{0,a_i\}\\ 1\le i\le n-1}} (-1)^{\delta(b_1,\dots,b_{n-1})} C_{1,\dots,n}(b_1,\dots,b_{n-1},u_n) \\
&\quad\overset{(**)}{=} C_{1,\dots,n}(a_1,\dots,a_{n-1},u_n) = \int_\Omega \mathbb{1}_{\{U_n\le u_n\}} \prod_{i=1}^{n-1} \mathbb{1}_{\{U_i\le a_i\}}\,\mathrm{dP} = \int_A \mathbb{1}_{\{U_n\le u_n\}}\,\mathrm{dP}
\end{aligned}
$$

where $\delta(b_1,\dots,b_{n-1}) := \sum_{i=1}^{n-1} \mathbb{1}_{\{0\}}(b_i)$ counts the number of zeros.
$(*)$ follows from Kamke (1960, Satz 8, p. 177) applied to $n-1$ dimensions.
$(**)$ is due to the fact that $C_{1,\dots,n}(b_1,\dots,b_{n-1},u_n) = 0$ if any of the b_i is 0 (groundedness, see (2.5)). $\square$

The following corollary is proven completely analogously so that we will omit its proof.

2.28 Corollary
Let $X_1,\dots,X_n$, $n\in\mathbb{N}$, be real-valued, continuous random variables with distribution functions $F_1,\dots,F_n$ and copula $C_{1,\dots,n}$. For $k\in\{1,\dots,n\}$, assume that $C_{1,\dots,n}$ is k-times partially differentiable w. r. t. the first k arguments and the partial derivative is continuous in these arguments, i. e., $D_{1,\dots,k}C_{1,\dots,n}(u_1,\dots,u_n)$ exists and is continuous in $u_1,\dots,u_k$.

Further, let $C_{1,\dots,k}$ be the copula of $(X_1,\dots,X_k)$ and $c_{1,\dots,k}$ the associated density function, i. e., $c_{1,\dots,k}(u_1,\dots,u_k) = D_{1,\dots,k}C_{1,\dots,k}(u_1,\dots,u_k)$. The marginal densities $f_i(x) :=$

$\frac{d}{dx}F_i(x)$ *are assumed to exist for* $1 \le i \le k$. *Then,*

$$\begin{gathered}\frac{D_{1,\dots,k}C_{1,\dots,n}\big(F_1(X_1),\dots,F_k(X_k),F_{k+1}(x_{k+1}),\dots,F_n(x_n)\big)}{c_{1,\dots,k}\big(F_1(X_1),\dots,F_k(X_k)\big)\prod_{i=1}^k f_i(X_i)} \\ \text{is a version of} \\ \mathsf{P}(X_{k+1} \le x_{k+1},\dots,X_n \le x_n \mid X_1,\dots,X_k)\end{gathered} \tag{2.33}$$

for all $x_{k+1},\dots,x_n \in \mathbb{R}$.

2.29 Remark
It would be possible to formulate an even more general form by allowing mixed partials. I. e., for $I \subset \{1,\dots,n\}$ we could express $\mathsf{P}(X_i \le x_i, i \in I \mid X_j, j \in \{1,\dots,n\} \setminus I)$ with partials of copulas (cf. equation (2.31) on page 19). As this is notationally inconvenient, we will omit this. ■

2.3 Archimedean Copulas

Archimedean copulas are a special class of copulas which find a wide range of applications for a number of reasons: They can be easily constructed, a great variety of families of copulas belongs to that class and they have many desirable properties.

Archimedean copulas originally appeared in the study of probabilistic metric spaces, where they were part of the development of a version of the triangle inequality. The term "Archimedean" seems to be derived from the fact that there is a very close connection between these copulas and t-norms used to define the triangle inequality. These t-norms induce a measure of distance between two points of the probabilistic metric space (see Schweizer, 1991, for details and the references cited therein).

The source of all Archimedean copulas is a so-called generator which is a continuous, strictly decreasing function φ from I to $[0,\infty]$ such that $\varphi(1) = 0$.[10] The following theorem is elementary for the construction of Archimedean copulas:

2.30 Theorem (Nelsen, 1999, Theorem 4.1.4, p. 91)
Let φ *be a continuous, strictly decreasing function from* I *to* $[0,\infty]$ *such that* $\varphi(1) = 0$, *and let* $\varphi^{[-1]}$ *be the generalized inverse of* φ *defined by*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \le t \le \varphi(0), \\ 0, & \varphi(0) < t \le \infty. \end{cases} \tag{2.34}$$

Then the function $C : I^2 \to I$ *given by*

$$C(u,v) = \varphi^{[-1]}\big(\varphi(u) + \varphi(v)\big) \tag{2.35}$$

is a copula if and only if φ *is convex.* φ *is called* a generator *of the copula.*

[10]Note that φ will also stand for the density of a standard normal random variable. As it is common to denote the generator by φ we will stick to that convention as the meaning is always clear from the context.

2.31 Remark

a) The original proof of this theorem was given by Alsina et al. (2001). As this monograph is not yet available, we refer the reader to the proof in Nelsen (1999).

b) If $\varphi(0) = \infty$, φ is called a *strict* generator and $\varphi^{[-1]} = \varphi^{-1}$, i. e., the usual inverse. The induced copula is also called *strict*.

c) The generator φ of an Archimedean copula *is not unique*, as $c\varphi$, $c > 0$, yields the same copula as φ.

■

2.32 Example

a) For $\varphi(t) = -\ln(t)$, $t \in (0, 1]$, we have $\varphi(0) = \lim_{t \searrow 0} \varphi(t) = \infty$, therefore φ is strict with $\varphi^{-1}(t) = \exp(-t)$. The induced copula is given by

$$C(u, v) = \exp(-[-\ln u - \ln v]) = uv = \Pi(u, v).$$

Thus, the independence copula is a strict Archimedean copula.

b) The lower Fréchet bound $W(u, v) = \max(0, u + v - 1)$, $u, v \in I$, is Archimedean: Choosing $\varphi(t) = 1 - t$, $0 \le t \le 1$, yields W.

However, the upper Fréchet bound M *is not* Archimedean. This is due to the result that the diagonal $\delta_C(u) := C(u, u)$ of any Archimedean copula C meets the condition $\delta_C(u) < u$ for all $u \in (0, 1)$ (which in combination with associativity[11] is characterizing for Archimedean copulas, see Ling, 1965). As $\delta_M(u) = u$, this condition cannot be met by M.

c) Setting $\varphi_\vartheta(t) = -\ln\left(\frac{\exp(-\vartheta t)-1}{\exp(-\vartheta)-1}\right)$, $\vartheta \in \mathbb{R} \setminus \{0\}$, yields $\varphi(0) = \infty$ so that a strict copula family given by

$$C_\vartheta(u, v) = -\frac{1}{\vartheta} \ln\left[1 + \frac{(\exp(-\vartheta u) - 1)(\exp(-\vartheta v) - 1)}{\exp(-\vartheta) - 1}\right], \quad \vartheta \in \mathbb{R} \setminus \{0\}, \tag{2.36}$$

is generated. This is the *Frank family* which will prove to be important in the context of symmetric Markov processes (see Section 5.4.2).

■

Many classes of Archimedean copulas and some of their properties are described in Nelsen (1999, pp. 94–97).

For $0 \le t \le \varphi(0)$ we have $\varphi^{[-1]}(t) = \varphi^{-1}(t)$, so that $C(u, v) = \varphi^{-1}\big(\varphi(u) + \varphi(v)\big)$ which is equivalent to $\varphi\big(C(u, v)\big) = \varphi(u) + \varphi(v)$ for $\varphi(u) + \varphi(v) \le \varphi(0)$. So we get

$$\frac{\partial}{\partial u}\varphi\big(C(u, v)\big) = \varphi'\big(C(u, v)\big) \cdot D_1 C(u, v) = \varphi'(u) \quad \text{and}$$
$$\frac{\partial}{\partial v}\varphi\big(C(u, v)\big) = \varphi'\big(C(u, v)\big) \cdot D_2 C(u, v) = \varphi'(v),$$

[11] A 2-copula C is associative if $C(C(u, v), w) = C(u, C(v, w))$ holds for all $u, v, w \in I$.

thus

$$D_1C(u,v) = \frac{\varphi'(u)}{\varphi'\big(\varphi^{-1}(\varphi(u)+\varphi(v))\big)},$$
$$D_2C(u,v) = \frac{\varphi'(v)}{\varphi'\big(\varphi^{-1}(\varphi(u)+\varphi(v))\big)}$$

for $\varphi(u)+\varphi(v) \le \varphi(0)$ (i. e., for all u, v if φ is strict).

If C is absolutely continuous, φ is twice differentiable almost everywhere, and we can derive the density of C:

$$\begin{aligned} c(u,v) = D_{1,2}C(u,v) = D_2D_1C(u,v) = \frac{\partial}{\partial v}\frac{\varphi'(u)}{\varphi'\big(C(u,v)\big)} &= -\frac{\varphi'(u)\varphi''\big(C(u,v)\big)D_2C(u,v)}{\big(\varphi'(C(u,v))\big)^2} \\ = -\frac{\varphi'(u)\varphi''\big(C(u,v)\big)\varphi'(v)}{\big(\varphi'(C(u,v))\big)^3} &= -\frac{\varphi'(u)\varphi'(v)\varphi''\left(\varphi^{-1}(\varphi(u)+\varphi(v))\right)}{\left(\varphi'(\varphi^{-1}(\varphi(u)+\varphi(v)))\right)^3} \end{aligned}$$

for $u, v \in (0,1)$.

The *level curve* of a copula is the set $L_t \equiv L(t) := \{(u,v) \in I^2 : C(u,v) = t\}$, $t \in I$. For an Archimedean copula, $L(t) = \{(u,v) \in I^2 : \varphi(t) = \varphi(u)+\varphi(v)\} \subset I^2$ which connects the points $(1,t)$ and $(t,1)$ (as $\varphi(1) = 0$). We can write the level curve as $L_t(u) = v$ since solving for v as a function of u yields

$$v = L_t(u) = \varphi^{[-1]}\big(\varphi(t)-\varphi(u)\big) = \varphi^{-1}\big(\varphi(t)-\varphi(u)\big)$$

where the last step is justified because $\varphi(t)-\varphi(u)$ is in the interval $[0,\varphi(0))$.

It can be shown (Nelsen, 1999, Theorem 4.3.2, p. 39) that the level curves $L_t(u)$ are convex for all $t \in [0,1)$.

An important theorem concerning applications such as estimation (cf. Genest and Rivest, 1993) and random variate generation (see Section B.2) can be obtained by finding the C-measure of the region in I^2 lying on or below and to the left of each level curve.

2.33 Theorem (Nelsen, 1999, Theorem 4.3.4, pp. 101)
Let C be an Archimedean copula generated by $\varphi \in \Omega := \{f : I \to [0,\infty],\ f(1) = 0,\ f$ continuous, strictly monotone decreasing and convex$\}$. For $t \in I$, let $K_C(t) \equiv K_\varphi(t)$ denote the C-measure of the set $\{(u,v) \in I^2 : C(u,v) \le t\} = \bigcup_{s\in[0,t]} L_s$, or equivalently, of the set $\{(u,v) \in I^2 : \varphi(u)+\varphi(v) \ge \varphi(t)\}$. Then, for $t \in I$,

$$K_\varphi(t) = t - \frac{\varphi(t)}{\varphi'(t+)}. \tag{2.37}$$

The following corollary presents a probabilistic interpretation of the above theorem which will be useful in random variate generation (it is also useful for dealing with Kendall's τ for Archimedean copulas).

2.34 Corollary (Nelsen, 1999, Cor. 4.3.6, pp. 103)
Let $U, V \sim \mathsf{U}(0,1)$ whose joint distribution function is the Archimedean copula C generated by $\varphi \in \Omega$. Then the function K_φ given by (2.37) is the distribution function of the random variable $C(U,V)$. Furthermore, the joint distribution function of U and $C(U,V)$ is given by

$$K'_C(s,t) = \begin{cases} s, & s \leq t, \\ t - \frac{\varphi(t)-\varphi(s)}{\varphi'(t+)}, & s > t. \end{cases} \tag{2.38}$$

The next theorem which is due to Genest and Rivest (1993) is an extension of Corollary 2.34; an application is the algorithm for generating random variates from distributions with Archimedean copulas (see Section B.2).

2.35 Theorem
Under the hypotheses of Corollary 2.34, the joint distribution function $H(s,t)$ of $S = \varphi(U)/\big(\varphi(U)+\varphi(V)\big)$ and $T = C(U,V)$ is given by $s \cdot K_C(t)$ for all $s,t \in I$. Hence S and T are independent, and S is uniformly distributed on $(0,1)$.

2.36 Remark

a) The function $K_\varphi(z) = \mathsf{P}(C(U,V) \leq z)$, $z \in I$, can be estimated by means of its empirical distribution function $K_\varphi^{(n)}(z)$: Given a sample $\{(x_1,y_1),\ldots,(x_n,y_n)\}$ of size n from (X,Y), construct the empirical bivariate distribution function $H_n(x,y) := n^{-1}\sum_{i=1}^n \mathbb{1}_{(-\infty,x]}(x_i)\cdot\mathbb{1}_{(-\infty,y]}(y_i)$, $x,y \in \mathbb{R}$, of the sample. Then, calculate $H_n(x_i,y_i)$, $1 \leq i \leq n$, and use these values to construct the one-dimensional empirical distribution function $K_\varphi^{(n)}(z)$. However, the first step is unnecessary, as $H_n(x_i,y_i)$ is the proportion of sample observations which are less than or equal to (x_i,y_i) componentwise. For details, see Genest and Rivest (1993, pp. 1036).

b) $K_\varphi^{(n)}(z)$ can be fitted by the distribution function $K_\vartheta(z)$ of any family of Archimedean copulas, where the parameter ϑ is estimated in such a manner that the fitted distribution has a coefficient of concordance (Kendall's τ) equal to the empirical coefficient (see Genest and Rivest, 1993).

■

2.4 Survival Copula and Symmetry Concepts

In this section we introduce the survival copula and its relation to a symmetry concept called *radial symmetry* which will prove to be useful when characterizing symmetric stochastic processes later on.

2.4.1 Symmetry Concepts

In a univariate setting the meaning of symmetry of X about $a \in \mathbb{R}$ is clear:

$$X - a \sim a - X, \quad \text{i.e.,} \quad \mathsf{P}(X - a \le x) = \mathsf{P}(a - X \le x) \quad \text{for all } x \in \mathbb{R}$$

so that

$$\overline{F}(a - x) = F(a + x) \quad \text{for all } x \in \mathbb{R},\ F \text{ continuous.} \tag{2.39}$$

But what does it mean to say that (X, Y) is "symmetric" about some point $(a, b) \in \mathbb{R}^2$ or more generally that $(X_1, \dots, X_n)$ is "symmetric" about $(a_1, \dots, a_n) \in \mathbb{R}^n$? In the multivariate setting there is a number of possibilities in which sense random vectors are symmetric. We will shortly introduce three of them for the bivariate case and then generalize the most important one (radial symmetry) to n dimensions which will be a key tool for characterizing symmetric processes.

2.37 Definition (e. g., Nelsen, 1999, p. 31)
Let X and Y be real-valued random variables and $(a, b) \in \mathbb{R}^2$.

a) (X, Y) *is* marginally symmetric about (a, b) *if X and Y are univariately symmetric about a and b, respectively, according to (2.39).*

b) (X, Y) *is* radially symmetric about (a, b) *if the joint distribution function of $(X - a)$ and $(Y - b)$ is the same as the joint distribution function of $(a - X)$ and $(b - Y)$, i. e.,*

$$(X - a, Y - b) \sim (a - X, b - Y). \tag{2.40}$$

c) (X, Y) *is* jointly symmetric about (a, b) *if the following four pairs of random vectors have a common joint distribution:*

$$(X - a, Y - b), \quad (X - a, b - Y), \quad (a - X, Y - b), \quad (a - X, b - Y).$$

2.38 Remark

a) For X and Y we can express radial symmetry in terms of the joint distribution and the univariate survival functions of X and Y in a manner analogous to (2.39) (see below).

b) In the above definition we have *c)⇒b)⇒a)*. It can be shown that jointly symmetric random variables must be uncorrelated when the second order moments exist (Randles and Wolfe, 1979). Thus, the concept is too strong for many applications and we will focus on radial symmetry.

c) Marginal symmetry does not imply an intuitive bivariate symmetry as can be seen from Figure 2.2. Although the margins are symmetric $\mathsf{N}(0, 1)$ and t_3-variables, respectively, the bivariate density does not seem to be symmetric due to the unsymmetric dependence structure.

■

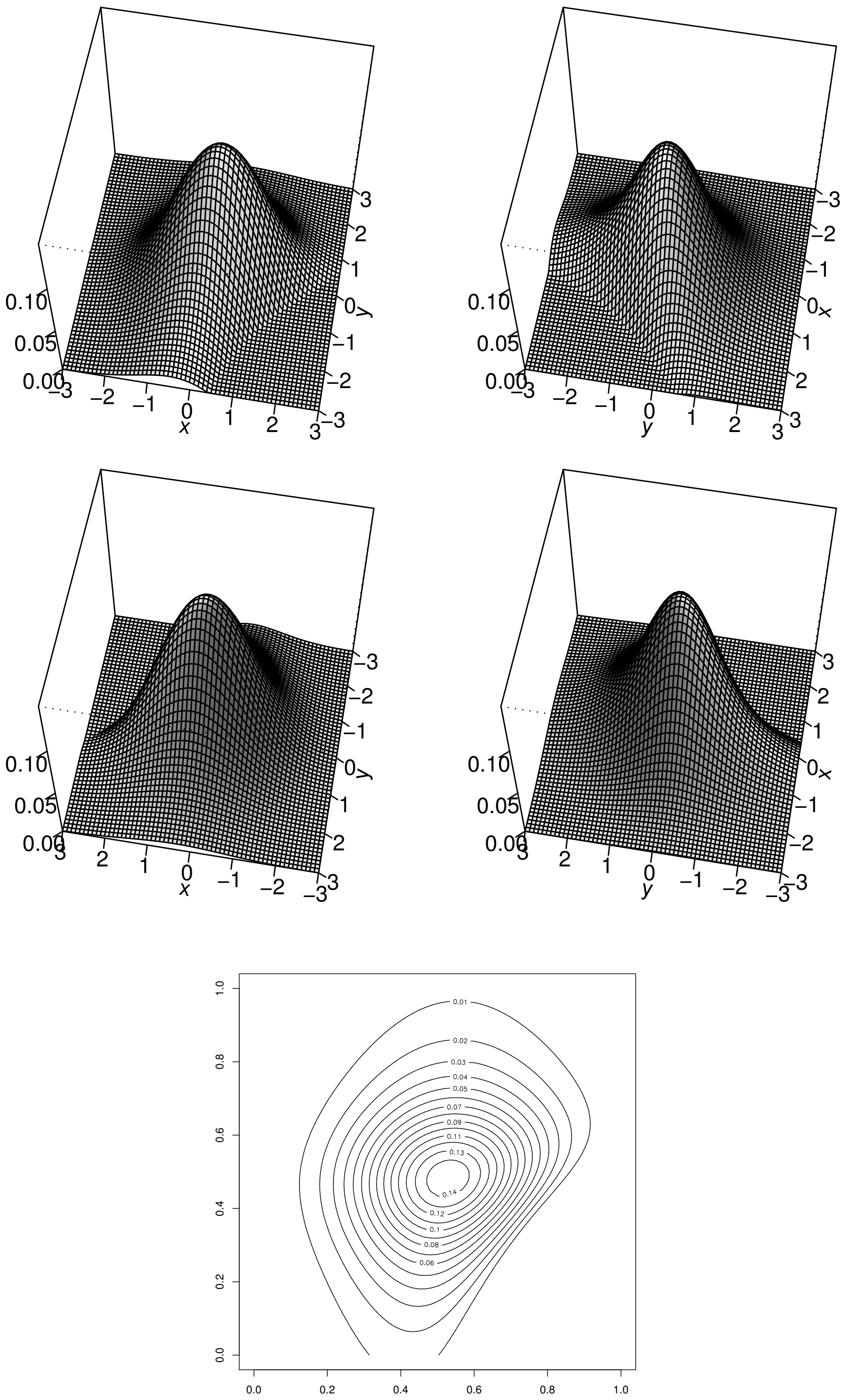

FIGURE 2.2. Bivariate distribution based on the min-max copula with parameter 3 and $\mathsf{N}(0,1)$- and t_3-distributed marginals (perspective and contour plots).

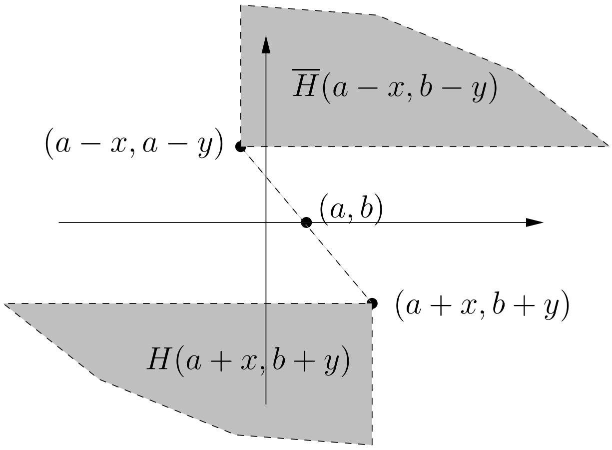

FIGURE 2.3. Regions of equal probability for radially symmetric random variables.

Let us generalize Definition 2.37 *b)*:

2.39 Definition (radial symmetry)
Let $X_1,\dots,X_n$ *be real-valued random variables,* $n \in \mathbb{N}$*, and* $(a_1,\dots,a_n) \in \mathbb{R}^n$*. The vector* $(X_1,\dots,X_n)$ *is said to be* radially symmetric about $(a_1,\dots,a_n)$ *if the joint distribution function of* $X_1-a_1,\dots,X_n-a_n$ *is the same as the one of* $a_1-X_1,\dots,a_n-X_n$*, i. e.,*

$$(X_1-a_1,\dots,X_n-a_n) \sim (a_1-X_1,\dots,a_n-X_n). \tag{2.41}$$

The following theorem easily connects radial symmetry with the survival function of the random vector $(X_1,\dots,X_n)$:

2.40 Theorem
A random vector $(X_1,\dots,X_n)$ *with joint distribution function* H *is radially symmetric about* $(a_1,\dots,a_n) \in \mathbb{R}^n$ *iff*

$$H(a_1+x_1,\dots,a_n+x_n) = \overline{H}(a_1-x_1,\dots,a_n-x_n) \qquad \text{for all } x_i \in \mathbb{R},\ 1 \le i \le n, \tag{2.42}$$

where $\overline{H}(x_1,\dots,x_n) := \mathsf{P}(X_1 > x_1,\dots,X_n > x_n)$ *is the survival function.*

The proof is obvious.

We see that “radial” means that $(a_1+x_1,\dots,a_n+x_n)$ and $(a_1-x_1,\dots,a_n-x_n)$ lie on rays emanating in opposite directions from $(a_1,\dots,a_n)$. For the bivariate case this is visualized in Figure 2.3.

2.4.2 Survival Copula

The univariate survival function of a random variable X is given by $\overline{F}(x) = 1-F(x)$, $x \in \mathbb{R}$. Analogously, the bivariate survival function of two random variables X and Y with distribution functions F and G, respectively, and joint distribution function H and copula C can be defined to be

$$\begin{aligned}\overline{H}(x,y) :&= \mathsf{P}(X > x, Y > y) = 1 - F(x) - G(y) + H(x,y)\\ &= \overline{F}(x) + \overline{G}(y) - 1 + C\big(1-\overline{F}(x), 1-\overline{G}(y)\big).\end{aligned}$$

Note that the marginals of $\overline{H}$ are the univariate survival functions, i. e., $\overline{H}(x, -\infty) = \overline{F}(x)$ and $\overline{H}(-\infty, y) = \overline{G}(y)$.

If we define

$$\widehat{C}(u, v) := u + v - 1 + C(1 - u, 1 - v), \qquad u, v \in I, \tag{2.43}$$

we have

$$\overline{H}(x, y) = \widehat{C}\big(\overline{F}(x), \overline{G}(y)\big).$$

$\widehat{C}$ is called the *survival copula* of C (it is easy to check that $\widehat{C}$ is indeed a copula).

Note that $\widehat{C}$ is **not** the joint survival function $\overline{C}$ of two $\mathsf{U}(0,1)$ distributed random variables with copula C as we have $\overline{C}(u, v) := \mathsf{P}(U > u, V > v) = 1 - u - v + C(u, v) = \widehat{C}(1 - u, 1 - v)$.

Now we want to define the n-variate survival copula $\widehat{C}$ of $(X_1, \ldots, X_n)$ such that

$$\overline{H}(x_1, \ldots, x_n) = \mathsf{P}(X_1 > x_1, \ldots, X_n > x_n) = \widehat{C}\big(\overline{F}_1(x_1), \ldots, \overline{F}_n(x_n)\big). \tag{2.44}$$

We can derive it with the help of the inclusion-exclusion formula (Billingsley, 1995, p. 24):

Let $(\Omega, \mathcal{A}, \mathsf{P})$ be a probability space and $A_1, \ldots, A_n \in \mathcal{A}$. Then,

$$\begin{aligned}\mathsf{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathsf{P}(A_k) - \sum_{1 \le i < j \le n} \mathsf{P}(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} \mathsf{P}(A_i \cap A_j \cap A_k) + \cdots \\ + (-1)^{n+1} \mathsf{P}(A_1 \cap \cdots \cap A_n).\end{aligned} \tag{2.45}$$

With $A_i := \{X_i \le x_i\}$ we have

$$\begin{aligned}\overline{H}(x_1, \ldots, x_n) &:= \mathsf{P}(X_1 > x_1, \ldots, X_n > x_n) = 1 - \mathsf{P}\left(\bigcup_{i=1}^{n} A_i\right) \\ &= 1 - \left(\sum_{i=1}^{n} \mathsf{P}(A_i) - \sum_{1 \le i < j \le n} \mathsf{P}(A_i \cap A_j) + \cdots + (-1)^{n+1} \mathsf{P}(A_1 \cap \cdots \cap A_n)\right) \\ &= 1 - \Big(\sum_{i=1}^{n} F_i(x_i) - \sum_{1 \le i < j \le n} C_{i,j}\big(F_i(x_i), F_j(x_j)\big) + \cdots \\ &\qquad + (-1)^{n+1} C_{1,\ldots,n}\big(F_1(x_1), \ldots, F_n(x_n)\big)\Big) \\ &= \sum_{i=1}^{n} \overline{F}_i(x_i) - (n-1) + \sum_{1 \le i < j \le n} C_{i,j}\big(1 - \overline{F}_i(x_i), 1 - \overline{F}_j(x_j)\big) + \cdots \\ &\qquad + (-1)^n \cdot C_{1,\ldots,n}\big(1 - \overline{F}_1(x_1), \ldots, 1 - \overline{F}_n(x_n)\big)\end{aligned}$$

so that

$$\begin{aligned}\widehat{C}(u_1, \ldots, u_n) &:= \sum_{i=1}^{n} u_i - (n-1) + \sum_{1 \le i < j \le n} C_{i,j}(1 - u_i, 1 - u_j) \\ &- \sum_{1 \le i < j < k \le n} C_{i,j,k}(1 - u_i, 1 - u_j, 1 - u_k) + \cdots + (-1)^n \cdot C_{1,\ldots,n}(1 - u_1, \ldots, 1 - u_n)\end{aligned} \tag{2.46}$$

is the n-dimensional survival copula of $(X_1, \ldots, X_n)$.

An equivalent representation due to the fact that all marginals $C_{i_1,\ldots,i_k}$ are derivable from $C = C_{1,\ldots,n}$ is given in Georges et al. (2001, Theorem 2), which states that

$$\widehat{C}(u_1, \ldots, u_n) = \overline{C}(1 - u_1, \ldots, 1 - u_n) \tag{2.47}$$

with

$$\begin{aligned}\overline{C}(u_1, \ldots, u_n) &= \sum_{k=0}^{n} \left[(-1)^k \sum_{v \in \mathcal{Z}(n-k,n,1;\boldsymbol{u})} C(v_1, \ldots, v_n) \right] \\ &= \sum_{(\varepsilon_1,\ldots,\varepsilon_n)\in\{0,1\}^n} (-1)^{\sum_{i=1}^n \varepsilon_i} C\big(\varepsilon_1 u_1 + (1-\varepsilon_1), \ldots, \varepsilon_n u_n + (1-\varepsilon_n)\big).\end{aligned}$$

where $\mathcal{Z}(m, n, \varepsilon; \boldsymbol{u}) := \{\boldsymbol{v} \in I^n : v_i \in \{u_i, \varepsilon\},\ 1 \le i \le n,\ \sum_{k=1}^n \mathbb{1}_{\{\varepsilon\}}(v_k) = m\}$.

For example, the three-variate version ($C = C_{1,2,3}$) is given by

$$\begin{aligned}\widehat{C}(u_1, u_2, u_3) &\overset{(2.46)}{=} u_1 + u_2 + u_3 - 2 + C_{1,2}(1-u_1, 1-u_2) + C_{1,3}(1-u_1, 1-u_3) \\ &\quad + C_{2,3}(1-u_2, 1-u_3) - C_{1,2,3}(1-u_1, 1-u_2, 1-u_3) \\ &\overset{(2.47)}{=} C(1,1,1) - C(1-u_1,1,1) - C(1,1-u_2,1) - C(1,1,1-u_3) \\ &\quad + C(1-u_1, 1-u_2, 1) + C(1, 1-u_2, 1-u_3) + C(1-u_1, 1, 1-u_3) \\ &\quad - C(1-u_1, 1-u_2, 1-u_3).\end{aligned}$$

We are now ready to derive the following corollary from Theorem 2.40.

