

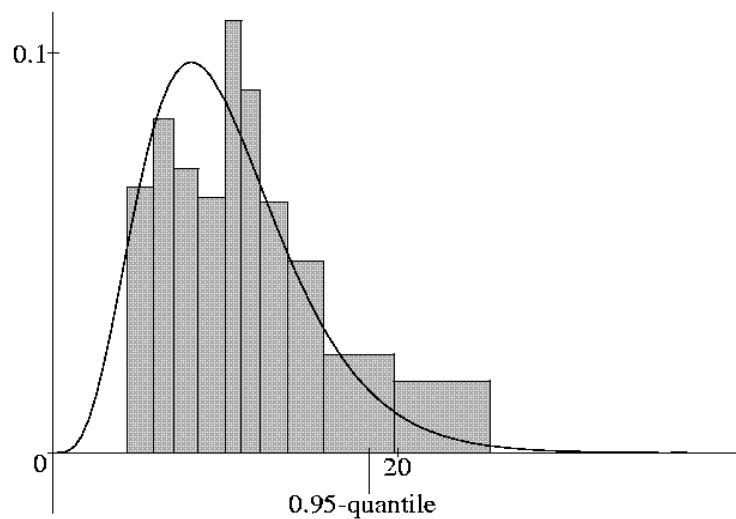


UNIVERSITÄT POTSDAM
Institut für Mathematik

Parametrische Schätzungen von elliptischen
Copulafunktionen

Diplomarbeit

Olga Kuxhaus



Mathematische Statistik und
Wahrscheinlichkeitstheorie

Universität Potsdam – Institut für Mathematik

Mathematische Statistik und Wahrscheinlichkeitstheorie

Parametrische Schätzungen von elliptischen
Copulafunktionen

Olga Kuxhaus

Institut für Mathematik der Universität Potsdam

e-mail: Olga.Kuxhaus@gmx.net

Preprint 2010/09

April 2010

Impressum

© Institut für Mathematik Potsdam, April 2010

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie
am Institut für Mathematik

Adresse: Universität Potsdam
Am Neuen Palais 10
14469 Potsdam

Telefon: +49-331-977 1500

Fax: +49-331-977 1578

E-mail: neisse@math.uni-potsdam.de

ISSN 1613-3307

UNIVERSITÄT POTSDAM
INSTITUT FÜR MATHEMATIK



Parametrische Schätzungen von elliptischen Copulafunktionen

Diplomarbeit
von
Olga Kuxhaus

Gutachter: Prof. Dr. Henning Läter
apl. Prof. Dr. Hannelore Liero
Lehrstuhl der Mathematischen Statistik
Universität Potsdam

Datum: 23. April 2010

Danksagung

Ganz besonders bedanken möchte ich mich bei meinem Betreuer, Herrn Prof. Dr. Läuter, für das interessante Thema sowie die tatkräftige und lehrreiche Unterstützung bei der Erstellung meiner Diplomarbeit. Vielen Dank für die vielen Freiräume, die mir bei der Gestaltung der Arbeit gelassen wurden, die hilfreichen Anregungen und die außerordentliche Geduld.

Bei Uwe Rüßbüldt möchte ich mich für die moralische und finanzielle Unterstützung bedanken. Danke auch dafür, dass Du mir bei den Problemen mit den Formulierungen und Lösungen technischer Natur immer zur Seite standest.

Ein herzlicher Dank geht an meine Eltern, die mir das Studium überhaupt erst ermöglicht haben. Danke, dass Ihr mich immer bestmöglich unterstützt und mir den Rücken gestärkt habt.

Inhaltsverzeichnis

Abbildungsverzeichnis	iii
Tabellenverzeichnis	v
1 Einleitung und Motivation	1
2 Multivariate Copulafunktionen	5
2.1 Einleitung	5
2.2 Satz von Sklar	7
2.3 Eigenschaften von Copulafunktionen	10
3 Abhängigkeitskonzepte	15
3.1 Lineare Korrelation	15
3.2 Copulabasierte Abhängigkeitsmaße	20
3.2.1 Konkordanz	21
3.2.2 Kendall's τ und Spearman's ρ	22
3.2.3 Asymptotische Randabhängigkeit	26
4 Elliptische Copulaklasse	33
4.1 Sphärische und elliptische Verteilungen	34
4.2 Normal-Copula	47
4.3 t-Copula	51
5 Parametrische Schätzverfahren	55
5.1 Maximum-Likelihood-Methode	55
5.1.1 Exakte Maximum-Likelihood-Methode	61
5.1.2 2-stufige parametrische Maximum-Likelihood-Methode	62

5.1.3	2-stufige semiparametrische Maximum-Likelihood-Methode	65
5.2	Momentenmethode	66
5.3	Kendall's τ -Momentenmethode	69
6	Parameterschätzungen für Normal- und t-Copula	73
6.1	Normal-Copula	74
6.1.1	Maximum-Likelihood-Methode	74
6.1.2	Momentenmethode	75
6.1.3	Kendall's τ -Momentenmethode	76
6.1.4	Spearman's ρ -Momentenmethode	77
6.2	t -Copula	78
6.2.1	Verfahren 1 (exakte ML-Methode)	78
6.2.2	Verfahren 2 (2-stufige rekursive ML-Methode)	80
6.2.3	Verfahren 3 (2-stufige KM-ML-Methode)	81
6.2.4	Verfahren 4 (3-stufige M-ML-Methode)	83
7	Simulationen	85
7.1	Grundlagen	86
7.2	Parametrischer Fall	86
7.3	Nichtparametrischer Fall	93
7.4	Fazit	98
A	Programmausschnitt	99
	Literaturverzeichnis	103

Abbildungsverzeichnis

2.1	Graphen und Konturdiagramme der bivariaten Copulae W , M und P	12
3.1	Iso-Kurven von zwei bivariaten Dichtefunktionen mit jeweils identischen und standardnormalverteilten Randverteilungen, jedoch unterschiedlicher Abhängigkeitsstruktur.	19
3.2	10.000 simulierte Datenpaare der Normal-Copula auf der linken Seite und t-Copula mit $\nu = 2$ Freiheitsgraden auf der rechten Seite. Die Randverteilungen sind jeweils identisch und standardnormalverteilt mit Korrelation von 0,9.	28
7.1	Parameterschätzungen für $C_{0,2;7}^t$ bei 40 Stichproben und 30 Wiederholungen. . .	88
7.2	Parameterschätzungen für $C_{0,5;7}^t$ bei 40 Stichproben und 30 Wiederholungen. . .	89
7.3	Parameterschätzungen für $C_{0,7;7}^t$ bei 40 Stichproben und 30 Wiederholungen. . .	90
7.4	Parameterschätzungen für $C_{0,8;7}^t$ bei 40 Stichproben und 30 Wiederholungen. . .	91
7.5	Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen. . .	92
7.6	Parameterschätzungen für $C_{0,2;7}^t$ und $C_{0,5;7}^t$ bei 40 Stichproben und 30 Wiederholungen.	94
7.7	Parameterschätzungen für $C_{0,7;7}^t$ und $C_{0,8;7}^t$ bei 40 Stichproben und 30 Wiederholungen.	95
7.8	Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen. . .	97

Abbildungsverzeichnis

Tabellenverzeichnis

4.1	Werte von λ_O für verschiedene Werte von ν und R_{12}	52
7.1	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,2;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.	88
7.2	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,5;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.	89
7.3	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.	90
7.4	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,8;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.	91
7.5	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen, parametrischer Fall.	92
7.6	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,2;7}^t$	96
7.7	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,5;7}^t$	96
7.8	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$	96
7.9	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,8;7}^t$	96
7.10	Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen, nichtparametrischer Fall.	97

Kapitel 1

Einleitung und Motivation

Aufgrund der rasanten Entwicklung der Finanzmärkte in den letzten Jahrzehnten ist die adäquate Modellierung von Abhängigkeiten ein zentraler Aspekt der modernen multivariaten Statistik geworden. Insbesondere im Bereich der versicherungstechnischen und finanzmathematischen Risiken besteht die Notwendigkeit, die Zusammenhänge angemessen zu modellieren, um mögliche fatale Folgen zu vermeiden. So sind beispielsweise die Ausfallwahrscheinlichkeiten von Krediten und damit verbundene Ausfallabhängigkeiten im Risikomanagementbereich vom besonderen Interesse.

In der Vergangenheit wurde oft die multivariate Normalverteilung zur Modellierung der Abhängigkeiten angewandt. Die Erfahrung zeigt jedoch, dass diese Verteilung in der Praxis einige Defizite aufweist und zur approximativen Modellierung von komplexen Abhängigkeitsstrukturen wenig geeignet ist. Zum Einen wird von der Normalverteilung das Auftreten von gemeinsamen Extremereignissen nicht ausreichend gewichtet. Studien im finanzmathematischen Bereich belegen, dass die für den Finanzmarkt relevanten empirischen Verteilungen mehr Wahrscheinlichkeitsmasse an den Rändern konzentrieren als es bei Normalverteilung der Fall ist. Zum Anderen ist die Abhängigkeitsstruktur innerhalb eines multivariat normalverteilten Vektors linearer Natur und kann mittels des Korrelationskoeffizienten nach Pearson gut gemessen werden. Jedoch sind die Abhängigkeitsstrukturen im

Risikomanagementbereich viel komplexer und benötigen häufig andere Beschreibungsinstrumente als die lineare Korrelation.

Seit den 90er Jahren werden zur Modellierung von Abhängigkeitsstrukturen Copulafunktionen eingesetzt. Die Grundlagen für dieses Konzept wurden bereits in den 40er und 50er Jahren von Hoeffding, Fréchet und Sklar erarbeitet. Von fundamentaler Bedeutung ist die Aussage des Satzes von Sklar, dass jede multivariate Verteilungsfunktion in eine Copulafunktion, welche die Abhängigkeitsstruktur beschreibt, und die univariaten Randverteilungen separiert werden kann. Mit diesem Ansatz kann die Konstruktion von multivariaten Verteilungen flexibel gestaltet werden. Die Modellierung der Abhängigkeitsstrukturen kann damit unabhängig von den Randverteilungen realisiert werden.

Die meisten wahrscheinlichkeitstheoretischen und mathematischen Grundlagen der Copula-Theorie sind seit den 40er Jahren gut analysiert worden und werden in wenigen Standardwerken wie beispielsweise Nelsen (2006) ausführlich beschrieben. Dennoch ist die Forschung zum Thema Copula insbesondere in den Anwendungsbereichen noch lange nicht am Ende. Die stetig wachsende Anzahl der Veröffentlichungen in diesem Bereich zeigt, dass das Interesse an Copulas nach wie vor sehr groß ist.

Es gibt eine große Anzahl an Copulafunktionen, von denen jede eine andere Abhängigkeitsstruktur darstellt. Copulafunktionen, die ähnliche Abhängigkeitsstrukturen beschreiben, können zu Copulaklassen zusammengefasst werden, wie z.B. die archimedischen oder die elliptischen Copulas. Das Hauptaugenmerk dieser Arbeit liegt auf der elliptischen Copulaklasse. Zur Zielsetzung gehört es, einige Eigenschaften dieser Klasse, die für die Konstruktion der Copulafunktionen entscheidend sind, zu erarbeiten. Die Umsetzung wird am Beispiel von zwei bekanntesten Vertretern, der Normal- und der t -Copula, aufgezeigt. Außerdem werden für die Parameter dieser Copulafunktionen geeignete Schätzverfahren angegeben. Für die bivariate t -Copula werden die hergeleiteten Verfahren anhand einer Simulationsstudie verglichen.

Im Folgenden wird der Aufbau dieser Arbeit dargelegt.

Im zweiten Kapitel werden die Grundlagen der Copula-Theorie vorgestellt. Dazu werden die Definition der Copula sowie ihre Bedeutung erarbeitet. Mithilfe des Satzes von Sklar können die Fragen der Existenz und Eindeutigkeit von Copulas geklärt sowie eine Methode zur Konstruktion von Copulas angegeben werden. Die hergeleitete Copuladichte spielt bei der Entwicklung von Schätzverfahren für die Copulaparameter eine wichtige Rolle. Es werden außerdem einige Eigenschaften von Copulas, unter anderem die Fréchet-Hoeffding-Schranken und die Invarianz unter den streng monoton wachsenden Transformationen, betrachtet.

Im dritten Kapitel werden verschiedene Abhängigkeitskonzepte genauer untersucht. Zunächst wird die lineare Korrelation nach Pearson vorgestellt und insbesondere ihre Nachteile ausführlich diskutiert. Anschließend wird ein alternatives Konzept der Konkordanzmaße eingeführt, welche die monotone Abhängigkeit von Zufallsvariablen messen und gegenüber der linearen Korrelation wesentliche Vorteile aufweisen. Das Hauptinteresse gilt dabei den Rangkorrelationsmaßen Kendall's τ und Spearman's ρ . Diese lassen sich als Funktionale der Copula darstellen und werden aus diesem Grund oft als copulabasierte Abhängigkeitsmaße bezeichnet. Zum Ende des Kapitels wird die asymptotische Randabhängigkeit von Zufallsvariablen eingeführt und auf ihre Eigenschaften untersucht.

Im vierten Kapitel wird die elliptische Copulaklasse vorgestellt. Dazu werden erst die sphärischen und die elliptischen Verteilungen eingeführt sowie ihre wesentlichen Eigenschaften erläutert. Es werden zwei wichtige Vertreter der elliptischen Copulaklasse, die Normal- und die t -Copula - mit Hilfe des Satzes von Sklar hergeleitet. Für diese Copulafunktionen werden die Abhängigkeitsmaße, die im vorigen Kapitel angegeben wurden, berechnet.

Im fünften Kapitel werden einige klassische parametrische Schätzverfahren wie die Maximum-Likelihood- und die Momentenmethode dargestellt. Hier geht es in ers-

ter Linie um eine allgemeine Herausarbeitung der Verfahren für den multivariaten Fall sowie die Analyse ihrer statistischen Eigenschaften.

Im sechsten Kapitel geht es um die Umsetzung der im vorigen Kapitel beschriebenen Verfahren für den Fall der elliptischen Copulafunktionen. Für die Parameter der Normal-Copula können die Schätzer nach den klassischen Verfahren exakt angegeben werden. Das ist für die Parameter der t -Copula nicht möglich. Hier müssen statistische Verfahren in Kombination mit numerischer Optimierung angewandt werden. Aus diesem Grund wurden zum Schätzen der Parameter einer t -Copula neue Verfahren entwickelt, die in diesem Kapitel ausführlich diskutiert werden.

Im siebten Kapitel werden die Ergebnisse einer Simulationsstudie, die die Zielsetzung verfolgt, mögliche Unterschiede in der Arbeitsweise der Verfahren herauszufinden, präsentiert. Hierzu wird ausschließlich die Anwendung für die t -Copula diskutiert. Da die Parameterschätzer in diesem Fall keine explizite Darstellung besitzen, ist nur ein Vergleich der Verfahren über die Simulation möglich. Die wichtigsten Aussagen und Ergebnisse der Simulation werden zusammengefasst. Zusätzlich soll hier ein Ausblick sowie Vorschläge geboten werden, an welchen Punkten eine Weiterentwicklung bzw. Verbesserung der Simulation als sinnvoll erscheint.

Kapitel 2

Multivariate Copulafunktionen

Zu verstehen, was Copulas sind und welche Bedeutung sie in der stochastischen Theorie tragen, ist das Hauptziel dieses Kapitels. In den drei Abschnitten sollen die Grundlagen der Copulatheorie dargestellt werden, die gleichzeitig die Grundlage für diese Arbeit bilden. Als erstes wird die Copula auf funktionelle und analytische Weisen definiert und im zweiten Abschnitt wird der Satz von Sklar sowie seine Bedeutung für die Copulatheorie diskutiert. Es wird außerdem eine Methode zur Konstruktion von Copulafunktionen aus einer multivariaten Verteilungsfunktion und ihren Randverteilungen vorgestellt. Im letzten Teil des Kapitels werden einige wesentliche Eigenschaften von Copulafunktionen analysiert.

2.1 Einleitung

Am kürzesten lässt sich wohl die Copula als (multivariate) Verteilungsfunktion auf dem Einheitsquader $[0, 1]^n$ mit gleichverteilten univariaten Randverteilungen beschreiben. Diese Charakterisierung sowie die Bezeichnung gehen auf die Arbeit von Abe Sklar, die er 1959 veröffentlichte, zurück. Doch viele grundlegende Ideen und Ergebnisse der Copulatheorie wurden bereits von Wassily Hoeffding in seinen Publikationen von 1940 bis 1942 erarbeitet. Auch der Mathematiker Maurice Fréchet kam 1951, ohne die Arbeiten von Hoeffding zu kennen, zu einer Vielzahl gleicher Erkenntnisse. Nach 1959 entdeckten verschiedene Autoren das Konzept der Copula

neu und so findet man sie teilweise unter den Bezeichnungen wie "uniform representations" oder "dependence functions" wieder. Eine Zeit lang sind Copulas dann in Vergessenheit geraten, bis die Theorie in den 80er Jahren neuen Aufschwung bekam. So wurden Copulas für die Untersuchung der Abhängigkeit zwischen Zufallsvariablen von Schweizer und Wolff im Jahre 1981 zum ersten Mal explizit benutzt, siehe Schweizer and Wolff (1981). Seit den 90er Jahren verläuft die Entwicklung der theoretischen sowie praktischen Copulakonzepte rasant, was in erster Linie dem großen Interesse von Finanz- und Versicherungsbranchen zuzuschreiben ist.

Es wird gezeigt, dass die oben angegebene Definition der Copula, ihren funktionalen Kern in präziser Weise trifft. Denn die Grundidee der Copulatheorie ist, die Information über die Abhängigkeitsstruktur zwischen den einzelnen Komponenten eines Zufallsvektors, welche implizit in der gemeinsamen Verteilungsfunktion enthalten ist, zu isolieren. Jedoch ist diese Darstellung in Bezug auf ihren praktischen Nutzen nicht unproblematisch. Aus diesem Grund soll auch die allgemeine (analytische) Definition der Copula vorgestellt werden.

Im Folgenden soll $\text{dom}H$ den Definitions- und $\text{ran}H$ den Wertebereich einer Funktion H bezeichnen. Ausserdem stellt $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty]$ die Abschließung der reellen Achse dar.

Definition 2.1.1

Eine reellwertige Funktion $C : [0, 1]^n \rightarrow [0, 1]$ heißt *multivariate Copulafunktion* oder kurz *n - Copula*, falls sie für alle $\mathbf{u}, \mathbf{v} \in [0, 1]^n$ und $\mathbf{u} \leq \mathbf{v}$ ($u_k \leq v_k$ für alle k) folgende Eigenschaften erfüllt:

1. $C(\mathbf{u}) = 0$, falls mindestens ein Element von u gleich 0 ist,
2. $C(\mathbf{u}) = u_k$, falls bis auf u_k alle Elemente von \mathbf{u} gleich 1 sind,
3. $\sum_{i_1=1}^2 \cdots \sum_{i_n=1}^2 (-1)^{i_1+\cdots+i_n} C(x_{1i_1}, \dots, x_{ni_n}) \geq 0$,
wobei $x_{j1} = u_j$ und $x_{j2} = v_j$ für alle $j \in 1, \dots, n$.

□

Diese Definition findet man bei (Embrechts et al., 1999, S. 4).

Die Randbedingungen und die Positivitätseigenschaft aus der Definition 2.1.1 lassen bereits den Schluss zu, dass es sich bei der Copula um eine multivariate Verteilungsfunktion handelt. Dieser Zusammenhang wird in dem Satz von Sklar, der im nächsten Abschnitt eingeführt wird, verdeutlicht.

2.2 Satz von Sklar

Es wird nun der von Sklar 1959 veröffentlichte Satz präsentiert, der auf Grund seiner Bedeutung wohl als Fundamentalsatz der Copulatheorie bezeichnet werden kann. Dieses Resultat spielt eine ganz zentrale Rolle, da er die Grundlage für die meisten copulabasierten Konzepte und Anwendungen in der mathematischen Statistik bildet.