2.41 Corollary

A random vector $(X_1, \ldots, X_n)$ with copula C and univariate survival functions $\overline{F}_i$, $1 \le i \le n$, is radially symmetric about $(a_1, \ldots, a_n) \in \mathbb{R}^n$ iff

a) *$\overline{F}_i(a_i - x_i) = F_i(a_i + x_i)$ for all $x_i \in \mathbb{R}$, $1 \le i \le n$, i. e., all the X_i are symmetric about a_i, and*

b) *$C(u_1, \ldots, u_n) = \widehat{C}(u_1, \ldots, u_n)$ for all $(u_1, \ldots, u_n) \in \operatorname{ran}(F_1) \times \cdots \times \operatorname{ran}(F_n)$.*

Proof. By Theorem 2.40, radial symmetry is equivalent to

$$\begin{gathered}H(a_1 + x_1, \ldots, a_n + x_n) = \overline{H}(a_1 - x_1, \ldots, a_n - x_n) \quad \Longleftrightarrow \\ C\big(F_1(a_1 + x_1), \ldots, F_n(a_n + x_n)\big) = \widehat{C}\big(\overline{F}_1(a_1 - x_1), \ldots, \overline{F}_n(a_n - x_n)\big).\end{gathered}$$

Thus, the sufficiency of *a)* and *b)* is clear. For the necessity, we observe that in the case of radial symmetry

$$\begin{aligned}F_i(a_i + x_i) &= H(\infty, \ldots, \infty, a_i + x_i, \infty, \ldots, \infty) \\ &\overset{\text{radial symm.}}{=} \overline{H}(-\infty, \ldots, -\infty, a_i - x_i, -\infty, \ldots, -\infty) = \overline{F}_i(a_i - x_i)\end{aligned} \tag{2.48}$$

for all $x_i \in \mathbb{R}$.

Now, let $(u_1, \ldots, u_n) \in \operatorname{ran}(F_1) \times \cdots \times \operatorname{ran}(F_n)$. Then there exist $x_i \in \mathbb{R}$ such that $u_i = F_i(a_i + x_i)$, $1 \leq i \leq n$. Thus,

$$\begin{aligned} C(u_1, \ldots, u_n) &= C\big(F_1(a_1 + x_1), \ldots, F_n(a_n + x_n)\big) = H(a_1 + x_1, \ldots, a_n + x_n) \\ &\overset{\text{radial symm.}}{=} \overline{H}(a_1 - x_1, \ldots, a_n - x_n) \overset{(2.44)}{=} \widehat{C}\big(\overline{F}_1(a_1 - x_1), \ldots, \overline{F}_n(a_n - x_n)\big) \\ &\overset{(2.48)}{=} \widehat{C}\big(F_1(a_1 + x_1), \ldots, F_n(a_n + x_n)\big) = \widehat{C}(u_1, \ldots, u_n) \end{aligned}$$

which completes the proof. □

This corollary will later be used to characterize symmetric stochastic processes.

3
Dependence Measures

In this chapter we will give a short introduction to some dependence measures commonly used in multivariate settings. Basic references are Nelsen (1999) and Joe (1997). It is obvious that we are interested in dependence measures which are closely related to copulas.

3.1 Comonotonicity

Two random variables X and Y are said to be *comonotonic* if a random variable Z exists such that

$$X = f(Z), \qquad Y = g(Z) \quad \text{a. s.},$$

where f and g are two increasing real-valued functions.

Comonotonicity is an extension of the concept of (perfect) positive correlation to variables with arbitrary distributions, i. e., while correlation of two random variables is often understood as *linear dependence*, we may here have any monotone relation between the variables.

We note that the upper Fréchet bound for copulas is reached if X and Y are comonotonic, the lower bound is reached if X and Y are countermonotonic, i.e., if X and $-Y$ are comonotonic (see Remark 2.14 and Embrechts et al., 1999). If X and Y are perfectly positively correlated, then X and Y are comonotonic; the converse is generally not true.

3.2 Concordance

In the following we will need the notion of *concordance*. Informally, a pair of random variables X and Y is concordant if "large" values of one tend to be associated to "large" values of the other variable, and analogously for "small" values. To be precise, let (X, Y) be a vector of two random variables and (x_1, y_1), (x_2, y_2) two samples from (X, Y). We will say that (x_1, y_1) and (x_2, y_2) are *concordant* if $(x_1 - x_2)(y_1 - y_2) > 0$, the samples are *discordant* if $(x_1 - x_2)(y_1 - y_2) < 0$. It is clear that we could define concordance of two vectors (x_1, y_1) and (x_2, y_2) without using random variables. However, in our context there is no need for such a generalization.

Note that $\mathsf{P}(X_1 = X_2) = 0$ if X_1 and X_2 are continuous. Therefore, if X and Y above are continuous, the regions of concordance and discordance split the sample space as a subset of $\mathbb{R}^2$ into two non-intersecting regions whose P-measure is 1.

3.3 Kendall's τ

3.1 Definition (Kendall's $\boldsymbol{\tau}$)
Let (X_1, Y_1) and (X_2, Y_2) be two independent and identically distributed random vectors on some probability space $(\Omega, \mathcal{A}, \mathsf{P})$. Kendall's τ is defined as

$$\tau \equiv \tau(X_1, Y_1) \equiv \tau_{X_1,Y_1} := \mathsf{P}\big((X_1 - X_2)(Y_1 - Y_2) > 0\big) - \mathsf{P}\big((X_1 - X_2)(Y_1 - Y_2) < 0\big). \quad (3.1)$$

We will also call this the population version *of Kendall's τ.*

Thus, Kendall's τ is just the difference between the probability of concordance and the probability of discordance.

As one may suppose, there is also a "sample version" of this dependence measure, based on a sample from $(X, Y) \sim H$. Let $\{(x_1, y_1), \ldots, (x_n, y_n)\}$, $n \geq 2$, denote a sample of n observations from a vector (X, Y) of continuous random variables. Each pair $\{(x_i, y_i), (x_j, y_j)\}$, $i, j \in \{1, \ldots, n\}$, $i \neq j$, is either discordant or concordant as there are no ties in the sample with probability one. There are obviously $\binom{n}{2}$ distinct pairs of observations in the sample. Let c denote the number of concordant pairs, d the number of discordant pairs. Then Kendall's τ for the sample (the above mentioned "sample version") is defined to be

$$\hat{\tau} \equiv t := \frac{c - d}{c + d} = \frac{c - d}{\binom{n}{2}}. \quad (3.2)$$

If we choose a pair of observations (x_i, y_i) and (x_j, y_j) randomly from the sample (for which there are $\binom{n}{2}$ possibilities which are equally likely) $\hat{\tau}$ can be interpreted as the probability that we get a concordant pair minus the probability that we get a discordant pair.

We can express concordance and discordance with indicator functions $\mathbb{1}_A(x)$ for Borel sets A in the following way:

$$\begin{aligned} (x_i, y_i) \text{ and } (x_j, y_j) \text{ are concordant } &\Leftrightarrow \mathbb{1}_{(0,\infty)}\big((x_i - x_j)(y_i - y_j)\big) = 1, \\ (x_i, y_i) \text{ and } (x_j, y_j) \text{ are discordant } &\Leftrightarrow \mathbb{1}_{(-\infty,0)}\big((x_i - x_j)(y_i - y_j)\big) = 1. \end{aligned}$$

When X or Y is continuous, we can also write

$$\mathbb{1}_{(0,\infty)}\big((X_i - X_j)(Y_i - Y_j)\big) = 0 \qquad [\mathsf{P}]$$

for discordance. Therefore, the number of concordant values $c\big((\boldsymbol{x}, \boldsymbol{y})\big)$ in a sample $(\boldsymbol{x}, \boldsymbol{y}) = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ is given by

$$c\big((\boldsymbol{x}, \boldsymbol{y})\big) = \frac{1}{2} \sum_{1 \leq i < j \leq n}^{n} \mathbb{1}_{(0,\infty)}\big((x_i - x_j)(y_i - y_j)\big).$$

The next theorem is the basis for connecting copulas with Kendall's τ.

3.2 Theorem
Let (X, Y) be a random vector with continuous marginals, i. e., X is a continuous random variable with distribution function F and Y is continuous with distribution function G. Let (X_1, Y_1) and (X_2, Y_2) be two random vectors with distribution function H_1 and H_2, respectively, where $X_i \sim F$ and $Y_i \sim G$, $i = 1, 2$. Assume that all random variables are defined on a probability space $(\Omega, \mathcal{A}, \mathsf{P})$.

Further, $H_1(x, y) = C_1\big(F(x), G(y)\big)$ and $H_2(x, y) = C_2\big(F(x), G(y)\big)$, C_1 and C_2 denoting the underlying copulas. If Q denotes the difference between the probabilities of concordance and discordance of (X_1, Y_1) and (X_2, Y_2),

$$Q := \mathsf{P}\big((X_1 - X_2)(Y_1 - Y_2) > 0\big) - \mathsf{P}\big((X_1 - X_2)(Y_1 - Y_2) < 0\big),$$

we have:

$$Q = Q(C_1, C_2) = 4 \iint_{I^2} C_2(u, v)\, \mathrm{d}C_1(u, v) - 1. \tag{3.3}$$

For a proof see Nelsen (1999, pp. 127).

Obviously, Q does not depend on the marginal distribution functions. Kendall's τ can then easily be expressed by the underlying copula of two random variables X and Y:

3.3 Theorem
Let X and Y be continuous real-valued random variables whose copula is C. Then Kendall's τ for X and Y is given by

$$\tau_{X,Y} \equiv \tau_C = Q(C, C) = 4 \iint_{I^2} C(u, v)\, \mathrm{d}C(u, v) - 1 = 4 \cdot \mathsf{E}_C\left(C(U, V)\right) - 1 \tag{3.4}$$

where U and V are $\mathsf{U}(0, 1)$ random variables with copula C and E_C denotes the expectation w. r. t. the measure P_C induced by the copula.

Proof. Nelsen (1999, p. 129). □

3.4 Spearman's ρ

Spearman's ρ is another nonparametric measure of dependence defined in terms of concordance and discordance.

3.4 Definition (Spearman's ρ)
Let (X_1, Y_1), X_2, and Y_2 be independent random variables and vector, respectively, where $X_i \sim F$, $Y_i \sim G$, $i = 1, 2$, and (X_1, Y_1) has joint distribution function H. All random variables are defined on $(\Omega, \mathcal{A}, \mathsf{P})$.

The population version *of Spearman's ρ is defined to be proportional to the probability of concordance minus the probability of discordance for the two vectors (X_1, Y_1) and (X_2, Y_2):*

$$\rho \equiv \rho(X_1, Y_1) \equiv \rho^s_{X_1,Y_1} := 3 \cdot \Big(\mathsf{P}\big((X_1 - X_2)(Y_1 - Y_2) > 0\big) - \mathsf{P}\big((X_1 - X_2)(Y_1 - Y_2) < 0\big)\Big). \tag{3.5}$$

We can then formulate the following relationship between the underlying copula C and Spearman's ρ. For a proof see Nelsen (1999, pp. 135).

3.5 Theorem
Let X and Y be continuous random variables with copula C. Then

$$\rho^s_{X,Y} \equiv \rho^s_C = 12 \cdot \iint_{I^2} uv \, \mathrm{d}C(u,v) - 3 = 12 \iint_{I^2} C(u,v) \, \mathrm{d}u \, \mathrm{d}v - 3 = 12 \cdot \mathsf{E}[C(U,V)] - 3, \tag{3.6}$$

where U and V are two iid $\mathsf{U}(0,1)$ random variables.

3.5 Tail Dependencies

In this section we will give definitions of common measures of dependence of two random variables X and Y w.r.t. their tail behaviour. Especially the concept of upper and lower tail dependence is useful for modeling clustering of extreme values in time series (see Joe, 1997, pp. 249).

The following concept was firstly established in the works of Lehmann (1966) and Esary et al. (1967).

3.6 Definition (LTD, RTI)
Let X and Y be two real-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathsf{P})$. X is said to be left-tail decreasing *in the variable Y, denoted* LTD$(X \mid Y)$*, if and only if*

$$\mathsf{P}(X \leq x \mid Y \leq y)$$

is a decreasing function in y for all $x \in \mathbb{R}$.
Likewise, X is declared to be right-tail increasing *in Y, denoted* RTI$(X \mid Y)$*, if and only if*

$$\mathsf{P}(X > x \mid Y > y)$$

is an increasing function of y for all $x \in \mathbb{R}$.

3.7 Remark
Note that we have the following equivalent relations. They directly follow from elementary calculations:

$$\begin{aligned} \mathrm{LTD}(X \mid Y) &\iff \mathsf{P}(X > x \mid Y \leq y) \text{ is increasing in } y \text{ for all } x, \text{ and} \\ \mathrm{RTI}(X \mid Y) &\iff \mathsf{P}(X \leq x \mid Y > y) \text{ is decreasing in } y \text{ for all } x. \end{aligned}$$

■

We can give a sufficient condition for $\rho(X,Y) \geq \tau(X,Y) \geq 0$, which is a feature often observed for positively dependent random variables.

3.8 Theorem (Capéraà and Genest, 1993)
Let X and Y be real-valued random variables with joint cumulative distribution function $H(x,y)$ and continuous marginals F and G, respectively. If $\mathsf{P}(X \le x \mid Y \le y)$ and $\mathsf{P}(X \le x \mid Y > y)$ are decreasing functions of y for all $x \in \mathbb{R}$, then $\rho(X,Y) \ge \tau(X,Y) \ge 0$. Symbolically,

$$\mathrm{RTI}(X \mid Y) \quad \text{and} \quad \mathrm{LTD}(X \mid Y) \quad \Longrightarrow \quad \rho(X,Y) \ge \tau(X,Y) \ge 0, \tag{3.7}$$

where ρ and τ denote Spearman's ρ and Kendall's τ, respectively.

It is now obvious to ask for a translation of these definitions and results into the "copula-language". We must observe that these definitions are not necessarily independent of the marginals which can be seen from the form $\mathsf{P}(X \le x \mid Y \le y) = C(F(x), G(y))/G(y)$. However, we are able to give sufficient conditions for two random variables X and Y to be RTI and LTD based only on their copula(s)[1] by using the next lemma and showing the relations *a)* or *b)* for all $u, v \in (0,1)$.

3.9 Lemma
Let X and Y be two (not necessarily continuous) random variables on some probability space $(\Omega, \mathcal{A}, \mathsf{P})$ with marginals F and G, respectively, and copula C.[2] Let $I_F := (0,1) \cap \operatorname{ran}(F)$ and $I_G := (0,1) \cap \operatorname{ran}(G)$. Then we have the following relations:

a) $\mathrm{LTD}(X \mid Y) \quad \Longleftrightarrow \quad C(u,v)/v$ *is decreasing in $v \in I_G$ for all $u \in I_F$.*

b) $\mathrm{RTI}(X \mid Y) \quad \Longleftrightarrow \quad (C(u,v) - u)/(1-v)$ *is increasing in $v \in I_G$ for all $u \in I_F$.*

Before we provide the proof, let us remark that we do not need to include the points $u = 0$ or $u = 1$ as the copula is either 0 or v which results in constant and insofar increasing functions.

Proof. We will only show the first part as the second one can be shown analogously. For the "if"-part, assume $C(u,v)/v$ to be a decreasing function in $v \in I_G$ for all $u \in I_F$. Take $x \in \mathbb{R}$, $y_1, y_2 \in \mathbb{R}$ with $y_1 < y_2$. Define $u := F(x)$, $v_1 := G(y_1)$ and $v_2 := G(y_2)$. Then $v_2 \ge v_1$ due to the monotonicity of G. By assumption, we have

$$\begin{aligned} \frac{C(u,v_1)}{v_1} \ge \frac{C(u,v_2)}{v_2} \quad &\Leftrightarrow \quad \frac{C\big(F(x), G(y_1)\big)}{G(y_1)} \ge \frac{C\big(F(x), G(y_2)\big)}{G(y_2)} \\ &\Leftrightarrow \quad \mathsf{P}(X \le x \mid Y \le y_1) \ge \mathsf{P}(X \le x \mid Y \le y_2). \end{aligned}$$

As x, y_1 and y_2 are arbitrary, it follows that $\mathrm{LTD}(X \mid Y)$.

For the "only if"-part, assume $\mathrm{LTD}(X \mid Y)$ and $u \in I_F$, $v_1, v_2 \in I_G$ with $v_1 < v_2$. Then there exist $x \in \mathbb{R}$ with $F(x) = u$ and $y_1, y_2 \in \mathbb{R}$ with $G(y_1) = v_1$, $G(y_2) = v_2$. As G is

[1] It turns out that we do not need uniqueness of the copula.

[2] Note that no assumption of uniqueness of the copula is needed. Cf. Remark 3.10

increasing we must have $y_2 > y_1$. (For, if $y_1 \geq y_2$, then $v_1 = G(y_1) \geq G(y_2) = v_2$ due to the monotonicity of G). But for $y_1 < y_2$ we have

$$\begin{aligned}\mathsf{P}(X \leq x \mid Y \leq y_2) = \frac{C\big(F(x), G(y_2)\big)}{G(y_2)} = \frac{C(u, v_2)}{v_2} &\leq \mathsf{P}(X \leq x \mid Y \leq y_1) \\ &= \frac{C\big(F(x), G(y_1)\big)}{G(y_1)} = \frac{C(u, v_1)}{v_1}.\end{aligned}$$

Therefore, $C(u, v)/v$ is decreasing in $v \in I_G$ for all $u \in I_F$. □

3.10 Remark

a) If the marginals F and G are continuous, we have $\operatorname{ran}(F) = \operatorname{ran}(G) = I$ and therefore a unique copula C. This is due to the fact that in the proof of Sklar's theorem, the copula is *defined* on the points $\operatorname{ran}(F) \times \operatorname{ran}(G)$ and interpolated for all other points. However, the interpolation is not unique. Thus, if $\operatorname{ran}(F) \times \operatorname{ran}(G) = I^2$, the copula can be uniquely defined.

b) If $C(u, v)$ is partially differentiable w. r. t. v for all $u \in (0, 1)$, i. e., D_2C exists, we have the following relations:

$$\begin{aligned} &\text{(i)}\ \ C(u, v) \geq v D_2 C(u, v) \ \text{ for all } u, v \in (0, 1) \quad \Longleftrightarrow \quad \mathrm{LTD}(X \mid Y), \\ &\text{(ii)}\ \ D_2 C(u, v)(1 - v) + C(u, v) - u \geq 0 \text{ for all } u, v \in (0, 1) \quad \Longleftrightarrow \quad \mathrm{RTI}(X \mid Y). \end{aligned}$$

■

The next dependence concept relates to the amount of dependence in the upper quadrant or lower quadrant tail of a bivariate distribution. It is particularly relevant to dependence in extreme values and for the modeling of consecutive extremal events. In contrast to τ and ρ, *upper and lower tail dependence* is a local dependence measure in the sense of being defined for some interesting areas (the upper and lower quadrants) instead of being defined globally.

Like Kendall's τ and Spearman's ρ tail dependence is a copula property. The amount of tail dependence is therefore invariant under strictly increasing transformations of X and Y.

3.11 Definition (upper and lower tail dependence)
Let X and Y be random variables with cumulative distribution functions (cdf) F_X and F_Y, respectively. The coefficient of upper tail dependence *is then defined as*

$$\lim_{u \nearrow 1} \mathsf{P}\big(Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u)\big) =: \lambda_U \tag{3.8}$$

if $\lambda_U \in [0, 1]$ exists. If $\lambda_U \in (0, 1]$, X and Y are said to be asymptotically dependent in the upper tail. If C is the copula associated to X and Y we obtain

$$\lambda_U = \lim_{u \nearrow 1} \frac{\overline{C}(u, u)}{1 - u} = \lim_{u \nearrow 1} \frac{1 - 2u + C(u, u)}{1 - u}, \tag{3.9}$$

where $\overline{C}(u,v) = 1 - u - v + C(u,v)$ *is the joint survival function. The* coefficient of lower tail dependence *is analogously defined as*

$$\lambda_L := \lim_{u \searrow 0} \mathsf{P}\big(Y \le F_Y^{-1}(u) \mid X \le F_X^{-1}(u)\big) = \lim_{u \searrow 0} \frac{C(u,u)}{u} \tag{3.10}$$

where the limit exists.

$\lambda_{\{L,U\}}$ is extensively used in extreme value theory for describing the property that one variable is extreme given the other one is extreme. Note that $\lambda_U = \lim_{u \nearrow 1} \mathsf{P}(U_1 > u \mid U_2 > u) = \lim_{u \nearrow 1} \mathsf{P}(U_2 > u \mid U_1 > u)$.

3.6 Quadrant Dependence and Positive Likelihood Ratio Dependence

A very prominent dependence property between random variables is a "lack of dependence" which means independence. Thus, it makes sense to compare the observed dependence with that of independent variables.

The following definition will be needed for the result in Theorem 4.5:

3.12 Definition (PQD; Lehmann, 1966)
Let X and Y be real-valued random variables. X and Y are positively quadrant dependent (PQD) *if for all* $(x,y) \in \mathbb{R}^2$

$$\mathsf{P}(X \le x, Y \le y) \ge \mathsf{P}(X \le x) \cdot \mathsf{P}(Y \le y). \tag{3.11}$$

We will write PQD(X,Y).

In words, X and Y are PQD if the probability of them being small simultaneously (or being large simultaneously) is at least as great as it would be were they independent.

It is clear that a sufficient condition for X and Y to be PQD is that

$$C(u,v) \ge uv \quad \text{for all } u, v \in I \tag{3.12}$$

holds for their copula.

For two positively quadrant dependent random variables, the following theorem holds for the relation between Kendall's τ and Spearman's ρ:

3.13 Theorem
Let X and Y be continuous random variables. If X and Y are PQD, *then*

$$3\tau_{X,Y} \ge \rho_{X,Y} \ge 0. \tag{3.13}$$

For a proof, see Nelsen (1999, Theorem 5.2.2).

A very strong dependence concept which will be sufficient for many other dependence relationships to hold is "positive likelihood ratio dependence" which is closely related to functions being "totally positive of order two". A function $f : \overline{\mathbb{R}}^2 \to \mathbb{R}$ is *totally positive of order two* (TP$_2$) if $f \ge 0$ and whenever $x \le x'$ and $y \le y'$,

$$\begin{vmatrix} f(x,y) & f(x,y') \\ f(x',y) & f(x',y') \end{vmatrix} \ge 0, \tag{3.14}$$

where $|\cdot|$ denotes the determinant.

3.14 Definition (PLR; Lehmann, 1966)
Let X and Y be continuous random variables with joint density function $h(x,y)$, $x,y \in \mathbb{R}$. Then X and Y are positively likelihood ratio dependent *(denoted by* $\mathrm{PLR}(X,Y)$*) if h satisfies*

$$h(x,y) \cdot h(x',y') \geq h(x,y') \cdot h(x',y) \tag{3.15}$$

for all $x,x',y,y' \in \overline{\mathbb{R}}$ such that $x \leq x'$ and $y \leq y'$.

This property derives its name from the fact that the inequality in (3.15) is equivalent to the requirement that the conditional density of Y given $X = x$ has a monotone likelihood ratio.

There are many relationships between all dependence concepts, however, we will only make use of the following:

3.15 Theorem
Let X and Y be random variables with an absolutely continuous distribution function. If $\mathrm{PLR}(X,Y)$*, then* $\mathrm{RTI}(X \mid Y)$, $\mathrm{LTD}(X \mid Y)$ *and* $\mathrm{PQD}(X,Y)$.

Proof. The assertion follows immediately with Theorem 5.2.19 and the remarks after Corollary 5.2.17 in Nelsen (1999). $\square$

4

Two Examples: Analysis of Spatial Dependence by Using Copulas

In this chapter we will give two examples of the use of copulas for the analysis of spatial dependence, i. e., of (at least) two random variables at the same time.

The first application deals with the depencence between the minimum and the maximum of n iid random variables. We will derive the corresponding copula and calculate Kendall's τ and Spearman's ρ for $(X_{(1)}, X_{(n)})$. We can finally show that the relation $3\tau_n \geq \rho_n \geq \tau_n > 0$ holds in this situation.

As a second application, we will take a closer look at the joint distribution of Brownian motion $(B_t)_{t\in[0,\infty)}$ (see Definition 5.22) and its supremum process $(S_t)_{t\in[0,\infty)}$ where $S_t := \sup_{0\leq s\leq t} B_t$. By deriving the copula we see that the dependence structure of B_t and S_t *does not* depend on time, i. e., although we have to start with the derivation of C_t as a function of time, the copula will prove to be time-independent.

4.1 The Min-Max Copula

In this section we are interested in the dependence structure of the minimum and the maximum of n independent and identically distributed random variables $X_1, \ldots, X_n$ with continuous distribution function F. Let $X_{(i)}$ denote the i-th order statistic of $(X_1, \ldots, X_n)$, i. e., $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$, then we would like to determine the copula of $(X_{(1)}, X_{(n)})$. If $F_{r,s}(x, y) := \mathsf{P}(X_{(r)} \leq x, X_{(s)} \leq y)$ denotes the joint distribution function of the r-th and s-th order statistic ($r < s$), then it is well known (e. g., David, 1981, pp. 8–11) that

$$F_{r,s}(x,y) = \begin{cases} \sum_{j=s}^{n} \sum_{i=r}^{j} \frac{n!}{i!(j-i)!(n-j)!} F^i(x)\left[F(y)-F(x)\right]^{j-i}\left[1-F(y)\right]^{n-j}, & x < y, \\ F_s(y), & x \geq y, \end{cases} \tag{4.1}$$

where $F_s(x) := \mathsf{P}(X_{(s)} \leq x)$ is the distribution function of the s-th order statistic:

$$F_s(x) = \sum_{i=s}^{n} \binom{n}{i} F^i(x)\bigl(1-F(x)\bigr)^{n-i}. \tag{4.2}$$

For the marginal distributions of $F_{1,n}$, we have in particular

$$F_1(x) = 1 - (1 - F(x))^n, \quad F_n(y) = F^n(y), \quad x, y \in \mathbb{R}$$

so that in the special case of $r = 1$ and $s = n$ we get

$$\begin{aligned} F_{1,n}(x, y) &= \begin{cases} \sum_{i=1}^n \binom{n}{i} F^i(x)(F(y) - F(x))^{n-i}, & x < y, \\ F_n(y), & x \geq y, \end{cases} \\ &= \begin{cases} F^n(y) - (F(y) - F(x))^n, & x < y, \\ F^n(y), & x \geq y. \end{cases} \end{aligned}$$

Solving the equation $C\big(F_1(x), F_n(y)\big) = F_{1,n}(x, y)$ for the function C by setting $C(u, v) = F_{1,n}\big(F_1^{-1}(u), F_n^{-1}(v)\big)$, $u, v \in (0, 1)$, with $F_1^{-1}(u) = F^{-1}\big(1 - (1 - u)^{1/n}\big)$ and $F_n^{-1}(v) = F^{-1}(v^{1/n})$ yields the underlying copula C_n of $(X_{(1)}, X_{(n)})$:

$$C_n(u, v) = \begin{cases} v - \big(v^{1/n} + (1 - u)^{1/n} - 1\big)^n, & 1 - (1 - u)^{1/n} < v^{1/n}, \\ v, & 1 - (1 - u)^{1/n} \geq v^{1/n}. \end{cases} \tag{4.3}$$

Note that the conditions for the domain of the copula stem from solving $F_1(x) > F_n(y)$. We thus found the family of copulas describing the dependence structure of the minimum and maximum of n independent random variables. To get a visual impression we provide some plots of these copulas for different values of n in Figure 4.1.

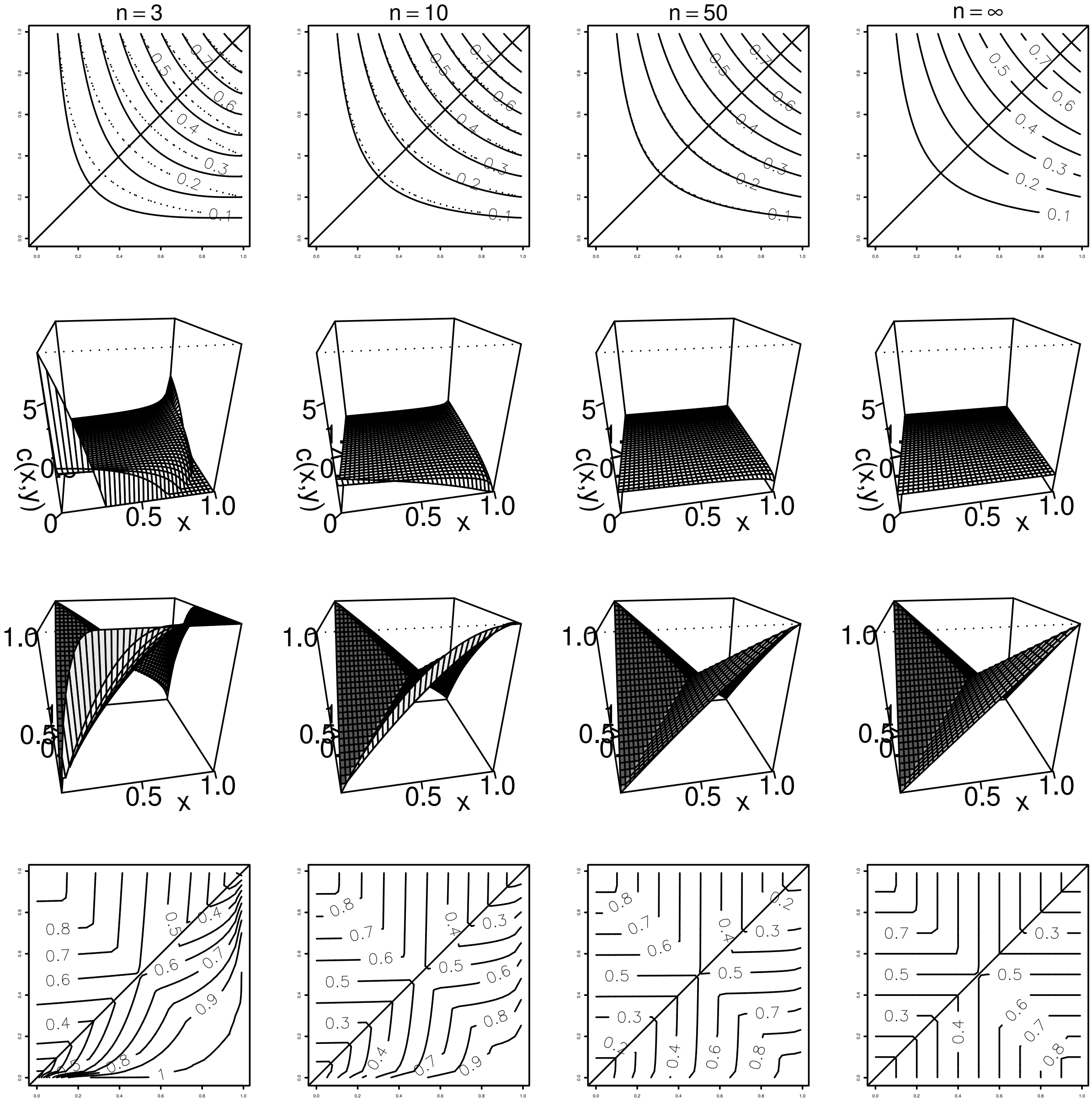

FIGURE 4.1. First row: contourplots of C_n (dotted line: contours of independence copula; main diagonal added to provide better orientation), second row: corresponding densities; third and fourth row: surface- and contourplots of $(C_n - W)/(M - W)$.

We also note the interesting fact that this copula is closely related to a well known copula family. From Theorem 2.15 we know that for a strictly decreasing function f and a strictly increasing function g, the copula of $(f(X_{(1)}), g(X_{(n)}))$ is given by

$$\begin{aligned} C_{f(X_{(1)}),g(X_{(n)})}(u,v) &= v - C_{X_{(1)},X_{(n)}}(1-u,v) = v - C_n(1-u,v) \\ &= \begin{cases} (v^{1/n}+u^{1/n}-1)^n, & v^{1/n}+u^{1/n}-1>0, \\ 0, & \text{elsewhere.} \end{cases} \end{aligned}$$

However, this is a member of the family of copulas discussed by Kimeldorf and Sampson (1975), Clayton (1978), Cook and Johnson (1981, 1986), Oakes (1982, 1986), and implicitly already by Mardia (1962) which is usually called *Clayton* or *Cook and Johnson* copula. For $\vartheta \in [-1,\infty) \setminus \{0\}$ it is defined as

$$C_\vartheta^{Cl}(u,v) := \max\bigl((u^{-\vartheta}+v^{-\vartheta}-1)^{-1/\vartheta}, 0\bigr), \quad u,v \in (0,1). \tag{4.4}$$

Note that it is an Archimedean copula with generator $\varphi_\vartheta(t) = (t^{-\vartheta}-1)/\vartheta$ (see Nelsen, 1999, pp. 94).

This observation can also be interpreted stochastically: Let $X_1,\dots,X_n$ be iid random variables and define $\boldsymbol{Y}_i = (Y_{1,i}, Y_{2,i}) := (-X_i, X_i)$, $1 \le i \le n$, and

$$\boldsymbol{X_{(n)}} := (\max_{1\le i\le n}(Y_{1,i}), \max_{1\le i\le n} Y_{2,i}).$$

Then

$$\boldsymbol{X_{(n)}} = \bigl(\max_{1\le i\le n}(-X_i), \max_{1\le i\le n} X_i\bigr) = \bigl(-X_{(1)}, X_{(n)}\bigr).$$

Note that $\boldsymbol{X_{(n)}} = \bigl(f(X_{(1)}), g(X_{(n)})\bigr)$ for $f(x)=-x$, $g(x)=x$. In view of the above remarks the copula of the distribution of $\boldsymbol{X_{(n)}}$ is given by

$$C_{\boldsymbol{X_{(n)}}}(u,v) = v - C_n(1-u,v) = C_\vartheta^{Cl}(u,v), \quad \vartheta = -1/n. \tag{4.5}$$

For further investigations, set $A_n := \{(u,v) \in (0,1)^2 : 1-(1-u)^{1/n} < v^{1/n}\}$ and $B_n := \{(u,v) \in (0,1)^2 : 1-(1-u)^{1/n} > v^{1/n}\}$. On these open sets, C_n is partially differentiable twice, and elementary calculation yields

$$D_1C_n(u,v) := \frac{\partial}{\partial u}C_n(u,v) = \begin{cases} (v^{1/n}+(1-u)^{1/n}-1)^{n-1}\cdot(1-u)^{(1/n)-1}, & (u,v)\in A_n, \\ 0, & (u,v)\in B_n, \end{cases} \tag{4.6}$$

$$D_2C_n(u,v) := \frac{\partial}{\partial v}C_n(u,v) = \begin{cases} 1-(v^{1/n}+(1-u)^{1/n}-1)^{n-1}\cdot v^{(1/n)-1}, & (u,v)\in A_n, \\ 1, & (u,v)\in B_n, \end{cases} \tag{4.7}$$

and

$$D_{12}C_n(u,v) = \frac{\partial^2}{\partial u \partial v} C_n(u,v) \tag{4.8}$$

$$= \begin{cases} \frac{n-1}{n}(v(1-u))^{(1/n)-1}(v^{1/n} + (1-u)^{1/n} - 1)^{n-2}, & (u,v) \in A_n, \\ 0, & (u,v) \in B_n. \end{cases} \tag{4.9}$$

We thus have the following (Riemann-integrable) probability density for C_n:

$$c_n(u,v) := \begin{cases} \frac{n-1}{n}\big(v(1-u)\big)^{1/n-1}\big(v^{1/n} + (1-u)^{1/n} - 1\big)^{n-2}, & (u,v) \in A_n, \\ 0, & \text{elsewhere.} \end{cases} \tag{4.10}$$

The next result has been known for a long time (Walsh, 1969, or David, 1981, p. 267), but it is interesting to see its derivation with the help of copulas. Additionally, its proof is simpler as, e. g., the ones in Reiss (1989, Section 4.2) or Galambos (1978, Theorem 2.9.1). From an intuitive point of view, it is clear that the minimum and maximum of n iid random variables should be asymptotically independent, i. e., we expect the limiting copula to be the independence copula.

4.1 Lemma
Let C_n be as in (4.3). Then

$$\lim_{n\to\infty} C_n(u,v) = \Pi(u,v) = uv \quad \textit{for all } u,v \in I.$$

Proof. For $v = 0$ or $u = 0$, $C_n(u,v) = 0 = uv$ as C_n is a copula; analogously, for $u = 1$ or $v = 1$ we have $C_n(1,v) = v = uv$ and $C_n(u,1) = u = uv$.
Let $(u,v) \in (0,1)^2$. Since $v^{1/n} \xrightarrow{n\to\infty} 1$ and $(1-u)^{1/n} \xrightarrow{n\to\infty} 1$, there exists a $N = N(u,v) \in \mathbb{N}$ such that $1-(1-u)^{1/n} < v^{1/n}$ for all $n \geq N$ so that we only need to consider $(u,v) \in A_n$, i. e.,

$$C_n(u,v) = v - \big(v^{1/n} + (1-u)^{1/n} - 1\big)^n.$$

Set $f(x) := \big(v^{1/x} + (1-u)^{1/x} - 1\big)^x$ for $x \in (0,\infty)$. Then

$$\ln f(x) = x \cdot \ln\big(v^{1/x} + (1-u)^{1/x} - 1\big) = \frac{\ln\big(v^{1/x} + (1-u)^{1/x} - 1\big)}{x^{-1}}$$

whose limit for $x \to \infty$ is of the form $0/0$. With the rule of de l'Hospital (e. g., Walter, 1992, 10.11) we have

$$\begin{aligned} \lim_{x\to\infty} \ln f(x) &= \lim_{x\to\infty} \frac{\big(-\ln v \cdot v^{1/x} - \ln(1-u)\cdot(1-u)^{1/x}\big)\, x^{-2}}{-x^{-2}\,(v^{1/x} + (1-u)^{1/x} - 1)} \\ &= \lim_{x\to\infty} \frac{v^{1/x}\ln v + (1-u)^{1/x}\ln(1-u)}{v^{1/x} + (1-u)^{1/x} - 1} = \ln v + \ln(1-u) = \ln\big(v(1-u)\big). \end{aligned}$$

Therefore,

$$\lim_{x\to\infty} f(x) = \exp\big(\ln(v(1-u))\big) = v(1-u),$$

and

$$\lim_{n\to\infty} C_n(u,v) = v - v(1-u) = uv = \Pi(u,v) \quad \text{for } (u,v) \in (0,1)^2.$$

□

4.2 Remark

The above proof is already short and non-technical in the sense that no other results and only elementary calculus are used. However, it is known for the Clayton copula that $\lim_{\vartheta\to 0} C_\vartheta^{Cl}(u,v) = uv$ for all $u, v \in (0,1)$ (e. g., Cook and Johnson, 1981). As mentioned above, $v - C_n(1-u,v) = C_{-1/n}^{Cl}(u,v)$ so that

$$\begin{aligned}\lim_{n\to\infty} C_n(u,v) &= v - \lim_{n\to\infty} C_{-1/n}^{Cl}(1-u,v) = v - \lim_{\vartheta\to 0} C_\vartheta^{Cl}(1-u,v)\\ &= v - (1-u)v = uv.\end{aligned}$$

This makes the proof even shorter. ■

Based on these findings we can start investigating some of the properties of the above defined min-max copula related to dependence. We will start with the calculation of Kendall's τ (see page 34) with the help of Theorem 3.3.

4.3 Theorem

Let $X_1, \ldots, X_n$, $n \in \mathbb{N}$, be iid random variables. Then Kendall's τ for $X_{(1)}$ and $X_{(n)}$ is given by

$$\tau_n(X_{(1)}, X_{(n)}) \equiv \tau_n = \frac{1}{2n-1}. \tag{4.11}$$

Proof. From (4.3) we know that the corresponding copula for $(X_{(1)}, X_{(n)})$ is given by C_n. Thus, with Theorem 3.3 we know that

$$\tau_n = 4 \cdot \iint_{I^2} C_n(u,v)\, \mathrm{d}C_n(u,v) - 1.$$

With (4.10) we have

$$\begin{aligned}\tau_n &= 4 \cdot \iint_{I^2} \frac{n-1}{n} \left[v - \left(v^{1/n} + (1-u)^{1/n} - 1\right)^n\right] \left(v(1-u)\right)^{(1/n)-1}\\ &\qquad\qquad \cdot \left(v^{1/n} + (1-u)^{1/n} - 1\right)^{n-2} \cdot \mathbb{1}_{A_n}\big((u,v)\big)\, \mathrm{d}u\, \mathrm{d}v - 1\\ &= 4\big(I_1(n) - I_2(n)\big) - 1\end{aligned}$$

with

$$I_1(n) := \iint_{I^2} \frac{n-1}{n} \cdot v^{1/n}(1-u)^{(1/n)-1}\left(v^{1/n} + (1-u)^{1/n} - 1\right)^{n-2} \cdot \mathbb{1}_{A_n}\big((u,v)\big)\, \mathrm{d}u\, \mathrm{d}v$$

and

$$I_2(n) := \iint\limits_{I^2} \frac{n-1}{n} \cdot v^{(1/n)-1}(1-u)^{(1/n)-1}\big(v^{1/n}+(1-u)^{1/n}-1\big)^{2n-2} \cdot \mathbb{1}_{A_n}\big((u,v)\big)\,\mathrm{d}u\,\mathrm{d}v.$$

The substitutions $s = v^{1/n}$ and $t = (1-u)^{1/n}$ yield for $I_1(n)$:

$$\begin{aligned}
I_1(n) &\overset{(*)}{=} \int\limits_0^1 \int\limits_{\left(1-(1-u)^{1/n}\right)^n}^1 \frac{n-1}{n} v^{1/n}(1-u)^{(1/n)-1}\big(v^{1/n}+(1-u)^{1/n}-1\big)^{n-2}\,\mathrm{d}v\,\mathrm{d}u \\
&= \int\limits_0^1 \int\limits_{\left(1-(1-u)^{1/n}\right)}^1 (n-1)s(1-u)^{(1/n)-1}\big(s+(1-u)^{1/n}-1\big)^{n-2} s^{n-1}\,\mathrm{d}s\,\mathrm{d}u \\
&= \int\limits_0^1 \int\limits_{1-t}^1 (n-1)ns^n(s+t-1)^{n-2}t^{n-1}t^{1-n}\,\mathrm{d}s\,\mathrm{d}t \\
&= (n-1)n \int\limits_0^1 \int\limits_{1-t}^1 s^n(s+t-1)^{n-2}\,\mathrm{d}s\,\mathrm{d}t \overset{(*)}{=} (n-1)n \int\limits_0^1 s^n \int\limits_{1-s}^1 (s+t-1)^{n-2}\,\mathrm{d}t\,\mathrm{d}s \\
&= n \int\limits_0^1 s^n \left[(s+t-1)^{n-1}\right]_{t=1-s}^{t=1}\,\mathrm{d}s = n \int\limits_0^1 s^{2n-1}\,\mathrm{d}s = \frac{1}{2}.
\end{aligned}$$

$(*)$ is valid by Fubini's theorem (e. g., Rogers and Williams, 2000a, II. 12).

With the same substitutions, $I_2(n)$ can be evaluated to

$$I_2(n) = \frac{n-1}{2(2n-1)}$$

so that

$$\tau_n = 4\left(\frac{1}{2} - \frac{n-1}{2(2n-1)}\right) - 1 = \frac{1}{2n-1}.$$

□

Analogously, we can derive the following result for Spearman's ρ (see p. 35) based on Theorem 3.5.

4.4 Theorem
Let $X_1, \dots, X_n$, $n \in \mathbb{N}$, be iid random variables. Then Spearman's ρ for $X_{(1)}$ and $X_{(n)}$ is given by

$$\rho_n(X_{(1)}, X_{(n)}) \equiv \rho_n = 3 - \frac{12n}{\binom{2n}{n}} \sum_{k=0}^{n} \frac{(-1)^k}{2n-k}\binom{2n}{n+k} + 12\frac{(n!)^3}{(3n)!}(-1)^n. \qquad (4.12)$$

Proof. With Theorem 3.5 and the sets A_n and B_n (see p. 44) we have

$$\begin{aligned}
\rho_n &= 12 \iint\limits_{I^2} C_n(u,v)\,\mathrm{d}u\,\mathrm{d}v - 3 \\
&= 12\Big\{\iint\limits_{A_n} \big[v - \big(v^{1/n} + (1-u)^{1/n} - 1\big)^n\big]\,\mathrm{d}u\,\mathrm{d}v + \iint\limits_{B_n} v\,\mathrm{d}u\,\mathrm{d}v\Big\} - 3 \\
&= 12\Big\{\iint\limits_{A_n\cup B_n} v\,\mathrm{d}u\,\mathrm{d}v - \underbrace{\iint\limits_{A_n} \big(v^{1/n} + (1-u)^{1/n} - 1\big)^n\,\mathrm{d}u\,\mathrm{d}v}_{=:I(n)}\Big\} - 3.
\end{aligned}$$

It is easy to see that the first integral yields $1/2$. With the substitutions $s = (1-u)^{1/n}$ and $t = v^{1/n}$, the integral $I(n)$ can be calculated as follows:

$$\begin{aligned}
I(n) &= \int_0^1 \int_0^{1-(1-v^{1/n})^n} \big(v^{1/n} + (1-u)^{1/n} - 1\big)^n\,\mathrm{d}u\,\mathrm{d}v = n\int_0^1 \int_{1-v^{1/n}}^1 \big(v^{1/n} + s - 1\big)^n s^{n-1}\,\mathrm{d}s\,\mathrm{d}v \\
&= n^2 \int_0^1 \int_{1-t}^1 (t+s-1)^n s^{n-1} t^{n-1}\,\mathrm{d}s\,\mathrm{d}t = n^2 \int_0^1 t^{n-1} \int_{1-t}^1 (t+s-1)^n s^{n-1}\,\mathrm{d}s\,\mathrm{d}t \\
&= n^2 \sum_{k=0}^n \binom{n}{k} \int_0^1 (t-1)^k t^{n-1} \int_{1-t}^1 s^{n-k} s^{n-1}\,\mathrm{d}s\,\mathrm{d}t \\
&= n^2 \sum_{k=0}^n \binom{n}{k} \int_0^1 (t-1)^k t^{n-1} \left[\frac{1}{2n-k}\big(1-(1-t)^{2n-k}\big)\right] \mathrm{d}t \\
&= n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} \left[\int_0^1 t^{n-1}(1-t)^k\,\mathrm{d}t - \int_0^1 t^{n-1}(1-t)^{2n}\,\mathrm{d}t\right] \\
&= n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} \left[B(n,k+1) - B(n,2n+1)\right],
\end{aligned}$$

where B denotes the Beta function $B(x,y) := \int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t$. With Gould (1972, formula (1.43)) the second sum is reduced to

$$\begin{aligned}
B(n,2n+1)\,n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} &= \frac{(n-1)!\,(2n)!}{(3n)!} n^2 \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{2n-k} = \frac{n\,(2n)!\,n!}{(3n)!} \frac{(-1)^n}{n\binom{2n}{n}} \\
&= (-1)^n \frac{(n!)^3}{(3n)!}.
\end{aligned}$$

Using $B(n,k+1) = (n-1)!\,k!/(n+k)!$ for the first sum proves the result. $\square$

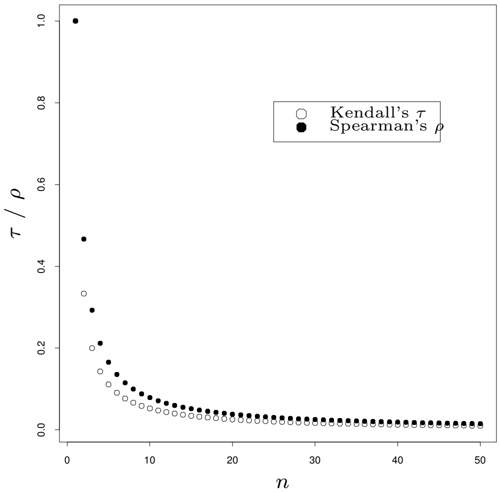

FIGURE 4.2. Comparison of Kendall's τ and Spearman's ρ for $(X_{(1)}, X_{(n)})$.

n	ρ_n	τ_n	ρ_n/τ_n
2	$0.4\overline{6}$	$0.\overline{3}$	1.4
3	0.2929	0.2	1.4643
5	0.1677	$0.\overline{1}$	1.4897
10	0.0788	0.0526	1.4978
50	0.0151	0.0101	1.4999
100	0.0075	0.0050	1.499998

TABLE 4.1. Values for ρ_n, τ_n and ρ_n/τ_n.

Figure 4.2 and Table 4.1 reveal the fact that τ_n is smaller than ρ_n which is not surprising in virtue of the results from Capéraà and Genest (1993) who showed that Spearman's ρ is generally larger than Kendall's τ for positively dependent random variables in the sense of left-tail decrease and right-tail increase (see Theorem 3.8). To be precise, they showed that $\rho \geq \tau \geq 0$ for two random variables X and Y whenever one of X or Y is simultaneously left-tail decreasing (LTD) and right tail-increasing (RTI) in the other variable (cf. Definition 3.6). Additionally, ρ_n/τ_n seems to converge to $3/2$; however, this has not yet been proven due to the unwieldy expression for ρ_n in equation (4.12). The convergence to the value $3/2$ of ρ_n/τ_n must be seen in a more general context: Nelsen (1991) showed this relation for some copula families and posed the (yet unresolved) question, if it is valid for all families of copulas $\{C_\vartheta(u,v)\}$ indexed by a (possibly multidimensional) parameter ϑ such that $C_{\vartheta_0} = \Pi$ and C_ϑ is a continuous function of ϑ at ϑ_0.

As it is not easy to see that $\rho_n \geq \tau_n$ in the case of $(X_{(1)}, X_{(n)})$ by the above results, we will make use of Definition 3.14 and Theorem 3.15 to show that $X_{(1)}$ and $X_{(n)}$ are positively likelihood ratio dependent so that this part of the the assertion follows.

4.5 Theorem
Let τ_n and ρ_n denote Kendall's τ and Spearman's ρ for $(X_{(1)}, X_{(n)})$ in the above situation. Then

$$3\tau_n \geq \rho_n \geq \tau_n > 0.$$

Proof. $\tau_n > 0$ follows trivially from Theorem 4.3. If we show that $\mathrm{PLR}(X,Y)$, the other inequalities follow from Theorem 3.8, Theorem 3.13 and Theorem 3.15. In virtue of Definition 3.14, let $u_1, u_2, v_1, v_2 \in (0,1)$ and $u_1 \leq u_2$, $v_1 \leq v_2$. It is clear from (4.10) that $c_n(x,y) = 0$ for $(x,y) \in \overline{\mathbb{R}} \setminus \{(0,1)^2\}$. We have to show

$$c_n(u_1,v_1)c_n(u_2,v_2) \geq c_n(u_1,v_2)c_n(u_2,v_1). \tag{4.13}$$

Recall that c_n is given by

$$c_n(u,v) = \begin{cases} \frac{n-1}{n}\big(v(1-u)\big)^{1/n-1}\big(v^{1/n}+(1-u)^{1/n}-1\big)^{n-2}, & (u,v)\in A_n,\\ 0, & \text{elsewhere,}\end{cases}$$

with $A_n = \{(u,v)\in(0,1)^2 : 1-(1-u)^{1/n} < v^{1/n}\}$. We may therefore assume (u_1,v_2), $(u_2,v_1)\in A_n$, for if this is not the case, the right hand side of (4.13) is zero and the inequality trivially true. As $u_1\le u_2$ and $v_1\le v_2$, we must also have $(u_1,v_1),(u_2,v_2)\in A_n$. But then (4.13) is equivalent to

$$\begin{aligned}
&\big(v_1(1-u_1)v_2(1-u_2)\big)^{1/n-1}\Big[\big(v_1^{1/n}+(1-u_1)^{1/n}-1\big)\big(v_2^{1/n}+(1-u_2)^{1/n}-1\big)\Big]^n\\
&\qquad\ge \big(v_1(1-u_2)v_2(1-u_1)\big)^{1/n-1}\Big[\big(v_1^{1/n}+(1-u_2)^{1/n}-1\big)\big(v_2^{1/n}+(1-u_1)^{1/n}-1\big)\Big]^n\\
\Leftrightarrow\ &\big(v_1^{1/n}+(1-u_1)^{1/n}-1\big)\big(v_2^{1/n}+(1-u_2)^{1/n}-1\big)\\
&\qquad\ge \big(v_1^{1/n}+(1-u_2)^{1/n}-1\big)\big(v_2^{1/n}+(1-u_1)^{1/n}-1\big).
\end{aligned}$$

Setting $x_1 := u_1^{1/n}$, $x_2 := u_2^{1/n}$, $y_1 := 1-(1-v_1)^{1/n}$ and $y_2 := 1-(1-v_2)^{1/n}$ so that $x_1\le x_2$ and $y_1\le y_2$, this inequality is equivalent to

$$\begin{aligned}
&(y_1-x_1)(y_2-x_2)\ge(y_1-x_2)(y_2-x_1)\\
\Leftrightarrow\quad & -y_1x_2-x_1y_2\ge -y_1x_1-x_2y_2\\
\Leftrightarrow\quad & y_2(x_2-x_1)-y_1(x_2-y_1)\ge 0\\
\Leftrightarrow\quad & (y_2-y_1)(x_2-x_1)\ge 0
\end{aligned}$$

which is clearly the case. The theorem is proved.

□

4.6 Remark

The question may arise if it is not possible to derive similar results for $(X_{(r)},X_{(s)})$, $(r,s)\neq(1,n)$. However, for the derivation of C_n it was of fundamental importance to be able to determine the generalized inverses of F_1 and F_n which could be done via (4.2). It does not seem to be possible to determine the generalized inverses of F_s for $1<s<n$ analytically. This is an unsolved problem and prevents the copula approach from being used for $(X_{(r)},X_{(s)})$. ■

4.2 The Copula of Brownian Motion and Its Supremum

In this section we will shortly derive the copula of (standard) Brownian motion $(B_t)_{t\geq 0}$ and its supremum process $(S_t)_{t\geq 0}$ with $S_t := \sup_{0\leq s\leq t} B_t$ for $t \geq 0$. Note that due to $B_0 = 0$ we have $S_t \geq 0$ for all $t \geq 0$.

Please refer to Rogers and Williams (2000a, Chapter I) or Karatzas and Shreve (2000, Chapter 2) for a detailed introduction to Brownian motion.

We will be interested in the derivation of the copula C_t^{BSup} of (B_t, S_t), $t > 0$. At this stage we do not have any clues about its form; particularly, we do not have reason *not* to expect the dependence structure to depend on the time-parameter t.

As $B_t \sim \mathsf{N}(0,t)$, we have $F_{B_t}(x) = \Phi(x/\sqrt{t})$, $x \in \mathbb{R}$, such that $F_{B_t}^{-1}(u) = \sqrt{t}\cdot\Phi^{-1}(u)$, $0 < u < 1$, where Φ denotes the cumulative distribution function of a standard normal random variable. For the distribution function F_{S_t} of S_t, we have (e. g., Rogers and Williams, 2000a, Corollary I.13.3 and the remark thereafter):

$$F_{S_t}(y) = \mathsf{P}(S_t \leq y) = 1 - 2\cdot\mathsf{P}(B_t \geq y) = 2\cdot\Phi(y/\sqrt{t}) - 1, \quad y \geq 0, \tag{4.14}$$

such that

$$F_{S_t}^{-1}(v) = \sqrt{t}\cdot\Phi^{-1}\left(\frac{v+1}{2}\right), \quad v \in (0,1). \tag{4.15}$$

Let $H_t(x,y) := \mathsf{P}(B_t \leq x, S_t \leq y)$, $x, y \in \mathbb{R}$, denote the joint distribution function of (B_t, S_t), $t > 0$. From Rogers and Williams (2000a, Corollary I.13.3) (stating the reflection principle) and using continuity of B_t and S_t, we have for $x \leq y$:

$$\begin{aligned} H_t(x,y) &= \mathsf{P}(B_t \leq x) - \mathsf{P}(B_t \leq x, S_t \geq y) \\ &= \Phi(x/\sqrt{t}) - \mathsf{P}\big(B_t \leq y - (y-x), S_t \geq y\big) \\ &= \Phi(x/\sqrt{t}) - \mathsf{P}(B_t \geq 2y - x) \\ &= \Phi(x/\sqrt{t}) - \left(1 - \left\{\Phi\left(\frac{2y-x}{\sqrt{t}}\right)\right\}\right) \\ &= \Phi(x/\sqrt{t}) - \Phi\left(\frac{x-2y}{\sqrt{t}}\right) \end{aligned} \tag{4.16}$$

for the joint distribution of (B_t, S_t). For the case $x > y$ we obtain

$$H_t(x,y) = \mathsf{P}(S_t \leq y) = 2\cdot\Phi(y/\sqrt{t}) - 1. \tag{4.17}$$

Then the unique[1] copula C_t^{BSup} is given by

$$C_t^{\text{BSup}}(u,v) = H_t\big(F_{B_t}^{-1}(u), F_{S_t}^{-1}(v)\big), \quad u, v \in (0,1). \tag{4.18}$$

This yields

[1] F_{B_t} and F_{S_t} are continuous.

a) for the case $\boxed{u \leq (v+1)/2}$ (which comes from $F_{B_t}^{-1}(u) \leq F_{S_t}^{-1}(v) \Leftrightarrow \sqrt{t} \cdot \Phi^{-1}(u) \leq \sqrt{t} \cdot \Phi^{-1}\big((v+1)/2\big)$):

$$\begin{aligned} C_t^{\mathrm{BSup}}(u,v) &= \Phi\left[\frac{1}{\sqrt{t}} \cdot \sqrt{t} \cdot \Phi^{-1}(u)\right] \\ &\qquad - \Phi\left[\frac{1}{\sqrt{t}}\left(\sqrt{t} \cdot \Phi^{-1}(u) - 2 \cdot \sqrt{t} \cdot \Phi^{-1}\big((u+1)/2\big)\right)\right] \\ &= u - \Phi\big(\Phi^{-1}(u) - 2 \cdot \Phi^{-1}\big((v+1)/2\big)\big), \end{aligned} \tag{4.19}$$

b) and for the case $\boxed{u > (v+1)/2}$:

$$C_t^{\mathrm{BSup}}(u,v) = 2 \cdot \Phi\left(\sqrt{t} \cdot \Phi^{-1}\big((v+1)/2\big)\frac{1}{\sqrt{t}}\right) - 1 = v. \tag{4.20}$$

We see that neither of these expressions depends on the point of time t. We can thus say that the dependence between Brownian motion and its supremum is entirely determined by the marginal distributions and the time-independent copula

$$C^{\mathrm{BSup}}(u,v) = \begin{cases} u - \Phi\left(\Phi^{-1}(u) - 2 \cdot \Phi^{-1}\big((v+1)/2\big)\right), & u \leq (v+1)/2, \\ v, & u > (v+1)/2. \end{cases} \tag{4.21}$$

What could be the use of these findings? One possible application could be found in mathematical finance where geometric Brownian motion is a stochastic process (still) used to model stock prices. Now, if this model were adequate, we could calculate the supremum process of the stock price process and transform both processes by the probability integral transform. For the resulting two uniform processes, we could develop a test for homogeneity w. r. t. time. Under the null hypothesis that the underlying model is geometric Brownian motion, we should not be able to detect dependencies on time; conservative tests could be developed by testing for a trend in the data (as a trend is a special case of time dependence). We will leave these aspects to future research.