Satz 2.2.1 (Satz von Sklar)

Sei $\mathbf{X} = (X_1, \dots, X_n)^T$ ein n -dimensionaler Zufallsvektor mit der multivariaten Verteilungsfunktion F und den Randverteilungen F_1, \dots, F_n . Dann existiert eine Copulafunktion C , so dass für alle $\mathbf{x} \in \overline{\mathbb{R}}^n$

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (2.1)$$

gilt.

Wenn alle F_1, \dots, F_n stetig sind, so ist die Copula C eindeutig. Anderenfalls wird die Copula C auf $\text{ran}F_1 \times \dots \times \text{ran}F_n$ eindeutig bestimmt.

Es gilt auch die Umkehrung des Satzes.

Satz 2.2.2

Sind F_1, \dots, F_n beliebige univariate Verteilungsfunktionen und C eine beliebige n -Copula. Dann ist die Funktion F , die gemäß der Gleichung (2.1) definiert ist, eine n -dimensionale Verteilungsfunktion mit den Randverteilungen F_1, \dots, F_n .

Über den Satz (2.2.1) wird die Bedeutung einer Copulafunktion vermittelt. Wie der Begriff "Copula" (lat.: "Band", "Verbindung", siehe Pertsch and Lange-Kowal (1999))

schon vermuten lässt, verknüpft sie eindimensionale Verteilungsfunktionen zu einer gemeinsamen multivariaten Verteilung. Dabei trägt sie die ganze Information über die Abhängigkeitsstruktur innerhalb eines Zufallsvektors und ihre Variablen können als Platzhalter für beliebige eindimensionale Verteilungen gesehen werden. Auf diese Weise bietet der Copulaansatz ganz neue Möglichkeiten zur Modellierung von Verteilungen und Abhängigkeiten in der multivariaten Statistik. Damit ist es möglich, die Abhängigkeitsstruktur eines Zufallsvektors hervorzuheben und diese unabhängig von den einzelnen Randverteilungen zu definieren und zu analysieren.

Als eine weitere Konsequenz des Satzes von Sklar erhält man eine Methode zur Konstruktion von Copulas ausgehend von einer multivariaten Verteilungsfunktion. Die Gleichung (2.1) wird zu einem Term über die gemeinsame Verteilungsfunktion und die Inversen der Randverteilungen umgeformt und liefert auf diese Weise einen neuen Ausdruck für die Copula. Dies ist insbesondere für den nichtparametrischen Ansatz zum Schätzen von Copulafunktionen vom großen Nutzen. In diesem Zusammenhang tritt das Problem auf, dass die Umkehrfunktion im streng mathematischen Sinne nur für streng monoton wachsende Verteilungsfunktionen existiert. Sie wird in diesem Fall mit F^{-1} für eine Verteilungsfunktion F bezeichnet und erfüllt die Bedingungen:

$$F(F^{-1}(u)) = u \quad \text{und} \quad F^{-1}(F(x)) = x.$$

Die übliche Verallgemeinerung ist durch die Quasi-Inverse gegeben, siehe die nachfolgende Definition.

Definition 2.2.1

Sei F eine rechtsstetige Verteilungsfunktion. Eine Funktion F^- mit dem Definitionsbereich $[0, 1]$ ist die *Quasi-Inverse* von F , wenn gilt:

$$F^-(u) = \inf\{x \mid F(x) \geq u\}. \tag{2.2}$$

□

Dann erfüllt F^- die Bedingungen:

$$F(F^-(u)) \geq u \quad \text{und} \quad F^-(F(x)) \leq x.$$

Für streng monoton wachsende Verteilungsfunktionen stimmt das Quasi-Inverse mit der Umkehrfunktion überein. Durch (2.2) ist die Quasi-Inverse eindeutig bestimmt und liefert das nötige Instrument, um eine Methode zur Konstruktion der Copula anzugeben.

Korollar 2.2.1

Sei F eine n -dimensionale Verteilungsfunktion mit Randverteilungen F_1, \dots, F_n und ihren Quasi-Inversen F_1^-, \dots, F_n^- . C sei eine Copulafunktion, die der Gleichung (2.1) genügt. Dann gilt für alle $\mathbf{u} \in [0, 1]^n$

$$C(u_1, \dots, u_n) = F(F_1^-(u_1), \dots, F_n^-(u_n)).$$

Nicht selten wird in der Statistik mit der Dichtefunktion einer Verteilung gearbeitet, zum Beispiel in der Schätztheorie und der Maximum-Likelihood-Methode. Da Copula einer multivariaten Verteilungsfunktion entspricht und viele Verfahren auf sie übertragbar sind, liegt es nahe, eine Copuladichte zu definieren, vergleiche (Embrechts et al., 2001, S. 8).

Definition 2.2.2

Die *Copuladichte* einer absolutstetigen n -dimensionalen Copulafunktion C ist fast überall auf $[0, 1]^n$ definiert und gegeben durch

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \cdots \partial u_n}.$$

□

In einem Teil der Statistik, der Zuverlässigkeitsanalyse, interessiert man sich für die Lebensdauer eines Systems oder einer Population. Man bestimmt die Wahrscheinlichkeit, mit der ein Individuum oder ein Teil des Systems ein gewisses Zeitintervall überlebt, die sogenannte *Überlebensfunktion* $\bar{F}(x) = \mathbb{P}(X > x) = 1 - F(x)$, wobei F die Verteilungsfunktion der Zufallsvariablen X ist. Die Überlebensfunktion lässt sich auch durch eine Copula darstellen, die im Folgenden mit *Überlebenscopula* bezeichnet werden soll.

Definition 2.2.3

Seien U_1, \dots, U_n auf $[0, 1]$ gleichverteilte Zufallsvariablen, so ist ihre gemeinsame Überlebensfunktion oder die *Überlebenscopula* \bar{C} gegeben durch

$$\bar{C}(u_1, \dots, u_n) := \mathbb{P}(U_1 > u_1, \dots, U_n > u_n).$$

□

Wegen

$$\mathbb{P}(U_1 > u_1, U_2 > u_2) = 1 - [\mathbb{P}(U_1 \leq u_1) + \mathbb{P}(U_2 \leq u_2) - \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2)]$$

erhält man im bivariaten Fall

$$\bar{C}(u_1, u_2) = 1 - u_1 - u_2 + C(u_1, u_2),$$

wobei die Copula C die gemeinsame Verteilungsfunktion von U_1 und U_2 ist.

Im folgenden Beispiel soll eine der wichtigsten Copulas, die *Produkt-* oder *Unabhängigkeitscopula*, eingeführt werden.

Beispiel 2.2.1 (Produktcopula)

Seien X_1, \dots, X_n unabhängige Zufallsvariablen mit den Verteilungsfunktionen F_1, \dots, F_n . Dann ist bekannt, dass die gemeinsame Verteilungsfunktion F als Produkt der einzelnen Verteilungsfunktionen berechnet werden kann. Bezeichnen außerdem $U_1 = F_1(X_1), \dots, U_n = F_n(X_n)$ auf $[0, 1]$ gleichverteilte Zufallsvariablen, so lässt sich die Abhängigkeitsstruktur folgendermaßen in eine zugehörige Copula P^n übertragen:

$$\begin{aligned} F(x_1, \dots, x_n) &= \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) \\ &= \mathbb{P}(X_1 \leq x_1) \cdots \mathbb{P}(X_n \leq x_n) \\ &= F_1(x_1) \cdots F_n(x_n) \\ &= u_1 \cdots u_n \\ &= P^n(u_1, \dots, u_n), \quad \text{für alle } u_1, \dots, u_n \in [0, 1]. \end{aligned}$$

□

2.3 Eigenschaften von Copulafunktionen

In diesem Abschnitt werden einige wichtige Eigenschaften der Copulafunktionen betrachtet werden. Folgender Satz, den man bei (Nelsen, 2006, S. 46) oder (Embrechts et al., 2001, S. 4) findet, liefert eine Aussage zur Stetigkeit von Copulas.

Satz 2.3.1

Sei C eine n -Copula, dann gilt für alle $\mathbf{u}, \mathbf{v} \in [0, 1]^n$

$$\left| C(\mathbf{v}) - C(\mathbf{u}) \right| \leq \sum_{k=1}^n |v_k - u_k|.$$

Damit ist C gleichmäßig stetig auf $[0, 1]^n$.

Ein weiterer Satz, siehe (Nelsen, 2006, S. 47) oder (Embrechts et al., 2001, S. 5), der auf Fréchet und Hoeffding zurück geht, stellt eine Ungleichung, die jede Copulafunktion erfüllt, dar.

Satz 2.3.2 (Fréchet-Hoeffding-Ungleichung)

Sei C eine n -dimensionale Copulafunktion, dann gilt für alle $\mathbf{u} \in [0, 1]^n$

$$\max(u_1 + \dots + u_n - n + 1, 0) \leq C(\mathbf{u}) \leq \min(u_1, \dots, u_n). \quad (2.3)$$

Die beiden Schranken aus (2.3) werden nach ihren Entdeckern als *Fréchet-Hoeffding-Schranken* bezeichnet. Die untere Fréchet-Hoeffding-Schranke notiert man oft mit $W^n(u_1, \dots, u_n) := \max(u_1 + \dots + u_n - n + 1, 0)$ und die obere Fréchet-Hoeffding-Schranke mit $M^n(u_1, \dots, u_n) := \min(u_1, \dots, u_n)$.

Für den bivariaten Fall sollen die beiden Schranken $W := W^2$, $M := M^2$ sowie Produktcopula $P := P^2$ aus dem Beispiel 2.2.1 veranschaulicht werden. Die Abbildung 2.1 zeigt auf der linken Seite die Graphen der Copulae und auf der rechten Seite die entsprechenden Konturdiagramme, die den Verlauf der Graphen verdeutlichen sollen. Es ist in der Darstellung gut erkennbar, dass alle drei Funktionen die Randbedingungen aus der Definition 2.1.1 erfüllen. Ausserdem liegt der Graph der Produktcopula P augenscheinlich zwischen den Graphen von W und M , so wie es für jede bivariate Copula nach der Ungleichung (2.3) zutreffen muss.

Die Funktion M^n ist für alle $n \geq 2$ eine n -Copula. Die untere Schranke W^n dagegen erfüllt die Bedingungen aus der Definition (2.1.1) ab $n > 2$ nicht und ist somit nur im Fall $n = 2$ eine Copula, siehe (Nelsen, 2006, S. 47-48) oder (Embrechts et al., 2001, S. 5). Sie ist also die bestmögliche untere Schranke in dem Sinne, dass sich für einen festen Wert $\mathbf{u} \in [0, 1]^n$ mit $n \geq 3$ eine n -Copula C finden lässt, die den gleichen Funktionswert $C(\mathbf{u})$ wie die Funktion W^n aufweist. In der Regel wird es jedoch für jeden Wert $\mathbf{u} \in [0, 1]^n$ eine andere Copula sein.

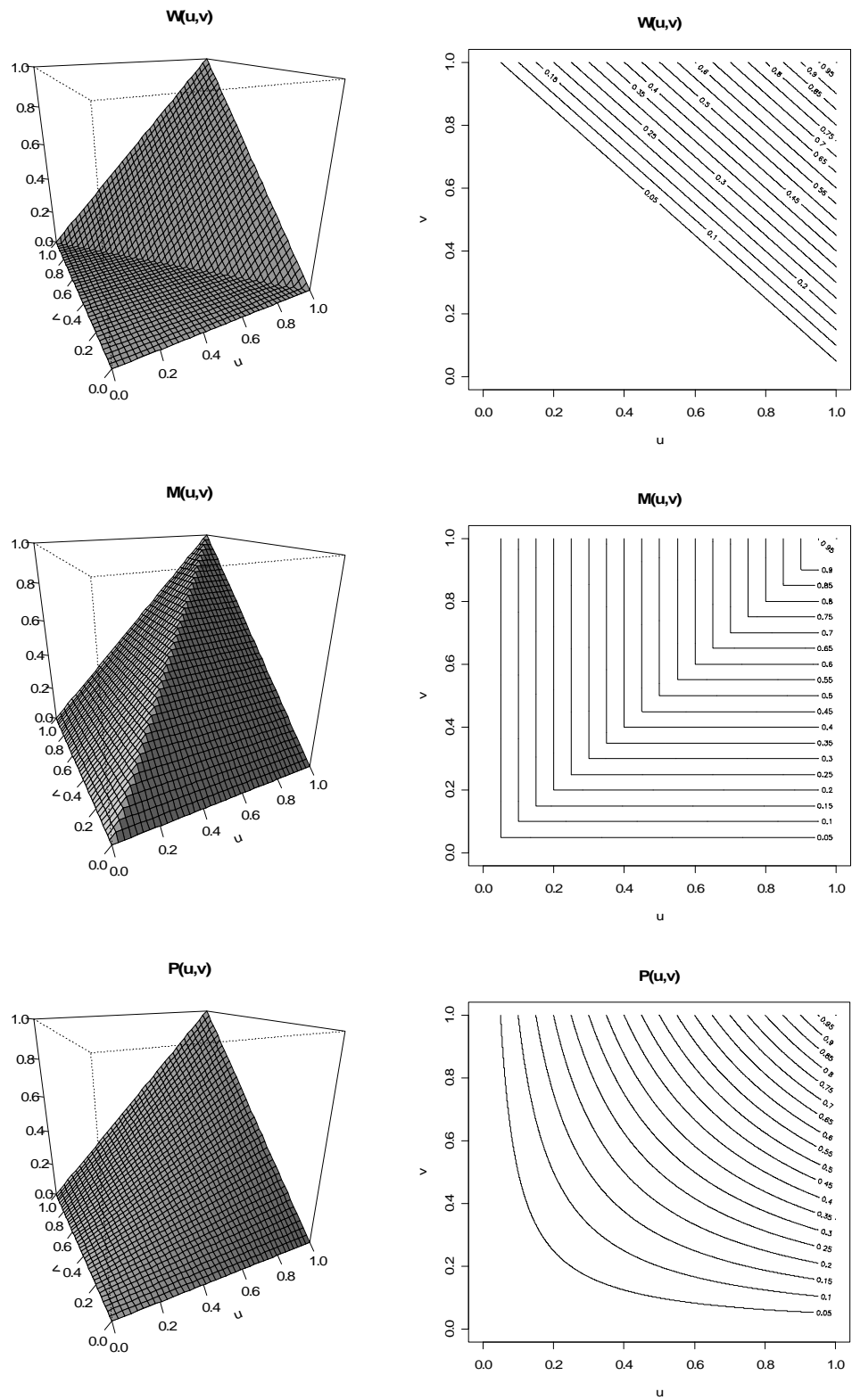


Abbildung 2.1: Graphen und Konturdiagramme der bivariaten Copulae W , M und P .

Satz 2.3.3

Für alle $n \geq 3$ und alle $\mathbf{u} \in [0, 1]^n$ existiert eine n -Copula $C_{\mathbf{u}}$, abhängig von \mathbf{u} , so dass

$$C_{\mathbf{u}} = W^n(\mathbf{u})$$

gilt.

Beweis.

Siehe (Nelsen, 2006, S. 48).

Satz 2.3.4

Für $n \geq 2$ seien X_1, \dots, X_n stetige Zufallsvariablen. Dann gilt:

1. X_1, \dots, X_n sind genau dann stochastisch unabhängig, wenn die n -Copula von X_1, \dots, X_n gleich der Produktcopula P^n ist.
2. Jede der Zufallsvariablen X_1, \dots, X_n ist genau dann eine fast sicher streng monoton wachsende Funktion von allen anderen X_i , wenn die n -Copula von X_1, \dots, X_n gleich der oberen Fréchet-Hoeffding-Schranke M^n ist.

Dieser Satz, siehe (Nelsen, 2006, S. 48), liefert neben der allgemeinen Definition der stochastischen Unabhängigkeit, noch die Aussage, dass auch die Funktion P^n für alle $n \geq 2$ eine n -Copula ist. Ausserdem sei angemerkt, dass zumindest im Fall $n = 2$ zwei Zufallsvariablen X und Y als *perfekt negativ abhängig* bezeichnet werden, wenn die zugehörige Copula gleich der unteren Fréchet-Hoeffding-Schranke W ist. Anderenfalls als *perfekt positiv abhängig*, wenn die Copula gleich der oberen Fréchet-Hoeffding-Schranke M ist. Die Funktion M ist dabei gerade die Verteilungsfunktion des Zufallsvektors $\mathbf{X} = (U, U)$ mit U gleichverteilt auf $[0, 1]$ (kurz: $U \sim \mathcal{U}[0, 1]$). Sie wird in diesem Zusammenhang auch als *komonotone Copula* bezeichnet. Entsprechend ist W die Verteilungsfunktion von $\mathbf{X} = (U, 1 - U)$ mit $U \sim \mathcal{U}[0, 1]$, und wird *kontramotone Copula* genannt, vergleiche (Embrechts et al., 2001, S. 11).

Eine weitere wichtige Eigenschaft von Copulas ist ihre Invarianz unter streng monoton wachsenden Transformationen.

Satz 2.3.5

Sei $\mathbf{X} = (X_1, \dots, X_n)^T$ ein Zufallsvektor mit stetigen Randverteilungen und Copula C , weiter seien T_1, \dots, T_n streng monoton wachsende Funktionen auf $\text{ran}X_1, \dots, \text{ran}X_n$. Dann besitzt der Zufallsvektor $(T_1(X_1), \dots, T_n(X_n))$ ebenfalls die Copula C .

Beweis.

Siehe (Embrechts et al., 2001, S. 6).

Im Folgenden soll eine Ordnung auf der Menge der Copulas eingeführt werden.

Definition 2.3.1

Seien C_1 und C_2 zwei Copulas, dann heißt C_1 *kleiner als* C_2 oder kurz $C_1 \prec C_2$, wenn

$$C_1(\mathbf{u}) \leq C_2(\mathbf{u}) \quad \text{für alle } \mathbf{u} \in [0, 1]^n \quad (2.4)$$

gilt. Analog sagt man, C_1 ist *größer als* C_2 oder kurz $C_1 \succ C_2$, wenn $C_1(u, v) \geq C_2(u, v)$ für alle $u, v \in [0, 1]$ ist.

□

Nicht jedes Paar von Copulafunktionen ist in dieser Ordnung vergleichbar, sie ist somit unvollständig auf der Menge von Copulafunktionen. Es gibt jedoch viele wichtige parametrische Familien, die im Sinne von (2.4) vollständig geordnet sind. So sei $\mathcal{C} := \{C_\theta \mid \theta \in \Theta\}$ eine einparametrische Copulafamilie. Man bezeichnet \mathcal{C} als *positiv geordnet*, wenn $C_{\theta_1} \prec C_{\theta_2}$ ist, wenn $\theta_1 \leq \theta_2$, siehe (Embrechts et al., 2001, S. 5).

Kapitel 3

Abhängigkeitskonzepte

Im ersten Teil des 3. Kapitels sollen der lineare Korrelationskoeffizient und seine Eigenschaften beleuchtet werden. Auch wenn dieses Maß sehr beliebt ist und häufig als Standardmittel zur Abhängigkeitsmodellierung in der multivariaten Statistik eingesetzt wird, verfügt nicht jede Verteilung über eine lineare Abhängigkeitsstruktur. Die Anwendung des linearen Korrelationskoeffizienten ist daher nicht unproblematisch und kann sogar irreführend sein. Es sollen aus diesem Grund alternative Abhängigkeitsmaße vorgestellt werden, die über den Begriff der Konkordanz definiert werden. Ziel ist es jedoch zu zeigen, dass diese Maße als Funktionale der Copulafunktionen dargestellt werden können, und damit die Zusammenhängestrukturen unabhängig von den jeweiligen Randverteilungen beschreiben.

3.1 Lineare Korrelation

Definition 3.1.1

Sei $(X, Y)^T$ ein bivariater Zufallsvektor mit $0 < \text{Var}(X), \text{Var}(Y) < \infty$, dann ist der *lineare Korrelationskoeffizient* oder der *Bravais-Pearson-Korrelationskoeffizient* gegeben durch

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}. \quad (3.1)$$

□

Es wird im Folgenden häufig nur von linearer Korrelation oder Korrelation gesprochen.