5

Stochastic Processes And Copulas

In this chapter we will investigate the relationship between copulas and univariate stochastic processes. Before we explore these relationships we first give the basic definitions and properties used in the following sections concerning stochastic processes. For a detailed introduction, please refer to the "standard literature", e. g., Ross (1996) or Todorovic (1992) for a general account to stochastic processes, Karatzas and Shreve (2000) or Revuz and Yor (1991) for an emphasis on Brownian motion and stochastic calculus; Rogers and Williams (2000a,b) may deal as a general account to diffusions, Markov processes and martingales.

The basic idea is that many properties of stochastic processes can be characterized by their finite-dimensional distributions so that copulas can be used for analysis. However, many concepts in the theory of stochastic processes are stronger than the finite-dimensional distribution approach; examples are martingales and stochastic differential equations which rely heavily on pathwise structures. Nevertheless, it is possible to derive the copula structure in some special cases.

5.1 Stochastic Processes

We will always assume an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a parameter set $T \subset \mathbb{R}$ (usually $T = [0, \infty)$).

5.1 Definition (stochastic process)
*A n-*dimensional stochastic process *is a collection of random variables $(X_t)_{t \in T}$ defined on $(\Omega, \mathcal{F}, \mathsf{P})$ and taking values in $\mathbb{R}^n$, $n \in \mathbb{N}$. Symbolically, $X_t : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}^n)$.*
If the "n-dimensional" is omitted, we mean a one-dimensional process.

5.2 Definition (finite-dimensional distribution)
For a stochastic process $X = (X_t)_{t \in T}$ the set of finite-dimensional distributions *is given by*

$$\mathcal{D}_X := \{\mathsf{P}^{X_{t_1}, \dots, X_{t_n}} : t_1 < \cdots < t_n,\ n \in \mathbb{N},\ t_i \in T\}. \tag{5.1}$$

From a stochastic point of view we are often only interested in the finite-dimensional distributions of processes so that we declare two processes with the same finite-dimensional distributions to be equivalent:

5.3 Definition (stochastic equivalence)
We will call two processes $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$ stochastically equivalent *and write $X \stackrel{d}{=} Y$ if their finite-dimensional distributions are the same, that is, $\mathcal{D}_X = \mathcal{D}_Y$.*

It is important to realize that this is only one version of a number of equivalence concepts for stochastic processes. Others are *modification* and *indistinguishability* which both are stronger than the finite-dimensional distribution concept (cf. Karatzas and Shreve, 2000, p. 2). However, as copulas are distribution functions, we will focus on the finite-dimensional distributions.

A valuable source for the relationship between these concepts and sufficient conditions for the existence of certain modified versions of a given process with desired properties (such as continuity or differentiability) is Cramér and Leadbetter (1967).

A certain flow of information is revealed by the concept of filtrations and adapted processes:

5.4 Definition (filtration, adapted process)
A filtration *on $(\Omega, \mathcal{F})$ is a family $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$ such that for $s < t$ we have $\mathcal{F}_s \subset \mathcal{F}_t$. We set $\mathcal{F}_\infty := \sigma(\bigcup_{s\in T} \mathcal{F}_s)$. A measurable space $(\Omega, \mathcal{F})$ endowed with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ is said to be a* filtered space *and will sometimes be denoted by $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$.*

A process $X = (X_t)_{t\in T}$ is said to be adapted *to $\mathbb{F}$ if X_t is $\mathcal{F}_t$-measurable for all $t \in T$.*

With a filtration, one can associate two other filtrations by setting

$$\mathcal{F}_{t-} := \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right) \quad \text{and} \quad \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s.$$

The σ-algebra $\mathcal{F}_{0-}$ is not defined and, by convention, we set $\mathcal{F}_{0-} = \mathcal{F}_0$.

5.5 Definition (right-continuous filtration)
The filtration $\mathbb{F}$ is right-continuous *if $\mathcal{F}_t = \mathcal{F}_{t+}$ for all $t \in T$.*

We will need right-continuity of a filtration for the "usual conditions" which will be used for technical reasons in Section 5.5.

5.6 Definition (usual conditions)
A filtration $(\mathcal{F}_t)$ is said to satisfy the usual conditions *if it is right continuous and $\mathcal{F}_0$ contains all the P-negligible events in $\mathcal{F}$.*

5.7 Definition (stopping time)
Let $\mathbb{F}$ *be a filtration on* $(\Omega, \mathcal{F}, \mathsf{P})$ *and assume* $T = [0, \infty)$*. A random variable* $\tau : \Omega \to [0, \infty)$ *is called* stopping time *w. r. t.* $\mathbb{F}$ *if*

$$\{\tau \leq t\} \in \mathcal{F}_t \qquad \text{for all } t \geq 0. \tag{5.2}$$

The stopping time σ-algebra *is defined to be*

$$\sigma_\tau := \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \text{for all } t \in T\}. \tag{5.3}$$

Stopping times play a prominent role in the technique of localizing a stochastic process (see Revuz and Yor, 1991, pp. 117). I. e., many theorems in the context of stochastic analysis are valid if some conditions are met locally by the process, meaning essentially that the stopped process has the desired properties.

5.8 Definition (stopped process)
For a stopping time τ *and a process* $X = (X_t)_{t \in T}$*, the* stopped process $X^\tau = (X_t^\tau)_{t \in T}$ *is defined as*

$$X_t^\tau(\omega) := X_{t \wedge \tau(\omega)}(\omega) = \begin{cases} X_t(\omega), & t \leq \tau(\omega), \\ X_{\tau(\omega)}(\omega), & t > \tau(\omega). \end{cases} \tag{5.4}$$

A class of widely studied processes is the class of martingales. They may be thought of as representing "fair games" as the expected value in the future based on "the present's point of view" is exactly the current value of the process.

5.9 Definition ((sub-/super-) martingale)
A real-valued process $X = (X_t)_{t \in T}$ *adapted to* $(\mathcal{F}_t)_{t \in T}$ *is a* submartingale *w. r. t.* $(\mathcal{F}_t)$ *if*

a) $E(X_t^+) < \infty$ *for every* $t \in T$,

b) $E(X_t \mid \mathcal{F}_s) \geq X_s$ *almost surely for all* $s < t$.

X_t^+ *denotes the positive part of* X_t*, i. e.,* $X_t^+ = \max(X_t, 0)$.
A process X *such that* $-X$ *is a submartingale is called a* supermartingale *and a process which is both a sub- and a supermartingale is a* martingale.

5.2 General Properties

Although the main use of copulas is presently seen in modelling "vertical" or "spatial" dependencies in the sense of the dependence between n random variables *at the same time*, it is (in view of the following classic theorem) also possible and beneficial to use copulas when constructing stochastic processes. The basic idea is that, under certain regularity conditions, the finite-dimensional distributions of a stochastic process determine the probabilistic behaviour, i. e., the distributional aspects, completely.

In the following, we will assume $T \subset \mathbb{R}$ to be a parameter set. It often has the interpretation of the time space for the process.

5.10 Theorem (Kolmogorov)
Let $\{F_{t_1,\ldots,t_n} : t_i \in T,\ n \in \mathbb{N},\ 1 \leq i \leq n,\ t_1 < t_2 < \cdots < t_n\}$ be a given family of finite-dimensional distribution functions, satisfying the consistency condition

$$\lim_{x_k \nearrow \infty} F_{t_1,\ldots,t_n}(x_1,\ldots,x_n) = F_{t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_n}(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n) \text{ for all } 1 \leq k \leq n, \tag{5.5}$$

for all $n \in \mathbb{N}$, $x_1,\ldots,x_n \in \mathbb{R}$, $t_1,\ldots,t_n \in T$. Then there exist a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$ and a stochastic process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$ such that

$$\mathsf{P}(X_{t_1} \leq x_1,\ldots,X_{t_n} \leq x_n) = F_{t_1,\ldots,t_n}(x_1,\ldots,x_n) \tag{5.6}$$

for all $x_i \in \mathbb{R}$, $t_i \in T$, $1 \leq i \leq n$, $n \in \mathbb{N}$, and X_t is $\mathcal{F}_t$-measurable for all $t \in T$.

For a proof, see Shiryaev (1996, Theorem II.9.1, pp. 246).

It is now easy to see that a collection of copulas and marginal distributions also define a stochastic process.

5.11 Corollary
Let $\mathcal{C} := \{C_{t_1,\ldots,t_n} : t_i \in T, t_1 < \cdots < t_n, n \in \mathbb{N}\}$ be a collection of copulas satisfying the consistency condition

$$\lim_{u_k \nearrow 1} C_{t_1,\ldots,t_n}(u_1,\ldots,u_n) = C_{t_1,\ldots,t_{k-1},t_{k+1},\ldots,t_n}(u_1,\ldots,u_{k-1},u_{k+1},\ldots,u_n) \tag{5.7}$$

for all $u_i \in (0,1)$, $1 \leq k \leq n$, and $\mathcal{D} = \{F_t : t \in T\}$ a collection of one-dimensional distribution functions.

Then there exist a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$ and a stochastic process $X = (X_t)_{t\in T}$ on $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$ such that

$$\mathsf{P}(X_{t_1} \leq x_1,\ldots,X_{t_n} \leq x_n) = C_{t_1,\ldots,t_n}\big(F_{t_1}(x_1),\ldots,F_{t_n}(x_n)\big) \tag{5.8}$$

for all $x_i \in \mathbb{R}$, $t_i \in T$, $1 \leq i \leq n$, $n \in \mathbb{N}$, and X_t is $\mathcal{F}_t$-measurable for all $t \in T$.

For a given process $X = (X_t)_{t\in T}$, we will denote the associated family of copulas by $C\langle X\rangle$. In the same sense, $C\langle X_{t_1},\ldots,X_{t_n}\rangle \equiv C^X_{t_1,\ldots,t_n}$ denotes a copula of $(X_{t_1},\ldots,X_{t_n})$.

Proof. The assertion follows directly from Theorem 5.10 and Theorem 2.13. □

So, from a copula point of view, it is possible to interpret a stochastic process as a process of uniform variables transformed by the marginals, i. e., we have a process $(U_t)_{t\in T}$ where each U_t is $\mathsf{U}(0,1)$-distributed with an intertemporal dependence structure defined by a family of copulas. We will call such a process a *uniform process*. This process is then transformed by the family of marginals $\mathcal{D}$ via a quantile transformation, i. e., $X_t := F_t^{-1}(U_t)$, $t \in T$.

5.12 Example

a) **Independence**
If we take $C_{t_1,\dots,t_n}(u_1,\dots,u_n) := \prod_{i=1}^n u_i$ (the independence copula family on I^n) we see that condition (5.7) is true. Therefore, for any family $\mathcal{D} = \{F_t : t \in T\}$ of one-dimensional distribution functions, we have $\mathsf{P}(X_{t_1} \le x_1,\dots,X_{t_n} \le x_n) = \prod_{i=1}^n F_{t_i}(x_i)$ so that the process $(X_t)_{t\in T}$ (which exists due to Corollary 5.11) consists of independent random variables. This is exactly Corollary II.9.1 of Shiryaev (1996, p. 247). The dependence structure is trivially homogeneous (in time), i. e., it does not depend on the t_i's.

b) **Archimedean Family**
Let $\mathcal{C} := \{C_{t_1,\dots,t_n}(u_1,\dots,u_n) := \varphi^{[-1]}\bigl(\sum_{i=1}^n \varphi(u_i)\bigr)\}$ be a set of Archimedean copulas with generator φ. $\varphi^{[-1]}$ denotes the generalized inverse of φ (see Definition A.1). Note that this is the most general case to ensure that $\mathcal{C}$ induces a stochastic process based on an Archimedean family as it is not possible to use different generators φ_1 and φ_2 due to the consistency condition (5.7). Then

$$\begin{aligned}\lim_{u_k \nearrow 1} C_{t_1,\dots,t_n}(u_1,\dots,u_n) &= \varphi^{[-1]}\Bigl(\sum_{\substack{i=1\\ i\neq k}}^n \varphi(u_i)\Bigr)\\ &= C_{t_1,\dots,t_{k-1},t_{k+1},\dots,t_k}(u_1,\dots,u_{k-1},u_{k+1},\dots,u_n),\end{aligned}$$

which is due to the fact that φ is continuous, strictly decreasing and convex with $\varphi(1) = 0$. Note that $\varphi^{[-1]}$ is also continuous. Therefore, a uniform process (U_t) exists with

$$\mathsf{P}(U_{t_1} \le u_1,\dots,U_{t_n} \le u_n) = C_{t_1,\dots,t_n}(u_1,\dots,u_n).$$

■

5.13 Remark

There is little overlap between the usual concepts of stochastic processes (Markov processes, martingales, time series analysis or diffusion processes) and the concept of modelling processes via copulas. For example, time series analysis heavily relies on the concept of autocorrelations whereas it is of great importance to see that the copula approach tries to capture the *full* intertemporal dependence structure in the data. However, Darsow, Nguyen, and Olsen (1992) managed to relate Markov processes to copulas and even to derive necessary and sufficient conditions for a process being Markov based on its copula family. We will come to this interesting point later. ■

5.2.1 Strict Stationarity

Let us take a look at strict stationarity of a process which is a well known concept although not as widely used as weak stationarity or stationarity of the second kind. Before commenting on this fact let us give the definition.

5.14 Definition (strict stationarity)
Let $X = (X_t)_{t\in T}$ be a real-valued stochastic process on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Then X is said to be strictly stationary*, if for any $n \in \mathbb{N}$, $\{t_1, \ldots, t_n\} \subset T$ and each $h > 0$ with $\{t_1 + h, \ldots, t_n + h\} \subset T$, we have*

$$\mathsf{P}(X_{t_1+h} \le x_1, \ldots, X_{t_n+h} \le x_n) = \mathsf{P}(X_{t_1} \le x_1, \ldots, X_{t_n} \le x_n) \tag{5.9}$$

for all $x_1, \ldots, x_n \in \mathbb{R}$.

5.15 Remark
It is clear (set $n = 1$ in the above definition) that all marginal distributions must be the same, i. e., $X_t \sim F$ for all $t \in T$ for some one-dimensional distribution function F. ■

As an immediate consequence we can give the following equivalent definition of a strictly stationary process:

5.16 Corollary
A real-valued stochastic process $(X_t)_{t\in T}$ is strictly stationary iff

- *$X_t \sim F$ for all $t \in T$ for some distribution function F and*
- *for all $n \in \mathbb{N}$ and $h > 0$ with $t_1, \ldots, t_n, t_1 + h, \ldots, t_n + h \in T$, we have*

$$C\langle X_{t_1}, \ldots, X_{t_n}\rangle = C\langle X_{t_1+h}, \ldots, X_{t_n+h}\rangle \text{ on } \operatorname{ran}(F), \text{ i. e.,}$$

$$C_{t_1,\ldots,t_n}(u_1, \ldots, u_n) = C_{t_1+h,\ldots,t_n+h}(u_1, \ldots, u_n) \text{ for all } u_1, \ldots, u_n \in \operatorname{ran}(F).$$

The processes in Example 5.12 clearly yield strictly stationary processes for $X_t \sim F$ for some arbitrary distribution function F.

5.2.2 Equivalence of Stochastic Processes

We already raised the question when processes are equivalent from a stochastic point of view. There are a few equivalence concepts providing possible answers to this question, such as stochastic equivalence in the wide sense or indistinguishability. As stochastic equivalence relies on the finite-dimensional distributions of a process, it should be related to copulas.

We will assume $(X_t)_{t\in T}$ and $(Y_t)_{t\in T}$ to be stochastic processes on the same common probability space $(\Omega, \mathcal{F}, \mathsf{P})$ with state space $(S, \mathcal{S})$ and parameter space T.

The definition of stochastic equivalence was given in Definition 5.3. It means that

$$\mathsf{P}(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n) = \mathsf{P}(Y_{t_1} \in B_1, \ldots, Y_{t_n} \in B_n) \tag{5.10}$$

for all $n \in \mathbb{N}$, $\{t_1, \ldots, t_n\} \subset T$ and $B_1, \ldots, B_n \in \mathcal{S}$.

5.17 Remark
If $(S, \mathcal{S}) = (\mathbb{R}^1, \mathcal{B}^1)$, (5.10) is equivalent to

$$\mathsf{P}(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n) = \mathsf{P}(Y_{t_1} \le x_1, \dots, Y_{t_n} \le x_n) \tag{5.11}$$

for all $n \in \mathbb{N}$, $x_i \in \mathbb{R}$, $1 \le i \le n$.

A fortiori, X_t and Y_t must have the same distribution function F_t for all $t \in T$. ■

If $C^X_{t_1,\dots,t_n}$ and $C^Y_{t_1,\dots,t_n}$ denote the copulas of $(X_{t_1}, \dots, X_{t_n})$ and $(Y_{t_1}, \dots, Y_{t_n})$, respectively, we can give an equivalent characterization of stochastically equivalent processes:

5.18 Theorem
Let (X_t) and (Y_t) be real-valued stochastic processes on $(\Omega, \mathcal{F}, \mathsf{P})$ with $X_t \sim F_t$, $Y_t \sim G_t$, $t \in T$, where F_t and G_t denote the distribution functions of X_t and Y_t. For $t_1 < \dots < t_n$, $n \in \mathbb{N}$, set $C^X_{t_1,\dots,t_n} = C\langle X_{t_1}, \dots, X_{t_n}\rangle$ and $C^Y_{t_1,\dots,t_n} = C\langle Y_{t_1}, \dots, Y_{t_n}\rangle$.
Then, (X_t) and (Y_t) are stochastically equivalent if and only if

a) $F_t = G_t$ for all $t \in T$,

b) $C^X_{t_1,\dots,t_n}(u_1, \dots, u_n) = C^Y_{t_1,\dots,t_n}(u_1, \dots, u_n)$ for all $n \in \mathbb{N}$, $u_i \in \operatorname{ran}(F_{t_i})$, $1 \le i \le n$.

Proof. For the "only if"-part, Remark 5.17 yields $F_t = G_t$ for all $t \in T$. Further, assume $(u_1, \dots, u_n) \in \operatorname{ran}(F_{t_1}) \times \dots \times \operatorname{ran}(F_{t_n})$ so that there exists $(x_1, \dots, x_n) \in \mathbb{R}^n$ such that $(u_1, \dots, u_n) = \big(F_{t_1}(x_1), \dots, F_{t_n}(x_n)\big)$. Thus,

$$\begin{aligned}
C^X_{t_1,\dots,t_n}(u_1, \dots, u_n) &= C^X_{t_1,\dots,t_n}\big(F_{t_1}(x_1), \dots, F_{t_n}(x_n)\big) \\
&= \mathsf{P}(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n) \overset{\text{equivalence}}{=} \mathsf{P}(Y_{t_1} \le x_1, \dots, Y_{t_n} \le x_n) \\
&= C^Y_{t_1,\dots,t_n}\big(G_{t_1}(x_1), \dots, G_{t_n}(x_n)\big) \overset{F_t = G_t}{=} C^Y_{t_1,\dots,t_n}\big(F_{t_1}(x_1), \dots, F_{t_n}(x_n)\big) \\
&= C^Y_{t_1,\dots,t_n}(u_1, \dots, u_n).
\end{aligned}$$

For the "if"-part, we have

$$\begin{aligned}
\mathsf{P}(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n) &= C^X_{t_1,\dots,t_n}\big(F_{t_1}(x_1), \dots, F_{t_n}(x_n)\big) \\
&\overset{a),b)}{=} C^Y_{t_1,\dots,t_n}\big(G_{t_1}(x_1), \dots, G_{t_n}(x_n)\big) = \mathsf{P}(Y_{t_1} \le x_1, \dots, Y_{t_n} \le x_n).
\end{aligned}$$

□

It should be clear that stronger concepts of equivalence do not relate to copulas in the same fashion: For example, (Y_t) is a *modification* of (X_t) if $\mathsf{P}(X_t = Y_t) = 1$ for all $t \in T$, but there is no connection to copulas for this relation. The same is true for *indistinguishability*, i. e., $\mathsf{P}(X_t = Y_t \text{ for all } t \in T) = 1$, which is the strongest form of equivalence (almost all sample paths coincide). For details, cf. Todorovic (1992, pp. 5).

5.2.3 Symmetric Processes

In this section we will use the results of Section 2.4 to characterize symmetric stochastic processes. Let us first give a slightly more general definition of a symmetric process. Let $T = [0, \infty)$ be the parameter space (w. l. o. g.).

5.19 Definition
Let $X = (X_t)_{t\in T}$ be a real-valued stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$ and $\mu_t \equiv \mu(t) : T \to \mathbb{R}$ a deterministic function. We will say that the process X is symmetric about μ_t for all $t \in T$ if

$$(X_t - \mu_t)_{t\in T} \stackrel{d}{=} (\mu_t - X_t)_{t\in T}, \tag{5.12}$$

i. e., all finite-dimensional distributions coincide.

5.20 Remark
The usual definition is the case $\mu(t) = 0$ for all t. It is easy to see that (X_t) is symmetric about (μ_t) iff $(X_t) = (\mu_t + Y_t)$ and (Y_t) is symmetric about 0 (as $(X_t - \mu_t) \stackrel{d}{=} (\mu_t - X_t) = -(X_t - \mu_t)$). $\mu(t)$ is usually called the "drift" of the process.

We can therefore assume $\mu(t) = 0$ for all t and immediately give a characterization of a symmetric process. ■

5.21 Theorem
Let $X = (X_t)_{t\in T}$ be a stochastic process, $X_t \sim F_t$, and denote the copulas by $C_{t_1,\dots,t_n} = C\langle X_{t_1}, \dots, X_{t_n}\rangle$ for all $t_1 < \dots < t_n$, $n \in \mathbb{N}$.
Then X is symmetric about 0 iff

a) $F_t(x) = \overline{F}_t(-x)$ for all $x \in \mathbb{R}$ ("marginals are symmetric"), and

b) $C_{t_1,\dots,t_n}(u_1, \dots, u_n) = \widehat{C}_{t_1,\dots,t_n}(u_1, \dots, u_n)$ for all $t_1 < \dots < t_n$, $t_i \in T$, $u_i \in \operatorname{ran}(F_{t_i})$, $1 \le i \le n$, $n \in \mathbb{N}$ ("all copulas are radially symmetric").

Proof. We must have $(X_t) \stackrel{d}{=} (-X_t)$. Take $t_1 < \dots < t_n$ arbitrarily. Then the proof is completely analogous to the one of Corollary 2.41. □

5.3 Continuous Time Stochastic Processes

In this section we will look at some properties relating to stochastic processes in continuous time, i. e., we assume T to be an interval in $\mathbb{R}$.

Let us start with an example how to extract the copula of well-known processes such as Brownian motion. There is a vast amount of literature dealing with Brownian motion, and we take the definition of the process from a very popular one:

5.22 Definition (Karatzas and Shreve, 2000, Definition 1.1, p. 47)
A (standard, one-dimensional) Brownian motion is a continuous, adapted[1] process $X = (X_t)_{t\geq 0}$, defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathsf{P})$ with the following properties:

[1]Recall that a process (X_t) is adapted to a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ if each X_t is $\mathcal{F}_t$-measurable for all t.

a) $X_0 = 0$ *almost surely,*

b) $X_t - X_s$ *is independent of* $\mathcal{F}_s$ *for all* $0 \le s < t$,

c) $X_t - X_s \sim \mathsf{N}(0, t-s)$, $0 \le s < t$.

It follows from Karatzas and Shreve (2000, Theorem 5.12, p. 75) that this process is a Markov process. We can now extract the underlying bivariate copula family $C\langle X\rangle$.

5.23 Example (copula of Brownian motion)
Let $(B_t)_{t\in\mathbb{R}^+}$ be a standard Brownian motion. In the sequel, let Φ denote the distribution function of a standard normal random variable and φ the corresponding density.

The transition probabilities for a standard Brownian motion are given by

$$\mathsf{P}(x,s;y,t) := \mathsf{P}(B_t \le y \mid B_s = x) = \Phi\left(\frac{y-x}{\sqrt{t-s}}\right), \ t > s, \ x, y \in \mathbb{R}. \tag{5.13}$$

From Lemma 2.22 we have

$$\mathsf{P}(x,s;y,t) = D_1 C^B_{s,t}\big(F_s(x), F_t(y)\big),$$

where $C^B_{s,t}$ denotes the copula of B_s and B_t, F_s and F_t are the corresponding marginals. Thus, we get

$$\begin{aligned} C^B_{s,t}\big(F_s(x), F_t(y)\big) &= \int_{-\infty}^{x} D_1 C^B_{s,t}\big(F_s(z), F_t(y)\big)\,\mathrm{d}F_s(z) \\ &= \int_{-\infty}^{x} \Phi\left(\frac{y-z}{\sqrt{t-s}}\right)\mathrm{d}F_s(z) \quad \text{for } 0 < s < t. \end{aligned} \tag{5.14}$$

From the assumption $B_0 = 0$ we have $B_t - B_0 \sim \mathsf{N}(0,t)$ so that $F_t(x) = \Phi(x/\sqrt{t})$ which is equivalent to $x = \sqrt{t}\cdot\Phi^{-1}(F_t(x))$. Substitution into (5.14) yields

$$C^B_{s,t}(u,v) = \int_0^u \Phi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(w)}{\sqrt{t-s}}\right)\mathrm{d}w. \tag{5.15}$$

It is now easy to derive the corresponding partial derivatives and the copula density. Elementary calculus yields

$$D_1 C^B_{s,t}(u,v) = \Phi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u)}{\sqrt{t-s}}\right), \qquad (u,v) \in (0,1)^2, \tag{5.16}$$

$$D_2 C^B_{s,t}(u,v) = \frac{1}{\varphi\big(\Phi^{-1}(v)\big)}\sqrt{\frac{t}{t-s}}\int_0^u \varphi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(w)}{\sqrt{t-s}}\right)\mathrm{d}w, \quad (u,v) \in (0,1)^2. \tag{5.17}$$

For the last result, an interchange of integration and differentiation is necessary which holds from the "differentiation lemma" (e. g., Bauer, 1992, Lemma 16.2, p. 102). Differentiation of (5.16) yields the density

$$c^B_{s,t}(u,v) = \sqrt{\frac{t}{t-s}}\,\frac{\varphi\left((\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u))/\sqrt{t-s}\right)}{\varphi\big(\Phi^{-1}(v)\big)}, \quad (u,v) \in (0,1)^2, \tag{5.18}$$

for the Brownian copula. ■

After having derived the copula we are free to generate a new stochastic process (again of Markovian nature as the Markov structure is characterized by the bivariate copulas, see Theorem 5.44) with the same intertemporal dependence structure as Brownian motion but with different marginals. An empirical example for this is given in Figure 5.1. Every path is generated the following way:

- Choose a time grid $(t_1, \dots, t_n)$ where $t_i > 0$, $t_i < t_{i+1}$.
- A realization $(u_1, \dots, u_n)$ from the Brownian copula process is generated on the time grid (cf. Section B.3).
- For the path of standard Brownian motion, set $b_n := \sqrt{t_n}\Phi^{-1}(u_n)$ such that b_n can be interpreted as a realization of a normal random variable with $\mu = 0$ and $\sigma^2 = t_n$.
- For the path of the generalized Brownian motion with scaled t_ν-marginals, set $d_n := \sqrt{t_n \frac{\nu-2}{\nu}} \cdot F_\nu^{-1}(u_n)$ where F_ν^{-1} is the quantile function of a t_ν-distributed random variable (cf. Section A.2, p. 94) and take $\nu = 3$. The d_n can then be interpreted as realizations from a rescaled t_3 distributed random variable D_n such that $\mathsf{E}(D_n) = 0$, $\mathsf{Var}(D_n) = t_n$.

Note that it is not possible to distinguish the standard Brownian motion from the generalized one only considering first and second moments of the marginals and the dependence structure. This is a widely unrecognized danger in practice.

5.24 Remark

a) It is now straightforward to see that the copula of any strictly monotone increasing transformation of a Brownian motion is exactly the Brownian copula (see Theorem 2.15).

For example, a *geometric Brownian motion* $X = (X_t)_{t\in T}$ satisfies the stochastic integral equation

$$X_t = X_0 + \mu \int_0^t X_s \,\mathrm{d}s + \nu \int_0^t X_s \,\mathrm{d}B_s \tag{5.19}$$

whose solution is given by

$$X_t = X_0 \cdot \exp\left((\mu - \frac{1}{2}\nu^2)t + \nu B_t\right) \tag{5.20}$$

(Karatzas and Shreve, 2000, pp. 349) where $\mu \in \mathbb{R}$, $\nu > 0$, $X_0 > 0$. Thus, X_t is just a strictly increasing transformation of B_t by $f(z) := X_0 \cdot \exp\big((\mu - \frac{1}{2}\nu^2)t + \nu z\big)$. The copula of geometric Brownian motion is therefore given by equation (5.15).

However, this is only a special case of a related question: How can we derive the family of copulas for general transformations of diffusions or—more general—local (semi-)martingales? We will deal with this interesting point in Section 5.5.

b) As Brownian motion is also a Gaussian process[2], $Z := (B_{t_1}, \ldots, B_{t_n})$ has a multivariate normal-distribution. Therefore, the n-dimensional copula must be the n-dimensional Gaussian copula which is derived from the multivariate normal distribution.

■

[2] A Gaussian process is a process whose finite-dimensional distributions are all multivariate normal distributions.

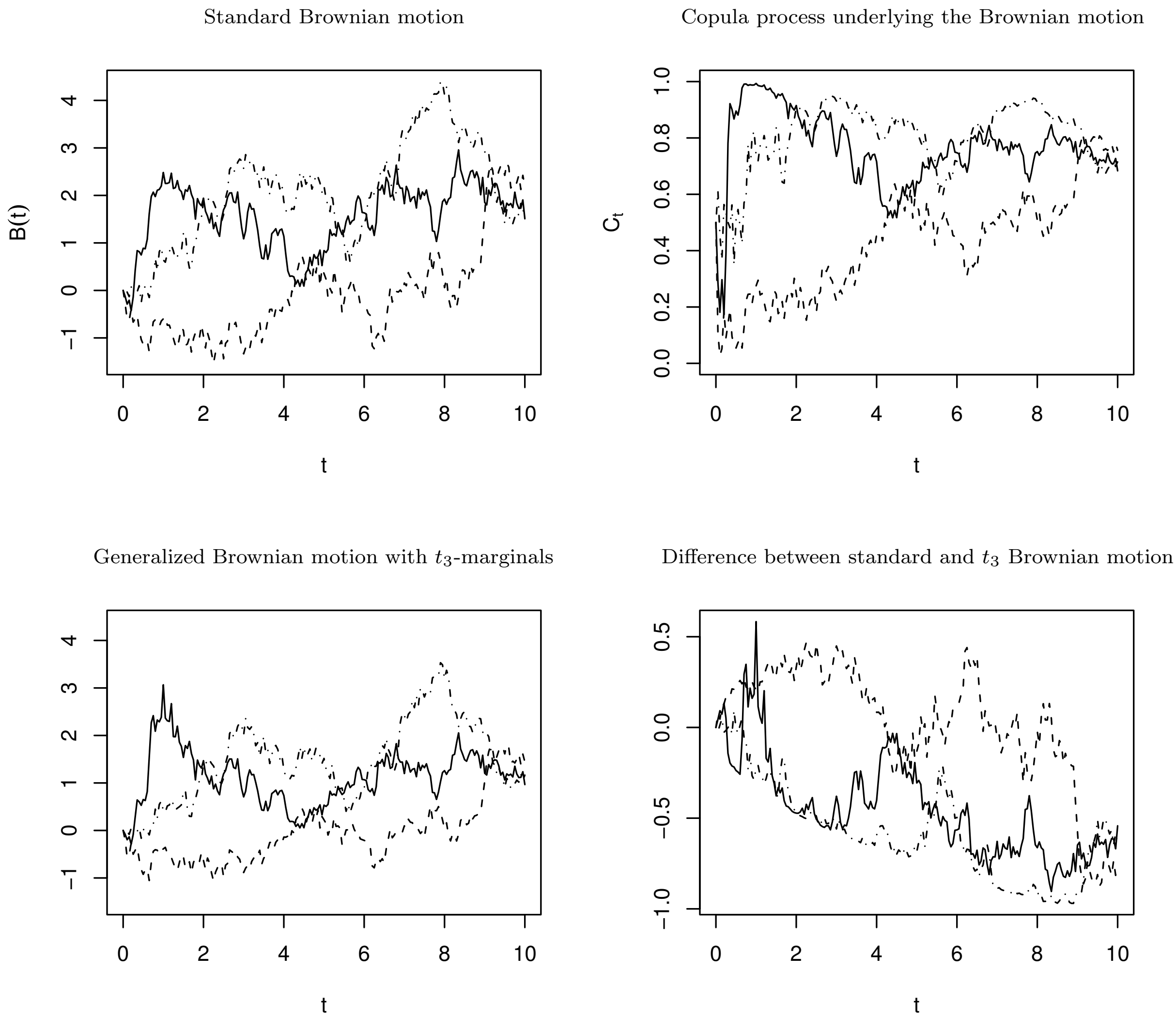

FIGURE 5.1. Standard Brownian motion and underlying copula process (upper row); "generalized Brownian motion" with scaled marginals such that $\mathsf{E}(X_t) = 0$, $\mathsf{Var}(X_t) = t$ as in the Brownian case.

5.3.1 Stochastic Continuity

In this section we will investigate the concept of stochastic continuity of processes and how it is related to copulas. When constructing processes from copulas it is essential to have sufficient conditions for the copulas so that the constructed process is continuous in a sense yet to be defined.

One possible concept is stochastic continuity which is an analogue to stochastic convergence of random variables. For the following, let $(X_t)_{t\in T}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$ where the index set T is a dense subset of $\mathbb{R}$. F_t denotes the distribution function of X_t, $H_{s,t}$ is the bivariate distribution function of (X_s, X_t) and $C_{s,t} = C\langle X_s, X_t\rangle$ the associated copula.

We will make the general assumption that the marginal distributions F_t are continuous and strictly increasing for all $t \in T$.

5.25 Definition (stochastic continuity)
(X_t) is said to be stochastically continuous *(or* continuous in probability*) at a point $t_0 \in T$ if, for any $\varepsilon > 0$,*

$$\lim_{\substack{t\to t_0\\ t\in T}} \mathsf{P}(\,|X_t - X_{t_0}| > \varepsilon\,) = 0. \tag{5.21}$$

If (5.21) holds for all $t \in T$, we say that the process is stochastically continuous on T.

Following Todorovic (1992, Remark 1.9.1), we note that "stochastic continuity is a regularity condition on bivariate marginal distributions of the process". Having in mind the goal of constructing a stochastically continuous process, we will examine the relationship between (5.21) and copulas.

First, note that

$$\begin{aligned}\mathsf{P}(\,|X_t - X_{t_0}| > \varepsilon\,) &= 1 - \mathsf{P}(\,|X_t - X_{t_0}| \le \varepsilon\,)\\ &= 1 - \int_{\mathbb{R}} \mathsf{P}(|X_t - X_{t_0}| \le \varepsilon \mid X_{t_0} = x)\,\mathrm{d}F_{t_0}(x)\\ &= 1 - \int_{\mathbb{R}}\int_{x-\varepsilon}^{x+\varepsilon} \mathrm{d}\mathsf{P}^{X_t|X_{t_0}=x}(y)\,\mathrm{d}\mathsf{P}^{X_{t_0}}(x),\end{aligned} \tag{5.22}$$

and

$$\mathsf{P}(\,|X_t - X_{t_0}| > \varepsilon) = \iint_{|x-y|>\varepsilon} H_{t_0,t}(\,\mathrm{d}x,\,\mathrm{d}y) = \iint_{|x-y|>\varepsilon} C_{t_0,t}\big(F_{t_0}(\,\mathrm{d}x), F_t(\,\mathrm{d}y)\big).$$

Denoting the conditional distribution function by $F_{t|t_0}(y \mid x) := \mathsf{P}(X_t \le y \mid X_{t_0} = x)$ and using Lebesgue-Stieltjes notation, we can write (5.22) as

$$\begin{aligned}
\mathsf{P}(\,|X_t - X_{t_0}| > \varepsilon) &= 1 - \int_{\mathbb{R}} \int_{x-\varepsilon}^{x+\varepsilon} \mathrm{d}F_{t|t_0}(y \mid x)\, \mathrm{d}F_{t_0}(x) \\
&= 1 - \int_{\mathbb{R}} \big[F_{t|t_0}(x+\varepsilon \mid x) - F_{t|t_0}(x-\varepsilon \mid x)\big]\, \mathrm{d}F_{t_0}(x) \\
&= 1 - \int_{\mathbb{R}} \big[D_1 C_{t_0,t}\big(F_{t_0}(x), F_t(x+\varepsilon)\big) - D_1 C_{t_0,t}\big(F_{t_0}(x), F_t(x-\varepsilon)\big)\big]\, \mathrm{d}F_{t_0}(x) \\
&= 1 - \int_{(0,1)} \big[D_1 C_{t_0,t}\big(u, F_t(F_{t_0}^{-1}(u)+\varepsilon)\big) - D_1 C_{t_0,t}\big(u, F_t(F_{t_0}^{-1}(u)-\varepsilon)\big)\big]\, \mathrm{d}u.
\end{aligned} \tag{5.23}$$

For the process to be stochastically continuous in t_0, this expression must tend to 0 for $t \to t_0$, which is equivalent to

$$\lim_{t\to t_0} \int_{(0,1)} \underbrace{\big[D_1 C_{t_0,t}\big(u, F_t(F_{t_0}^{-1}(u)+\varepsilon)\big) - D_1 C_{t_0,t}\big(u, F_t(F_{t_0}^{-1}(u)-\varepsilon)\big)\big]}_{=:f(u;t_0,t,\varepsilon)}\, \mathrm{d}u = 1.$$

Now, from (2.22) we have $0 \le f(u; t_0, t, \varepsilon) \le 1$ for all $u \in (0, 1)$. Further, assume that the limit $\lim_{t\to t_0} f(u; t_0, t, \varepsilon)$ exists for almost all $u \in (0, 1)$ and is a measurable function. With Lebesgue's dominated convergence theorem (e. g., Billingsley, 1995, Theorem 16.4) we obtain the condition

$$\lim_{t\to t_0} \int_{(0,1)} f(u; t_0, t, \varepsilon)\, \mathrm{d}u = \int_{(0,1)} \lim_{t\to t_0} f(u; t_0, t, \varepsilon)\, \mathrm{d}u = 1.$$

But then necessarily

$$\lim_{t\to t_0} f(u; t_0, t, \varepsilon) = 1 \qquad [\lambda^1 \mid_{(0,1)}]$$

which follows from Bauer (1992, Theorem 13.2). Moreover, we see that the following is equivalent to stochastic continuity in t_0:

$$\begin{aligned}
&\lim_{t\to t_0} D_1 C_{t_0,t}\big(u, F_t(F_{t_0}^{-1}(u)+\varepsilon)\big) = 1 \quad [\lambda], \qquad \text{and} \\
&\lim_{t\to t_0} D_1 C_{t_0,t}\big(u, F_t(F_{t_0}^{-1}(u)-\varepsilon)\big) = 0 \quad [\lambda]
\end{aligned} \tag{5.24}$$

for any $\varepsilon > 0$. This is due to the fact that $0 \le D_1 C(u, v) \le 1$ for any copula (see (2.21)) and $F_t\big(F_{t_0}^{-1}(u)+\varepsilon\big) \ge F_t\big(F_{t_0}^{-1}(u)-\varepsilon\big)$ for all $\varepsilon > 0$.

The next goal should be to provide sufficient conditions for (5.24) to hold.

Let us remark that for the upper Fréchet bound $M(u, v) = \min(u, v)$, $u, v \in (0, 1)$, we have $D_1 M(u, v) = \mathbb{1}_{(0,v)}(u)$ for $v \in (0, 1)$. If we compare this with (5.24), we see that the limit should yield strictly monotone dependence. This will be formalized in the following Lemma.

5.26 Lemma
Let $(C_n)_{n\in\mathbb{N}} \subset \mathcal{C}_2$ be a sequence of 2-copulas and $M(u,v) = \min(u,v)$, $u, v \in (0,1)$, the upper Fréchet bound. Assume that $C_n(u,v)$ is differentiable on $(0,1)$ w. r. t. u and

$$\lim_{n\to\infty} D_1 C_n(u,v) = D_1 M(u,v) \quad \left(= \mathbb{1}_{(0,v)}(u) = \begin{cases} 1, & u < v, \\ 0, & u > v, \end{cases} \right) \quad \text{for } u \in (0,1) \setminus \{v\}, \tag{5.25}$$

for all $v \in (0,1)$.

Then $\lim_{n\to\infty} C_n(u,v) = M(u,v)$ for all $u, v \in (0,1)$.

Proof. First, assume $u < v$. Then,

$$\begin{aligned} |C_n(u,v) - M(u,v)| &\overset{(2.11)}{=} M(u,v) - C_n(u,v) = \min(u,v) - C_n(u,v) = u - C_n(u,v) \\ &= u - [C_n(u,v) - \underbrace{C_n(0,v)}_{=0}] \overset{(*)}{=} u - D_1 C_n(\xi_n, v) \cdot u \end{aligned}$$

for some $\xi_n \in (0,u)$. $(*)$ is valid due to the existence of the partial derivative and the continuity of C_n (mean value theorem of differential calculus, e. g., Walter, 1992, Theorem 10.10). As $\xi_n < u < v$ for all $n \in \mathbb{N}$, the assumption yields $\lim_{n\to\infty} D_1 C_n(\xi_n, v) = 1$ so that

$$\lim_{n\to\infty} |C_n(u,v) - M(u,v)| = u - \lim_{n\to\infty} D_1 C_n(\xi_n, v) \cdot u = u - u = 0.$$

The case $u > v$ follows analogously:

$$\begin{aligned} |C_n(u,v) - M(u,v)| &= v - C_n(u,v) = v - C_n(u,v) + \underbrace{C_n(1,v)}_{=v} - v \\ &= (1-u) D_1 C_n(\xi_n, v) \end{aligned}$$

for some $u < \xi_n < 1$. As $v < u < \xi_n$, the assumption yields $D_1 C_n(\xi_n, v) \xrightarrow{n\to\infty} 0$.
The remaining case $u = v$ follows from

$$0 \leq M(u,u) - C_n(u,u) \overset{\text{Cor. 2.9}}{\leq} u - C_n(u - \varepsilon, u) \xrightarrow{n\to\infty} u - (u - \varepsilon) = \varepsilon$$

for all $0 < \varepsilon < u$ so that $\lim_{n\to\infty} C_n(u,u) = M(u,u)$ for all $u \in (0,1)$. □

If we know that C_n converges to M, we can provide the following lemma:

5.27 Lemma
Let $(C_n)_{n\in\mathbb{N}} \subset \mathcal{C}_2$ and $M(u,v) := \min(u,v)$. If $C_n(u,v) \xrightarrow{n\to\infty} M(u,v)$ on $(0,1)^2$, $\lim_{n\to\infty} D_1 C_n(u,v)$ exists for all $u \in (0,1) \setminus \{v\}$ and $\lim_{n\to\infty} D_1 C_n(u,v)$ is continuous in u for all v, then, for any $v \in (0,1)$,

$$\lim_{n\to\infty} D_1 C_n(u,v) = D_1 M(u,v) = \begin{cases} 1, & u < v, \\ 0, & u > v, \end{cases} \quad \text{for } u \in (0,1) \setminus \{v\}.$$

Proof. Let $v \in (0,1)$ and $u \in (0,1) \setminus \{v\}$. W. l. o. g. $u < v$. Assume

$$\lim_{n\to\infty} D_1 C_n(u,v) = q(u,v) < 1 = D_1 M(u,v). \tag{5.26}$$

Due to the continuity of the limit, there exists an $\varepsilon > 0$ such that $\lim_{n\to\infty} D_1 C_n(x,v) < 1$ for all $x \in (u-\varepsilon, u+\varepsilon) \cap (0,1)$. Choose $\delta = \min\{\varepsilon, 1-u, u, v-u\}$ so that $(u-\delta, u+\delta) \subset (0,v)$. Then also $\lim_{n\to\infty} D_1 C_n(x,v) < 1$ for all $x \in (u-\delta, u+\delta)$. Now, observe $C_n(u+\delta,v) - C_n(u,v)$. First,

$$\begin{aligned}\lim_{n\to\infty}[C_n(u+\delta,v) - C_n(u,v)] &= \lim_{n\to\infty} C_n(u+\delta,v) - \lim_{n\to\infty} C_n(u,v)\\ &= M(u+\delta,v) - M(u,v) = u+\delta-u = \delta,\end{aligned} \tag{$*$}$$

as $u, u+\delta < v$. But from the mean value theorem, we also have

$$C_n(u+\delta,v) - C_n(u,v) = D_1 C_n(\xi_n, v)\cdot\delta$$

for some $\xi_n \in (u, u+\delta)$ so that

$$\lim_{n\to\infty}[C_n(u+\delta,v) - C_n(u,v)] = \underbrace{\lim_{n\to\infty} D_1 C_n(\xi_n,v)}_{<1} \cdot\delta < \delta,$$

which is clearly a contradiction to the above equation $(*)$. The assertion follows. □

From these two Lemmas, we can immediately derive the following corollary.

5.28 Corollary
Let $(C_n)_{n\in\mathbb{N}} \subset \mathcal{C}_2$ be a sequence of 2-copulas and $M(u,v) = \min(u,v)$. For any $v \in (0,1)$, let $\lim_{n\to\infty} D_1 C_n(u,v)$ exist for all $u \in (0,1)\setminus\{v\}$ and be continuous in u. Then the following properties are equivalent:

a) $\lim_{n\to\infty} C_n(u,v) = M(u,v)$ *for all* $u, v \in (0,1)$.

b) $\lim_{n\to\infty} D_1 C_n(u,v) = D_1 M(u,v) = \mathbb{1}_{(0,v)}(u)$ *for any* $v \in (0,1)$.

Proof. *a)*⇒*b)* follows from Lemma 5.27, *b)*⇒*a)* from Lemma 5.26. □

5.29 Remark
Lemmas 5.26 and 5.27 and Corollary 5.28 could equivalently have been formulated for a family $\{C_{s,t} : s,t \in T\}$ of bivariate copulas where $\lim_{t\to s} C_{s,t}(u,v)$ must be considered. This is an important aspect when dealing with continuous time stochastic processes. ■

With these preparations, we are able to give sufficient conditions for the stochastic continuity of a process generated by a family of copulas:

5.30 Theorem
Let $(X_t)_{t \in T}$ be a stochastic process, F_t the cumulative distribution function of X_t, $t \in T$, and $C_{s,t} = C\langle X_s, X_t \rangle$ the copula of (X_s, X_t), $s, t \in T$.
If

a) $\lim_{t \to s} F_t^{-1}(u) = F_s^{-1}(u)$ *for almost all* $u \in (0,1)$, *and*

b) $\lim_{t \to s} D_1 C_{s,t}(u,v) = \mathbb{1}_{(0,v)}(u)$ *exists for all* $u \in (0,1) \setminus \{v\}$,

then (X_t) is stochastically continuous at time s.

Proof. Let $\varepsilon > 0$ be given. Let $N \in \mathcal{B}^1$ be a Lebesgue null-set such that

$$\lim_{t \to s} F_t^{-1}(u) = F_s^{-1}(u) \quad \text{for all } u \in N^c.$$

For a fixed $u \in N^c$, there exists $\delta = \delta(\varepsilon, u)$ such that

$$F_s^{-1}(u) - \varepsilon < F_t^{-1}(u) < F_s^{-1}(u) + \varepsilon \quad \text{if } |t-s| < \delta.$$

Now, $\lim_{t \to s} D_1 C_{s,t}(u,v) = \mathbb{1}_{(0,v)}(u)$, and therefore

$$\lim_{t \to s} D_1 C_{s,t}(u,v)\Big|_{\left(u, F_t(F_s^{-1}(u)-\varepsilon)\right)} = 0 \quad \text{and} \quad \lim_{t \to s} D_1 C_{s,t}(u,v)\Big|_{\left(u, F_t(F_s^{-1}(u)+\varepsilon)\right)} = 1.$$

The assertion follows with (5.24). □

5.31 Remark

a) We cannot expect to enhance the above result w. r. t. stronger continuity concepts as they are usually based on other concepts than finite-dimensional distributions.

b) The finite-dimensional distributions alone do not give all the information regarding the continuity properties of a process. For example, let $(\Omega, \mathcal{F}, \mathsf{P}) = \big([0,\infty), \mathcal{B}, \mu\big)$ where $\mathcal{B}$ denotes the Borel σ-algebra on $[0,\infty)$, and μ is a probability measure on $[0,\infty)$ with no mass on single points. It is easy to see that $X_t(\omega) = \mathbb{1}_{\{t\}}(\omega)$ and $Y_t(\omega) = 0$ for all $(t,\omega) \in [0,\infty) \times [0,\infty)$ have the same finite-dimensional distributions and that (X_t) is a version of (Y_t),[3] but $t \to Y_t(\omega)$ is continuous for all ω while $t \to X_t(\omega)$ is discontinuous for all ω.

c) An important consequence of stochastic continuity is given by the following fact (Todorovic, 1992, Proposition 1.11.1, p. 29): If $(X_t)_{t \in T}$ is a real-valued process and $T \subset \mathbb{R}$ a compact interval such that (X_t) is stochastically continuous (or continuous in probability) on T, then there exists a version $(\tilde{X}_t)$ of (X_t) which is separable and measurable.

■

[3] Recall that (X_t) is a version of (Y_t) if for all $t \in T$: $\mathsf{P}(X_t = Y_t) = 1$.

Let us give an example for the use of the theorem. We have already determined the copula of a Brownian motion (see Example 5.23). It is known that Brownian motion almost surely has continuous sample paths, from which the stochastic continuity on $\mathbb{R}^+$ follows. We should be able to check the latter property via the copula approach:

5.32 Example (stochastic continuity of Brownian motion)
Let $(B_t)_{t\in\mathbb{R}^+}$ be a standard Brownian motion (i. e., $B_0 = 0$). Then $B_t \sim \mathsf{N}(0,t)$ so that $F_t(x) = \Phi(x/\sqrt{t})$, $x \in \mathbb{R}$, which is equivalent to $F_t^{-1}(y) = \sqrt{t}\Phi^{-1}(y)$, $y \in (0,1)$. Then

$$F_t^{-1}(y) - F_s^{-1}(y) = (\sqrt{t} - \sqrt{s})\Phi^{-1}(y) \xrightarrow{t\to s} 0$$

for $y \in (0,1)$. From Example 5.23 we know that the copula $C_{s,t}^B$ is given by

$$C_{s,t}^B(u,v) = \int_0^u \Phi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(w)}{\sqrt{t-s}}\right) \mathrm{d}w,$$

i. e.,

$$D_1 C_{s,t}^B(u,v) = \Phi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u)}{\sqrt{t-s}}\right), \quad t > s.$$

(Note that the copula would be $\Phi\big((\sqrt{s}\Phi^{-1}(u) - \sqrt{t}\Phi^{-1}(v))/\sqrt{s-t}\big)$ for $s > t$.) Clearly, $D_1 C_{s,t}^B(u,v)$ is continuous in u. Now,

$$\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u) \xrightarrow{t\to s} \sqrt{s}\big(\Phi^{-1}(v) - \Phi^{-1}(u)\big)$$

and $\sqrt{t-s} \xrightarrow{t\to s} 0$; therefore

$$\lim_{t\to s} \frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u)}{\sqrt{t-s}} = \begin{cases} \infty, & u < v, \\ -\infty, & u > v, \end{cases}$$

so that $\lim_{t\to s} D_1 C_{s,t}^B(u,v) = \mathbb{1}_{(0,v)}(u)$. The stochastic continuity now follows from Theorem 5.30. ■

We will show in the next section that Theorem 5.30 can be used advantageously to shorten existing proofs.

5.3.2 EFGM processes

In the following, we will give an example of a family of copulas whose induced uniform process cannot be stochastically continuous. We will use a particulary simple class of multivariate distributions, the so-called Eyraud-Farlie-Gumbel-Morgenstern distributions (EFGM-distributions), which were discussed by Morgenstern (1956), Gumbel (1958) and Farlie (1960), but it seems that the earliest publication of the functional form is Eyraud (1938).

Although their form makes them natural candidates for multivariate distributions, it was discovered that there are fundamental problems inhibiting the definition of smooth continuous-parameter stationary EFGM processes. Cambanis (1991) showed that these processes are not stochastically continuous and even not measurable. To be precise, he showed this for *strictly stationary* EFGM processes. He did not use an approach via the underlying copulas but considered the complete parametric form (including the marginals), showing the discontinuity of any such process by using a set of inequalities. These results were obtained under a weak symmetry-type condition on the support of the marginal distribution[4]. Due to this approach, Cambanis was not able to derive results for the non-stationary case. (Note: "non-stationarity" can mean time-varying marginals, different copulas $C_{s,t}$ and $C_{s+h,t+h}$, or both!)

In the usual form, a bivariate EFGM-distribution H with marginal distributions F_1 and F_2 is given by

$$H(x,y) = F_1(x)F_2(y)\Big(1-\alpha\big(1-F_1(x)\big)\big(1-F_2(y)\big)\Big), \quad \alpha \in [\alpha_{\min},\alpha_{\max}],\ x,y\in\mathbb{R}. \tag{5.27}$$

When both marginals are absolutely continuous, $\alpha_{\min} = -1$, $\alpha_{\max} = 1$, but generally,

$$\begin{aligned} \alpha_{\min} &= -\min\big\{(M_1M_2)^{-1}, ((1-m_1)(1-m_2))^{-1}\big\}, \\ \alpha_{\max} &= \min\big\{(M_1(1-m_2))^{-1}, ((1-m_1)M_2)^{-1}\big\}, \end{aligned} \tag{5.28}$$

where

$$m_k = \inf\left(\{F_k(x) : x\in\mathbb{R}\}\setminus\{0,1\}\right), \quad M_k = \sup\left(\{F_k(x) : x\in\mathbb{R}\}\setminus\{0,1\}\right), \quad k=1,2.$$

For a proof of these assertions, see Cambanis (1977, Theorem 1).
A simple multivariate version discussed by Johnson and Kotz (1975) is

$$H(x_1,\dots,x_n) = \prod_{i=1}^n F_i(x)\left(1+\sum_{1\le k<j\le n}\alpha_{k,j}\big(1-F_k(x_k)\big)\big(1-F_j(x_j)\big)\right), \tag{5.29}$$

where

$$1+\sum_{1\le k<j\le n}\varepsilon_k\varepsilon_j\alpha_{k,j}\ge 0 \quad \text{for all } \varepsilon_k\in\{-M_k, 1-m_k\} \tag{5.30}$$

must be valid.

Based on this multivariate distribution, one can now introduce EFGM random processes $(X_t)_{t\in T}$ by requiring that all finite-dimensional distributions are multivariate EFGM, i. e., for the distribution function of $(X_{t_1},\dots,X_{t_n})$, $t_1<\dots<t_n$, $t_i\in T$, $n\in\mathbb{N}$, we have

$$H_{t_1,\dots,t_n}(x_1,\dots,x_n) = \prod_{i=1}^n F_{t_i}(x_i)\left(1+\sum_{1\le k<j\le n}\alpha(t_k,t_j)\big(1-F_{t_k}(x_k)\big)\big(1-F_{t_j}(x_j)\big)\right), \tag{5.31}$$

where F_t is the cdf of X_t.[5] The coefficient function $\alpha(s,t)$, $s\ne t$, is assumed to be symmetric ($\alpha(s,t)=\alpha(t,s)$) and the admissible values are given by $1+\sum_{1\le k<j\le n}\varepsilon_k\varepsilon_j\alpha(t_k,t_j)\ge 0$ for

[4]Note that we deliberately write "marginal distribution" instead of "marginal distributions" as Cambanis considers strictly stationary processes.

[5]The consistency condition (5.5) can be easily checked.

all $\varepsilon_k \in \{-M_{t_k}, 1-m_{t_k}\}$ according to (5.30), or $\varepsilon_k \in \{-1,1\}$ if F_{t_k} is absolutely continuous. The dependence structure is fully determined by $\alpha(\cdot,\cdot)$ via (5.31) which will prove to be the main limitation of these processes w. r. t. stochastic continuity.

In the case of strict stationarity, we have $F_t \equiv F$ for all $t \in T$ on the one hand and on the other hand, $\alpha(s,t) = \tilde{\alpha}(t-s)$ for all $t \neq s$ for a function $\tilde{\alpha}$.