Der Ausdruck in (3.1) beschreibt ein Maß für lineare Abhängigkeit. Demnach sind zwei Zufallsvariablen X und Y *perfekt linear abhängig*, wenn sich die eine fast sicher als lineare Abbildung der anderen darstellen lässt, $Y \stackrel{\text{f.s.}}{=} aX + b$ für $a \in \mathbb{R} \setminus \{0\}$ und $b \in \mathbb{R}$. Der Korrelationskoeffizient nimmt dabei die Werte 1 und -1 , d.h.

$|\text{Corr}(X, Y)| = 1$. In sonstigen Fällen (der nichtperfekten Abhängigkeit) ist $|\text{Corr}(X, Y)| < 1$, vergleiche (Embrechts et al., 1999, S. 6).

Es soll nun der bivariate Fall genutzt werden, um eine allgemeine Definition für die Korrelation zwischen zwei multivariat verteilten Zufallsvariablen anzugeben. Man konstruiert dazu für $\mathbf{X} = (X_1, \dots, X_n)^T$ und $\mathbf{Y} = (Y_1, \dots, Y_m)^T$ aus jeweils paarweise ermittelten Kovarianzen und Korrelationskoeffizienten $n \times m$ -Matrizen $\text{Cov}(\mathbf{X}, \mathbf{Y})$ und $\text{Corr}(\mathbf{X}, \mathbf{Y})$:

$$\text{Cov}(\mathbf{X}, \mathbf{Y})_{ij} := \text{Cov}(X_i, Y_j),$$

$$\text{Corr}(\mathbf{X}, \mathbf{Y})_{ij} := \text{Corr}(X_i, Y_j), \quad 1 \leq i \leq n, 1 \leq j \leq m.$$

Nicht selten braucht man die Kovarianz oder die Korrelation innerhalb eines Vektors X . In diesem Fall setzt man $\text{Var}(\mathbf{X}) := \text{Cov}(\mathbf{X}, \mathbf{X})$. Aus der multivariaten Statistik ist bekannt, dass diese Matrizen symmetrisch und positiv semi-definit sind.

Sei $(X, Y)^T$ ein bivariater Zufallsvektor, $\alpha, \gamma \in \mathbb{R} \setminus \{0\}$ und $\beta, \delta \in \mathbb{R}$. Die folgende Eigenschaft zeigt, dass der lineare Korrelationskoeffizient invariant unter den streng monoton wachsenden *linearen* Transformationen ist

$$\text{Corr}(\alpha X + \beta, \gamma Y + \delta) = \text{sgn}(\alpha\gamma) \text{Corr}(X, Y).$$

Auch im allgemeinen Fall lassen sich die Kovarianz und die Korrelation unter der Anwendung von linearen Operationen leicht berechnen. Seien $x \rightarrow Ax + a$ und $y \rightarrow By + b$ lineare Transformationen mit $A \in \mathbb{R}^{k \times n}$, $B \in \mathbb{R}^{l \times m}$, $a \in \mathbb{R}^k$, $b \in \mathbb{R}^l$, X n -dimensionale und Y m -dimensionale Zufallsvariablen, dann ist

$$\text{Cov}(AX + a, BY + b) := A \text{Cov}(X, Y) B^T.$$

3.1 Lineare Korrelation

Speziell für jede lineare Kombination $\alpha^T X$ mit $\alpha \in \mathbb{R}^n$ erhält man

$$\text{Var}(\alpha^T X) = \alpha^T \text{Cov}(X)\alpha,$$

siehe (Embrechts et al., 1999, S. 7).

Die leichte Handhabung und die einfache Manipulation unter den linearen Transformationen sind die bedeutendsten Gründe für die große Popularität und weite Verbreitung der Kovarianz und Korrelation. Insbesondere in der praxisnahen Literatur findet diese Methode zur Abhängigkeitsbestimmung großen Anklang. In der Tat ist für viele bivariate Verteilungen die Berechnung der zweiten Momente sogar simpler als die des Erwartungswertes. Und auf der Grundlage einer Stichprobe lassen sich mit geringem Rechenaufwand gute Näherungen für die Kovarianz und Korrelation herleiten. Trotz dieser Vorteile sollte die Korrelation nicht uneingeschränkt verwendet werden, denn nichtlineare Zusammenhänge, auch wenn sie sehr stark sind, werden im Allgemeinen von ihr nicht erfasst und können in der Anwendung zu falschen oder verzerrten Auswertungen führen. Aus diesem Grund soll im Folgenden auf die Nachteile und mögliche Fehlschlüsse bezüglich der linearen Korrelation genauer eingegangen werden.

- Ein Defizit der linearen Korrelation ist, dass sie für Verteilungen mit unendlicher Varianz (engl. heavy-tailed distributions) nicht definiert ist. Das betrifft eine Vielzahl von Verteilungen und ist somit für ein Abhängigkeitsmaß nicht ideal. Beispielsweise, sind die Kovarianz und Korrelation einer bivariaten t -Verteilung mit ν Freiheitsgraden für $\nu \leq 2$ nicht definiert, siehe (Embrechts et al., 1999, S. 7-8).
- Ein möglicher Fehlschluss tritt in Bezug auf die folgende Aussage auf. Zwei Zufallsvariablen sind dann unkorreliert, wenn sie unabhängig sind. Diese Aussage ist uneingeschränkt richtig, doch die Umkehrung, auch wenn es fälschlicherweise oft angenommen wird, ist im allgemeinen Fall falsch. Nimmt man beispielsweise $X \sim N(0, 1)$ und $Y = X^2$, so ist klar, dass die beiden Zufallsgrößen nicht unabhängig sind. Jedoch sind sie wegen $\text{Cov}(X, X^2) = \mathbb{E}(X^3)$ und $\mathbb{E}(X^3) = 0$ unkorreliert, vergleiche (Embrechts et al., 1999, S. 8). Die einzige Ausnahme sind die Komponenten eines multivariat normalverteilten Zufallsvektors. Diese sind genau dann unabhängig, wenn sie unkorreliert sind. Dabei

ist die Bedingung, dass die gemeinsame Verteilung des Vektors normal ist, entscheidend. Es soll im nächsten Punkt genauer darauf eingegangen werden.

- Ein weiterer möglicher Fehlschluss ist die Behauptung, die Randverteilungen und Korrelation bestimmen die gemeinsame Verteilung eines Zufallsvektors eindeutig. Im Allgemeinen können jedoch zu den vorgegebenen Randverteilungen und der Kovarianzmatrix eines Zufallsvektors unendlich viele gemeinsame Verteilungen existieren. Eine Ausnahme bildet die Klasse der elliptischen Verteilungen. Diese Klasse zeichnet die Eigenschaft aus, dass die gemeinsame elliptische Verteilung über die elliptischen Randverteilungen gleichen Typs verfügt. Umgekehrt bestimmen die elliptischen Randverteilungen gleichen Typs und gegebene Kovarianzmatrix ihre gemeinsame Verteilung, falls bekannt ist, dass diese elliptisch ist, eindeutig. Aus diesem Grund kann Korrelation in der elliptischen Verteilungsklasse als natürliches Abhängigkeitsmaß angesehen werden. Lässt man jedoch die Bedingung, dass die gemeinsame Verteilung elliptisch sein muss, fallen, so gibt es auch für elliptische Randverteilungen gleiche Typs und vorgegebene Korrelation unendlich viele Möglichkeiten für die gemeinsame Verteilung. Dieser Sachverhalt soll anhand eines Beispiels demonstriert werden.

Es sollen dazu zwei Zufallsvariablen X und Y betrachtet werden, die jeweils standardnormalverteilt und unkorreliert sind. Unter der Annahme, dass die gemeinsame Verteilung elliptisch sein soll, ist der Vektor $(X, Y)^T$ bivariat standardnormalverteilt mit der Dichtefunktion

$$f(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}.$$

Die Iso-Kurven von $f(x, y)$ sind in der linken Graphik der Abbildung 3.1 dargestellt. Die rechte Graphik zeigt Iso-Kurven der Dichtefunktion

$$g(x, y) = \frac{1}{2\pi} \left[(\sqrt{2} e^{-\frac{x^2}{2}} - e^{-x^2}) e^{-y^2} + (\sqrt{2} e^{-\frac{y^2}{2}} - e^{-y^2}) e^{-x^2} \right].$$

3.1 Lineare Korrelation

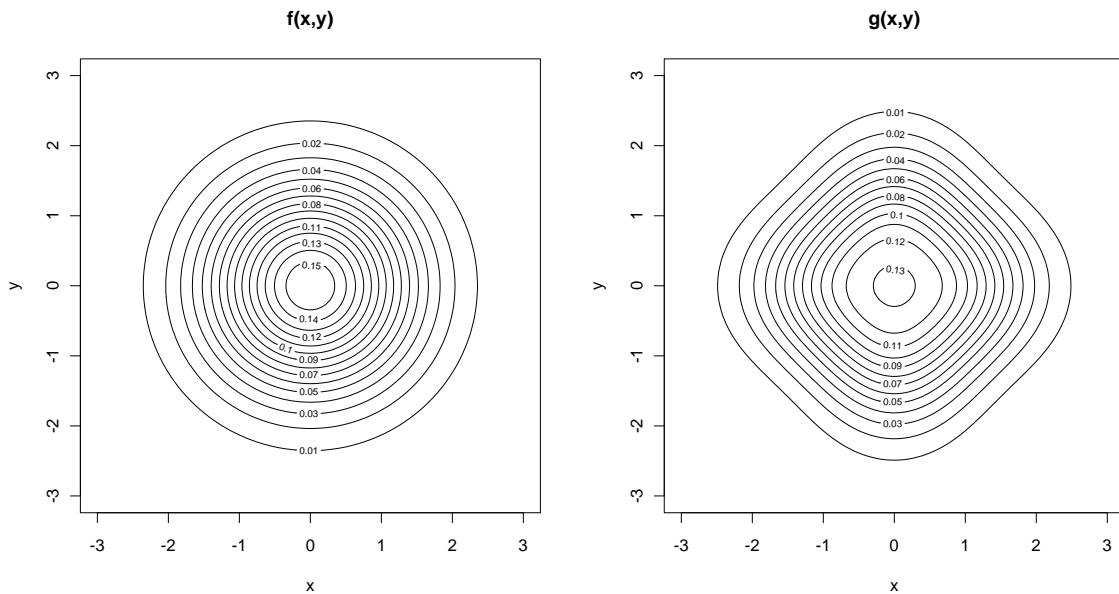


Abbildung 3.1: Iso-Kurven von zwei bivariaten Dichtefunktionen mit jeweils identischen und standardnormalverteilten Randverteilungen, jedoch unterschiedlicher Abhängigkeitsstruktur.

Anhand dieser Darstellung ist gut erkennbar, dass die Iso-Kurven keine elliptische Form haben und die Dichtefunktion $g(x, y)$ somit nicht zu der elliptischen Verteilungsklasse gehört. Trotzdem erfüllt sie wegen

$$\begin{aligned}
 g_X(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[(\sqrt{2} e^{-\frac{x^2}{2}} - e^{-x^2}) e^{-y^2} + (\sqrt{2} e^{-\frac{y^2}{2}} - e^{-y^2}) e^{-x^2} \right] dy \\
 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\
 &= g_Y(x)
 \end{aligned}$$

und

$$\begin{aligned}\text{Corr}(X, Y) &= \mathbb{E}(XY) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy g(x, y) dx dy \\ &= 0\end{aligned}$$

die oben genannten Bedingungen.

- Ein weiterer Nachteil der linearen Korrelation ist, dass sie nur unter *linearen* streng monoton wachsenden Transformationen invariant ist. Im Allgemeinen gilt also für $T : \mathbb{R} \rightarrow \mathbb{R}$ nichtlinear streng monoton wachsend und zwei Zufallsvariablen X und Y :

$$\text{Corr}(T(X), T(Y)) \neq \text{Corr}(X, Y).$$

Betrachtet man zum Beispiel einen bivariaten standardnormalverteilten Zufallsvektor $(X, Y)^T$ mit der Korrelation $\text{Corr}(X, Y)$ und wendet darauf als Transformation die Standardnormalverteilung Φ , so erhält man

$$\text{Corr}(T(X), T(Y)) = \frac{6}{\pi} \arcsin \left(\frac{\text{Corr}(X, Y)}{2} \right),$$

vergleiche (Embrechts et al., 1999, S. 8) oder (Schweizer, 1991, S. 31).

3.2 Copulabasierte Abhängigkeitsmaße

Im Abschnitt 3.2 soll der Abhängigkeitskonzept der Konkordanz angeschaut werden. Dazu werden die Rangkorrelationskoeffizienten von Kendall und Spearman vorgestellt, die einige im Abschnitt 3.1 erläuterte Schwächen der Korrelation überwinden. Entscheidend für die weitere Betrachtung ist außerdem, dass die Konkordanzmaße als Funktionale der Copulafunktionen dargestellt werden können, weshalb sie hier auch als copulabasierte Abhängigkeitsmaße bezeichnet werden.

3.2.1 Konkordanz

Es soll nun der Begriff der Konkordanz eingeführt werden, der vom Lateinischen "concordare" abgeleitet wird und "übereinstimmen" (Pertsch and Lange-Kowal (1999)) bedeutet. Anschaulich lässt sich die Konkordanz so beschreiben: zwei Zufallsvariablen X und Y sind konkordant, wenn die großen Beobachtungswerte von X mit den großen Beobachtungswerten von Y und entsprechend auch die kleinen Werte von X mit den kleinen Werten von Y zusammenhängen, siehe (Nelsen, 2006, S. 157-158)

Definition 3.2.1

Zwei Beobachtungen $(x_1, y_1)^T$ und $(x_2, y_2)^T$ eines stetigen Zufallsvektors $(X, Y)^T$ heißen

- *konkordant*, wenn

$$(x_1 - x_2)(y_1 - y_2) > 0,$$

- *diskordant*, wenn

$$(x_1 - x_2)(y_1 - y_2) < 0.$$

□

Basierend auf der Konkordanz definiert man ein Maß zur Abhängigkeitsbestimmung, der in seinen Eigenschaften viele Defizite der Korrelation aufhebt, siehe (Nelsen, 2006, S. 168-169).

Definition 3.2.2

Ein reellwertiges Zusammenhangsmaß κ_C zwischen zwei stetig verteilten Zufallsvariablen X und Y , deren Copula C ist, wird *Konkordanzmaß* genannt, wenn es folgende Bedingungen erfüllt:

1. κ_C existiert für jedes Paar von stetig verteilten Zufallsvariablen X und Y ;
2. $-1 \leq \kappa_C(X, Y) \leq 1$, $\kappa_C(X, X) = 1$ und $\kappa_C(X, -X) = -1$;
3. $\kappa_C(X, Y) = \kappa_C(Y, X)$;
4. wenn X und Y unabhängig, so ist $\kappa_C(X, Y) = \kappa_P = 0$;

5. $\kappa_C(-X, Y) = \kappa_C(X, -Y) = -\kappa_C(X, Y)$;
6. wenn C_1 und C_2 Copulas mit $C_1 \prec C_2$ sind, so ist $\kappa_{C_1} \leq \kappa_{C_2}$;
7. ist (X_n, Y_n) eine Folge von stetig verteilten Zufallsvariablen mit den Copulas C_n und konvergiert die Folge C_n punktweise gegen C , so gilt: $\lim_{n \rightarrow \infty} \kappa_{C_n} = \kappa_C$.

□

Als Auswirkung der Definition 3.2.2 erhält man den folgenden Satz, siehe (Nelsen, 2006, s. 169).

Satz 3.2.1

Sei κ_C ein Konkordanzmaß für zwei stetig verteilte Zufallsvariablen X und Y , deren Copula C ist, dann gelten:

1. wenn $Y = T(X)$ f.s. und T streng monoton wachsend, dann ist

$$\kappa_C(X, Y) = \kappa_M = 1;$$

2. wenn $Y = T(X)$ f.s. und T streng monoton fallend, dann ist

$$\kappa_C(X, Y) = \kappa_W = -1;$$

3. sind α und β streng monotone Funktionen auf $\text{ran}X$ und $\text{ran}Y$, so ist

$$\kappa_C(\alpha(X), \beta(Y)) = \kappa_C(X, Y).$$

3.2.2 Kendall's τ und Spearman's ρ

Definition 3.2.3

Seien $(X_1, Y_1)^T$ und $(X_2, Y_2)^T$ zwei unabhängige Kopien eines Zufallsvektors $(X, Y)^T$. Dann wird durch

$$\tau(X, Y) = \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) < 0\right) \quad (3.2)$$

der *Rangkorrelationskoeffizient von Kendall* oder *Kendall's τ* definiert, vergleiche (Nelsen, 2006, S. 158).

□

Kendall's τ ist also die Differenz der Wahrscheinlichkeiten für Konkordanz und Diskordanz. Der folgende Satz gibt die bereits erwähnte Darstellung des Kendall's τ als Funktional der Copula an, siehe (Nelsen, 2006, S. 161).

Satz 3.2.2

Seien X und Y zwei stetige Zufallsvariablen, deren zugehörige Copula C ist. Dann lässt sich Kendall's τ für $(X, Y)^T$ schreiben als

$$\tau(X, Y) = 4 \int \int_{[0,1]^2} C(u, v) dC(u, v) - 1. \quad (3.3)$$

Man bemerke, dass das Integral in (3.3) sich auch als Erwartungswert einer Funktion $C(U, V)$ interpretieren lässt, also

$$\tau(X, Y) = 4\mathbb{E}(C(U, V)) - 1.$$

Dabei sind U und V auf $[0, 1]$ gleichverteilt und haben die gemeinsame Verteilungsfunktion C .

Ähnlich Kendall's τ wird Spearman's ρ basierend auf den Wahrscheinlichkeiten für Konkordanz und Diskordanz eingeführt.

Definition 3.2.4

Seien $(X_1, Y_1)^T$, $(X_2, Y_2)^T$ und $(X_3, Y_3)^T$ drei unabhängige Kopien eines Zufallsvektors $(X, Y)^T$. Dann definiert

$$\rho(X, Y) = 3 \left[\mathbb{P} \left((X_1 - X_2)(Y_1 - Y_3) > 0 \right) - \mathbb{P} \left((X_1 - X_2)(Y_1 - Y_3) < 0 \right) \right] \quad (3.4)$$

den *Rangkorrelationskoeffizienten von Spearman* oder *Spearman's ρ* , vergleiche (Nelsen, 2006, S. 167).

□

Der folgende Satz soll zeigen, dass auch Spearman's ρ als Funktional der Copula geschrieben werden kann, vergleiche (Nelsen, 2006, S. 167).

Satz 3.2.3

Seien X und Y zwei stetige Zufallsvariablen, deren zugehörige Copula C ist. Dann lässt sich Spearman's ρ für $(X, Y)^T$ schreiben als

$$\rho(X, Y) = 12 \int \int_{[0,1]^2} C(u, v) \, du \, dv - 3 = 12 \int \int_{[0,1]^2} uv \, dC(u, v) - 3. \quad (3.5)$$

Man beachte, dass der Ausdruck in (3.5) als Erwartungswert einer bivariaten Funktion UV mit auf $[0, 1]$ gleichverteilten Zufallsvariablen U und V und der gemeinsamen Verteilungsfunktion C gedeutet werden kann. Benutzt man weiter, dass auf $[0, 1]$ gleichverteilte Zufallsvariablen den Erwartungswert $\frac{1}{2}$ und die Varianz $\frac{1}{12}$ haben, so lässt sich das Integral umformen in

$$\begin{aligned} \rho(X, Y) &= 12 \int \int_{[0,1]^2} uv \, dC(u, v) - 3 = 12\mathbb{E}(UV) - 3 \\ &= \frac{\mathbb{E}(UV) - \frac{1}{4}}{\frac{1}{12}} = \frac{\mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V)}{\sqrt{\text{Var}(U)}\sqrt{\text{Var}(V)}} \\ &= \text{Corr}(U, V). \end{aligned}$$

Damit ist Spearman's ρ für X und Y gleich der Bravais-Pearson-Korrelation für U und V . Setzt man außerdem $U = F_1(X)$ und $V = F_2(Y)$ mit stetigen Verteilungsfunktionen G und H an, so erhält man:

$$\rho(X, Y) = \text{Corr}(F_1(X), F_2(Y)),$$

vergleiche (Nelsen, 2006, S. 169-170).