With the help of several inequalities, Cambanis (1991, Proposition 4) was able to show:

5.33 Theorem
Let $(X_t)_{t\in T}$ be a real-valued, strictly stationary EFGM process with nondegenerate marginal distribution F and dependence function α. If $M(F) = 1 - m(F)$, where $m(F) := \inf(\{F(x) : x \in \mathbb{R}\} \setminus \{0,1\})$, $M(F) := \sup(\{F(x) : x \in \mathbb{R}\} \setminus \{0,1\})$, then the process is not stochastically continuous.[6]

From Section 5.3 we know that stochastic continuity is a matter of bivariate distributions. It is therefore clear and advantageous for the analysis of stochastic continuity that we only need the bivariate distributions of the process, i. e., a stochastic process form of (5.27):

$$H_{s,t}(x_s,x_t) = F_s(x_s)F_t(x_t)\big(1+\alpha(s,t)\big(1-F_s(x_s)\big)\big(1-F_t(x_t)\big)\big), \quad x_s,x_t \in \mathbb{R}, \tag{5.32}$$

where $H_{s,t}$ is the distribution function of (X_s,X_t), $s<t$. We can easily identify the copula as

$$C_{s,t}(u,v) = uv\big(1+\alpha(s,t)(1-u)(1-v)\big), \quad u,v \in [0,1]. \tag{5.33}$$

Note that the conditions for $\alpha(s,t)$ yield that $C_{s,t}$ is 2-increasing for all $s<t$; generally, only $\alpha(s,t) \in [-1,1]$ is required for $C_{s,t}$ to be a copula. Now,

$$D_1C_{s,t}(u,v) = v + \alpha(s,t)(1-v)(1-2u), \quad u,v \in (0,1). \tag{5.34}$$

From this, we can easily formulate the following lemma:

5.34 Lemma
A uniform process $(U_t)_{t\in T}$ generated from the EFGM family is not continuous in probability in any point $t \in T$.

Proof. From (5.24) and (5.34), we have ($F_t(u) = F_t^{-1}(u) = u$ for all $u \in [0,1]$) for all $\varepsilon > 0$:

$$\begin{aligned}\lim_{t\to s} D_1C_{s,t}(u,u+\varepsilon) &= \lim_{t\to s}[u+\varepsilon+\alpha(s,t)(1-(u+\varepsilon))(1-2u)] = 1\\ &\iff \lim_{s\to t}\alpha(s,t) = \frac{1-u-\varepsilon}{(1-u-\varepsilon)(1-2u)} \quad \text{for all } \varepsilon < 1-u,\ u \neq 1/2 \quad [\lambda].\end{aligned}$$

This can clearly not be the case as $\alpha(s,t)$ is independent from u. Thus, the process cannot be continuous in t. $\square$

In the following theorem we will present a result similar to Theorem 5.33.

[6] In the sense of: "For any $t \in T$, X is not stochastically continuous in t."

5.35 Theorem
Let $(X_t)_{t \in T}$ be a real-valued, strictly stationary EFGM process with continuous marginal(s) F. Then (X_t) is not stochastically continuous in any $t \in T$.

Proof. For the process to be continuous in t, (5.24) must hold for all $\varepsilon > 0$ and almost all $u \in (0,1)$. Taking the second equation and (5.34), we must particularly have

$$D_1 C_{s,t}\big(u, F(F^{-1}(u)-\varepsilon)\big) = F\big(F^{-1}(u)-\varepsilon\big)+\alpha(s,t)\Big(1-F\big(F^{-1}(u)-\varepsilon\big)\Big)(1-2u) \xrightarrow{t \to s} 0 \quad [\lambda^1].$$

Without loss of generality, assume $u \neq 1/2$, as $\{1/2\}$ is a λ^1 null set. Let $\varepsilon > 0$ be given.

The case $F(F^{-1}(u) - \varepsilon) = 1$ cannot occur as this is equivalent to $F^{-1}(u) \geq \omega(F) + \varepsilon$.[7] As $u \in (0,1)$, this is impossible. For the same reason, $g(u,\varepsilon) := F(F^{-1}(u) - \varepsilon) < 1$ for all $u \in (0,1)$. So the above equation is equivalent to

$$\alpha(s,t) \xrightarrow{s \to t} \frac{g(u,\varepsilon)}{g(u,\varepsilon) - 1} \cdot \frac{1}{1-2u} \qquad [\lambda^1],$$

which evidently cannot be the case, as any Borel set with positive measure consists of uncountably many points. (The left side is independent of the right side.) Therefore, (X_t) is not stochastically continuous in t. □

5.4 Markov Processes

Markov processes are a certain class of stochastic processes with a special kind of dependence structure: The "future" is independent of the "past" given the "present". They are important in many disciplines. For example, in mathematical finance using Markov processes to model stock price behaviour is compatible with the assumption of (weak) market efficiency. This states that the present price of a stock impounds all the information contained in a record of past prices.

We assume $T \subset \mathbb{R}^+$.

5.36 Definition (Markov process)
A real-valued process $(X_t)_{t \in T}$ on some probability space $(\Omega, \mathcal{F}, \mathsf{P})$ is called a Markov process *if, for any $n \in \mathbb{N}$, $t_1 < \cdots < t_n$, $t_1, \ldots, t_n \in T$, and $B \in \mathcal{B}$ (where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$), we have*

$$\mathsf{P}(X_{t_n} \in B \mid X_{t_1}, \ldots, X_{t_{n-1}}) = \mathsf{P}(X_{t_n} \in B \mid X_{t_{n-1}}) \qquad [\mathsf{P}]. \tag{5.35}$$

5.37 Remark
(5.35) is equivalent to $\mathsf{P}(X_{t_n} \leq x \mid X_{t_1}, \ldots, X_{t_{n-1}}) = \mathsf{P}(X_{t_n} \leq x \mid X_{t_{n-1}})$ [P] for all $x \in \mathbb{R}$ as $\{(-\infty, x] : x \in \mathbb{R}\}$ is a generator of $\mathcal{B}$ closed under intersections. ■

---

[7] $\omega(F)$ is the right endpoint of the support of F; see p. 94.

We will use the previously introduced notation for the conditional distribution functions:

$$\mathsf{P}(x,s;y,t) := \mathsf{P}(X_t \le y \mid X_s = x), \quad x,y \in \mathbb{R},\ s,t \in T,\ s < t. \tag{5.36}$$

It is straightforward to see that we are actually dealing with transition kernels:

5.38 Definition (transition probability)
Let $\mathsf{P}(x,s;y,t)$ be a version of $\mathsf{P}(X_t \le y \mid X_s = x)$ having the properties:

a) For all $s < t$, $x \in \mathbb{R}$, $\mathsf{P}(x,s;\,\cdot\,,t)$ is a distribution function on $\mathbb{R}$ (and therefore induces a unique probability measure on $\mathcal{B}$);

b) for all $s < t$, $y \in \mathbb{R}$, $\mathsf{P}(\,\cdot\,,s;y,t)$ is $\mathcal{B}$-measurable.

Then $\mathsf{P}(x,s;y,t)$ is called the transition probability *of the Markov process.*

As a consequence of the Markov property, the conditional probabilities satisfy the *Chapman-Kolmogorov equations* (e. g., Meyn and Tweedie, 1996, Theorem 3.4.2, for the form used here), which relate the state of the process at time t with that at an earlier time s through an intermediate time u:

$$\mathsf{P}(x,s;y,t) = \int_{\mathbb{R}} \mathsf{P}(z,u;y,t)\,\mathsf{P}(x,s;\,\mathrm{d}z,u) \quad (s < u < t,\ x,y \in \mathbb{R}). \tag{5.37}$$

If the conditional densities $p(x,s;y,t) := \mathsf{P}(X_t = y \mid X_s = x)$ exist, we can write

$$p(x,s;y,t) = \int_{\mathbb{R}} p(x,s;z,u)p(z,u;y,t)\,\mathrm{d}z \quad (s < u < t,\ x,y \in \mathbb{R}).$$

This can be interpreted as a continuous version of the law of total probability, modified by the Markov property.

We will now give an outline of the work of Darsow, Nguyen, and Olsen (1992) who were the first to relate copulas to Markov processes. The key to the understanding of this relationship was the introduction of a product for 2-copulas which essentially captures the dependence model of the Chapman-Kolmogorov equations.

5.39 Definition (product of copulas)
Let $C_1, C_2 \in \mathcal{C}_2$ be 2-copulas. The product *of C_1 and C_2 is the function $C_1 * C_2 \colon I^2 \to I$, given by*

$$(C_1 * C_2)(u,v) := \int_{(0,1)} D_2C_1(u,z) \cdot D_1C_2(z,v)\,\mathrm{d}z. \tag{5.38}$$

Note that the $*$-product is well-defined due to (2.21).

The next theorem shows that $\mathcal{C}_2$ is closed under the $*$-operation. We will give a detailed version as the proof in Darsow et al. (1992) is only sketched.

5.40 Theorem (Darsow et al., 1992, Theorem 2.1)
For $C_1, C_2 \in \mathcal{C}_2$, $C_1 * C_2 \in \mathcal{C}_2$, i. e., $C_1 * C_2$ is again a copula.

Proof. For the boundary conditions (2.5) and (2.6) we have

$$(C_1 * C_2)(0, v) = \int\limits_{(0,1)} D_2C_1(0, z) \cdot D_1C_2(z, v)\,\mathrm{d}z,$$

but $D_2C_1(0, z) = \lim_{h\searrow 0}(C_1(0, z + h) - C_1(0, z))/h = 0$ as copulas are grounded so that $(C_1 * C_2)(0, v) = 0$ for $v \in I$; the case $(C_1 * C_2)(u, 0) = 0$ for all $u \in I$ follows analogously. Further,

$$\begin{aligned}(C_1 * C_2)(1, v) &= \int\limits_{(0,1)} D_2C_1(1, z) \cdot D_1C_2(z, v)\,\mathrm{d}z = \int\limits_{(0,1)} D_1C_2(z, v)\,\mathrm{d}z = C_2(1, v) - C_2(0, v)\\ &= v - 0 = v\end{aligned}$$

for $v \in I$, as $D_2C_1(1, z) = \lim_{h\searrow 0}(C_1(1, z + h) - C_1(1, z))/h = \lim_{h\searrow 0}(z + h - z)/h = 1$. The case $(C_1 * C_2)(u, 1) = u$ for all $u \in I$ is analogous.

It remains to show that $C_1 * C_2$ is 2-increasing, i. e., the $(C_1 * C_2)$-volume of any rectangle $(\boldsymbol{u}, \boldsymbol{v}]$ ($\boldsymbol{u} = (u_1, u_2)$, $\boldsymbol{v} = (v_1, v_2)$, $u_i, v_i \in I$, $u_1 \le u_2$, $v_1 \le v_2$) must be non-negative:

$$\begin{aligned}&V_{C_1 * C_2}\big((\boldsymbol{u}, \boldsymbol{v}]\big)\\ &\quad= (C_1 * C_2)(u_2, v_2) - (C_1 * C_2)(u_1, v_2) - (C_1 * C_2)(u_2, v_1) + (C_1 * C_2)(u_1, v_1)\\ &\quad= \int\limits_{(0,1)} \big[D_2C_1(u_2, z)D_1C_2(z, v_2) - D_2C_1(u_1, z)D_1C_2(z, v_2)\\ &\qquad\qquad - D_2C_1(u_2, z)D_1C_2(z, v_1) + D_2C_1(u_1, z)D_1C_2(z, v_1)\big]\,\mathrm{d}z\\ &\quad= \int\limits_{(0,1)} \underbrace{[D_2C_1(u_2, z) - D_2C_1(u_1, z)]}_{\ge 0,\text{ see (2.22)}}\underbrace{[D_1C_2(z, v_2) - D_1C_2(z, v_1)]}_{\ge 0,\text{ see (2.22)}}\,\mathrm{d}z \ge 0.\end{aligned}$$

Thus, $C_1 * C_2$ is a copula. □

The $*$-product of copulas can be seen as a continuous analogue of the multiplication operator for transition matrices. Some algebraic properties are preserved:

5.41 Lemma (Darsow et al., 1992, pp. 605)
Let $C \in \mathcal{C}_2$ be any 2-copula, Π the independence copula and M, W the upper and lower Fréchet bounds. Then

$$\begin{aligned}\Pi * C = C * \Pi &= \Pi,\\ M * C = C * M &= C,\\ (W * C)(x, y) &= y - C(1 - x, y), \quad x, y \in I,\\ (C * W)(x, y) &= x - C(x, 1 - y), \quad x, y \in I.\end{aligned}$$

In particular, Π is a null element in $\mathcal{C}_2$ and M is an identity w. r. t. the $*$-operator.

Additionally, we have the following continuity property of "$*$":

5.42 Theorem (Darsow et al., 1992, Theorems 2.3 and 2.4)
Let $A, B, C \in \mathcal{C}_2$, $(A_n)_{n\in\mathbb{N}} \subset \mathcal{C}_2$ with $A_n \xrightarrow{n\to\infty} A$ (pointwise in I^2). Then the following hold:

a) $A_n * B \xrightarrow{n\to\infty} A * B$ *and* $B * A_n \xrightarrow{n\to\infty} B * A$.

b) *"$*$" is associative, i. e.,*
$$A * (B * C) = (A * B) * C.$$

The main result of Darsow, Nguyen, and Olsen is the following:

5.43 Theorem (Darsow et al., 1992, Theorem 3.2)
Let $(X_t)_{t\in T}$ be a real-valued stochastic process, and for each $s, t \in T$, let $C_{s,t}$ denote the copula of (X_s, X_t). Then the following are equivalent:

a) The conditional distribution functions $\mathsf{P}(x, s; y, t)$ satisfy the Chapman-Kolmogorov conditions (5.37) for all $s < u < t$, $s, u, t \in T$, and almost all $x, y \in \mathbb{R}$.

b) For all $s, u, t \in T$ with $s < u < t$,
$$C_{s,t} = C_{s,u} * C_{u,t}. \tag{5.39}$$

Satisfaction of the Chapman-Kolmogorov equations is a necessary but not sufficient condition for $(X_t)_{t\in T}$ to be a Markov process. If we introduce an additional operator (a generalized version of the $*$-operator), we can give equivalent conditions in terms of copulas.

Let $C_1 \in \mathcal{C}_m$, $C_2 \in \mathcal{C}_n$, $m, n \geq 2$. Define $C_1 \star C_2 : I^{m+n-1} \to I$ by
$$(C_1 \star C_2)(u_1, \ldots, u_{m+n-1}) := \int\limits_{(0,u_m)} D_m C_1(u_1, \ldots, u_{m-1}, z) D_1 C_2(z, u_{m+1}, \ldots, u_{m+n-1})\, \mathrm{d}z. \tag{5.40}$$
If $m = n = 2$, the $*$ and $\star$ products are related by
$$(C_1 * C_2)(u, v) = (C_1 \star C_2)(u, 1, v), \quad u, v \in I.$$

In analogy to the above theorems, $C_1 \star C_2$ is a $(m + n - 1)$-copula, the $\star$-product is associative and continuous in each coordinate.

We can finally give the following characterization of Markov processes:

5.44 Theorem (Darsow et al., 1992, Theorem 3.3)
A real-valued stochastic process $(X_t)_{t\in T}$ is a Markov process if and only if for all $n \geq 3$, $t_1, \ldots, t_n \in T$ with $t_k < t_{k+1}$, $k = 1, \ldots, n-1$,
$$C_{t_1,\ldots,t_n} = C_{t_1,t_2} \star C_{t_2,t_3} \star \cdots \star C_{t_{n-1},t_n} \tag{5.41}$$
where $C_{t_1,\ldots,t_n} = C\langle X_{t_1}, \ldots, X_{t_n}\rangle$ and $C_{t_k,t_{k+1}} = C\langle X_{t_k}, X_{t_{k+1}}\rangle$.

These findings create a new approach to the theory of Markov processes and yield a new technique for constructing such processes. We cite Darsow et al. (1992):

> *In the conventional approach, one specifies a Markov process by giving the initial distribution F_{t_0} and a family of transition probabilities $\mathsf{P}(x, s; y, t)$ satisfying the Chapman-Kolmogorov equations. In our approach, one specifies a Markov process by giving all of the marginal distributions and a family of 2-copula satisfying (5.41). Ours is accordingly an alternative approach to the study of Markov processes which is different in principle from the conventional one. Holding the transition probabilities of a Markov process fixed and varying the initial distribution necessarily varies all of the marginal distributions, but holding the copulas of the process fixed and varying the initial distribution does not affect any other marginal distribution.*

5.4.1 An Application: Structure of Records

It might be interesting to see the use of the findings of Section 5.4. Let us therefore consider the following setting:

Let X_i, $i \in \mathbb{N}$, be a sequence of iid random variables with common continuous distribution function F, and denote the maximum up to time n by

$$M_n := \bigvee_{i=1}^{n} X_i := \max_{1 \le i \le n} X_i,$$

$n \in \mathbb{N}$. Thus, $M_n \sim F^n$. If n is such that $M_n > M_{n-1}$, n is called "record time". Define

$$L(1) := 1, \quad L(n+1) := \inf\{j > L(n) : M_j > M_{L(n)}\}$$

with the convention that $L(n) = \infty$ if the set is empty. $\{L(n) : n \ge 1\}$ are the times where the Markov process[8] (M_n) jumps.
The set $\{X_{L(n)} : n \ge 1\} = \{M_{L(n)} : n \ge 1\}$ is called the "record values".

For $t_1, t_2 \in \mathbb{N}$, $t_1 < t_2$, the bivariate distribution function of (M_{t_1}, M_{t_2}) is given by

$$H_{t_1,t_2}(x_1, x_2) = F^{t_1}(x_1 \wedge x_2) \cdot F^{t_2-t_1}(x_2), \quad x_1, x_2 \in \mathbb{R},$$

(see Resnick, 1987, p. 165) so that the copula and its partial derivatives are given by

$$\begin{aligned} C_{t_1,t_2}(u_1, u_2) &= F^{t_1}\big(F^{-1}(u_1^{1/t_1}) \wedge F^{-1}(u_2^{1/t_2})\big) \cdot F^{t_2-t_1}\big(F^{-1}(u_2^{1/t_2})\big) \\ &= \left(u_1^{1/t_1} \wedge u_2^{1/t_2}\right)^{t_1} \cdot \left(u_2^{1/t_2}\right)^{t_2-t_1} = \begin{cases} u_1 \cdot u_2^{1-t_1/t_2}, & u_1^{1/t_1} \le u_2^{1/t_2}, \\ u_2, & u_1^{1/t_1} > u_2^{1/t_2}, \end{cases} \end{aligned} \tag{5.42}$$

$$D_1 C_{t_1,t_2}(u_1, u_2) = \begin{cases} u_2^{1-t_1/t_2}, & 0 < u_1^{1/t_1} < u_2^{1/t_2} < 1, \\ 0, & 0 < u_2^{1/t_2} < u_1^{1/t_1} < 1, \end{cases} \tag{5.43}$$

[8] This is a known fact (see Resnick, 1987, Proposition 4.1); we will derive this result with the help of the new tools based on the theory of copulas.

and

$$D_2C_{t_1,t_2}(u_1,u_2) = \begin{cases} u_1\left(1-\frac{t_1}{t_2}\right)u_2^{-t_1/t_2}, & 0 < u_1^{1/t_1} < u_2^{1/t_2} < 1, \\ 1, & 0 < u_2^{1/t_2} < u_1^{1/t_1} < 1. \end{cases} \tag{5.44}$$

Note that "$\wedge$" and "$\vee$" denote the minimum and maximum, respectively. For $t_1 = t_2$ we have (clearly) $C(u,v) = M(u,v) = \min(u,v)$, $u,v \in I$, so this case is consistent with the findings. For $t_2 \to \infty$, we have $C_{t_1,t_2}(u,v) \to \Pi(u,v)$, $u,v \in I$, as can be easily seen. M_k and M_l are asymptotically independent.

To derive the finite-dimensional distributions, i. e. for $(M_{t_1},\dots,M_{t_k})$, we note that for $t_1 < t_2 < \cdots < t_k$,

$$H_{t_1,\dots,t_k}(x_1,\dots,x_k) = F^{t_1}\left(\bigwedge_{i=1}^k x_i\right)\cdot F^{t_2-t_1}\left(\bigwedge_{i=2}^k x_i\right)\cdots F^{t_k-t_{k-1}}(x_k)$$

(see Resnick, 1987, p. 165). The copula is derived to be

$$\begin{aligned} C_{t_1,\dots,t_k}&(u_1,\dots,u_k) \\ &= F^{t_1}\Big(\bigwedge_{i=1}^k F^{-1}(u_i^{1/t_i})\Big)\cdot F^{t_2-t_1}\Big(\bigwedge_{i=2}^k F^{-1}(u_i^{1/t_i})\Big)\cdots F^{t_k-t_{k-1}}\Big(F^{-1}(u_k^{1/t_k})\Big) \\ &= \Big(\bigwedge_{i=1}^k u_i^{1/t_i}\Big)^{t_1}\cdot\Big(\bigwedge_{i=2}^k u_i^{1/t_i}\Big)^{t_2-t_1}\cdots\Big(u_k^{1/t_k}\Big)^{t_k-t_{k-1}} \\ &\overset{t_0:=0}{=} \prod_{j=1}^k\Big(\bigwedge_{i=j}^k u_i^{1/t_i}\Big)^{t_j-t_{j-1}}. \end{aligned} \tag{5.45}$$

Let us see if the Chapman-Kolmogorov equations hold. Assume $t_1 < t_2 < t_3$, $u_1,u_3 \in (0,1)$. Then, by (5.42)–(5.44),

$$\begin{aligned} (C_{t_1,t_2}{*}C_{t_2,t_3})(u_1,u_3) &:= \int_{(0,1)} D_2C_{t_1,t_2}(u_1,v)\cdot D_1C_{t_2,t_3}(v,u_3)\,\mathrm{d}v \\ &= \int_0^{u_1^{t_2/t_1}} D_1C_{t_2,t_3}(v,u_3)\,\mathrm{d}v + \int_{u_1^{t_2/t_1}}^1 u_1(1-t_1/t_2)v^{-t_1/t_2}\underbrace{D_1C_{t_2,t_3}(v,u_3)}_{=0\text{ for }v>u_3^{t_2/t_3}}\,\mathrm{d}v, \\ &= C_{t_2,t_3}(u_1^{t_2/t_1},u_3) + u_1(1-t_1/t_2)\cdot\int_{u_1^{t_2/t_1}\wedge u_3^{t_2/t_3}}^{u_3^{t_2/t_3}} v^{-t_1/t_2}\cdot u_3^{1-t_2/t_3}\,\mathrm{d}v \end{aligned}$$

$$= C_{t_2,t_3}(u_1^{t_2/t_1}, u_3) + u_1 u_3^{1-t_2/t_3} \Big[v^{1-t_1/t_2} \Big]_{u_1^{t_2/t_1} \wedge u_3^{t_2/t_3}}^{u_3^{t_2/t_3}}$$

$$= \left(u_1^{1/t_1} \wedge u_3^{1/t_3}\right)^{t_2} \left(u_3^{1/t_3}\right)^{t_3-t_2} + u_1 u_3^{1-t_2/t_3} \left[u_3^{t_2/t_3 - t_1/t_3} - \big(u_1^{1/t_1} \wedge u_3^{1/t_3}\big)^{t_2-t_1} \right]$$

$$= \big(u_1^{1/t_1} \wedge u_3^{1/t_3}\big)^{t_2} u_3^{1-t_2/t_3} + u_1 u_3^{1-t_1/t_3} - u_1 u_3^{1-t_2/t_3} \big(u_1^{1/t_1} \wedge u_3^{1/t_3}\big)^{t_2-t_1}$$

$$= \begin{cases} u_1^{t_2/t_1} u_3^{1-t_2/t_3} + u_1 u_3^{1-t_1/t_3} - u_1 u_3^{1-t_2/t_3} u_1^{(t_2-t_1)/t_1} = u_1 u_3^{1-t_1/t_3}, & u_1^{1/t_1} < u_3^{1/t_3}, \\ u_3 + u_1 u_3^{1-t_1/t_3} - u_1 u_3^{1-t_1/t_3} = u_3, & u_1^{1/t_1} \geq u_3^{1/t_3}. \end{cases}$$

As we had expected, this yields $C_{t_1,t_2} * C_{t_2,t_3} = C_{t_1,t_3}$ when comparing with equation (5.42).

We can now prove the above mentioned theorem within the copula framework:

5.45 Theorem
Let the situation be as above. Then $\{X_{L(n)} : n \in \mathbb{N}\}$ is a Markov process.

Proof. We will use the characterization provided by Theorem 5.44.
For $t_1 < t_2 < \cdots < t_n$, $t_i \in \mathbb{N}$, $n \geq 3$, we have to show

$$C_{t_1,\ldots,t_n} = C_{t_1,t_2} \star C_{t_2,t_3} \star \cdots \star C_{t_{n-1},t_n}.$$

We use induction w. r. t. n starting with $n = 3$. Take $u_1, u_2, u_3 \in (0,1)$. Then,

$$\big(C_{t_1,t_2} \star C_{t_2,t_3}\big)(u_1,u_2,u_3) := \int\limits_{(0,u_2)} D_2 C_{t_1,t_2}(u_1,z) \cdot \underbrace{D_1 C_{t_2,t_3}(z,u_3)}_{=0 \text{ for } z > u_3^{t_2/t_3}} \,\mathrm{d}z$$

$$= \int\limits_0^{u_2 \wedge u_3^{t_2/t_3}} D_2 C_{t_1,t_2}(u_1,z) \cdot u_3^{1-t_2/t_3} \,\mathrm{d}z = u_3^{1-t_2/t_3} \cdot \Big[C_{t_1,t_2}(u_1,z) \Big]_{z=0}^{z=u_2 \wedge u_3^{t_2/t_3}}$$

$$= u_3^{1-t_2/t_3} \left(u_1^{1/t_1} \wedge \big(u_2 \wedge u_2^{t_2/t_3}\big)^{1/t_2} \right)^{t_1} \left(\big(u_2 \wedge u_3^{t_2/t_1}\big)^{1/t_2} \right)^{t_2-t_1}$$

$$= u_3^{1-t_2/t_3} \big(u_1^{1/t_1} \wedge u_2^{1/t_2} \wedge u_3^{1/t_3}\big)^{t_1} \big(u_2^{1/t_2} \wedge u_3^{1/t_3}\big)^{t_2-t_1}$$

$$\overset{(5.45)}{=} C_{t_1,t_2,t_3}(u_1,u_2,u_3).$$

Assuming the validity for some $n \in \mathbb{N}$, $n \geq 3$, we get the validity for the case $n+1$:

$$\begin{aligned}
&\big(C_{t_1,t_2} \star \cdots \star C_{t_{n-1},t_n} \star C_{t_n,t_{n+1}}\big)(u_1,\dots,u_n,u_{n+1}) \\
&\overset{\text{“}\star\text{” associative}}{=} \big(C_{t_1,t_2} \star \cdots \star C_{t_{n-1},t_n}\big)(u_1,\dots,u_n) \;\star\; C_{t_n,t_{n+1}}(u_n,u_{n+1}) \\
&\overset{\text{assumption}}{=} C_{t_1,\dots,t_n}(u_1,\dots,u_n) \star C_{t_n,t_{n+1}}(u_n,u_{n+1}) \\
&\overset{\text{Def.}}{=} \int\limits_{(0,u_n)} D_n C_{t_1,\dots,t_n}(u_1,\dots,u_{n-1},z) \cdot D_1 C_{t_n,t_{n+1}}(z,u_{n+1})\,\mathrm{d}z \\
&= \int\limits_0^{u_n \wedge u_{n+1}^{t_n/t_{n+1}}} u_{n+1}^{1-t_n/t_{n+1}} D_n C_{t_1,\dots,t_n}(u_1,\dots,u_{n-1},z)\,\mathrm{d}z \\
&= u_{n+1}^{1-t_n/t_{n+1}} \Big[C_{t_1,\dots,t_n}(u_1,\dots,u_{n-1},z)\Big]_{z=0}^{z=u_n \wedge u_{n+1}^{t_n/t_{n+1}}} \\
&= u_{n+1}^{1-t_n/t_{n+1}} \Big(\bigwedge_{i=1}^{n-1} u_i^{1/t_i} \wedge \big(u_n \wedge u_{n+1}^{t_n/t_{n+1}}\big)^{1/t_n}\Big)^{t_1} \\
&\qquad \cdot \Big(\bigwedge_{i=2}^{n-1} u_i^{1/t_i} \wedge \big(u_n \wedge u_{n+1}^{t_n/t_{n+1}}\big)^{1/t_n}\Big)^{t_2-t_1} \cdots \big(u_n \wedge u_{n+1}^{t_n/t_{n+1}}\big)^{1-t_{n-1}/t_n} \\
&= u_{n+1}^{1-t_n/t_{n+1}} \big(\bigwedge_{i=1}^{n+1} u_i^{1/t_i}\big)^{t_1} \big(\bigwedge_{i=2}^{n+1} u_i^{1/t_i}\big)^{t_2-t_1} \cdots \big(\bigwedge_{i=n}^{n+1} u_i^{1/t_i}\big)^{t_n-t_{n-1}} \\
&= C_{t_1,\dots,t_{n+1}}(u_1,\dots,u_{n+1}).
\end{aligned}$$

This finishes the proof. □

5.46 Remark

This result easily yields Proposition 4.1 in Resnick (1987, p. 165) which says that the stationary transition probabilities of $\{X_{L(n)}\}$ are given by

$$\Pi\big(x,(y,\infty)\big) := \mathsf{P}(X_{L(n+1)} > y \mid X_{L(n)} = x) = \begin{cases} \frac{1-F(y)}{1-F(x)}, & y > x, \\ 1, & y \leq x, \end{cases} \tag{5.46}$$

as $\Pi\big(x,(y,\infty)\big) = 1 - D_1 C_{L(n),L(n+1)}\big(F^{L(n)}(x), F^{L(n+1)}(y)\big)$. F denotes the distribution function of the X_i as usual. ■

However, for absolutely continuous cdf F we have

$$F_{X_{L(r)}}(x) = 1 - \big(1 - F(x)\big) \sum_{j=0}^{r-1} \frac{1}{j!} \left(\log \frac{1}{1-F(x)}\right)^j$$

(Kamps, 1995, p. 32). Thus, an analytic derivation of the generalized inverse of $F_{X_{L(r)}}$ and therefore the copula of $(X_{L(r)}, X_{L(s)})$ does not seem to be possible.

5.4.2 Symmetry

A well-known property of standard Brownian motion is its symmetry in the sense of $(B_t) \stackrel{d}{=} (-B_t)$. We can now ask under which condition(s) a “generalized Brownian motion”, which is a process (X_t) with $C\langle(X_t)\rangle = C\langle(B_t)\rangle$ (copulas coincide but possibly other marginals) will be symmetric. Recalling that Brownian motion is a Markov process we can generalize this question:

Which condition(s) must be fulfilled by the copulas and the marginals of the process for X_t to be symmetric? In Section 5.2.3 a characterization for symmetry of general processes was given so that we can make use of it in the current context.

We will tackle the problem by considering a general Markov process $(X_t)_{t\in T}$ where $T = [0, \infty)$ without loss of generality.

To this end, the Markov property yields

$$\begin{aligned} \mathsf{P}(X_{t_1} \le x_1, \dots, X_{t_n} \le x_n) \\ &= \prod_{i=2}^{n} \mathsf{P}(X_{t_i} \le x_i \mid X_{t_1} \le x_1, \dots, X_{t_{i-1}} \le x_{i-1}) \cdot \mathsf{P}(X_{t_1} \le x_1) \\ &\stackrel{\text{Markov property}}{=} \prod_{i=2}^{n} \mathsf{P}(X_{t_i} \le x_i \mid X_{t_{i-1}} \le x_{i-1}) \cdot \mathsf{P}(X_{t_1} \le x_1) \\ &= \frac{\prod_{i=2}^{n} C_{t_{i-1},t_i}\big(F_{t_{i-1}}(x_{i-1}), F_{t_i}(x_i)\big)}{\prod_{i=2}^{n-1} F_{t_i}(x_i)} \end{aligned} \tag{5.47}$$

for all $t_1, \dots, t_n \in T$, $n \ge 2$, $x_1, \dots, x_n \in \mathbb{R}$. We see that the finite-dimensional distributions of the process are given by the bivariate copulas and the marginals.

On the other hand, we have

$$\begin{aligned} \mathsf{P}(X_{t_1} > x_1, \dots, X_{t_n} > x_n) \\ &= \prod_{i=2}^{n} \mathsf{P}(X_{t_i} > x_i \mid X_{t_1} > x_1, \dots, X_{t_{i-1}} > x_{i-1}) \cdot \mathsf{P}(X_{t_1} > x_1) \\ &\stackrel{\text{Markov property}}{=} \prod_{i=2}^{n} \mathsf{P}(X_{t_i} > x_i \mid X_{t_{i-1}} > x_{i-1}) \cdot \mathsf{P}(X_{t_1} > x_1) \\ &= \frac{\prod_{i=2}^{n} \widehat{C}_{t_{i-1},t_i}\big(\overline{F}_{t_{i-1}}(x_{i-1}), \overline{F}_{t_i}(x_i)\big)}{\prod_{i=2}^{n-1} \overline{F}_{t_i}(x_i)}. \end{aligned} \tag{5.48}$$

In view of Theorem 5.21 it is clear that the conditions for symmetry of a Markov process only depend on the marginals and the *bivariate* copulas $C_{s,t}$. Thus, the process is symmetric iff

$$F_t(x) = \overline{F}_t(-x) \tag{5.49}$$

for all $t \in T$, $x \in \mathbb{R}$ (marginal symmetry) and for all $s < t$:

$$C_{s,t}(u, v) \stackrel{!}{=} \widehat{C}_{s,t}(u, v) = u + v - 1 + C_{s,t}(1 - u, 1 - v) \text{ for all } u, v \in (0, 1), \tag{5.50}$$

where $\widehat{C}_{s,t}$ is the "survival copula" of (X_s, X_t) (see Section 2.4.2).

5.47 Example
For the EFGM copula (see Section 5.3.2)

$$C_\alpha(u,v) = uv(1-\alpha(1-u)(1-v)), \quad u,v \in I,\ \alpha \in [-1,1],$$

it is easyily shown by elementary calculations that $C_\alpha = \widehat{C}_\alpha$ so that this family generates symmetric Markov processes. ■

It is well known that Brownian motion is symmetric about 0 (Karatzas and Shreve, 2000, Lemma 9.4, p. 104). In the following example (which ends on page 84) we will give an alternative proof for the symmetry with the help of Theorem 5.21. Note that we only need to check condition (5.50) for the bivariate copulas due to the above remarks.

5.48 Example (symmetry of Brownian motion)
We have $F_t(x) = \Phi(x/\sqrt{t})$ and with $\Phi(x) = 1-\Phi(-x)$ property *a)* of Theorem 5.21 follows, i. e., $\Phi(x/\sqrt{t}) = \overline{\Phi}(-x/\sqrt{t})$.

To check condition *b)*, we need the following result from Heuser (1988, Satz 167.5, p. 279):

5.49 Theorem
Let $G \subset \mathbb{R}^p$ be an open connected set and $f = (f_1, \ldots, f_q) : G \to \mathbb{R}^q$, $p, q \in \mathbb{N}$. If the partial derivatives of all component functions $f_1, \ldots, f_q$ of f exist and vanish on G, then f is constant.

We can easily deduce:

5.50 Corollary
Let $G \subset \mathbb{R}^p$ be an open connected set and $p, q \in \mathbb{N}$. If $f = (f_1, \ldots, f_q), g = (g_1, \ldots, g_q) : G \to \mathbb{R}^q$ and all partial derivatives are the same, i. e., $\frac{\partial}{\partial x_i} f_j(x_1, \ldots, x_p) = \frac{\partial}{\partial x_i} g_j(x_1, \ldots, x_p)$ for all $1 \le i \le p$, $1 \le j \le q$, $(x_1, \ldots, x_p) \in G$, and there exists $\boldsymbol{x_0} \in G$ with $f(\boldsymbol{x_0}) = g(\boldsymbol{x_0})$, then $f \equiv g$ on G.

Proof. Set $h = f - g$. Then $\frac{\partial}{\partial x_j} h(x_1, \ldots, x_p) = 0$, $1 \le j \le p$ so that by Theorem 5.49 $f(\boldsymbol{x}) = g(\boldsymbol{x}) + \boldsymbol{c}$ for some $\boldsymbol{c} \in \mathbb{R}^q$ for all $\boldsymbol{x} \in \mathbb{R}^p$. As $f(\boldsymbol{x_0}) = g(\boldsymbol{x_0})$ we have $\boldsymbol{c} = \boldsymbol{0}$. □

5.51 Remark
If f and g are continuous on $\overline{G}$ (the closure of G), then $f(\boldsymbol{x_0}) = g(\boldsymbol{x_0})$ for some $\boldsymbol{x_0} \in \partial G := \overline{G} \setminus \operatorname{int} G$ (the boundary, "int" denotes the interior of G) will be sufficient for the above corollary to hold. ■

Due to Corollary 5.50 and the following Remark 5.51 it is sufficient for proving symmetry to check if the partial derivatives of $C^B_{s,t}$ and $\widehat{C}^B_{s,t}$ coincide on $(0,1)^2$ where $C^B_{s,t}$ is the copula of Brownian motion (see Example 5.23).[9]

Let $(u,v) \in (0,1)^2$, $0 < s < t < \infty$. It was shown in (5.16) and (5.17) that

$$D_1 C^B_{s,t}(u,v) = \Phi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u)}{\sqrt{t-s}}\right)$$

and

$$D_2 C^B_{s,t}(u,v) = \sqrt{\frac{t}{t-s}} \cdot \frac{1}{\varphi\big(\Phi^{-1}(v)\big)} \int\limits_0^u \varphi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(w)}{\sqrt{t-s}}\right) \mathrm{d}w.$$

Now, from (5.50), we have

$$\begin{aligned}
D_1 \widehat{C}^B_{s,t}(u,v) &= 1 - D_1 C^B_{s,t}(1-u, 1-v) = 1 - \Phi\left(\frac{\sqrt{t}\Phi^{-1}(1-v) - \sqrt{s}\Phi^{-1}(1-u)}{\sqrt{t-s}}\right) \\
&\overset{\Phi^{-1}(u) = -\Phi^{-1}(1-u)}{=} 1 - \Phi\left(-\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u)}{\sqrt{t-s}}\right) \\
&= \Phi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(u)}{\sqrt{t-s}}\right) = D_1 C^B_{s,t}(u,v) \quad \text{as required.}
\end{aligned}$$

For the second partial derivative we have

$$\begin{aligned}
D_2 \widehat{C}^B_{s,t}(u,v) &= 1 - D_2 C^B_{s,t}(1-u, 1-v) \\
&= 1 - \sqrt{\frac{t}{t-s}} \cdot \frac{1}{\varphi\big(\Phi^{-1}(1-v)\big)} \int\limits_0^{1-u} \varphi\left(\frac{\sqrt{t}\Phi^{-1}(1-v) - \sqrt{s}\Phi^{-1}(w)}{\sqrt{t-s}}\right) \mathrm{d}w \\
&\overset{\substack{\Phi^{-1}(1-v) = -\Phi^{-1}(v) \\ \text{and } \varphi \text{ symmetric}}}{=} 1 - \sqrt{\frac{t}{t-s}} \cdot \frac{1}{\varphi\big(\Phi^{-1}(v)\big)} \int\limits_0^{1-u} \varphi\left(\frac{-\sqrt{t}\Phi^{-1}(v) + \sqrt{s}\Phi^{-1}(1-w)}{\sqrt{t-s}}\right) \mathrm{d}w \\
&= 1 - \sqrt{\frac{t}{t-s}} \cdot \frac{1}{\varphi\big(\Phi^{-1}(v)\big)} \int\limits_0^{1-u} \varphi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(1-w)}{\sqrt{t-s}}\right) \mathrm{d}w \\
&\overset{1-w=x}{=} 1 - \sqrt{\frac{t}{t-s}} \cdot \frac{1}{\varphi\big(\Phi^{-1}(v)\big)} \int\limits_u^{1} \varphi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(x)}{\sqrt{t-s}}\right) \mathrm{d}x.
\end{aligned}$$

[9] Note that both $C^B_{s,t}$ and $\widehat{C}^B_{s,t}$ are copulas so that they are continuous on I^2 and all values $C^B_{s,t}(u,v)$ and $\widehat{C}^B_{s,t}(u,v)$ are the same for $u \in \{0,1\}$ or $v \in \{0,1\}$.

Thus,

$$D_2 C^B_{s,t}(u,v) - D_2 \widehat{C}^B_{s,t}(u,v) = \sqrt{\frac{t}{t-s}} \cdot \frac{1}{\varphi\big(\Phi^{-1}(v)\big)} \int_0^1 \varphi\left(\frac{\sqrt{t}\Phi^{-1}(v) - \sqrt{s}\Phi^{-1}(x)}{\sqrt{t-s}}\right) \mathrm{d}x - 1$$
$$\overset{(5.18)}{=} \int_0^1 c^B_{s,t}(x,v)\,\mathrm{d}x - 1 = \mathbb{1}_{(0,1)}(v) - 1 = 0 \tag{5.51}$$

for $v \in (0,1)$.

We have shown that *b)* of Theorem 5.21 holds so that we have an alternative proof for the symmetry of Brownian motion. (End of Example 5.48) ■

The question arises if there is a characterization of those classes of bivariate copulas, i. e., of a set $\mathcal{C}^{sym} \subset \mathcal{C}_2$, which generate symmetric Markov processes by solving the functional equation $C = \widehat{C}$.

For a wide class of copulas, the Archimedean family (see Section 2.3), we can give a positive answer due to the work of Frank (1979). The answer leads to a special class of Archimedean copulas, the Frank family which is the only Archimedean family with $\widehat{C}(u,v) = C(u,v)$ for all $u, v \in I$. This is exactly the desired property.

5.52 Definition (Frank, 1979, or Nelsen, 1999, p. 94)
For $\vartheta \in \overline{\mathbb{R}}$ define

$$C_\vartheta(u,v) := \begin{cases} -\frac{1}{\vartheta} \ln\left[1 + \frac{(\exp(-\vartheta u)-1)(\exp(-\vartheta v)-1)}{\exp(-\vartheta)-1}\right], & \vartheta \in \mathbb{R} \setminus \{0\}, \\ \Pi(u,v), & \vartheta = 0, \\ M(u,v), & \vartheta = \infty, \\ W(u,v), & \vartheta = -\infty. \end{cases} \tag{5.52}$$

Then this family of copulas is called Frank family, *each member is called a* Frank copula.

Please refer to Figure 5.2 for a visual impression of the density when the margins are $\mathsf{N}(0,1)$.

5.53 Remark

a) The characteristic generator of C_ϑ for $\vartheta \in \mathbb{R} \setminus \{0\}$ is $\varphi_\vartheta(t) = -\ln\left(\frac{e^{-\vartheta t}-1}{e^{-\vartheta}-1}\right)$ (see Section 2.3).

b) The definition of $C_{-\infty}$, C_0 and C_∞ is motivated by the fact that these copulas are the limits of C_ϑ, e. g., $\lim_{\vartheta \to 0} C_\vartheta(u,v) = \Pi(u,v)$. They have been added for completeness as they also generate symmetric Markov processes.

c) Some of the statistical properties of this family were discussed in Nelsen (1986) and Genest (1987).

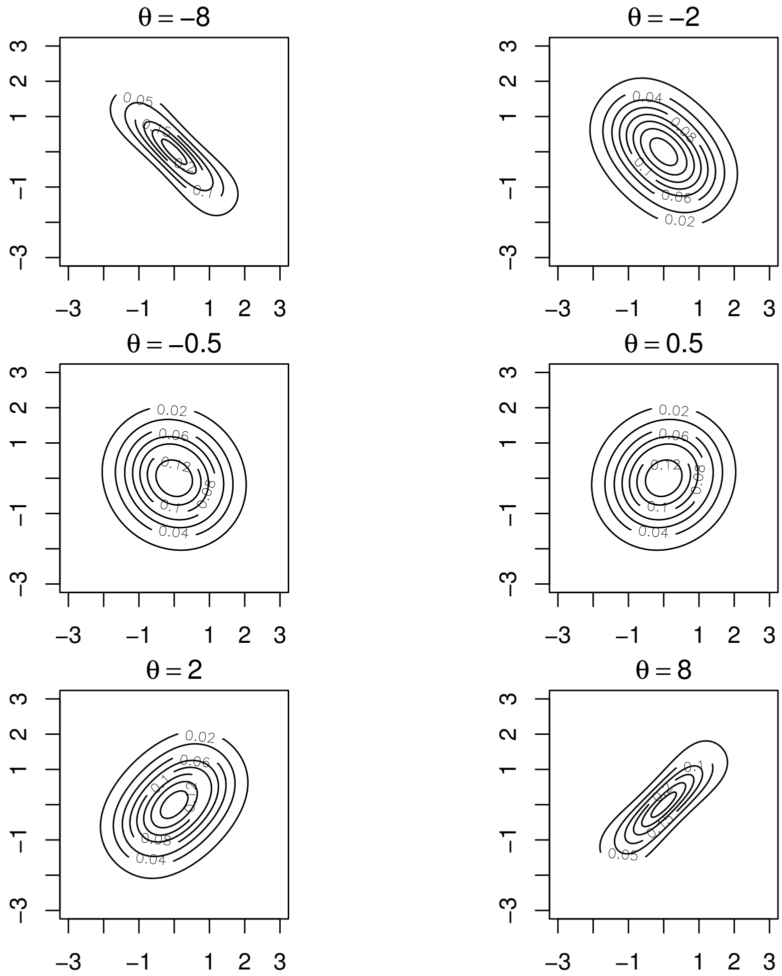

FIGURE 5.2. Contour plots of the Frank copula density with standard normal margins.

d) C_ϑ is radially symmetric for all $\vartheta \in \overline{\mathbb{R}}$. This could be checked tediously, but we get this from Theorem 5.21 and the following Theorem 5.54.

■

By investigating associativity of copulas and certain functional equations, Frank (1979) derived the following theorem:

5.54 Theorem (Frank, 1979, Theorem 4.1)
The functions $\{C_\vartheta : \vartheta \in \mathbb{R} \setminus \{0\}\}$ *are the only bivariate Archimedean copulas which satisfy* $C = \widehat{C}$.

Thus, this result yields a characterization of symmetric Markov processes whose copulas are in the Archimedean class.

5.5 Continuous Local Martingales

In this section we will use the technique of time transformation to show how the copula of continuous local martingales can be determined. The idea of time transformation is a rather old one, already used by Dambis (1965) and Dubins and Schwarz (1965). Their results provide us with a very elegant way to derive the copulas of the process. We have to assume that the usual conditions hold.

5.55 Definition (continuous local martingale)
*An adapted, right continuous process X is a $(\mathcal{F}_t, \mathsf{P})$-*local martingale *if there exist stopping times T_n, $n \geq 1$, such that*

a) $T_{n+1} \geq T_n$ for all n and $\lim_{n\to\infty} T_n = \infty$ almost surely, and

b) for every n, the stopped process[10] *X^{T_n} is a $(\mathcal{F}_t, \mathsf{P})$-martingale.*

We write $X \in \mathcal{M}^{loc}$ for a local martingale and $X \in \mathcal{M}^{c,loc}$ if it is also continuous.

To understand and apply the following Theorem 5.58 we need the "quadratic variation" $\langle X \rangle$ of a process X. However, neither is it possible nor does it make sense to develop the complete theory here. Many excellent books have been written on the subject of stochastic integrals and differential equations. As already mentioned in the introduction the reader should take monographs such as Øksendal (1998), Rogers and Williams (2000a,b), Revuz and Yor (1991) or Karatzas and Shreve (2000) as a reference. Only the necessary definitions and results will be stated here.

5.56 Definition (square integrable)
Let $X = (X_t)$ be a right-continuous martingale. We say that X is square-integrable *if $\mathsf{E}(X_t^2) < \infty$ for all $t \in T$. If, in addition, $X_0 = 0$ a. s., we write $X \in \mathcal{M}_2$ (or $X \in \mathcal{M}_2^c$ if X is also pathwise continuous).*

5.57 Definition (quadratic variation, cross variation)
For $X \in \mathcal{M}_2$, the quadratic variation (process) *$\langle X \rangle = (\langle X \rangle_t)_{t\in T}$ of X is defined to be the unique (up to indistinguishability) adapted, increasing, natural process for which $\langle X \rangle_0 = 0$ a. s. and $X^2 - \langle X \rangle$ is a martingale.*

Here, a process is natural *if for P-a. e. $\omega \in \Omega$ the process is increasing and right-continuous.*

For any two martingales $X, Y \in \mathcal{M}_2$, we define their cross-variation process *$\langle X, Y \rangle$ by*

$$\langle X, Y \rangle_t := \frac{1}{4} \left[\langle X + Y \rangle_t - \langle X - Y \rangle_t \right], \quad t \in T. \tag{5.53}$$

For a martingale M we clearly have

$$\langle M, M \rangle_t := \frac{1}{4} \left[\langle M + M \rangle_t - \langle M - M \rangle_t \right] = \langle M \rangle_t \tag{5.54}$$

[10] see Definition 5.8

as $\langle 0 \rangle_t \equiv 0$ and $\langle 2M \rangle_t = 4\langle M \rangle_t$ for all $t \geq 0$ due to

$$(2M_t)^2 - \langle 2M \rangle_t = 4\,\underbrace{(M_t^2 - \langle M \rangle_t)}_{\text{martingale}}.$$

One can think of $\langle X \rangle$ as the increasing process of the Doob-Meyer-decomposition of X (see Karatzas and Shreve, 2000, Theorem 4.10, pp. 24).

It can be shown that cross variation is bilinear.

5.58 Theorem (Dambis, 1965; Dubins and Schwarz, 1965)
If M is a $(\mathcal{F}_t, \mathsf{P})$-local martingale vanishing at 0 and such that $\langle M \rangle_t \xrightarrow{t\to\infty} \infty$ a. s., and if we set

$$T_t := \inf\{s : \langle M \rangle_s > t\}, \tag{5.55}$$

then $B_t := M_{T_t}$ is a $(\mathcal{F}_{T_t})$-Brownian motion and $M_t = B_{\langle M \rangle_t}$ for all $t \in T$.

For a proof, see Karatzas and Shreve (2000, Theorem 4.6, pp. 174) or Revuz and Yor (1991).

This theorem states that any continuous local martingale can be time transformed into a Brownian motion. Thus, the problem of determining the copula is changed into the problem of determining $\langle M \rangle_t$. As we have already determined the copula of Brownian motion in Example 5.23, the next Corollary can be given without a proof as it is obvious from the above Theorem and (5.15).

5.59 Corollary
If M is a $(\mathcal{F}_t, \mathsf{P})$-local martingale vanishing at 0 and such that $h(t) := \langle M \rangle_t \xrightarrow{t\to\infty} \infty$, the bivariate copulas of the process are given by

$$C_{s,t}(u, v) = \int_0^u \Phi\left(\frac{\sqrt{h(t)}\Phi^{-1}(v) - \sqrt{h(s)}\Phi^{-1}(w)}{\sqrt{h(t) - h(s)}}\right) \mathrm{d}w, \quad 0 \leq s < t. \tag{5.56}$$

Let us give some examples for the application of Corollary 5.59. B_t always denotes standard Brownian motion.

5.60 Example (Ornstein-Uhlenbeck process)
Let $\alpha, \sigma > 0$ and X_0 be a random variable on $(\Omega, \mathcal{F})$. Then the *Ornstein-Uhlenbeck process* is given by the stochastic differential equation (SDE)

$$\mathrm{d}X_t = -\alpha X_t \,\mathrm{d}t + \sigma \,\mathrm{d}B_t, \quad t > 0, \tag{5.57}$$

(see, e. g., Karatzas and Shreve, 2000, p. 358, or Øksendal, 1998, p. 74). The solution is given by

$$X_t = X_0 \cdot \exp(-\alpha t) + \sigma \int_0^t \exp(-\alpha(t - s)) \,\mathrm{d}B_s, \qquad 0 \leq t < \infty. \tag{5.58}$$

However, (X_t) is not a martingale if $X_0 \neq 0$ on a set $A \in \mathcal{F}$ with $\mathsf{P}(A) > 0$. This is due to

$$\begin{aligned}
\mathsf{E}(X_t \mid \mathcal{F}_s) &= X_0 \cdot \exp(-\alpha t) + \sigma \cdot \mathsf{E}\Big(\int_0^t \exp(-\alpha(t-u))\,\mathrm{d}B_u \mid \mathcal{F}_s\Big) \\
&= X_0 \cdot \exp(-\alpha t) + \sigma \int_0^s \exp(-\alpha(t-u))\,\mathrm{d}B_u \\
&= X_s + X_0 \cdot \exp(-\alpha(t-s)), \quad 0 \le s < t,
\end{aligned}$$

where the second last equality is valid because $\int_0^t \exp(-\alpha(t-u))\,\mathrm{d}B_u$ is a martingale w. r. t. $(\mathcal{F}_t)$ (see Øksendal, 1998, Corollary 3.2.6 or Karatzas and Shreve, 2000, Proposition 3.2.10, p. 139). If we set

$$Y_t := \exp(\alpha t) X_t = X_0 + \sigma \int_0^t \exp(\alpha s)\,\mathrm{d}B_s \tag{5.59}$$

we see that $(Y_t)_{t \ge 0}$ is a martingale and

$$M_t := Y_t - X_0, \quad t \ge 0,$$

is a martingale vanishing at 0. We then get

$$\begin{aligned}
\langle M \rangle_t &\overset{(5.54)}{=} \langle Y - X_0, Y - X_0 \rangle_t = \langle \sigma \int_0^t \exp(\alpha s)\,\mathrm{d}B_s, \sigma \int_0^t \exp(\alpha s)\,\mathrm{d}B_s \rangle \\
&\overset{\langle \cdot,\cdot \rangle \text{ bilinear}}{=} \sigma^2 \langle \int_0^t \exp(\alpha s)\,\mathrm{d}B_s, \int_0^t \exp(\alpha s)\,\mathrm{d}B_s \rangle \overset{(*)}{=} \sigma^2 \int_0^t \exp(2\alpha s)\,\mathrm{d}s \\
&= \frac{\sigma^2}{2\alpha}(\exp(2\alpha t) - 1),
\end{aligned} \tag{5.60}$$

where $(*)$ is due to the so-called “Itô-isometry” (see Karatzas and Shreve, 2000, equation (2.19) on page 138 and Proposition 2.10 thereafter). Note that $\langle M \rangle_t \xrightarrow{t \to \infty} \infty$ as required by Theorem 5.58.

As $X_t = \exp(-\alpha t) Y_t$ is a strictly monotone increasing transformation of Y_t the copulas of (X_s, X_t) and (Y_s, Y_t) are equal (cf. Theorem 2.15). The same applies to $Y_t = f(M_t)$ with $f(x) := x + X_0$. The bivariate copulas for the Ornstein-Uhlenbeck process are therefore given by

$$\begin{aligned}
C^{\mathrm{OU}}_{s,t}(u,v) &= \int_0^u \Phi\left(\sqrt{\frac{h(t)}{h(t)-h(s)}} \cdot \Phi^{-1}(v) - \sqrt{\frac{h(s)}{h(t)-h(s)}} \cdot \Phi^{-1}(w)\right) \mathrm{d}w \\
&= \int_0^u \Phi\left(\sqrt{\frac{\exp(2\alpha t)-1}{\exp(2\alpha t)-\exp(2\alpha s)}} \cdot \Phi^{-1}(v) - \sqrt{\frac{\exp(2\alpha s)-1}{\exp(2\alpha t)-\exp(2\alpha s)}} \cdot \Phi^{-1}(w)\right) \mathrm{d}w
\end{aligned} \tag{5.61}$$

for $0 \le s < t$ and where $h(t) \equiv h(t;\alpha,\sigma) = \sigma^2(2\alpha)^{-1}(\exp(2\alpha t) - 1)$.

As in the case of the Brownian copula, we have

$$\lim_{t\to\infty} C^{\mathrm{OU}}_{s,t}(u,v) = \int_0^u \Phi\big(\Phi^{-1}(v)\big)\,\mathrm{d}z = uv = \Pi(u,v),$$

i.e., X_s and X_t are asymptotically independent. Note that Lebesgue's dominated convergence theorem must be used to interchange integration and the taking of the limit.

The copula is also independent of σ. This means that α alone determines the dependence between X_s and X_t.

For the parameter α, we have the following limiting copulas:

$\boxed{\alpha \to \infty :}$

We have

$$\frac{e^{2\alpha t} - 1}{e^{2\alpha t} - e^{2\alpha s}} = \frac{1 - e^{-2\alpha t}}{1 - e^{-2\alpha(t-s)}} \xrightarrow{\alpha\to\infty} 1$$

and

$$\frac{e^{2\alpha s} - 1}{e^{2\alpha t} - e^{2\alpha s}} = \frac{1 - e^{-2\alpha s}}{e^{2\alpha(t-s)} - 1} \xrightarrow{\alpha\to\infty} 0$$

from which we get

$$\lim_{\alpha\to\infty} C^{\mathrm{OU}}_{s,t}(u,v) = \int_0^u \Phi(\Phi^{-1}(v))\,\mathrm{d}s = uv = \Pi(u,v),$$

the independence copula.

$\boxed{\alpha \to 0 :}$

In this case, we use the rule of de l'Hospital (case "0/0") to obtain

$$\lim_{\alpha\to 0} \frac{1 - e^{-2\alpha t}}{1 - e^{-2\alpha(t-s)}} = \lim_{\alpha\to 0} \frac{2te^{-2\alpha t}}{2(t-s)e^{-2\alpha(t-s)}} = \frac{t}{t-s}$$

and

$$\lim_{\alpha\to 0} \frac{1 - e^{-2\alpha s}}{e^{2\alpha(t-s)} - 1} = \lim_{\alpha\to 0} \frac{2se^{-2\alpha s}}{2(t-s)e^{-2\alpha(t-s)}} = \frac{s}{t-s}$$

so that

$$\lim_{\alpha \to 0} C^{\mathrm{OU}}_{s,t} = C^{B}_{s,t}$$

where $C^B_{s,t}$ is the Brownian copula. This result is not surprising in view of the defining SDE (5.57) and its solution (5.58): Letting $\alpha \to 0$ in (5.58), we get

$$X_t = X_0 + \sigma \int_0^t \mathrm{d}B_s = X_0 + \sigma B_t.$$

This again is only a strictly monotone increasing transformation of the Brownian motion (B_t), so the copulas must be the same. The interchange of the integral and limit is justified by monotone convergence; it is standard to show it for elementary functions and then generalize it by monotone class theorems. ∎

The next example is somewhat different as the time parameter t is bounded.

5.61 Example (Brownian bridge)
Let $a, b, T \in \mathbb{R}$, $a < b$, $T > 0$, and consider the one-dimensional SDE

$$\mathrm{d}X_t = \frac{b - X_t}{T - t}\,\mathrm{d}t + \mathrm{d}B_t, \quad 0 \le t < T,\ X_0 = a. \tag{5.62}$$

It can be shown (Karatzas and Shreve, 2000, pp. 358 or Øksendal, 1998, p. 75) that the solution is given by

$$X_t = a\left(1 - \frac{t}{T}\right) + b\frac{t}{T} + (T - t)\int_0^t \frac{\mathrm{d}B_s}{T - s}, \qquad 0 \le t < T. \tag{5.63}$$

The process can be thought of to be a "linear bridge" between the points $(0, a)$ and (T, b) which is disturbed by a time-scaled Brownian motion.