Im Abschnitt 3.2.1 wurden die Eigenschaften eines Konkordanzmaßes beschrieben. Der nachfolgende Satz zeigt, dass Kendall's τ und Spearman's ρ diese Eigenschaften erfüllen und somit selbst Konkordanzmaße sind.

Satz 3.2.4

Sind X und Y zwei stetige Zufallsvariablen, deren zugehörige Copula C ist, so erfüllen Kendall's τ (3.2) und Spearman's ρ (3.4) alle Eigenschaften aus der Definition 3.2.2 und dem Satz 3.2.1 für die Konkordanzmaße.

Beweis.

Siehe (Nelsen, 2006, S. 169).

Mit dem Konkordanzmaß wurde ein Maß für die Abhängigkeit definiert, das viele wünschenswerte Eigenschaften erfüllt. Jedoch hat die Definition 3.2.2 in der 4. Eigenschaft eine Schwäche, denn zwei Zufallsvariablen können trotz $\kappa_C = 0$ abhängig sein. Dies sollte bei einem Abhängigkeitsmaß idealerweise nicht der Fall sein. Aus diesem Grund soll im Folgenden ein Maß für die Abhängigkeit angegeben werden, der diese Schwäche überwindet und den Wert 0 genau dann annimmt, wenn die Zufallsvariablen unabhängig sind, siehe (Nelsen, 2006, S. 208) oder (Schweizer, 1991, S. 30-31).

Definition 3.2.5

Ein reellwertiges Zusammenhangsmaß δ_C zwischen zwei stetig verteilten Zufallsvariablen X und Y , deren Copula C ist, wird *Abhängigkeitsmaß* genannt, wenn es folgende Bedingungen erfüllt:

1. δ_C existiert für jedes Paar von stetig verteilten Zufallsvariablen X und Y ;
2. $0 \leq \delta_C(X, Y) \leq 1$;
3. $\delta_C(X, Y) = \delta_C(Y, X)$;
4. $\delta_C(X, Y) = 0$ genau dann, wenn X und Y unabhängig sind;
5. $\delta_C(X, Y) = 1$ genau dann, wenn jede der Zufallsvariablen X und Y f.s. eine streng monotone Funktion der anderen ist;
6. sind α und β streng monotone Funktionen auf $\text{ran}X$ und $\text{ran}Y$, so ist $\delta_C(\alpha(X), \beta(Y)) = \delta_C(X, Y)$;
7. ist (X_n, Y_n) eine Folge von stetig verteilten Zufallsvariablen mit den Copulas C_n und konvergiert die Folge C_n punktweise gegen C , so gilt: $\lim_{n \rightarrow \infty} \delta_{C_n} = \delta_C$.

□

Um ein Beispiel für ein solches Abhängigkeitsmaß anzugeben, wird Spearman's ρ betrachtet und die bereits bekannte Gleichung (3.5) umgeschrieben

$$\begin{aligned} \rho(X, Y) &= 12 \int \int_{[0,1]^2} C(u, v) \, du \, dv - 3 \\ &= 12 \int \int_{[0,1]^2} [C(u, v) - uv] \, du \, dv \\ &= 12 \int \int_{[0,1]^2} [C(u, v) - P(u, v)] \, dP(u, v). \end{aligned}$$

Damit ist Spearman's ρ bis auf Vorzeichen proportional zu dem Volumen zwischen den Graphen der Copula C und der Produktcopula P . Ersetzt man die Differenz $[C(u, v) - P(u, v)]$ durch die absolute Differenz $|C(u, v) - P(u, v)|$, so erhält man ein Maß, das auf dem L_1 -Abstand zwischen den Graphen von C und P basiert. Dieses Maß, das nach seinen Entdeckern Schweizer und Wolff benannt wurde, erfüllt alle gewünschten Eigenschaften und ist somit gemäß der Definition 3.2.5 ein Abhängigkeitsmaß, siehe (Nelsen, 2006, S. 209) oder (Schweizer, 1991, S. 30).

Definition 3.2.6

Seien X und Y zwei stetige Zufallsvariablen, deren zugehörige Copula C ist. Dann ist das Maß von Schweizer und Wolff oder Schweizer und Wolff's σ gegeben durch

$$\sigma = 12 \int \int_{[0,1]^2} |C(u, v) - uv| \, du \, dv. \tag{3.6}$$

□

Satz 3.2.5

Seien X und Y stetig verteilte Zufallsvariablen, deren zugehörige Copula C ist. Dann ist das in (3.6) definierte Maß σ ein Abhängigkeitsmaß und erfüllt alle in der Definition 3.2.5 aufgeführten Eigenschaften.

Beweis.

Siehe Schweizer and Wolff (1981) oder (Nelsen, 2006, S. 209).

3.2.3 Asymptotische Randabhängigkeit

Es wurden bereits Abhängigkeitskonzepte vorgestellt, die auf der Beschreibung, wie die großen (oder kleinen) Werte einer Zufallsvariablen mit den großen (oder klei-

nen) Werten der anderen Zufallsvariablen zusammenhängen, basieren. In praxisorientierten Anwendungen ist oft interessant, wie sich solche Zusammenhänge in Extremsituationen entwickeln. Aus diesem Grund soll in diesem Abschnitt ein Konzept angeschaut werden, der die Abhängigkeiten zwischen Extremwerten zweier Zufallsvariablen untersucht.

Definition 3.2.7

Seien X und Y zwei stetig verteilte Zufallsvariablen mit Verteilungsfunktionen F_1 und F_2 . Dann bezeichnet

$$\lambda_O = \lim_{\alpha \rightarrow 1^-} \mathbb{P}(Y > F_2^-(\alpha) \mid X > F_1^-(\alpha)) \tag{3.7}$$

den *Koeffizienten für die obere asymptotische Randabhängigkeit* von X und Y , falls der linksseitiger Grenzwert $\lambda_O \in [0, 1]$ existiert. Infolgedessen heißen X und Y

- *asymptotisch abhängig im oberen Rand*, wenn $\lambda_O \in (0, 1]$,
- *asymptotisch unabhängig im oberen Rand*, wenn $\lambda_O = 0$.

□

Demnach sind X und Y asymptotisch abhängig im oberen Rand, wenn eine positive Wahrscheinlichkeit für das tendenziell gleichzeitige Auftreten der großen Realisierungen von X und Y besteht. Die Abbildung 3.2 zeigt die Darstellung von 10.000 simulierten Datenpaaren zweier bivariaten Verteilungen mit jeweils standardnormalverteilten Randverteilungen und Korrelation von 0,9. Auch wenn man über Simulation keine eindeutige Aussage über die asymptotische Randabhängigkeit machen kann, so lässt sich anhand der Bilder eine Vermutung anstellen. Das linke Diagramm zeigt eine kompakte Streuung in Form einer Ellipse und lässt sowohl oben als auch unten keine oder sehr geringe asymptotische Randabhängigkeit erkennen. Tatsächlich wird man im Kapitel 4 sehen, dass die der Simulation zugrunde liegende Normal-Copula mit dem Korrelationskoeffizienten $\text{Corr} = 0,9$ keine asymptotische Randabhängigkeit im oberen und unteren Rand besitzt. Die Punktwolke im rechten Diagramm ist auf der Grundlage einer t -Copula mit $\nu = 2$ Freiheitsgraden simuliert und zeigt erkennbare Ausschweifungen im oberen und unteren Bereich, was

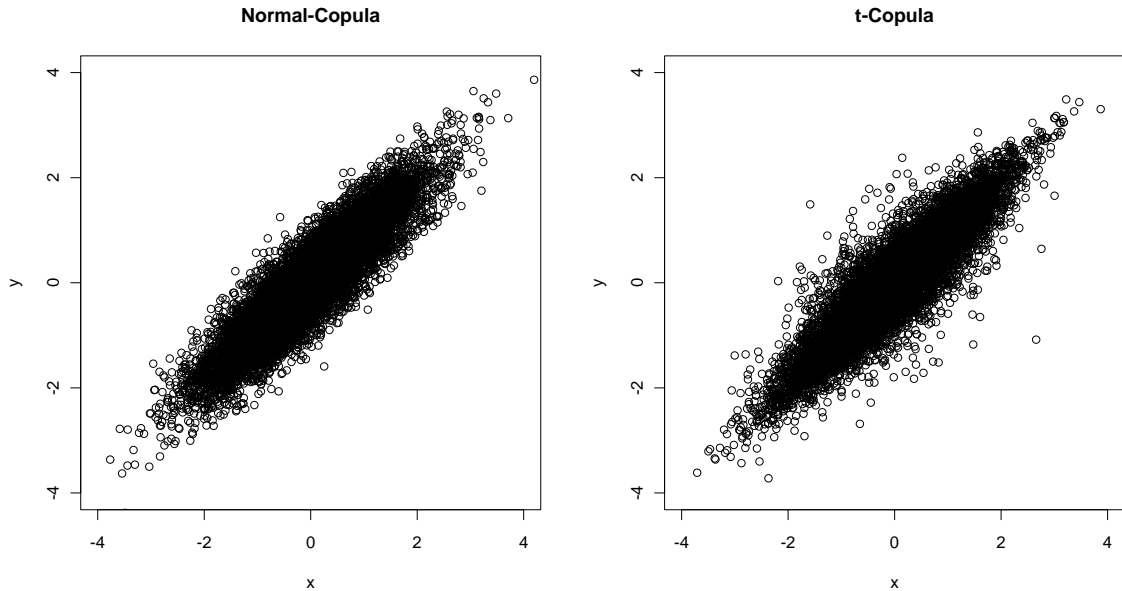


Abbildung 3.2: 10.000 simulierte Datenpaare der Normal-Copula auf der linken Seite und t -Copula mit $\nu = 2$ Freiheitsgraden auf der rechten Seite. Die Randverteilungen sind jeweils identisch und standardnormalverteilt mit Korrelation von 0,9.

auf entsprechende asymptotische Randabhängigkeiten hindeutet. Bei der ausführlichen Betrachtung der t -Copula im Kapitel 4 wird sich zeigen, dass im vorliegenden Fall $\lambda_O = \lambda_U = 0,72$ ist.

Das Konzept der unteren asymptotischen Randabhängigkeit lässt sich nun in analoger Weise einführen.

Definition 3.2.8

Seien X und Y zwei stetig verteilte Zufallsvariablen mit Verteilungsfunktionen F_1 und F_2 . Dann bezeichnet

$$\lambda_U = \lim_{\alpha \rightarrow 0^+} \mathbb{P}(Y \leq F_2^-(\alpha) \mid X \leq F_1^-(\alpha)) \tag{3.8}$$

3.2 Copulabasierte Abhängigkeitsmaße

den Koeffizienten für die untere asymptotische Randabhängigkeit von X und Y , falls der Grenzwert $\lambda_O \in [0, 1]$ existiert. Infolgedessen heißen X und Y

- *asymptotisch abhängig im unteren Rand*, wenn $\lambda_U \in (0, 1]$,
- *asymptotisch unabhängig im unteren Rand*, wenn $\lambda_U = 0$.

□

Gemäß der Definition für bedingte Wahrscheinlichkeit kann man den Ausdruck $\mathbb{P}(Y > F_2^-(u) \mid X > F_1^-(u))$ umformen in

$$\begin{aligned}
 & \mathbb{P}(Y > F_2^-(u) \mid X > F_1^-(u)) \\
 &= \frac{\mathbb{P}(F_1(X) > u, F_2(Y) > u)}{\mathbb{P}(F_1(X) > u)} \\
 &= \frac{1 - \mathbb{P}(F_1(X) > u \vee F_2(Y) > u)}{1 - \mathbb{P}(F_1(X) \leq u)} \\
 &= \frac{1 - \mathbb{P}(F_1(X) \leq u) - \mathbb{P}(F_2(Y) \leq u) + \mathbb{P}(F_1(X) \leq u, F_2(Y) \leq u)}{1 - \mathbb{P}(F_1(X) \leq u)}.
 \end{aligned}$$

Der Wahrscheinlichkeitsausdruck in (3.8) lässt sich auf gleiche Weise umschreiben in

$$\mathbb{P}(Y \leq F_2^-(u) \mid X \leq F_1^-(u)) = \frac{\mathbb{P}(F_1(X) \leq u, F_2(Y) \leq u)}{\mathbb{P}(F_1(X) \leq u)}.$$

Daraus wird der nachfolgende Satz abgeleitet, der die asymptotische Randabhängigkeit als eine Eigenschaft der Copula interpretieren lässt.

Satz 3.2.6

Seien X und Y zwei stetig verteilte Zufallsvariablen mit den zugehörigen Copula C und der Überlebenscopula \bar{C} . Wenn die in (3.7) und (3.8) definierte Grenzwerte existieren, so gilt:

$$\lambda_O = \lim_{u \rightarrow 1^-} \frac{\bar{C}(u, u)}{1 - u} = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u} = 2 - \lim_{u \rightarrow 1^-} \frac{1 - C(u, u)}{1 - u}$$

und

$$\lambda_U = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}.$$

Bei $\lambda_O \in (0, 1]$ ($\lambda_U \in (0, 1]$) spricht man davon, dass die Copula C die obere (untere) asymptotische Randabhängigkeit besitzt. Wenn $\lambda_O = 0$ ($\lambda_U = 0$) ist, so hat C keine obere (untere) asymptotische Randabhängigkeit, siehe (Nelsen, 2006, S. 215).

Für Copulas mit komplexer Struktur, wie beispielsweise die Normal- oder die t-Copula, eignet sich eine andere Formel für die asymptotische Randabhängigkeit besser, vergleiche (Embrechts et al., 1999, S. 18). Seien U_1 und U_2 auf $[0, 1]$ gleichverteilte Zufallsvariablen mit der gemeinsamen Verteilungsfunktion $C(u_1, u_2)$, die in beiden Variablen u_1 und u_2 differenzierbar ist. Wendet man die Regel von l'Hospital an, so lässt sich der Grenzwert λ_O umformen in

$$\lambda_O = - \lim_{u \rightarrow 1^-} \frac{d\bar{C}(u, u)}{du} = \lim_{u \rightarrow 1^-} \mathbb{P}(U_2 > u \mid U_1 = u) + \lim_{u \rightarrow 1^-} \mathbb{P}(U_1 > u \mid U_2 = u).$$

Lassen sich die Variablen der Copula C vertauschen in dem Sinne, dass $(U_1, U_2)^T \stackrel{d}{=} (U_2, U_1)^T$, dann folgt

$$\lambda_O = 2 \lim_{u \rightarrow 1^-} \mathbb{P}(U_2 > u \mid U_1 = u).$$

Sind außerdem U_1 und U_2 identisch verteilt, dann erhält man oft unter der Anwendung der Quantil-Transformation F_1^{-1} auf beide Randverteilungen eine bivariate Verteilung mit bekannten bedingten Randverteilungen, die die Berechnung des Grenzwertes möglich macht. Setzt man also voraus, dass F_1 eine stetige und streng monotone Verteilungsfunktion mit infinitem rechtem Endpunkt ist, dann ist

$$\begin{aligned} \lambda_O &= 2 \lim_{u \rightarrow 1^-} \mathbb{P}(U_2 > u \mid U_1 = u) \\ &= 2 \lim_{x \rightarrow \infty} \mathbb{P}(F_1^{-1}(U_2) > x \mid F_1^{-1}(U_1) = x) \\ &= 2 \lim_{x \rightarrow \infty} \mathbb{P}(X_2 > x \mid X_1 = x) \end{aligned}$$

für $(X_1, X_2)^T \sim C(F_1(x_1), F_2(x_2))$.

3.2 Copulabasierte Abhängigkeitsmaße

Unter allen getroffenen Voraussetzungen und mit gleichen Überlegungen kann für die untere asymptotische Randabhängigkeit λ_U folgende Formel hergeleitet werden:

$$\begin{aligned}\lambda_U &= \lim_{u \rightarrow 1^-} \mathbb{P}(U_2 \leq u \mid U_1 = u) + \lim_{u \rightarrow 1^-} \mathbb{P}(U_1 \leq u \mid U_2 = u) \\ &= 2 \lim_{u \rightarrow 1^-} \mathbb{P}(U_2 \leq u \mid U_1 = u) \\ &= 2 \lim_{x \rightarrow \infty} \mathbb{P}(F_1^{-1}(U_2) \leq x \mid F_1^{-1}(U_1) = x) \\ &= 2 \lim_{x \rightarrow \infty} \mathbb{P}(X_2 \leq x \mid X_1 = x),\end{aligned}$$

siehe (Glauser, 2003, S. 85-86).

Kapitel 4

Elliptische Copulaklasse

In den vorherigen Kapiteln wurden die allgemeine Copulatheorie, das Konzept der Abhängigkeit und einige copulabasierte Abhängigkeitsmaße betrachtet. In diesem Kapitel wird eine parametrische Copulafamilie, die elliptischen Copulafunktionen, untersucht. Parametrisch heißt eine Copulafunktion, wenn sie durch einen endlich-dimensionalen Parameter beschrieben wird.

Auf dem Weg zu den elliptischen Copulafunktionen soll die Klasse der sphärischen und elliptischen Verteilungen analysiert werden. Eine Zusammenfassung wichtiger und interessanter Eigenschaften soll ein besseres Verständnis über diese Art der Verteilungen schaffen.

Anschließend wird die elliptische Copulafamilie anhand von zwei wichtigen Vertretern, der multivariaten Normal-Copula und der multivariaten t -Copula, vorgestellt. Für jede Copulafunktion wird die funktionale Form und die Darstellung von einigen copulabasierten Abhängigkeitsmaßen hergeleitet. Es wird außerdem gezeigt, dass der Rangkorrelationskoeffizient von Kendall in direkter Beziehung zum Parameter der elliptischen Copulafunktionen steht und in diesem Zusammenhang für die Schätzung des Parameters eingesetzt werden kann. Zur Generierung der gemäß der jeweiligen Copula verteilten Realisierungen werden geeignete Algorithmen angeführt.

4.1 Sphärische und elliptische Verteilungen

Es gibt einige Ansätze, die sphärischen und elliptischen Verteilungen zu definieren. So stellt die sphärische Verteilung die Erweiterung der Standardnormalverteilung $N_n(\mathbf{0}, \mathbf{I}_n)$ dar und die elliptische Verteilung ist die Erweiterung der multivariaten Normalverteilung $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Bekannterweise kann die Normalverteilung $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ über die Standardnormalverteilung $N_n(\mathbf{0}, \mathbf{I}_n)$ definiert werden. Sei $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, dann gilt:

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{A}^T \mathbf{Y},$$

wobei $\mathbf{Y} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ und $\boldsymbol{\Sigma} = \mathbf{A}^T \mathbf{A}$. Aus dieser Darstellung heraus lassen sich viele Eigenschaften von $N_n(\mathbf{0}, \mathbf{I}_n)$ auf $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ übertragen, siehe (Fang et al., 1989, S. 27). Es kann gezeigt werden, dass einige interessante Eigenschaften und Ergebnisse, die uns von der Klasse der Normalverteilungen bekannt sind, in gleicher Weise für die sphärischen und elliptischen Verteilungsklassen gelten.

Als erstes sollen die Definition und Eigenschaften der sphärischen Verteilung betrachtet werden, um anschließend die elliptische Verteilungsklasse zu analysieren.

Definition 4.1.1

Ein n -dimensionaler Zufallsvektor \mathbf{X} ist *sphärisch verteilt*, wenn für jedes $\Gamma \in \mathcal{O}(n)$

$$\Gamma \mathbf{X} \stackrel{d}{=} \mathbf{X},$$

gilt. Dabei bezeichnet $\mathcal{O}(n)$ die Menge der orthogonalen $n \times n$ -Matrizen und das Symbol $\stackrel{d}{=}$ steht für die Verteilungsgleichheit von zwei Zufallsvariablen.

□

Also heißt X sphärisch verteilt, wenn die Verteilung von X invariant gegenüber orthogonalen Transformationen ist. Anschaulich bedeutet das, dass bei einer bivariaten sphärischen Verteilung die Konturlinien der Dichtefunktion Kreise sind. Eine äquivalente Art, sphärische Verteilungen zu definieren, zeigt der folgende Satz.

Satz 4.1.1

Ein n -dimensionaler Zufallsvektor \mathbf{X} ist genau dann sphärisch verteilt, wenn seine charakteristische Funktion $\psi(\mathbf{t})$ eine der folgenden äquivalenten Eigenschaften erfüllt:

1. $\psi(\mathbf{\Gamma}^T \mathbf{t}) = \psi(\mathbf{t})$ für ein $\mathbf{\Gamma} \in \mathcal{O}(n)$;
2. Es existiert eine Funktion $\phi(\cdot)$ von der Skalarvariablen $\mathbf{t}^T \mathbf{t}$, so dass

$$\psi(\mathbf{t}) = \phi(\mathbf{t}^T \mathbf{t}) = \phi(\|\mathbf{t}\|^2). \tag{4.1}$$

Beweis.

Siehe (Fang et al., 1989, S. 28).

Wenn die charakteristische Funktion von \mathbf{X} die Form wie in (4.1) besitzt, so wird die Funktion $\phi(\cdot)$ *charakteristische Generatorfunktion* genannt und die sphärische Verteilung von \mathbf{X} kurz mit $S_n(\phi)$ notiert.

Beispiel 4.1.1

Ein n -dimensionaler, auf der Einheitssphäre $S_n := \{\mathbf{u} \in \mathbb{R}^n \mid \|\mathbf{u}\|^2 = 1\}$ gleichverteilter Zufallsvektor \mathbf{U} ist gemäß der Definition 4.1.1 sphärisch verteilt, siehe (Fang et al., 1989, S. 28).

□

Zukünftig soll ein n -dimensionaler Zufallsvektor, der auf der Einheitssphäre S_n gleichverteilt ist, mit $\mathbf{U}^{(n)}$ bezeichnet werden.

Beispiel 4.1.2 (Multivariate Standardnormalverteilung)

Sei $\mathbf{X} \sim N_n(\mathbf{0}, \mathbf{I}_n)$, dann weiß man, dass die Komponenten $X_i \sim N(0, 1)$, $i = 0, \dots, n$, unabhängig sind und die charakteristische Funktion $\psi_i(t_i) = \exp(-t_i^2/2)$ besitzen. Die charakteristische Funktion von \mathbf{X} ist

$$\psi(\mathbf{t}) = \exp\left(-\frac{1}{2}(t_1^2 + \dots + t_n^2)\right) = \exp\left(-\frac{1}{2}\|\mathbf{t}\|^2\right).$$

Demnach weist ψ die Form wie in (4.1) mit der charakteristischen Generatorfunktion $\phi(u) = \exp(-u/2)$ auf. Die standardnormalverteilte Zufallsvariable X ist also sphärisch verteilt, siehe (Fang et al., 1989, S. 28).

□

Es wurde gezeigt, dass die charakteristische Funktion einer sphärisch verteilten Zufallsgröße sich über eine gewisse Skalarfunktion generieren lässt. Interessant ist die Frage, wie man solche Generatorfunktionen darstellen kann. Um die Frage zu beantworten, soll ein neuer Begriff eingeführt werden.

Sei Φ_n die Familie aller möglichen charakteristischen Generatorfunktionen für einen n -dimensionalen Zufallsvektor und definiert über

$$\Phi_n = \{\phi(\cdot) \mid \phi(t_1^2 + \dots + t_n^2) \text{ ist eine } n\text{-dimensionale charakteristische Funktion}\}.$$

Es ist leicht zu sehen, dass $\Phi_1 \supset \Phi_2 \supset \Phi_3 \dots$ gilt, siehe (Fang et al., 1989, S. 29).

Satz 4.1.2

Eine Funktion ϕ gehört genau dann zu Φ_n , wenn

$$\phi(x) = \int_0^\infty \psi_n(xr^2) dF(r), \tag{4.2}$$

wobei $F(\cdot)$ eine Verteilungsfunktion auf $[0, \infty)$ und

$$\psi_n(\mathbf{y}^T \mathbf{y}) = \int_{S_n} e^{i\mathbf{y}^T \mathbf{x}} \frac{d\mathbf{x}}{S}$$

mit S dem Volumeninhalt von der Einheitskugel S_n . Demnach ist $\psi_n(\mathbf{y}^T \mathbf{y})$ die charakteristische Funktion der Zufallsvariablen $\mathbf{U}^{(n)}$.

Beweis.

Siehe (Fang et al., 1989, S. 29-30).

In direkter Konsequenz aus dem Satz 4.1.2 lässt sich folgende Darstellung einer sphärisch verteilten Zufallsvariablen herleiten, vergleiche (Fang et al., 1989, S. 30-31).

Korollar 4.1.1

Sei \mathbf{X} ein n -dimensionaler sphärisch verteilter Zufallsvektor mit der charakteristischen Generatorfunktion $\phi(\mathbf{t}^T \mathbf{t})$ und $\phi \in \Phi_n$. Dann besitzt \mathbf{X} die Darstellung

$$\mathbf{X} \stackrel{d}{=} R_n \mathbf{U}^{(n)},$$

wobei $R_n \sim F(x)$ und unabhängig von $\mathbf{U}^{(n)}$ ist, und $F(x)$ in der Beziehung (4.2) zu ϕ steht.

4.1 Sphärische und elliptische Verteilungen

Wenn immer im Folgenden die Zerlegung $\mathbf{X} \stackrel{d}{=} R_n \mathbf{U}^{(n)}$ benutzt wird, so ist diese im Sinne des Korollars 4.1.1 zu verstehen. Der Satz 4.1.3 gibt die Darstellung der Variablen R_n und $\mathbf{U}^{(n)}$ für den Spezialfall, dass $\mathbf{X} \stackrel{d}{=} R_n \mathbf{U}^{(n)}$ und $\mathbb{P}(\mathbf{X} = 0) = 0$, an.

Satz 4.1.3

Angenommen $\mathbf{X} \stackrel{d}{=} R_n \mathbf{U}^{(n)}$ und $\mathbb{P}(\mathbf{X} = 0) = 0$, so gilt:

$$\|\mathbf{X}\| \stackrel{d}{=} R_n, \quad \frac{\mathbf{X}}{\|\mathbf{X}\|} \stackrel{d}{=} \mathbf{U}^{(n)}. \quad (4.3)$$

Außerdem sind $\|\mathbf{X}\|$ und $\mathbf{X}/\|\mathbf{X}\|$ unabhängig.

Beweis.

Siehe (Fang et al., 1989, S. 31).

Es wurde gezeigt, dass die sphärische Verteilung auf verschiedene Arten hergeleitet werden kann. Im Folgenden soll die Klasse der elliptischen Verteilungen betrachtet werden, die als Erweiterung der sphärischen Verteilungsklasse gesehen werden kann, siehe (Fang et al., 1989, S. 31).

Definition 4.1.2

Seien \mathbf{X} ein n -dimensionaler Zufallsvektor, $\boldsymbol{\mu} \in \mathbb{R}^n$ und $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ eine nichtnegativ definite Matrix. Dann sagt man, \mathbf{X} ist *elliptisch verteilt* mit den Parametern $\boldsymbol{\mu}$ und $\boldsymbol{\Sigma}$ oder kurz $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$, wenn

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{A}^T \mathbf{Y},$$

wobei $\mathbf{Y} \sim S_k(\phi)$, $\mathbf{A} \in \mathbb{R}^{k \times n}$ mit $\mathbf{A}^T \mathbf{A} = \boldsymbol{\Sigma}$ und $\text{Rang}(\boldsymbol{\Sigma}) = k$ sind.

□

Mit Hilfe der Definition 4.1.2 und des Korollars 4.1.1 lässt sich eine weitere stochastische Repräsentation von einer elliptisch verteilten Zufallsgröße herleiten.

Satz 4.1.4

Ein n -dimensionaler Zufallsvektor $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ mit $\text{Rang}(\boldsymbol{\Sigma}) = k$ besitzt die Darstellung

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + R_k \mathbf{A}^T \mathbf{U}^{(k)}$$

für ein $R_k \geq 0$ unabhängig von $\mathbf{U}^{(k)}$ und $\mathbf{A} \in \mathbb{R}^{k \times n}$ mit $\mathbf{A}^T \mathbf{A} = \boldsymbol{\Sigma}$.

Beispiel 4.1.3 (Multivariate Normalverteilung)

Besitzt ein n -dimensionaler Zufallsvektor \mathbf{X} die Darstellung

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{A}^T \mathbf{Y},$$

wobei $\boldsymbol{\mu} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m \times n}$ und $\mathbf{Y} \sim N_m(\mathbf{0}, \mathbf{I}_m)$, dann ist bekannt, dass \mathbf{X} multivariat normalverteilt ist, $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ mit $\boldsymbol{\Sigma} = \mathbf{A}^T \mathbf{A}$. Aus dem Beispiel 4.1.2 weiß man, dass $\mathbf{Y} \sim \mathcal{S}_m(\phi)$ mit $\phi(u) = \exp(-u/2)$. Und damit folgt nach der Definition 4.1.2, dass $\mathbf{X} \sim \mathcal{E}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$. Aufgrund des Korollars 4.1.1 und des Satzes 4.1.3 lässt sich \mathbf{X} äquivalent darstellen als

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \|\mathbf{Y}\| \mathbf{A}^T \mathbf{U}^{(m)},$$

wobei $\|\mathbf{Y}\|$ und $\mathbf{U}^{(m)}$ unabhängig sind.

□

Beispiel 4.1.4 (Multivariate t-Verteilung)

Seien $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ und $S \sim \chi_\nu^2$ unabhängige Zufallsvariablen. Dann ist

$$\mathbf{Y} = \frac{\mathbf{Z}}{\sqrt{\frac{S}{\nu}}}$$

multivariat t -verteilt mit ν Freiheitsgraden, kurz: $\mathbf{Y} \sim t_\nu^n(\mathbf{0}, \mathbf{I}_n)$. Nutzt man die Darstellung aus dem Satz 4.1.3, so erhält man

$$\mathbf{Y} \stackrel{d}{=} \frac{\|\mathbf{Z}\| \mathbf{U}^{(n)}}{\sqrt{\frac{S}{\nu}}} = R_n \mathbf{U}^{(n)},$$

wobei R_n und $\mathbf{U}^{(n)}$ unabhängig sind und $R_n = \frac{\sqrt{\nu} \|\mathbf{Z}\|}{\sqrt{S}}$ mit $\|\mathbf{Z}\|^2 \sim \chi_n^2$. Wegen des Korollars 4.1.1 ist \mathbf{Y} sphärisch verteilt. Sei nun

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{A}^T \mathbf{Y}$$

mit $\boldsymbol{\mu} \in \mathbb{R}^n$ und $\mathbf{A} \in \mathbb{R}^{m \times n}$, dann sagt man, \mathbf{X} ist multivariat t -verteilt zu den Parametern $\boldsymbol{\mu}$, $\boldsymbol{\Sigma} = \mathbf{A}^T \mathbf{A}$ und ν Freiheitsgraden. Man schreibt kurz $\mathbf{X} \sim t_\nu^n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Gemäß der Definition 4.1.2 ist \mathbf{X} elliptisch verteilt und es gilt $t_\nu^n(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{E}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ für eine spezielle Generatorfunktion ϕ . Mehr Details zur t -Verteilung und ϕ findet man bei Fang et al. (1989).

□

Mit den Beispielen 4.1.3 und 4.1.4 wurden zwei wichtige Vertreter der multivariaten elliptischen Verteilungsklasse präsentiert, die multivariate Normalverteilung und die multivariate t -Verteilung. Um noch mehr über die Verteilungen dieser Klasse zu erfahren, werden in den nächsten Sätzen einige interessante Ergebnisse dargestellt.

Satz 4.1.5

Sei $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ mit $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ vom $\text{Rang}(\boldsymbol{\Sigma}) = k$, dann besitzt die charakteristische Funktion von \mathbf{X} die Form

$$\psi_{\mathbf{X}}(\mathbf{t}) = e^{i\mathbf{t}^T \boldsymbol{\mu}} \phi(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}).$$

Beweis.

Laut Definition 4.1.2, existiert ein Zufallsvektor $\mathbf{Y} \sim \mathcal{S}_k(\phi)$, so dass $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{A}^T \mathbf{Y}$ mit $\mathbf{A}^T \mathbf{A} = \boldsymbol{\Sigma}$ und $\mathbf{A} \in \mathbb{R}^{k \times n}$. Wegen des Eindeutigkeitssatzes für charakteristische Funktionen gilt:

$$\begin{aligned} \psi_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E}\left(e^{i\mathbf{t}^T \mathbf{X}}\right) = \mathbb{E}\left(e^{i\mathbf{t}^T (\boldsymbol{\mu} + \mathbf{A}^T \mathbf{Y})}\right) \\ &= e^{i\mathbf{t}^T \boldsymbol{\mu}} \mathbb{E}\left(e^{i(\mathbf{A}\mathbf{t})^T \mathbf{Y}}\right) = e^{i\mathbf{t}^T \boldsymbol{\mu}} \phi((\mathbf{A}\mathbf{t})^T \mathbf{A}\mathbf{t}) \\ &= e^{i\mathbf{t}^T \boldsymbol{\mu}} \phi(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}). \end{aligned}$$

■

Die Zufallsvariable X muss nicht notwendigerweise eine Dichte besitzen. Falls jedoch die Dichtefunktion existiert, so muss sie die Form wie im Satz 4.1.6 haben.

Satz 4.1.6

Sei $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$, wobei $\boldsymbol{\mu} \in \mathbb{R}^n$ und $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ positiv definit mit $\text{Rang}(\boldsymbol{\Sigma}) = n$ sind. \mathbf{X} besitze die stochastische Zerlegung $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + R_n \mathbf{A}^T \mathbf{U}^{(n)}$ mit nichtsigulärer Matrix \mathbf{A} und $\mathbf{A}^T \mathbf{A} = \boldsymbol{\Sigma}$. Außerdem sei R_n absolut stetig verteilt mit der Dichte f_{R_n} . Dann besitzt die Dichtefunktion von \mathbf{X} folgende Form:

$$f_{\mathbf{X}}(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1/2} g\left((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad \mathbf{x} \neq \boldsymbol{\mu},$$

wobei

$$g_{\mathbf{X}}(t) := \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{n/2}} \sqrt{t}^{-(n-1)} \cdot f_{R_n}(\sqrt{t}), \quad t > 0.$$

Beweis.

Den vollständigen Beweis findet man bei (Frahm, 2004, S. 8-9) (oder in einer anderen Form bei Fang et al. (1989)). Hier wird eine Beweisskizze angeführt.

Zunächst wird die Dichte einer sphärisch verteilten Zufallsvariable $\mathbf{Y} \stackrel{d}{=} R_n \mathbf{U}^{(n)}$ hergeleitet. Da R_n und $\mathbf{U}^{(n)}$ stochastisch unabhängig sind und $\mathbf{U}^{(n)}$ auf der Einheitssphäre S^n gleichverteilt ist, lässt sich die gemeinsame Dichte $f_{(R_n, \mathbf{U}^{(n)})}$ bestimmen als

$$f_{(R_n, \mathbf{U}^{(n)})}(r, \mathbf{u}) = \frac{\Gamma(\frac{n}{2})}{2\pi^{n/2}} f_{R_n}(r).$$

Mit Hilfe einer geeigneten Variablentransformation kann folgende Darstellung der Dichte von Y erreicht werden:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma(\frac{n}{2})}{2\pi^{n/2}} \|\mathbf{y}\|^{-(n-1)} f_{R_n}(\|\mathbf{y}\|), \quad \mathbf{y} \neq 0.$$

Anschließend wird die Transformation $\mathbf{y} \mapsto \boldsymbol{\mu} + \mathbf{A}^T \mathbf{y} =: \mathbf{x}$ benutzt, um die Dichtefunktion für \mathbf{X} herzuleiten:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma(\frac{n}{2})}{2\pi^{n/2}} |\boldsymbol{\Sigma}|^{-1/2} \|(\mathbf{A}^T)^{-1}(\mathbf{x} - \boldsymbol{\mu})\|^{-(n-1)} f_{R_n}(\|(\mathbf{A}^T)^{-1}(\mathbf{x} - \boldsymbol{\mu})\|).$$

■

Die Funktion $g_{\mathbf{X}}$ wird oft als *Dichtegenerator* von \mathbf{X} oder von der Verteilung von \mathbf{X} bezeichnet, siehe (Fang et al., 1989, S. 35). In der graphischen Darstellung sind die Konturlinien der Dichtefunktion $f_{\mathbf{X}}$ Ellipsoide.

Beispiel 4.1.5 (Dichtegenerator der multivariaten Normalverteilung)

Sei $\mathbf{X} \sim N_n(\mathbf{0}, \mathbf{I}_n)$. Aus dem Beispiel 4.1.3 ist bekannt, dass $\mathbf{X} \stackrel{d}{=} R_n \mathbf{U}^{(n)}$, wobei $R_n = \sqrt{Z}$ und $Z \sim \chi_n^2$. Die Dichte von Z ist

$$f_Z(z) = \frac{z^{\frac{n}{2}-1} \cdot e^{-\frac{z}{2}}}{2^{n/2} \Gamma(\frac{n}{2})} \cdot \mathbb{1}_{z>0}.$$

Nach der Variablentransformation $z \mapsto \sqrt{z} =: r$, erhält man die Dichte von R_n

$$f_{R_n}(r) = 2r f_Z(r^2).$$

4.1 Sphärische und elliptische Verteilungen

Dann ist der Dichtegenerator von \mathbf{X}

$$\begin{aligned} g_{\mathbf{X}}(t) &= \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{n/2}} \sqrt{t}^{-(n-1)} \cdot 2\sqrt{t} \cdot \frac{t^{\frac{n}{2}-1} \cdot e^{-\frac{t}{2}}}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} \cdot \mathbb{1}_{t>0} \\ &= \frac{1}{(2\pi)^{n/2}} \cdot e^{-\frac{t}{2}} \cdot \mathbb{1}_{t>0}. \end{aligned}$$

□

Beispiel 4.1.6 (Dichtegenerator der multivariaten t -Verteilung)

Sei $\mathbf{X} \sim t_{\nu}^n(\mathbf{0}, \mathbf{I}_n)$. Aus dem Beispiel 4.1.4 ist bekannt, dass $\mathbf{X} \stackrel{d}{=} R_n \mathbf{U}^{(n)}$, wobei $R_n = \sqrt{\nu} \frac{\sqrt{Z}}{\sqrt{S}}$ mit unabhängigen Zufallsvariablen $Z \sim \chi_n^2$ und $S \sim \chi_{\nu}^2$. Für die gemeinsame Dichte von Z und S gilt:

$$f_{(Z,S)}(z, s) = \frac{z^{\frac{n}{2}-1} \cdot e^{-\frac{z}{2}}}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} \cdot \mathbb{1}_{z>0} \cdot \frac{s^{\frac{\nu}{2}-1} \cdot e^{-\frac{s}{2}}}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} \cdot \mathbb{1}_{s>0}.$$

Durch die Variablentransformation $(z, s) \mapsto (\sqrt{\nu} \frac{\sqrt{z}}{\sqrt{s}}, s) =: (x, y)$ kann die Dichte von R_n ermittelt werden:

$$\begin{aligned} f_{R_n}(x) &= \int_0^{\infty} \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} \left(\frac{x^2 y}{\nu}\right)^{\frac{n}{2}-1} e^{-\frac{x^2 y}{2\nu}} \cdot \frac{1}{2^{\nu/2}\Gamma\left(\frac{\nu}{2}\right)} y^{\frac{\nu}{2}-1} e^{-\frac{y}{2}} \cdot \frac{2}{\nu} x y \, dy \\ &= \frac{2 \cdot \Gamma\left(\frac{n+\nu}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{x^{n-1}}{\nu^{n/2}} \cdot \left(1 + \frac{x^2}{\nu}\right)^{-\frac{n+\nu}{2}}. \end{aligned}$$

Damit ergibt sich der Dichtegenerator von X :

$$g_{\mathbf{X}}(t) = \frac{\Gamma\left(\frac{n+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \frac{1}{(\nu\pi)^{n/2}} \cdot \left(1 + \frac{t}{\nu}\right)^{-\frac{n+\nu}{2}} \cdot \mathbb{1}_{t>0}.$$

□

Eine wichtige Eigenschaft der elliptischen Verteilungen ist, dass diese unter der linearen Transformation elliptisch und vom selben Verteilungstyp bleiben.