Setting $Y_t := \int_0^t \frac{\mathrm{d}B_s}{T-s}$, we see that, again, $X_t = f(Y_t)$ for $f(x) \equiv f(x; a, b, T, t) = a(1 - \frac{t}{T}) + b(\frac{t}{T}) + (T - t)x$ so that the copulas must coincide. $(Y_t)_{0 \le t < T}$ is a continuous martingale with $Y_0 = 0$. In analogy to Example 5.60 we calculate the quadratic variation process to be

$$\langle Y \rangle_t = \int_0^t \frac{1}{(T - s)^2}\,\mathrm{d}s = (T - s)^{-1}\Big|_0^t = \frac{1}{T - t} - \frac{1}{T}. \tag{5.64}$$

One problem remains: We do not have $\langle Y \rangle_t \xrightarrow{t \to \infty} \infty$, but we do not need this requirement. Taking a look at the proof of Theorem 5.58 in Karatzas and Shreve (2000, pp. 174) we see that in the case of a bounded time interval $[0, T)$, we need the condition $\lim_{t \to T} \langle Y \rangle_t = \infty$ which is met by (5.64). For $h(t) = \frac{1}{T-t} - \frac{1}{T}$, $0 \le t < T$, we get

$$\sqrt{\frac{h(t)}{h(t) - h(s)}} = \sqrt{\frac{t(T - s)}{T(t - s)}} \quad \text{and} \quad \sqrt{\frac{h(s)}{h(t) - h(s)}} = \sqrt{\frac{s(T - t)}{T(t - s)}}.$$

The copula of the Brownian bridge is therefore given by

$$C_{s,t}^{\mathrm{BB}}(u,v) = \int_0^u \Phi\left(\frac{\sqrt{t(T-s)}\Phi^{-1}(v) - \sqrt{s(T-t)}\Phi^{-1}(w)}{\sqrt{T(t-s)}}\right) \mathrm{d}w \tag{5.65}$$

for $0 \leq s < t < T$. ■

5.62 Remark

Note that $h(t) = \langle M, M \rangle_t$ is usually a *stochastic* process, i.e., one should write $h(t,\omega)$. However, in the above examples, $h(t)$ turned out to be a deterministic function which is due to the fact that the quadratic variation of Brownian motion is only a function of time which is a consequence of the Itô-isometry.

It is this fact which makes the theory of time transformation applicable to the derivation of copulas, for, if the time transformation were stochastic, we would end up with a copula with stochastic time. With reference to equation (5.56) we had to write $C_{s,t}(u,v;\omega)$ and $h(t,\omega)$, yielding a random copula. This would be of only theoretical interest.

Nonetheless, from a theoretical point of view, the central result is that of Monroe (1978) who showed that a process is equivalent to a time change of Brownian motion if and only if it is a local semimartingale. This is the most general result in that direction. Note that—for example—Lévy processes fall into this class. ■

It is well-known that stochastic integrals are Markov processes, and therefore, the above examples also yield bivariate copulas generating Markov processes. In accordance with the remarks on page 62 we are now free to construct new processes with the intertemporal dependence given by the above copulas but with arbitrary marginals.

Appendix A
General Results from Probability Theory

A.1 Generalized Inverse

A.1 Definition (generalized inverse)
Let $F : \mathbb{R} \to [0,1]$ *be a distribution function. Then the* generalized inverse *or* quantile function $F^{-1} : [0,1] \to \overline{\mathbb{R}}$ *of* F *is defined as*

$$F^{-1}(t) := \begin{cases} \sup\{x \in \mathbb{R} : F(x) = 0\}, & t = 0, \\ \inf\{x \in \mathbb{R} : F(x) \geq t\}, & 0 < t \leq 1, \end{cases} \tag{A.1}$$

where $\inf(\emptyset) := \infty$, $\sup(\emptyset) := -\infty$.

The generalized inverse has the following properties:

A.2 Lemma
Let F, F^{-1} *be as in Definition A.1.*

a) F^{-1} *is increasing on* $(0,1)$ *and left-continuous.*

b) $F\big(F^{-1}(y)\big) \geq y$ *for all* $0 < y < 1$.

c) If F *is continuous in* $F^{-1}(y)$ $(0 < y < 1)$, *then* $F\big(F^{-1}(y)\big) = y$.

d) $F^{-1}\big(F(x)\big) \leq x$, $x \in \mathbb{R}$.

e) If F^{-1} *is continuous in* $F(x)$ $(x \in \mathbb{R})$ *and* $0 < F(x) < 1$, *then* $F^{-1}\big(F(x)\big) = x$.

Proof. See Pfeifer (1989, Lemma 1.1). □

A useful consequence of Lemma A.2 is

A.3 Lemma
Let X *be a real-valued random variable with continuous distribution function* F. *Then* $F(X) \sim \mathsf{U}(0,1)$. *If* F *is an arbitrary distribution function and* $U \sim \mathsf{U}(0,1)$, *then* $F^{-1}(U)$ *has distribution function* F.

For a proof, see Pfeifer (1989, Lemma 1.2).

We introduce the following notation:

A.4 Notation
Let F be a distribution function. Then

$$\begin{aligned}\alpha(F) &\equiv \alpha_F := \inf\{x \in \mathbb{R} : F(x) > 0\},\\ \omega(F) &\equiv \omega_F := \sup\{x \in \mathbb{R} : F(x) < 1\},\end{aligned}$$

(with $\inf(\emptyset) := \infty$, $\sup(\emptyset) := -\infty$) denote the right and the left endpoint of the support of F.

A.2 Probability Distributions

Uniform Distribution

Symbol:	$\mathsf{U}(a,b)$, $a,b \in \mathbb{R}$, $a<b$
Density:	$f(x) = (b-a)^{-1} \cdot \mathbb{1}_{(a,b)}(x)$
Dominating measure:	Lebesgue measure on $[a,b]$
Mean:	$(a+b)/2$
Variance:	$(b-a)^2/12$

Normal Distribution

Symbol:	$\mathsf{N}(\mu,\sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$
Density:	$f(x) = (\sqrt{2\pi}\sigma)^{-1} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
Dominating measure:	Lebesgue measure on $\mathbb{R}$
Mean:	μ
Variance:	σ^2

t-Distribution

Symbol:	$t_\nu(\mu,\sigma^2)$; t_ν for $\mu=0$, $\sigma=1$; $\mu \in \mathbb{R}$, $\sigma>0$, $\nu \in \mathbb{N}$
Density:	$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})\sigma}\left(1+\frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}$
Dominating measure:	Lebesgue measure on $\mathbb{R}$
Mean:	μ, if $\nu > 1$
Variance:	$\sigma^2\frac{\nu}{\nu-2}$, if $\nu > 2$

Appendix B

Simulation/Random Variate Generation

In most cases of the application of copulas no closed form solutions exist (e. g., portfolio analysis and option pricing in mathematical finance) so that there is a great interest in simulation and Monte Carlo studies.

We will therefore give a short introduction to the generation of random samples from a given joint distribution. We assume that procedures for generating independent uniform variates are known to the reader (e. g., Lange, 1998, pp. 269, and Devroye, 1986, for alternative methods). From these uniform samples we can generate samples x of a random variable X with distribution function F by the following two steps:

1. Generate u from $\mathsf{U}(0, 1)$.
2. Set $x = F^{-1}(u)$ where F^{-1} is the generalized inverse of F (see page 93).

Here we will focus on generating random samples from a copula. No proofs will be given as they can be found in Nelsen (1999) or follow directly as in the case of the generation of time-series from bivariate copulas.

By virtue of Sklar's theorem, we only need to generate a pair (u, v) of $\mathsf{U}(0, 1)$ variables (U, V) with copula C (which is also the copula of $X := F^{-1}(U)$ and $Y := G^{-1}(V)$) and then transform this pair by the quantile transformation.

B.1 The Conditional Distribution Method

For this method we need the conditional distribution function of V given $U = u$ which we denote by $c_u(v)$ as a shorthand notation, $c_u(v) := \mathsf{P}(V \le v \mid U = u)$, $u, v \in (0, 1)$. From Lemma 2.22 we have

$$c_u(v) = D_1 C(u, v)$$

so that the following algorithm yields a sample (u, v) from a copula C (recall that $c_u(v)$ exists almost everywhere for any copula C):

1. Generate u and t independently from $\mathsf{U}(0,1)$
2. Set $v = c_u^{(-1)}(t)$ where $c_u^{(-1)}$ denotes the generalized inverse of c_u
3. The desired pair is (u, v)

More generally, for a n-copula we have:

Let $C_{1,\dots,k}(u_1,\dots,u_k)$ be the copula of $(U_1,\dots,U_k)$, $2 \le k \le n$, and set $C_1(u_1) := u_1$, $C_n(u_1,\dots,u_n) := C(u_1,\dots,u_n)$. From Theorem 2.27 we then have

$$C_k(u_k \mid u_1,\dots,u_{k-1}) := \mathsf{P}(U_k \le u_k \mid U_1 = u_1,\dots,U_{k-1} = u_{k-1}) \\ = \frac{D_{1,\dots,k-1}C_{1,\dots,k}(u_1,\dots,u_k)}{D_{1,\dots,k-1}C_{1,\dots,k-1}(u_1,\dots,u_{k-1})}.$$

We can deduce the following algorithm:

1. Simulate u_1 from $\mathsf{U}(0,1)$
2. Simulate u_2 from the conditional distribution function $C_2(u_2 \mid u_1)$
3. Simulate u_3 from the conditional distribution function $C_3(u_3 \mid u_1, u_2)$

$\vdots$

n. Simulate u_n from $C_n(u_n \mid u_1, \dots u_{n-1})$

The vector $(u_1,\dots,u_n)$ is the desired sample from C

Simulating u_k from $C_k(u_k \mid u_1,\dots,u_{k-1})$ can be done by simulating q from $\mathsf{U}(0,1)$ from which $u_k = C_k^{(-1)}(q \mid u_1,\dots,u_{k-1})$ can be obtained by solving $q = C_k(u_k \mid u_1,\dots,u_{k-1})$ w. r. t. u_k[1] if the inverse C_k^{-1} of C_k does not exist in closed form.

Because of the potential numeric root finding, this general algorithm is often not suitable. However, other and more efficient algorithms exist only in special cases such as the Archimedean family, as we will see in the next section.

B.2 Simulating from Archimedean Copulas

Recall (cf. Section 2.3) that an Archimedean copula is defined by its generator φ and that $K_\varphi(t) := t - \varphi(t)/\varphi'(t+)$, $t \in (0,1)$, is the cumulative distribution function of the variable $Z = C(U,V)$, $U, V \sim \mathsf{U}(0,1)$ (see Corollary 2.34 and Genest and Rivest, 1993). We can then use the following algorithm for random variate generation from C:

1. Simulate s and q independently from $\mathsf{U}(0,1)$

[1] E. g., by numerical root finding.

2. Set $t = \lambda_\varphi^{-1}(q)$ where λ_φ^{-1} is the generalized inverse of $\lambda_\varphi(v) = \varphi(v)/\varphi'(v+)$, $0 < v \leq 1$.
3. Set $u = \varphi^{-1}\big(s\varphi(t)\big)$, $v = \varphi^{-1}\big((1-s)\varphi(t)\big)$

   The desired pair is (u, v).

B.3 Generation of Time-Series from Bivariate Copulas

The following algorithm is a consequence of the conditional distribution method presented in Section B.1.

Let $0 \leq t_0 < t_1 < t_2 < \cdots < t_n < \infty$, $n \in \mathbb{N}$, be a (time-)grid on which we want to simulate a uniform process (U_t).[2] Let $C_{t_i,t_{i+1}}(u_i, u_{i+1})$ be a bivariate copula of $(U_{t_i}, U_{t_{i+1}})$ and $D_1 C_{t_i,t_{i+1}}$ the partial derivative w. r. t. the first argument.

Then the following algorithm yields a time-series $(u_0, u_1, \ldots, u_n)$ generated from $(U_{t_0}, U_{t_1}, \ldots, U_{t_n})$ with starting value u_{t_0} (which may be a realization of an $\mathsf{U}(0,1)$-variable or simply $F_{t_0}^{-1}(x_0)$ for some starting value $x_0 \in \mathbb{R}$) with bivariate copulas $C_{t_i,t_{i+1}}$:

1. Choose the starting value $u_0 \in [0, 1]$
2. Set $q_u(v; s, t) := D_1 C_{s,t}(u, v)$ and let $q_u^{-1}(z; s, t)$ denote the generalized inverse, $z \in (0, 1)$
3. Generate n iid $\mathsf{U}(0,1)$ samples $(r_1, \ldots, r_n)$
4. For each i in $\{1, \ldots, n\}$ set $u_i := q_{u_{i-1}}^{-1}(r_i; t_{i-1}, t_i)$

   The vector $(u_0, u_1, \ldots, u_n)$ contains the sample including the starting value.

A library for the computer language `R` is available from the author on request (email: `Volker.Schmitz@rwth-aachen.de`). `R` is an integrated suite of software facilities for data manipulation, calculation and graphical display not unlike S^{TM}(see Ihaka and Gentleman, 1996) and freely available under the GNU General Public License for many platforms from `http://www.r-project.org`.

The `copula`-library contains a set of fully documented functions for dealing with copulas and can be easily installed within a working `R`-environment by the command

`R CMD INSTALL copula*.tgz` (Linux)

or within the graphical user interface. Please consult the documentation.

[2]Recall that we get a general process from this by applying the quantile transformation $F_{t_i}^{-1}$ to the sample.

List of Symbols and Abbreviations

$\times$	cartesian product
$<, \leq$	less than (or equal to) componentwise
$\langle X \rangle$	quadratic variation process of X, page 86
$\stackrel{d}{=}$	equality of finite-dimensional distributions
$\sim$	"is distributed as"
$*$	product of copulas, see equation (5.37), page 74
$\star$	generalized product of copulas, see equation (5.39), page 76
$\mathbb{1}$, $\mathbb{1}_A(x)$	indicator function (of a set A): $\mathbb{1}_A(x) = 1$ if $x \in A$ and otherwise 0
a. s.	"almost surely"
$\alpha(F)$, $\omega(F)$	left and right endpoint of the support of F
$\mathcal{B}^n$	n-dimensional Borel-σ-algebra
B_t	Brownian motion
cdf	"cumulative distribution function"
$C, C_i, C_t, C_{t_1,\dots,t_n}$, etc.	copulas
$C\langle \cdot \rangle$	copula (family) of the argument
$\mathcal{C}_n$	set of all n-copulas
$\widehat{C}$	survival copula of C, see equation (2.43)
$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H(x_1, \dots, x_n)$	n-th difference of a function H w. r. t. $(\boldsymbol{a}, \boldsymbol{b})$, see equation (2.2), page 6
$\delta_C(u) := C(u, u)$	diagonal of the copula C
$\Delta_{(a_i,b_i)}^{(i)} H(x_1, \dots, x_n)$	first difference of the function H in the i-th component w. r. t. (a_i, b_i), see equation (2.3), page 6

D_k, $D_{i,j,k}$, etc.	partial derivatives of a copula, page 17
$\mathrm{dom}(H)$	domain of the function H
$\mathcal{D}_X$	finite-dimensional distributions of X
$\overline{F} = 1 - F$	univariate survival function
F, G, F_i, etc.	cumulative distribution functions (cdf)
Φ, φ	cdf and pdf of a standard normal random variable
I, I^n	unit interval, n-dimensional unit cube
iff	"if and only if"
iid	"independent and identically distributed"
inf, sup	infimum, supremum
λ^n	n-dimensional Lebesgue measure
λ_U, λ_L	coefficients of upper and lower tail dependence, page 38
L_t	level curve, page 24
LTD	"left-tail decreasing", page 36
M	upper Fréchet bound, see equation (2.11), page 9
$\mathcal{M}_2$	set of square-integrable martingales with $M_0 = 0$
min, max	minimum, maximum
$\mathsf{N}(\mu, \sigma^2)$	normal distribution with expecation μ and variance σ^2
$(\Omega, \mathcal{A}, \mathsf{P})$	probability space
P, Q	probability measures
P_C	probability measure induced by a copula C
pdf	"probability density function"
Π	independence copula, see equation (2.12), page 9
PLR	"positive likelihood ratio dependent"
PQD	"positively quadrant dependent"
$\mathbb{R}$, $\overline{\mathbb{R}}$, $\mathbb{R}^+$	real line, extended real line, non-negative real line
$\mathrm{ran}(H)$	range of the function H
$\mathcal{R}^n$	set of all n-dimensional rectangles
$R_{\boldsymbol{x}}^{\boldsymbol{y}}$	n-dimensional rectangle induced by $\boldsymbol{x}$ and $\boldsymbol{y}$, page 6
ρ, ρ^s, $\rho_{X,Y}$	Spearman's ρ
RTI	"right-tail increasing", page 36
SDE	"stochastic differential equation"
$\sigma(\,\cdot\,)$	generated σ-algebra
τ, $\tau_{X,Y}$	Kendall's τ

$\mathsf{U}(a,b)$	uniform distribution on (a,b)
$\mathrm{vert}(R_{\boldsymbol{x}}^{\boldsymbol{y}})$	vertices of the rectangle
$V_H(R_{\boldsymbol{x}}^{\boldsymbol{y}})$	H-volume of $R_{\boldsymbol{x}}^{\boldsymbol{y}}$, see equation (2.1), page 6
W	lower Fréchet bound, see equation (2.11), page 9
w. l. o. g.	"without loss of generality"
w. r. t.	"with respect to"
X, Y, X_i, etc.	random variables
$X_{(i)}$	i-th order statistic
$\blacksquare$	end of example or remark
$\square$	end of proof

Bibliography

Numbers at the end of each entry refer to the page(s) on which the work is cited.

Alsina, Claudi, Maurice J. Frank, and Berthold Schweizer (2001). Associative functions on intervals: A primer of triangular norms. Monograph, to appear in 2001. 23

Bauer, Heinz (1992). *Maß- und Integrationstheorie*. de Gruyter, Berlin, second edition. 18, 62, 66

Behnen, Konrad and Georg Neuhaus (1995). *Grundkurs Stochastik*. B. G. Teubner, Stuttgart, third edition. 7

Billingsley, Patrick (1995). *Probability and Measure*. Wiley, New York, third edition. 11, 17, 18, 29, 66

Cambanis, Stamatis (1977). Some properties and generalizations of multivariate Eyraud-Gumbel-Morgenstern distributions. *Journal of Multivariate Analysis*, **7**, 551–559. 71

Cambanis, Stamatis (1991). On Eyraud-Farlie-Gumbel-Morgenstern random processes. In: G. Dall'Aglio, S. Kotz, and G. Salinetti (editors), *Advances in Probability Distributions with Given Marginals: Beyond the Copulas*, volume 67 of *Mathematics and Its Applications*. Kluwer Academic Publishers, Dordrecht – Boston – London. 3, 71, 72

Capéraà, Philippe and Christian Genest (1993). Spearman's ρ is larger than Kendall's τ for positively dependent random variables. *Nonparametric Statistics*, **2**, 183–194. 37, 49

Clayton, David G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**(1), 141–151. 44

Cook, R. Dennis and Mark E. Johnson (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society B*, **43**(2), 210–218. 2, 44, 46

Cook, R. Dennis and Mark E. Johnson (1986). Generalized Burr-Pareto-logistic distributions with applications to a uranium exploration data set. *Technometrics*, **28**(2), 123–131. 44

Cramér, Harald and M. Ross Leadbetter (1967). *Stationary and Related Stochastic Processes: Sample Function Properties and Their Applications*. Wiley, New York. 54

Dambis, Karl E. (1965). On the decomposition of continuous submartingales. *Theor. Probab. Appl.*, **10**, 401–410. 4, 86, 87

Darsow, William F., Bao Nguyen, and Elwood T. Olsen (1992). Copulas and Markov processes. *Illinois Journal of Mathematics*, **36**(4), 600–642. 2, 3, 12, 17, 57, 74, 75, 76, 77

David, Herbert Aron (1981). *Order Statistics*. Wiley, New York. 41, 45

Devroye, Luc (1986). *Non-Uniform Random Variate Generation*. Springer, New York. 95

Dubins, Lester E. and Gideon Schwarz (1965). On continuous martingales. *Proc. Nat. Acad. Sciences U.S.A.*, **53**, 913–916. 4, 86, 87

Elstrodt, Jürgen (1996). *Maß- und Integrationstheorie*. Springer, Berlin. 17, 18

Embrechts, Paul, Alexander McNeil, and Daniel Straumann (1999). Correlation and dependence in risk management: Properties and pitfalls. Working Paper. Risklab: `http://www.risklab.ch/Papers.html`. 2, 33

Esary, James D., Frank Proschan, and David W. Walkup (1967). Association of random variables, with applications. *The Annals of Mathematical Statistics*, **38**(2), 1466–1474. 36

Eyraud, Henri (1938). Les principes de la mesure des corrélations. *Annales de l'Université de Lyon, Series A*, **1**, 30–47. 70

Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika*, **47**, 307–323. 70

Fisher, Nicholas I. (1997). Copulas. In: Samuel Kotz, Campbell B. Read, and David L. Banks (editors), *Encyclopedia of Statistical Sciences, Update Volume 1*. Wiley, New York, pp. 159–163. 2

Frank, Maurice J. (1979). On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. *Aequationes Mathematicae*, **19**, 194–226. 84, 85

Fréchet, Maurice (1957). Les tableauxs de corrélations dont les marges sont données. *Annales de l'Université de Lyon, Sciences Mathématiques et Astronomie*, **Serie A**(4), 13–31. 11

Galambos, János (1978). *The Asymptotic Theory of Extreme Order Statistics*. Wiley, New York. 1, 45

Genest, Christian (1987). Frank's family of bivariate distributions. *Biometrika*, **74**(3), 549–555. 84

Genest, Christian and Louis-Paul Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, **88**(423), 1034–1043. 24, 25, 96

Georges, P., A.-G. Lamy, E. Nicolas, G. Quibel, and Thierry Roncalli (2001). Multivariate survival modelling: A unified approach with copulas. working paper, Groupe de Recherche Opérationelle Crédit Lyonnais, `http://gro.creditlyonnais.fr/content/rd/rd_math/copulas.htm`. 30

Gould, Henry W. (1972). Combinatorial Identities. A Standardized Set Of Tables Listing 500 Binomial Coefficient Summations. Rev. Ed. Morgantown, W. Va. 48

Gumbel, Emile J. (1958). Distributions à plusieurs variables dont les marges sont données. *C. R. Acad. Sci. Paris*, **246**, 2717–2719. 70

Heuser, Harro (1988). *Lehrbuch der Analysis, Teil 2*. Teubner, Stuttgart, fourth edition. 82

Hoeffding, Wassily (1940). Maßstabinvariante Korrelationstheorie. *Schriften des Mathematischen Seminars und des Instituts für Angewandte Mathematik der Universität Berlin*, **5**(3), 179–233. 11

Ihaka, Ross and Robert Gentleman (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**(3), 299–314. 97

Joe, Harry (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London. 33, 36

Jogdeo, Kumar (1982). Dependence, concepts of. In: Samuel Kotz, Norman L. Johnson, and Campbell B. Read (editors), *Encyclopedia of Statistical Sciences*, volume 2. Wiley, New York, pp. 324–334. 1

Johnson, Norman L. and Samuel Kotz (1975). On some generalized Farlie-Gumbel-Morgenstern distributions. *Communications in Statistics*, **4**(5), 415–427. 71

Kamke, Erich (1960). *Das Lebesgue-Stieltjes-Integral*. B. G. Teubner, Leipzig, second edition. 18, 21

Kamps, Udo (1995). *A Concept of Generalized Order Statistics*. Teubner Skripten zur Mathematischen Stochastik. Teubner, Stuttgart. 80

Karatzas, Ioannis and Steven E. Shreve (2000). *Brownian Motion and Stochastic Calculus.* Graduate Texts in Mathematics. Springer, New York, second edition. 51, 53, 54, 60, 61, 63, 82, 86, 87, 88, 90

Kimeldorf, George and Allan Sampson (1975). Uniform representations of bivariate distributions. *Communications in Statistics*, **4**(7), 617–627. 2, 44

Kulpa, Tomasz (1999). On approximation of copulas. *International Journal of Mathematics & Mathematical Sciences*, **22**(2), 259–269. 15

Lange, Kenneth (1998). *Numerical Analysis for Statisticians.* Springer, New York. 95

Lehmann, Erich L. (1966). Some concepts of dependence. *The Annals of Mathematical Statistics*, **37**(2), 1137–1153. 36, 39, 40

Li, Xin, Piotr Mikusiński, and Michael D. Taylor (1998). Strong approximation of copulas. *Journal of Mathematical Analysis and Applications*, **225**, 608–623. 15

Ling, C. H. (1965). Representation of associative functions. *Publ. Math. Debrecen*, **12**, 189–212. 23

Mardia, Kanti V. (1962). Multivariate Pareto distributions. *Annals Math. Statist.*, **33**, 1008–1015. 44

Meyn, Sean P. and Richard L. Tweedie (1996). *Markov Chains and Stochastic Stability.* Springer, London, third edition. 74

Monroe, Itrel (1978). Processes that can be embedded in Brownian motion. *The Annals of Probability*, **6**(1), 42–56. 91

Morgenstern, Dietrich (1956). Einfache Beispiele zweidimensionaler Verteilungen. *Mitteilungsblatt für Mathematische Statistik*, **8**, 234–235. 70

Nelsen, Roger B. (1986). Properties of a one-parameter family of bivariate distributions with given marginals. *Comm. Statist. – Theory and Methods*, **15**, 3277–3285. 84

Nelsen, Roger B. (1991). Copulas and association. In: Giorgio Dall'Aglio, Samuel Kotz, and Gabriella Salinetti (editors), *Advances in Probability Distributions with Given Marginals: Beyond the Copulas*, volume 67 of *Mathematics and Its Applications.* Kluwer Academic Publishers, Dordrecht – Boston – London. 49

Nelsen, Roger B. (1999). *An Introduction to Copulas.* Lecture Notes in Statistics. Springer, New York. 9, 10, 11, 22, 23, 24, 25, 26, 33, 35, 36, 39, 40, 44, 84, 95

Oakes, David (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society B*, **44**(3), 414–422. 44

Oakes, David (1986). Semiparametric inference in a model for association in bivariate survival data. *Biometrika*, **73**(2), 353–361. 44

Øksendal, Bernt (1998). *Stochastic Differential Equations.* Springer, New York, fifth edition. 86, 87, 88, 90

Pfeifer, Dietmar (1989). *Einführung in die Extremwertstatistik.* Teubner, Stuttgart. 93, 94

Randles, Ronald H. and Douglas A. Wolfe (1979). *Introduction to the Theory of Nonparametric Statistics.* Wiley, New York. 26

Reiss, Rolf-Dieter (1989). *Approximate Distributions of Order Statistics.* Springer, New York. 45

Resnick, Sidney I. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer, New York. 77, 78, 80

Revuz, Daniel and Marc Yor (1991). *Continuous Martingales and Brownian Motion.* Springer, Berlin. 53, 55, 86, 87

Rogers, L. Chris G. and David Williams (2000a). *Diffusions, Markov Processes and Martingales, Volume 1: Foundations.* Cambridge University Press, Cambridge, second edition. 4, 17, 18, 47, 51, 53, 86

Rogers, L. Chris G. and David Williams (2000b). *Diffusions, Markov Processes and Martingales, Volume 2: Itô Calculus.* Cambridge University Press, Cambridge, second edition. 4, 53, 86

Ross, Sheldon M. (1996). *Stochastic Processes.* Wiley, New York, second edition. 53

Schmitz, Volker and Erwin Pier-Ribbert (2001). Copulas in finance. *derivatives week*, **10**(6), 6–7. 2

Schweizer, Berthold (1991). Thirty years of copulas. In: G. Dall'Aglio, S. Kotz, and G. Salinetti (editors), *Advances in Probability Distributions with Given Marginals*, Mathematics and Its Applications. Kluwer Academic Publishers, Dordrecht. 2, 22

Shiryaev, Albert N. (1996). *Probability.* Springer, New York, second edition. 11, 56, 57

Sklar, Abe (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231. 1, 9, 10

Sklar, Abe (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, **9**(6), 449–460. 9

Todorovic, Petar (1992). *An Introduction to Stochastic Processes and Their Applications.* Springer, New York. 53, 59, 65, 69

Walsh, John E. (1969). Asymptotic independence between largest and smallest of a set of independent observations. *Ann. Inst. Statist. Math.*, **21**, 287–289. 45

Walter, Wolfgang (1992). *Analysis 1.* Springer, Berlin, third edition. 45, 67

Index

Lebenslauf — Volker Schmitz

14. Mai 1973	In Düren geboren als Sohn von Irene und Klaus Schmitz
1979–1983	Besuch der Katholischen Grundschule in Kreuzau
1983–1992	Besuch des Burgau Gymnasiums in Düren
27. Juni 1992	Erwerb der allgemeinen Hochschulreife (Abitur) am Burgau Gymnasium in Düren
1992–1997	Studium der Mathematik mit Nebenfach Betriebswirtschaftslehre an der RWTH Aachen
1995–1997	Studentische Hilfskraft am Institut für Statistik und Wirtschaftsmathematik der RWTH Aachen
1997–1998	Ableisten des Zivildienstes
seit 1. Juli 1998	Wissenschaftlicher Angestellter am Institut für Statistik und Wirtschaftsmathematik der RWTH Aachen