Satz 4.1.7

Sei $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$, \mathbf{B} eine $q \times n$ -Matrix und $\mathbf{b} \in \mathbb{R}^q$. Dann gilt:

$$\mathbf{B}\mathbf{X} + \mathbf{b} \sim E_q(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T, \phi).$$

Beweis.

Für den Beweis nutzt man die stochastische Zerlegung von \mathbf{X} aus dem Satz 4.1.4, siehe (Fang et al., 1989, S. 43).

Eine direkte Folgerung aus dem Satz 4.1.7 ist das nachfolgende Korollar. Es zeigt, dass die Randverteilungen einer multivariaten elliptischen Verteilung selbst elliptisch sind und die gleiche charakteristische Generatorfunktion wie die gemeinsame Verteilung besitzen. Das bedeutet beispielsweise, dass die multivariate Standardnormalverteilung über standardnormalverteilte Randverteilungen verfügt.

Korollar 4.1.2

Sei $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)^T \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ mit $\mathbf{X}_1 \in \mathbb{R}^p$ und $\mathbf{X}_2 \in \mathbb{R}^{n-p}$. Es bezeichnen weiter

$$\mathbb{E}(\mathbf{X}) = \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \quad \text{und} \quad \text{Var}(\mathbf{X}) = \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

den Erwartungsvektor und die Varianzmatrix von \mathbf{X} mit $\boldsymbol{\mu}_1 \in \mathbb{R}^p$ und $\boldsymbol{\Sigma}_{11} \in \mathbb{R}^{p \times p}$. Dann gilt:

$$\mathbf{X}_1 \sim E_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, \phi) \quad \text{und} \quad \mathbf{X}_2 \sim E_{n-p}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22}, \phi).$$

Beweis.

Für den Beweis wähle man $\mathbf{B} = (\mathbf{I}_p \mathbf{0}_{p \times (n-p)})$ bzw. $\mathbf{B} = (\mathbf{0}_{(n-p) \times p} \mathbf{I}_{(n-p)})$, dann folgt die Behauptung aus dem Satz 4.1.7.

Sei $\mathbf{X} = (X_1, \dots, X_n)^T$. Haben alle X_i endliche Varianzen und ist $\boldsymbol{\Sigma}$ eine Diagonalmatrix, so weiß man, dass die Komponenten von \mathbf{X} unkorreliert sind. Der nachfolgende Satz, siehe (Muirhead, 1982, S. 35), zeigt, dass die unkorrelierten Zufallsvariablen X_1, \dots, X_n nur dann (paarweise) stochastisch unabhängig sind, wenn die gemeinsame Verteilung Normalverteilung ist. Das bedeutet, wenn der Vektor \mathbf{X} nicht multivariat normalverteilt ist, sind die Komponenten X_1, \dots, X_n zwangsläufig abhängig, auch wenn sie selbst normalverteilt sind. Aus diesem Grund kann die gemeinsame Verteilung von paarweise unabhängigen, aber nicht normalverteilten Zufallsvariablen nicht elliptisch sein, vergleiche (Kaiser, 2009, S. 17). Demnach ist die multivariate Normalverteilung (nicht nur unter den elliptischen Verteilungen) die einzige, bei der aus Unkorreliertheit der Komponenten ihre Unabhängigkeit folgt.

Satz 4.1.8

Sei $\mathbf{X} = (X_1, \dots, X_n)^T \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$, wobei $\boldsymbol{\Sigma}$ eine Diagonalmatrix ist. Falls X_1, \dots, X_n paarweise unabhängig sind, ist der Zufallsvektor \mathbf{X} normalverteilt, $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Beweis.

O.B.d.A. sei $\boldsymbol{\mu} = \mathbf{0}$ gewählt. Da $\boldsymbol{\Sigma}$ Diagonalmatrix ist, besitzt sie die Form

$$\boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{11} & 0 & \cdots & 0 \\ 0 & \Sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_{nn} \end{pmatrix}.$$

Für die charakteristische Funktion von \mathbf{X} gilt:

$$\psi_{\mathbf{X}}(\mathbf{t}) = \phi(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}) = \phi\left(\sum_{i=1}^n t_i^2 \Sigma_{ii}\right). \quad (4.4)$$

Andererseits gilt wegen des Korollars 4.1.2 und der paarweisen Unabhängigkeit der Komponenten:

$$\phi(\mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}) = \prod_{i=1}^n \phi(t_i^2 \Sigma_{ii}). \quad (4.5)$$

Sei $u_i := t_i \sqrt{\Sigma_{ii}}$, dann folgt aus dem Gleichsetzen von (4.4) und (4.5)

$$\phi\left(\sum_{i=1}^n u_i^2\right) = \prod_{i=1}^n \phi(u_i^2).$$

Diese unter der Bezeichnung *Hamel Gleichung* bekannte Gleichung liefert als einzige stetige Lösung

$$\phi(z) = e^{kz}$$

mit einer gewissen Konstante k , siehe (Feller, 1968, S. 459). Damit besitzt die charakteristische Funktion von \mathbf{X} die Form:

$$\psi_{\mathbf{X}}(\mathbf{t}) = e^{k \mathbf{t}^T \boldsymbol{\Sigma} \mathbf{t}}.$$

Sei $f_{\mathbf{X}}$ die Dichtefunktion von \mathbf{X} . Gemäß den Voraussetzungen

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-it^T \mathbf{x}} \psi_{\mathbf{X}}(\mathbf{t}) \, dt \quad \text{und} \quad \int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} = 1,$$

lässt sich $k = -\frac{1}{2}$ ermitteln. Das impliziert aber, dass \mathbf{X} normalverteilt ist.



Sei \mathbf{X} wie in Korollar 4.1.2, dann stellt das nachfolgende Ergebnis fest, dass die bedingte Verteilung von \mathbf{X}_1 für einen festen Wert von \mathbf{X}_2 auch elliptisch ist, in der Regel aber nicht die gleiche charakteristische Generatorfunktion wie \mathbf{X}_1 besitzt.

Satz 4.1.9

Sei $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + R_n \mathbf{A}^T \mathbf{U}^{(n)} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ mit $\boldsymbol{\Sigma}$ positiv definit. Darüber hinaus seien \mathbf{X} , $\boldsymbol{\mu}$ und $\boldsymbol{\Sigma}$ wie im Korollar 4.1.2. Dann ist

$$(\mathbf{X}_1 \mid \mathbf{X}_2 = \mathbf{x}_0) \stackrel{d}{=} \tilde{\boldsymbol{\mu}} + R_{q(\mathbf{x}_0)} \tilde{\mathbf{A}}^T \mathbf{U}^{(p)} \sim E_p(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}, \tilde{\phi}),$$

wobei

$$\begin{aligned} \tilde{\boldsymbol{\mu}} &= \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{x}_0 - \boldsymbol{\mu}_2), \\ \tilde{\boldsymbol{\Sigma}} &= \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21} = \tilde{\mathbf{A}}^T \tilde{\mathbf{A}}, \\ q(\mathbf{x}_0) &= (\mathbf{x}_0 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_{22}^{-1} (\mathbf{x}_0 - \boldsymbol{\mu}_2) \\ R_{q(\mathbf{x}_0)} &\stackrel{d}{=} \left((R_n^2 - q(\mathbf{x}_0))^{1/2} \mid \mathbf{X}_2 = \mathbf{x}_0 \right) \end{aligned}$$

und $R_{q(\mathbf{x}_0)}$ unabhängig von $\mathbf{U}^{(p)}$ ist.

Beweis.

Siehe (Fang et al., 1989, S. 45).

Das Lemma 4.1.1 zeigt, dass die lineare Kombination von unabhängigen elliptisch verteilten Zufallsvariablen mit der gleichen (bis auf eine positive Konstante c) Varianzmatrix $\boldsymbol{\Sigma}$ wieder elliptisch verteilt ist.

Lemma 4.1.1

Seien $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ und $\tilde{\mathbf{X}} \sim E_n(\tilde{\boldsymbol{\mu}}, c\boldsymbol{\Sigma}, \tilde{\phi})$ mit $c > 0$ unabhängige Zufallsvariablen. Für $a, b \in \mathbb{R}$ gilt

$$a\mathbf{X} + b\tilde{\mathbf{X}} \sim E_n(a\boldsymbol{\mu} + b\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}, \phi^*)$$

mit $\phi^*(u) = \phi(a^2 u) \tilde{\phi}(b^2 c u)$.

Beweis.

Wegen der Unabhängigkeit von \mathbf{X} und $\tilde{\mathbf{X}}$ gilt:

$$\begin{aligned} \psi_{a\mathbf{X}+b\tilde{\mathbf{X}}-a\boldsymbol{\mu}-b\tilde{\boldsymbol{\mu}}}(\mathbf{t}) &= \psi_{a(\mathbf{X}-\boldsymbol{\mu})}(\mathbf{t})\psi_{b(\tilde{\mathbf{X}}-\tilde{\boldsymbol{\mu}})}(\mathbf{t}) \\ &= \phi((a\mathbf{t})^T \boldsymbol{\Sigma}(a\mathbf{t}))\tilde{\phi}((b\mathbf{t})^T c\boldsymbol{\Sigma}(b\mathbf{t})) \\ &= \phi(a^2\mathbf{t}^T \boldsymbol{\Sigma}\mathbf{t})\tilde{\phi}(b^2c\mathbf{t}^T \boldsymbol{\Sigma}\mathbf{t}). \end{aligned}$$

Dann folgt die Behauptung aus dem Satz 4.1.1. ■

Wird die Verteilung von \mathbf{X} durch $E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ beschrieben, so ist diese Form der Präsentation nicht eindeutig. Der Parameter $\boldsymbol{\mu}$ ist immer eindeutig festgelegt, die Varianzmatrix $\boldsymbol{\Sigma}$ und die Generatorfunktion ϕ dagegen sind bis auf eine positive Konstante eindeutig.

Satz 4.1.10

Sei $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ und $\mathbf{X} \sim E_n(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*, \phi^*)$. Dann gilt:

$$\boldsymbol{\mu}^* = \boldsymbol{\mu}, \quad \boldsymbol{\Sigma}^* = c\boldsymbol{\Sigma} \quad \text{und} \quad \phi^*(\cdot) = \phi(\cdot/c)$$

für eine Konstante $c > 0$.

Beweis.

O.B.d.A. sei der $\text{Rang}(\boldsymbol{\Sigma}) = n$. In der Regel bezeichnet man mit $\boldsymbol{\Sigma}$ die Varianzmatrix von \mathbf{X} , $\text{Var}(\mathbf{X}) = \boldsymbol{\Sigma} = \mathbf{A}^T \mathbf{A}$. Mit der Darstellung aus dem Satz 4.1.1 erhält man:

$$\text{Var}(\mathbf{X}) = \text{Var}(\boldsymbol{\mu} + R_n \mathbf{A}^T \mathbf{U}^{(n)}) = \mathbf{A}^T \mathbb{E}(R_n^2) \text{Var}(\mathbf{U}^{(n)}) \mathbf{A},$$

vorausgesetzt $\mathbb{E}(R_n^2) < \infty$.

Laut dem Satz 4.1.3 ist $\mathbf{U}^{(n)} \stackrel{d}{=} \mathbf{Y}/\|\mathbf{Y}\|$ mit $\mathbf{Y} \sim N_n(\mathbf{0}, \mathbf{I}_n)$ und $\|\mathbf{Y}\|$ ist unabhängig von $\mathbf{U}^{(n)}$. Die Verteilung von $\|\mathbf{Y}\|^2$, $\|\mathbf{Y}\|^2 \sim \chi_n^2$ (Chi-Quadrat-Verteilung mit n Freiheitsgraden) ist bekannt, und somit auch die Varianz von $\|\mathbf{Y}\|$, $\text{Var}(\|\mathbf{Y}\|) = \mathbb{E}(\|\mathbf{Y}\|^2) = n$. Damit kann die Varianz von $\mathbf{U}^{(n)}$

$$\text{Var}(\mathbf{U}^{(n)}) = \text{Var}(\mathbf{Y})/\text{Var}(\|\mathbf{Y}\|) = \mathbf{I}_n/n$$

bestimmt werden und liefert für die Varianz von \mathbf{X}

$$\text{Var}(\mathbf{X}) = \mathbf{A}^T \mathbf{A} \mathbb{E}(R_n^2)/n.$$

Setzt man also die charakteristische Generatorfunktion $\phi^*(s) = \phi(s/c)$ mit $c = \mathbb{E}(R_n^2)/n$, so bekommt man $\text{Var}(\mathbf{X}) = \Sigma$. ■

Die elliptische Verteilung wird demnach über ihre Parameter $\boldsymbol{\mu}$, Σ und die Generatorfunktion ϕ vollständig beschrieben wird, wobei ϕ so ausgetauscht werden kann, dass $\text{Var}(\mathbf{X}) = \Sigma$, falls $\text{Var}(\mathbf{X})$ definiert ist. Sofern $\text{Var}(X_i) < \infty$ für alle $1 \leq i \leq n$ und die gemeinsame Verteilung nichtdegeneriert ($\Sigma_{ii} > 0$ für alle i) ist, kann die Korrelationsmatrix \mathbf{R} mit

$$R_{ij} = \text{Corr}(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Var}(X_i) \text{Var}(X_j)}} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii} \Sigma_{jj}}},$$

anstelle von Σ als Beschreibungsparameter genommen werden. Da die elliptische Verteilung über $\boldsymbol{\mu}$, Σ und ϕ eindeutig bestimmt ist, ist die Copulafunktion einer nichtdegenerierten elliptischen Verteilung über die Parameter \mathbf{R} und ϕ eindeutig bestimmt.

Ein Problem in Bezug auf die elliptischen Verteilungen ist die Eigenschaft, dass alle Randverteilungen vom selben Typ sind. Um jedoch realistische Risikoverteilungen modellieren zu können, muss es möglich sein, elliptische Copulafunktionen mit anders verteilten Randverteilungen zu konstruieren. Ein solches Modell scheint dann aber den großen Nachteil zu haben, dass der Copulaparameter \mathbf{R} nicht mehr direkt über die Daten geschätzt werden kann. Nur im Falle der nichtentarteten elliptischen Verteilung mit endlichen Varianzen ist \mathbf{R} die übliche Korrelationsmatrix und kann mittels eines linearen Korrelationsschätzers approximiert werden. Die Grundlage für eine alternative Schätzmethode, die auf Kendall's τ basiert, gibt der nachfolgende Satz an.

Satz 4.1.11

Sei $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ mit $\mathbb{P}(X_i = \mu_i) < 1$ und $\mathbb{P}(X_j = \mu_j) < 1$. Außerdem sei \mathbf{R} die lineare Korrelationsmatrix von \mathbf{X} . Dann gilt

$$\tau(X_i, X_j) = \left(1 - \sum_{x \in \mathbb{R}} (\mathbb{P}(X_i = x))^2\right) \frac{2}{\pi} \arcsin R_{ij}, \quad (4.6)$$

wobei die Summe fast überall auf dem Definitionsbereich der Verteilung von X_i existiert. Wenn $\text{Rang}(\boldsymbol{\Sigma}) \geq 2$, dann vereinfacht sich (4.6) zu

$$\tau(X_i, X_j) = \left(1 - (\mathbb{P}(X_i = \mu_i))^2\right) \frac{2}{\pi} \arcsin R_{ij}$$

Beweis.

Siehe (Lindskog et al., 2001, S. 3-6).

Im Falle von stetigen elliptischen Verteilungen, gilt also

$$\tau(X_i, X_j) = \frac{2}{\pi} \arcsin R_{ij}, \quad i, j = 1, \dots, n. \quad (4.7)$$

Wichtig bei diesem Ergebnis ist, dass Kendall's τ bei den elliptischen Verteilungen in keiner Weise von der Form der Verteilung abhängt, sondern einzig und allein von der Korrelation. Aufgrund der Beziehung (4.7), kann für jede stetige bivariate Randverteilung eine Korrelationsschätzung,

$$\hat{R}_{ij} = \sin(\pi \hat{\tau}(x_i, x_j)/2), \quad i, j = 1, \dots, n,$$

hergeleitet werden. Die entsprechende Schätzmethode und ihre Eigenschaften werden im Kapitel 5 und 6 näher erläutert.

4.2 Normal-Copula

Sei \mathbf{X} ein n -dimensionaler normalverteilter Zufallsvektor, $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ und $\boldsymbol{\Sigma}$ positiv definit (d.h. insbesondere $|\boldsymbol{\Sigma}| > 0$). In diesem Abschnitt soll die zu X gehörige Copulafunktion dargestellt werden. Nach dem Satz 4.1.6 und Beispiel 4.1.5 besitzt die Dichtefunktion einer multivariat normalverteilten Zufallsgröße die Form

$$f_{\mathbf{N}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right).$$

Sei \mathbf{R} die Korrelationsmatrix von \mathbf{X} ,

$$\mathbf{R} = \begin{pmatrix} 1 & R_{12} & \cdots & R_{1n} \\ R_{21} & 1 & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & 1 \end{pmatrix},$$

wobei R_{ij} die lineare Korrelation zwischen den Komponenten X_i und X_j , $i, j = 1 \dots n$, ist. Falls alle Komponenten von \mathbf{X} standardnormalverteilt sind, d.h.

$\mathbb{E}(X_i) = 0$ und $\text{Var}(X_i) = 1$ für alle i , ist $\Sigma = \mathbf{R}$. In diesem Fall soll die Dichtefunktion mit $f_{\mathbf{R}}$ bezeichnet werden und es gilt:

$$f_{\mathbf{R}} = \frac{1}{(2\pi)^{n/2} |\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right).$$

Da die Abhängigkeitsstruktur in einem normalverteilten Vektor gerade über die Korrelationsmatrix beschrieben wird, ist die Copulafunktion von $N_n(\boldsymbol{\mu}, \Sigma)$ mit der Copula von $N_n(\mathbf{0}, \mathbf{R})$ identisch. Aus diesem Grund soll nachfolgend zur Vereinfachung mit der Verteilung $N_n(\mathbf{0}, \mathbf{R})$ gearbeitet werden. Seien $\Phi_{\mathbf{R}}$ die Verteilungsfunktion von $N_n(\mathbf{0}, \mathbf{R})$ und Φ die Verteilungsfunktion einer univariaten Standardnormalverteilung, dann lässt sich die Normal-Copula nach dem Korollar 2.2.1 wie folgt bestimmen:

$$\begin{aligned} C_{\mathbf{R}}^{\mathbf{N}}(u) &= \Phi_{\mathbf{R}}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)) \\ &= \frac{1}{(2\pi)^{n/2} |\mathbf{R}|^{1/2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_n)} \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right) d\mathbf{x}. \end{aligned}$$

Die Bezeichnung $C_{\mathbf{R}}^{\mathbf{N}}$ verdeutlicht, dass die Korrelationsmatrix \mathbf{R} gerade die Parametermatrix einer Normal-Copula ist. In der Folge sollen einige Eigenschaften der Normal-Copula untersucht werden. Nach dem Satz 4.1.11 erschließt sich folgende Beziehung zwischen den Elementen der Rangkorrelationsmatrix von Kendall und den Elementen von \mathbf{R}

$$\tau_{ij} = \tau(X_i, X_j) = \frac{2}{\pi} \arcsin(R_{ij}), \quad i = 1 \dots n, j = 1 \dots n.$$

(Man beachte, dass bei der multivariaten Normalverteilung $\mathbb{P}(X_i = \mu_i) = 0$ für alle i gilt.)

4.2 Normal-Copula

Zwischen den Elementen der Rangkorrelationsmatrix von Spearman und den Elementen von R gilt die Beziehung

$$\rho_{ij} = \rho(X_i, X_j) = \frac{6}{\pi} \arcsin(R_{ij}), \quad i = 1 \dots n, j = 1 \dots n,$$

siehe (Lindskog, 2000, S. 47).

Sei $(X_1, X_2)^T$ ein bivariat normalverteilter Vektor mit

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & R_{12} \\ R_{12} & 1 \end{pmatrix} \right),$$

wobei R_{12} die lineare Korrelation zwischen X_1 und X_2 ist. Gemäß der Grundvorlesung der mathematischen Statistik ist bekannt, dass $(X_2 \mid X_1 = x)$ ebenfalls normalverteilt ist. Nach dem Satz 4.1.9 sind

$$\mathbb{E}(X_2 \mid X_1 = x) = R_{12}x \quad \text{und} \quad \text{Var}(X_2 \mid X_1 = x) = 1 - R_{12}^2.$$

Die Copulafunktion von $(X_1, X_2)^T$ ist die Normal-Copula mit dem Parameter R_{12} . Es soll die asymptotische Randabhängigkeit von X_1 und X_2 untersucht werden. Da die Bedingung $C(u_1, u_2) = C(u_2, u_1)$ in diesem Fall erfüllt ist, kann die Formel

$$\lambda_O = 2 \lim_{x \rightarrow \infty} \mathbb{P}(X_2 > x \mid X_1 = x)$$

zur Herleitung genutzt werden und liefert

$$\begin{aligned} \lambda_O &= 2 \lim_{x \rightarrow \infty} \left(1 - \Phi \left(\frac{x - R_{12}x}{\sqrt{1 - R_{12}^2}} \right) \right) \\ &= 2 \lim_{x \rightarrow \infty} \left(1 - \Phi \left(\frac{x\sqrt{1 - R_{12}^2}}{\sqrt{1 + R_{12}^2}} \right) \right). \end{aligned}$$

Demnach ist

$$\lambda_O = \begin{cases} 0 & : -1 < R_{12} < 1 \\ 1 & : R_{12} = 1 \\ \text{undefiniert} & : R_{12} = -1 \end{cases}$$

Man kann also sagen, dass zwei Zufallsvariablen X_1 und X_2 mit zugehöriger Normal-Copula asymptotisch unabhängig im oberen Rand für $-1 < R_{12} < 1$ und asymptotisch abhängig im oberen Rand für $R_{12} = 1$ sind.

Analog lässt sich die Formel für die asymptotische untere Randabhängigkeit herleiten

$$\lambda_U = 2 \lim_{x \rightarrow \infty} \Phi\left(\frac{x\sqrt{1-R_{12}}}{\sqrt{1+R_{12}}}\right).$$

Folglich sind zwei Zufallsvariablen X_1 und X_2 mit zugehöriger Normal-Copula asymptotisch unabhängig im unteren Rand für $-1 < R_{12} < 1$ und asymptotisch abhängig im unteren Rand für $R_{12} = 1$. Offensichtlich gilt: $\lambda_O = \lambda_U$. Das bedeutet, dass Normal-Copulas stets gleiche asymptotische obere und asymptotische untere Randabhängigkeiten nach sich ziehen, siehe (Glauser, 2003, S. 105).

Um Beobachtungen einer Normal-Copula mit dem Parameter \mathbf{R} zu generieren, bedient man sich einer einfachen Transformationsmethode. Dazu benötigt man einen multivariat normalverteilten Vektor mit der Korrelationsmatrix \mathbf{R} . Zunächst muss eine $n \times n$ -Matrix \mathbf{A} so bestimmt werden, dass $\mathbf{R} = \mathbf{A}^T \mathbf{A}$ gilt. Seien $Z_1, \dots, Z_n \sim \mathbf{N}(0, 1)$ unabhängige identisch verteilte Zufallsvariablen, dann liefert

$$\boldsymbol{\mu} + \mathbf{A}\mathbf{Z} \sim \mathbf{N}_n(\boldsymbol{\mu}, \mathbf{R})$$

den gewünschten Vektor. Eine häufig verwendete Wahl für \mathbf{A} ist die Cholesky-Zerlegung von \mathbf{R} . Diese Methode liefert eine eindeutige untere Dreiecksmatrix \mathbf{L} , so dass $\mathbf{L}^T \mathbf{L} = \mathbf{R}$ gilt, siehe Schwarz (1997). Die Cholesky-Zerlegung gehört in fast jeder mathematischen Software zu den standardmäßig implementierten Verfahren. Der Algorithmus zur Erzeugung von Zufallsrealisierungen, die gemäß der Normal-Copula mit dem Parameter \mathbf{R} verteilt sind, kann wie folgt angegeben werden, vergleiche (Embrechts et al., 2001, S. 26):

Algorithmus 4.2.1

- Bilde die Cholesky-Zerlegung \mathbf{A} von \mathbf{R} .
- Erzeuge n unabhängige Realisierungen $z_i \sim \mathbf{N}(0, 1)$, $i = 1, \dots, n$.
- Setze $\mathbf{x} = \mathbf{A}\mathbf{z}$.
- Berechne $u_i = \Phi(x_i)$, $i = 1, \dots, n$.
- $(u_1, \dots, u_n)^T \sim C_{\mathbf{R}}^{\mathbf{N}}$.

4.3 *t*-Copula

Sei \mathbf{X} ein n -dimensionaler t -verteilter Zufallsvektor mit ν Freiheitsgraden, $\mathbf{X} \sim t_\nu^n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ und $\boldsymbol{\Sigma}$ positiv definit. Nach dem Satz 4.1.6 und Beispiel 4.1.6 besitzt die Dichtefunktion einer multivariat t -verteilten Zufallsgröße die Form

$$f_t(\mathbf{x}) = \frac{\Gamma(\frac{\nu+n}{2})}{(\pi\nu)^{n/2}\Gamma(\frac{\nu}{2})|\boldsymbol{\Sigma}|^{1/2}} \left(1 + \frac{1}{\nu}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)^{-\frac{\nu+n}{2}}.$$

Sei \mathbf{R} die Korrelationsmatrix von \mathbf{X} . Mit der gleichen Überlegung wie bei der Normal-Copula kann aufgrund der einfacheren Handhabung die Verteilung $t_\nu^n(\mathbf{0}, \mathbf{R})$ zur Herleitung der t -Copula verwendet werden. Sei t_ν die Verteilungsfunktion der univariaten t -Verteilung, dann ist

$$C_{\mathbf{R},\nu}^t(\mathbf{u}) = \frac{\Gamma(\frac{\nu+n}{2})}{(\pi\nu)^{n/2}\Gamma(\frac{\nu}{2})|\boldsymbol{\Sigma}|^{1/2}} \int_{-\infty}^{t_\nu^{-1}(u_1)} \cdots \int_{-\infty}^{t_\nu^{-1}(u_n)} \left(1 + \frac{1}{\nu}\mathbf{x}^T \mathbf{R}^{-1}\mathbf{x}\right)^{-\frac{\nu+n}{2}} d\mathbf{x}$$

die gesuchte t -Copula.

Für die Elemente der Parametermatrix \mathbf{R} und die Elemente der Rangkorrelationsmatrix von Kendall gilt ebenfalls die Beziehung, die sich aus dem Satz 4.1.11 herleiten lässt,

$$\tau_{ij} = \frac{2}{\pi} \arcsin(R_{ij}), \quad i = 1 \dots n, j = 1 \dots n.$$

Um ein besseres Verständnis für die t -Copula zu gewinnen, soll auch hier die asymptotische Randabhängigkeit betrachtet werden. Sei $(X_1, X_2)^T$ ein bivariat t -verteilter Vektor, so dass

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim t_\nu^2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & R_{12} \\ R_{12} & 1 \end{pmatrix} \right),$$

wobei R_{12} die lineare Korrelation zwischen X_1 und X_2 und ν Anzahl der Freiheitsgrade sind. Gemäß der Grundvorlesung in der mathematischen Statistik ist bekannt, dass $(X_2 | X_1 = x)$ $t_{\nu+1}^2$ -verteilt ist. Nach dem Satz 4.1.9 sind

$$\mathbb{E}(X_2 | X_1 = x) = R_{12}x \quad \text{und} \quad \text{Var}(X_2 | X_1 = x) = \left(\frac{\nu + x^2}{\nu + 1}\right)(1 - R_{12}^2).$$

Das ist hilfreich, um die asymptotische Randabhängigkeit zu berechnen:

$$\begin{aligned}
 \lambda_O &= 2 \lim_{x \rightarrow \infty} \mathbb{P}(X_2 > x \mid X_1 = x) \\
 &= 2 \lim_{x \rightarrow \infty} \left[1 - t_{\nu+1} \left(\left(\frac{\nu+1}{\nu+x^2} \right)^{1/2} \left(\frac{x - R_{12}x}{\sqrt{1-R_{12}^2}} \right) \right) \right] \\
 &= 2 \lim_{x \rightarrow \infty} \left[1 - t_{\nu+1} \left(\left(\frac{\nu+1}{\frac{\nu}{x^2} + 1} \right)^{1/2} \left(\frac{\sqrt{1-R_{12}}}{\sqrt{1+R_{12}}} \right) \right) \right] \\
 &= 2 \left[1 - t_{\nu+1} \left(\frac{\sqrt{\nu+1} \sqrt{1-R_{12}}}{\sqrt{1+R_{12}}} \right) \right].
 \end{aligned}$$

Für $R_{12} = -1$ ist also λ_O nicht definiert, für $R_{12} > -1$ sind zwei Zufallsvariablen X_1 und X_2 mit der zugehörigen t -Copula asymptotisch abhängig im oberen Rand. Man kann außerdem erkennen, dass λ_O bei zunehmendem R_{12} steigt und bei zunehmendem ν fällt.

$\nu \backslash R_{12}$	-0.5	0.0	0.5	0.9	1.0
2	0.06	0.18	0.39	0.72	1.00
4	0.01	0.08	0.25	0.63	1.00
10	0.00	0.01	0.08	0.46	1.00
∞	0.00	0.00	0.00	0.00	1.00

Tabelle 4.1: Werte von λ_O für verschiedene Werte von ν und R_{12} .

Die Tabelle 4.1 zeigt die Werte von λ_O für verschiedene Werte von ν und R_{12} , vergleiche (Demarta and McNeil, 2004, S. 5). Wenn die Anzahl der Freiheitsgrade zu ∞ tendiert, konvergiert der Wert von λ_O gegen 0 für ein $-1 < R_{12} < 1$. Das entspricht der letzten Zeile in der Tabelle, die gleichzeitig die Werte des Koeffizienten für die obere asymptotische Randabhängigkeit der Normal-Copula darstellt. Mit anderen Worten, für ein $-1 < R_{12} < 1$ und $\nu \rightarrow \infty$ konvergiert der Wert von λ_O für die t -Copula gegen den Wert von λ_O für die Normal-Copula.

Wie bereits bei den Normal-Copulafunktionen gesehen, lässt sich ebenfalls für die t -Copulafunktionen nachweisen, dass $\lambda_O = \lambda_U$ gilt. Folglich sind zwei Zufallsva-

riablen mit der zugehörigen *t*-Copula und dem linearen Korrelationskoeffizienten $R_{12} > -1$ auch asymptotisch abhängig im unteren Rand.

Um Beobachtungen einer *t*-Copula mit den Parametern ν und \mathbf{R} zu generieren, kommt ähnlich wie bei der Normal-Copula eine Transformationsmethode zum Einsatz. Die Algorithmen sind fast identisch, nur wird anstelle eines multivariat normalverteilten Zufallsvektors ein *t*-verteilter Zufallsvektor mit ν Freiheitsgraden und der Korrelationsmatrix \mathbf{R} erzeugt.

Der Algorithmus zur Simulation von Zufallsrealisierungen, die gemäß der *t*-Copula mit den Parametern ν und \mathbf{R} verteilt sind, kann wie folgt angegeben werden, vergleiche (Embrechts et al., 2001, S. 27):

Algorithmus 4.3.1

- Bilde die Cholesky-Zerlegung \mathbf{A} von \mathbf{R} .
- Erzeuge n unabhängige Realisierungen $z_i \sim \mathbf{N}(0, 1)$, $i = 1, \dots, n$.
- Setze $\mathbf{y} = \mathbf{A}\mathbf{z}$.
- Erzeuge eine von z_1, \dots, z_n unabhängige Realisierung $s \sim \chi_\nu^2$.
- Setze $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{s}}\mathbf{y}$.
- Berechne $u_i = t_\nu(x_i)$, $i = 1, \dots, n$.
- $(u_1, \dots, u_n)^T \sim C_{\mathbf{R}, \nu}^t$.

Kapitel 5

Parametrische Schätzverfahren

Die elliptische Copulaklasse, welche im Kapitel 4 betrachtet wurde, ist parametrischer Natur. Da die Parameter einer Copulafunktion in den meisten praktischen Anwendungen unbekannt sind, besteht das Ziel diese mittels statistischer Verfahren zu schätzen. Über die möglichen Schätzmethoden soll in diesem Kapitel ein Überblick gegeben werden. Das Hauptaugenmerk liegt dabei auf Punkteschätzern, d.h. auf Verfahren, bei denen der gesuchte Parameter anhand eines beobachteten Datensatzes geschätzt wird. Es werden dabei sowohl parametrische, wie die exakte und sequentielle Maximum-Likelihood-Methode oder die Momentenmethode, als auch semiparametrische Verfahren wie die kanonische Maximum-Likelihood-Methode vorgestellt. Die Kendall's τ -Momentenmethode wird insbesondere bei der parametrischen Schätzung der t -Copula angewandt.

5.1 Maximum-Likelihood-Methode

Unter den Methoden zur Konstruktion von Punktschätzern ist die Maximum-Likelihood-Methode (ML-Methode) das am meisten verwendete Verfahren. Das liegt mitunter daran, dass die ML-Schätzer unter bestimmten Regularitätsbedingungen einige wünschenswerte asymptotische Eigenschaften vorweisen. Zunächst wird die allgemein übliche Vorgehensweise des Maximum-Likelihood-Prinzips, das auf R.A. Fisher zurück geht, vorgestellt. Dieses soll dann in den Teilabschnitten

5.1.1-5.1.3 zur Konstruktion der ML-Schätzer für die Copulaparameter genutzt werden. Auch Fragen der Existenz und Eindeutigkeit sowie asymptotische Eigenschaften einer ML-Schätzung sollen in diesem Abschnitt kurz diskutiert werden.

Sei $(\mathfrak{X}, \mathfrak{B}, \mathbb{P}_\vartheta \mid \vartheta \in \Theta)$ ein statistischer Raum mit der Ergebnismenge \mathfrak{X} , der auf \mathfrak{X} induzierten Borel- σ -Algebra \mathfrak{B} und den Wahrscheinlichkeitsmaßen \mathbb{P}_ϑ auf $(\mathfrak{X}, \mathfrak{B})$. X beschreibe eine Zufallsvariable auf $(\mathfrak{X}, \mathfrak{B}, \mathbb{P}_\vartheta \mid \vartheta \in \Theta)$. Es wird vorausgesetzt, dass ein σ -endliches Maß μ auf \mathfrak{B} existiert, so dass für alle $\vartheta \in \Theta$ das Maß \mathbb{P}_ϑ absolut stetig bezüglich μ ist, d.h. es gelte:

$$\mu(B) = 0 \quad \Rightarrow \quad \mathbb{P}_\vartheta(B) = 0 \quad \forall B \in \mathfrak{B}, \forall \vartheta \in \Theta.$$

Man sagt auch, die Verteilungsklasse $\{\mathbb{P}_\vartheta \mid \vartheta \in \Theta\}$ ist bezüglich des Maßes μ dominiert. Nach dem Satz von Radon-Nikodym existiert eine μ -Dichte $p(x; \vartheta)$, so dass

$$\mathbb{P}_\vartheta(B) = \int_B p(x; \vartheta) \mu(dx), \quad \forall B \in \mathfrak{B},$$

gilt, d.h. $p(x; \vartheta) = \left(d\mathbb{P}_\vartheta / d\mu \right)(x)$. Meistens ist μ ein Lebesgue-Maß oder ein Abzählmaß und $p(x; \vartheta)$ eine stetige oder diskrete Wahrscheinlichkeitsdichte. Sei x eine Realisierung von X , x wird beobachtet. Man möchte anhand dieser Beobachtung den unbekannt Parameter ϑ schätzen.

Definition 5.1.1

Eine Funktion $L : \Theta \rightarrow \mathbb{R}$ heißt *Likelihoodfunktion*, falls

$$L(\vartheta; x) = p(x; \vartheta).$$

□

Die Likelihoodfunktion ist zwar in ihrer funktionellen Form mit der Dichtefunktion der Realisierung identisch, wird jedoch als eine Funktion in ϑ zu einem festen Wert x betrachtet. Das *Maximum-Likelihood-Prinzip* besteht nun darin, als Schätzung für den Parameter ϑ einen Wert zu nehmen, der unter der gegebenen Realisierung am plausibelsten ist, d.h. der die Likelihoodfunktion maximiert.

Definition 5.1.2

Jeder Wert $\hat{\vartheta} \in \Theta$, für den

$$\hat{\vartheta} = \arg \max_{\vartheta \in \Theta} L(\vartheta; x)$$

gilt, wird als *Maximum-Likelihood-Schätzung* (ML-Schätzung) für den Parameter $\vartheta \in \Theta$ bezeichnet.

□

Im Allgemeinen existiert eine Lösung nur dann, wenn der Parameterraum Θ abgeschlossen ist. Das kann nicht in jeder Anwendung erreicht werden. Häufig lässt sich jedoch die Abbildung $\vartheta \mapsto p(x; \vartheta)$ stetig auf die abgeschlossene Hülle $\bar{\Theta}$ von Θ fortsetzen. In solchen Fällen wird man auch eine Lösung $\hat{\vartheta} \in \bar{\Theta} \setminus \Theta$ als Schätzung akzeptieren, siehe (Witting, 1985, S. 31). In vielen wichtigen Beispielen ist festzustellen, dass die Lösungen durch messbare Funktionen beschrieben werden. In so einem Fall kann man von einer Schätzfunktion oder einem Schätzer sprechen.

Definition 5.1.3

Ist die Abbildung $\hat{\vartheta} : \mathfrak{X} \rightarrow \bar{\Theta}$, die jedem $x \in \mathfrak{X}$ eine Maximum-Likelihood-Schätzung $\hat{\vartheta}(x)$ zuordnet, messbar, so wird sie als *Maximum-Likelihood-Schätzfunktion* (ML-Schätzfunktion) oder *Maximum-Likelihood-Schätzer* (ML-Schätzer) bezeichnet.

□

Unter welchen Bedingungen eine existierende Lösung messbare Abbildung ist, gibt das folgende Lemma an.

Korollar 5.1.1

Seien $\Theta \subseteq \mathbb{R}^k$ und $g : \mathfrak{X} \times \Theta \rightarrow (\mathbb{R}^l, \mathbb{B}^l)$ eine Abbildung, die für jedes $\vartheta \in \Theta$ messbar und für jedes $x \in \mathfrak{X}$ stetig auf Θ ist. Falls Θ eine σ -kompakte Teilmenge ist und eine Abbildung $\theta : \mathfrak{X} \rightarrow \Theta$ derart existiert, dass $h(x) := g(x, \theta(x))$ messbar ist, so existiert eine messbare Abbildung $\tilde{\vartheta} : (\mathfrak{X}, \mathfrak{B}) \rightarrow (\Theta, \Theta \cap \mathbb{B}^k)$ mit $g(x, \tilde{\vartheta}(x)) = h(x)$.

Beweis.

Siehe (Witting and Müller-Funk, 1995, S. 173-174).

Über allgemeinere Regularitätsbedingungen für die Existenz einer ML-Schätzfunktion kann man zum Beispiel bei Pfanzagl (1969) nachlesen. Es sei noch angemerkt, dass die ML-Schätzung unabhängig von der speziellen Wahl des dominierenden Maßes ist. Jedoch kann die Festlegung einer Radon-Nikodym Dichte $p(x; \vartheta)$ sich entscheidend auf die Existenz einer ML-Schätzung auswirken,

vergleiche Pfanzagl (1969).

Es wird jetzt der folgende Fall betrachtet: \mathbf{X} sei eine stetig verteilte Zufallsvariable mit der Dichte $f(\mathbf{x}; \vartheta)$ und $\mathbf{x} = (x_1, \dots, x_n)$ sei eine Stichprobe mit unabhängigen Beobachtungen $x_i, i = 1, \dots, n$, von \mathbf{X} . Dann weiß man, dass aufgrund der Unabhängigkeit die gemeinsame Dichte ein Produkt der Dichten der einzelnen Beobachtungen ist. Somit ergibt sich die Likelihoodfunktion

$$L(\vartheta; x_1, \dots, x_n) = p_n(\mathbf{x}; \vartheta) = \prod_{i=1}^n f(x_i; \vartheta).$$

Wegen der Produkte kann die Likelihoodfunktion dabei eine komplizierte Gestalt annehmen, was das Maximieren erheblich erschweren kann. In solchen Fällen geht man zu der logarithmierten Likelihoodfunktion, der sogenannten *Log-Likelihoodfunktion*, über.

$$l(\vartheta; x_1, \dots, x_n) = \ln L(\vartheta; x_1, \dots, x_n) = \sum_{i=1}^n \ln f(x_i; \vartheta).$$

Da bei einer Maximierung monotone Transformationen erlaubt sind und der natürliche Logarithmus eine streng monoton wachsende Transformation ist, gilt:

$$\hat{\vartheta}_n = \arg \max_{\vartheta \in \Theta} L(\vartheta; x_1, \dots, x_n) = \arg \max_{\vartheta \in \Theta} l(\vartheta; x_1, \dots, x_n).$$

Sei $\Theta \subseteq \mathbb{R}^k$ und $l(\vartheta; \cdot)$ partiell nach $\vartheta_1, \dots, \vartheta_k$ differenzierbar. In der Regel wird das lokale Maximum der Log-Likelihoodfunktion als Lösung der *Likelihood-Gleichungen*

$$\frac{\partial l(\vartheta; x_1, \dots, x_n)}{\partial \vartheta_i} = 0, \quad i = 1, \dots, k, \quad (5.1)$$

ermittelt. Ob eine Lösung tatsächlich zu einem Maximum führt, muss anhand der zweiten partiellen Ableitungen (falls diese existieren) überprüft werden. Es können auch mehrere Lösungen existieren, die alle zu einem lokalen Maximum führen. Andererseits können Lösungen auf dem Rand von Θ liegen, diese brauchen über (5.1) nicht erfasst zu werden, siehe (Witting and Müller-Funk, 1995, S. 169). Falls ein analytisches Auflösen der Gleichungen nicht möglich ist, muss ein numerisches

(z.B. Newton-) Verfahren herangezogen werden. Jedoch sind in vielen Fällen die Lösungen des nichtlinearen Gleichungssystem (5.1) auch numerisch nur schwer zu errechnen.

Über die Eindeutigkeit einer Lösung kann man keine allgemeine Aussagen treffen. Es lassen sich jedoch, wie oben bereits erwähnt, unter bestimmten Regularitätsbedingungen einige asymptotische Eigenschaften, wie die Konsistenz und die asymptotische Normalität, nachweisen.

Satz 5.1.1 (Konsistenz)

Sei $(\mathcal{X}, \mathfrak{B}, \mathbb{P}_\vartheta \mid \vartheta \in \Theta)$ ein μ -dominierter statistischer Raum mit $\Theta \subset \mathbb{R}^k$ abgeschlossen und $\mathbb{P}_\vartheta \neq \mathbb{P}_{\vartheta'}$ für $\vartheta \neq \vartheta'$. Seien $X_i, i \in \mathbb{N}$, unabhängige identisch verteilte Zufallsvariablen mit $X_i \sim \mathbb{P}_\vartheta$. Für jedes $\vartheta \in \Theta$ sei \mathbb{P}_ϑ über die μ -Dichten $f(\cdot; \vartheta)$ repräsentiert, die folgende Eigenschaften erfüllen:

1. $f(x; \cdot)$ ist für μ -fast alle x stetig;
2. Ist Θ beschränkt, so gelte mit $w(x, \vartheta, \varepsilon) := \sup\{f(x; \vartheta') \mid |\vartheta - \vartheta'| < \varepsilon\}$, dass für alle $\vartheta \in \Theta$ und $\vartheta \neq \vartheta'$ ein $\varepsilon > 0$ existiert mit

$$\mathbb{E}_\vartheta \ln \left[\frac{w(X_1, \vartheta, \varepsilon)}{f(X_1; \vartheta)} \right] < \infty;$$

3. Ist Θ unbeschränkt, so sei zusätzlich $f(x; \vartheta) \rightarrow 0$ für $|\vartheta| \rightarrow \infty$ außerhalb einer Nullmenge. Außerdem gelte mit $w(x, \infty, M) := \sup\{f(x; \vartheta') \mid \vartheta' \in \Theta, |\vartheta'| > M\}$, dass für alle $\vartheta \in \Theta$ ein $M > 0$ existiert mit

$$\mathbb{E}_\vartheta \ln \left[\frac{w(X_1, \infty, M)}{f(X_1; \vartheta)} \right] < \infty.$$

Dann gilt für jeden Maximum-Likelihood-Schätzer $\hat{\vartheta}_n$ und jedes $\vartheta \in \Theta$:

$$\hat{\vartheta}_n \xrightarrow{\mathbb{P}_\vartheta} \vartheta,$$

d.h. $\hat{\vartheta}_n$ ist ein (schwach) konsistenter Schätzer.

Beweis.

Siehe (Witting and Müller-Funk, 1995, S. 201-202).

Um zu zeigen, dass $\hat{\vartheta}_n$ asymptotisch normalverteilt ist, siehe (Satz 5.1.2), müssen neben der Konsistenz noch weitere Regularitätsvoraussetzungen an das vorliegende Modell gestellt werden, insbesondere die Existenz und Integrierbarkeit der Fisher-Information (siehe Definition 5.1.4). Gleichzeitig geht aus dem Satz 5.1.2 hervor, dass $\hat{\vartheta}_n$ die für asymptotische Effizienz gültigen Bedingungen erfüllt.

Definition 5.1.4 (Fisher-Information)

Sei $(\mathfrak{X}, \mathfrak{B}, \mathbb{P}_\vartheta \mid \vartheta \in \Theta)$ ein statistischer Raum und $\{\mathbb{P}_\vartheta \mid \vartheta \in \Theta\}$ eine k -parametrische μ -dominierte Verteilungsklasse. Die μ -Dichten $f(x; \vartheta)$ seien positiv und für jedes $x \in \mathfrak{X}$ als Funktion von ϑ zweimal stetig differenzierbar. Weiter seien die Vertauschbarkeitsbedingungen

$$\mathbb{E}_\vartheta \frac{1}{f(\cdot; \vartheta)} \nabla f(\cdot; \vartheta) = 0, \quad \mathbb{E}_\vartheta \frac{1}{f(\cdot; \vartheta)} \nabla \nabla^t f(\cdot; \vartheta) = 0$$

für jedes $\vartheta \in \overset{\circ}{\Theta}$ erfüllt. Dann ist

$$\mathcal{I}(\vartheta) = -\mathbb{E}_\vartheta \nabla \nabla^t \ln f(\cdot; \vartheta) \quad \forall \vartheta \in \overset{\circ}{\Theta} .$$

die *Fisher-Information* von $\{\mathbb{P}_\vartheta \mid \vartheta \in \Theta\}$ bezüglich ϑ , vergleiche (Witting, 1985, S. 155).

□

Satz 5.1.2 (Asymptotische Normalität)

Sei $(\mathfrak{X}, \mathfrak{B}, \mathbb{P}_\vartheta \mid \vartheta \in \Theta)$ ein statistischer Raum und $\{\mathbb{P}_\vartheta \mid \vartheta \in \Theta\}$ eine μ -dominierte Verteilungsklasse mit $\Theta \subset \mathbb{R}^k$. $X_i, i = 1, \dots, n$, seien unabhängige identisch verteilte Zufallsvariablen mit $X_i \sim f(x; \vartheta)$, wobei $f(x; \vartheta)$ eine μ -Dichte von \mathbb{P}_ϑ ist, die folgende Voraussetzungen erfülle:

1. $\forall x \in \mathfrak{X} \quad \forall i, j = 1, \dots, k$ existiert $\nabla_i \nabla_j f(x; \vartheta)$ und ist stetig auf $\overset{\circ}{\Theta}$;
2. $\forall \vartheta \in \overset{\circ}{\Theta}$ gilt: $\mathbb{E}_\vartheta \frac{1}{f(X_1; \vartheta)} \nabla_i f(X_1; \vartheta) = 0$ und $\mathbb{E}_\vartheta \frac{1}{f(X_1; \vartheta)} \nabla_i \nabla_j f(X_1; \vartheta) = 0$ für $i, j = 1, \dots, k$;
3. Für jedes $i, j = 1, \dots, k$ gibt es ein $\delta_\vartheta > 0$ mit einer Umgebung $U(\vartheta, \delta_\vartheta) \subset \overset{\circ}{\Theta}$ und eine messbare Funktion $M(\cdot, \vartheta)$ auf \mathfrak{X} mit $\mathbb{E}_\vartheta M(X_1, \vartheta) < \infty$ derart, dass gilt:

$$|\nabla_i \nabla_j \ln f(\cdot, \vartheta')| \leq M(\cdot, \vartheta) \quad \forall \vartheta' \in U(\vartheta, \delta_\vartheta), \quad i, j = 1, \dots, k;$$

4. $\forall \vartheta \in \overset{\circ}{\Theta}$ ist $|\mathcal{I}(\vartheta)| \neq 0$.

Dann ist jeder konsistente ML-Schätzer $\hat{\vartheta}_n$ asymptotisch normal und asymptotisch effizient:

$$\sqrt{n}(\hat{\vartheta}_n(X) - \vartheta) \xrightarrow{\mathcal{D}} \mathbf{N}(0, \mathcal{I}(\vartheta)^{-1}) \quad \forall \vartheta \in \overset{\circ}{\Theta} .$$

Beweis.

Siehe (Witting and Müller-Funk, 1995, S. 202-205).

5.1.1 Exakte Maximum-Likelihood-Methode

Das Maximum-Likelihood-Prinzip soll nun zur Schätzung der Copulaparameter verwendet werden. Wie bereits gesehen, lässt sich die Verteilungsfunktion eines d -dimensionalen Zufallsvektors \mathbf{X} nach dem Satz von Sklar als

$$F(\mathbf{x}; \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_d, \mathbf{R}) = C(F_1(x_1; \boldsymbol{\alpha}_1), \dots, F_d(x_d; \boldsymbol{\alpha}_d); \mathbf{R}) \quad (5.2)$$

darstellen, wobei C die zugehörige Copulafunktion mit dem Parametervektor \mathbf{R} und F_1, \dots, F_d univariate Randverteilungen mit den jeweiligen Parametervektoren $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_d$ bezeichnen. Der unbekannte Parameter $\boldsymbol{\vartheta} = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_d, \mathbf{R}) \in \Theta$ soll geschätzt werden. Sei angenommen, dass die Copula C über die Dichtefunktion c und die Randverteilungen F_1, \dots, F_d über die Randdichten f_1, \dots, f_d verfügen. Durch Ableiten von (5.2) erhält man folgende Darstellung der Dichtefunktion f von \mathbf{X} :

$$\begin{aligned} f(\mathbf{x}; \boldsymbol{\vartheta}) &= \frac{\partial^d}{\partial x_1 \cdots \partial x_d} F(x_1, \dots, x_d; \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_d, \mathbf{R}) \\ &= \frac{\partial^d}{\partial x_1 \cdots \partial x_d} C(F_1(x_1; \boldsymbol{\alpha}_1), \dots, F_d(x_d; \boldsymbol{\alpha}_d); \mathbf{R}) \\ &= c(F_1(x_1; \boldsymbol{\alpha}_1), \dots, F_d(x_d; \boldsymbol{\alpha}_d); \mathbf{R}) \prod_{i=1}^d f_i(x_i; \boldsymbol{\alpha}_i). \end{aligned} \quad (5.3)$$

Basierend auf n unabhängigen Beobachtungen $\mathbf{x}_1, \dots, \mathbf{x}_n$ von \mathbf{X} lautet die Likelihoodfunktion für $\boldsymbol{\vartheta}$:

$$L(\boldsymbol{\vartheta}; \mathbf{x}_1, \dots, \mathbf{x}_n) = \prod_{i=1}^n f(\mathbf{x}_i; \boldsymbol{\vartheta}).$$

Entsprechend kann die Log-Likelihoodfunktion geschrieben werden als

$$\begin{aligned} l(\boldsymbol{\vartheta}; \mathbf{x}_1, \dots, \mathbf{x}_n) &= \sum_{i=1}^n \ln f(\mathbf{x}_i; \boldsymbol{\vartheta}) \\ &= \sum_{i=1}^n \ln c(F_1(x_{i1}; \boldsymbol{\alpha}_1), \dots, F_d(x_{id}; \boldsymbol{\alpha}_d); \mathbf{R}) + \sum_{i=1}^n \sum_{j=1}^d \ln f_j(x_{ij}; \boldsymbol{\alpha}_j). \end{aligned}$$

Damit ist der ML-Schätzer für $\boldsymbol{\vartheta}$

$$\hat{\boldsymbol{\vartheta}}_n = \arg \max_{\boldsymbol{\vartheta} \in \Theta} l(\boldsymbol{\vartheta}; \mathbf{x}_1, \dots, \mathbf{x}_n).$$

Da bei dieser Methode in einem Schritt die Parameter der Copula und der Randverteilungen geschätzt werden, wird sie als *exakte Maximum-Likelihood-Methode* bezeichnet. Sind die Randverteilungen vollständig bekannt und ist man bestrebt den Copulaparameter \mathbf{R} zu schätzen, so wird nur die Summe

$$\sum_{i=1}^n \ln c(F_1(x_{i1}; \boldsymbol{\alpha}_1), \dots, F_d(x_{id}; \boldsymbol{\alpha}_d); \mathbf{R})$$

nach \mathbf{R} maximiert, da der Copulaparameter in den einzelnen Randverteilungen nicht vorkommt.

5.1.2 2-stufige parametrische Maximum-Likelihood-Methode

Ist bei einem Einschrittverfahren wie der exakten ML-Methode eine analytische Herleitung der Lösung nicht möglich, muss auf ein numerisches Verfahren zur Lösung des Gleichungssystems (5.1) ausgewichen werden. Dies kann aber im Allgemeinen zwei Schwierigkeiten beinhalten. Zum Einen kann es vorkommen, dass das numerische Verfahren nicht konvergiert, zum Anderen kann eine hohe Dimension d des angenommenen Modells die Berechnung sehr aufwändig und zeitintensiv werden lassen.

Eine Möglichkeit, die numerische Anwendung zu umgehen oder ihre Komplexität zu verringern, wird bei Joe and Xu (1996) als "Method of Inference Functions for

Margins" vorgestellt. Die Schätzidee basiert darauf, dass der gemeinsame Parameter ϑ aus den Parametern, die eindeutig den Randverteilungsdichten f_1, \dots, f_d zugeordnet sind, $\alpha_1, \dots, \alpha_d$, und dem Parameter, der eindeutig der Copulafunktion C zugeordnet ist, \mathbf{R} , besteht. Aufgrund dieser Tatsache lässt sich das Schätzproblem in zwei getrennten, weniger aufwändigen Stufen durchführen.

Im ersten Schritt werden für die Parameter $\alpha_1, \dots, \alpha_d$ d separate univariate Log-Likelihoodfunktionen

$$l_i(\alpha_i; x_{1i}, \dots, x_{ni}) = \sum_{j=1}^n \ln f_j(x_{ji}; \alpha_i), \quad i = 1, \dots, d,$$

aufgestellt. Das Maximieren liefert dann die Schätzwerte $\tilde{\alpha}_1, \dots, \tilde{\alpha}_d$.

Anschließend werden diese Schätzwerte in die multivariate Log-Likelihoodfunktion eingesetzt:

$$l(\mathbf{R}; \mathbf{x}_1, \dots, \mathbf{x}_n, \tilde{\alpha}_1, \dots, \tilde{\alpha}_d) = \sum_{i=1}^n \ln c(F_1(x_{i1}, \tilde{\alpha}_1), \dots, F_d(x_{id}, \tilde{\alpha}_d); \mathbf{R}) + \sum_{i=1}^n \sum_{j=1}^d \ln f_j(x_{ij}, \tilde{\alpha}_j).$$

Auf diese Weise erhält man eine Funktion, die nur noch den Copulaparameter als Unbekannte enthält. Da der Parameter \mathbf{R} in den Randverteilungsdichten nicht vorkommt, ist das Maximieren von $l(\mathbf{R}; \mathbf{x}_1, \dots, \mathbf{x}_n, \tilde{\alpha}_1, \dots, \tilde{\alpha}_d)$ gleichbedeutend mit dem Maximieren von

$$\sum_{i=1}^n \ln c(F_1(x_{i1}, \tilde{\alpha}_1), \dots, F_d(x_{id}, \tilde{\alpha}_d); \mathbf{R}).$$

Damit ist die Schätzung für den Parameter \mathbf{R}

$$\tilde{\mathbf{R}} = \arg \max_{\mathbf{R}} \sum_{i=1}^n \ln c(F_1(x_{i1}, \tilde{\alpha}_1), \dots, F_d(x_{id}, \tilde{\alpha}_d); \mathbf{R}).$$

Offensichtlich werden bei diesem Verfahren die ursprünglich beobachteten Daten mit Hilfe der parametrisch geschätzten Randverteilungen transformiert und

die transformierten Daten genutzt, um in der Folge eine Schätzung für den Copulaparameter zu gewinnen. Da das Verfahren in beiden Schritten auf parametrischen Schätzungen aufgebaut ist, soll es hier als *2-stufige parametrische Maximum-Likelihood-Methode* (PML-Methode) bezeichnet werden. Entsprechend wird die Schätzung des Parameters ϑ nach der PML-Methode mit

$$\hat{\vartheta}^{PML} = (\tilde{\alpha}_1, \dots, \tilde{\alpha}_d, \tilde{\mathbf{R}})$$

gekennzeichnet.

Wie oben bereits erwähnt, kommt die PML-Methode insbesondere dann zum Einsatz, wenn eine analytische Herleitung der exakten Lösungen nicht möglich ist. Im Folgenden soll gezeigt werden, dass eine Optimierung mit Hilfe der 2-stufigen parametrischen ML-Methode im Vergleich zu der exakten Methode einige Vorteile aufweist. So ist es nicht notwendig eine numerische Optimierung zur Schätzung der Parameter heranzuziehen, wenn die Schätzungen in den einzelnen Stufen auf analytischem Wege erreichbar sind. Es wurde auch bereits erwähnt, dass ein numerisches Verfahren nicht zwangsweise konvergieren muss und so nicht immer zu einer Lösung führt. Lässt sich eine numerische Optimierung nicht umgehen, so wirkt sich das Zwei-Schritt-Verfahren komplexitätsreduzierend aus. In bestimmten Fällen kann sogar erst durch Aufteilen der Parameterschätzungen die Konvergenz der eingesetzten numerischen Verfahren und damit verbundene Lösung eines Schätzproblems erreicht werden, siehe (Glauser, 2003, S. 92). Letztendlich stellen Joe and Xu (1996) in ihren Studien fest, dass das gleichzeitige Optimieren aller Modellparameter zusammen zeitaufwändiger ist als die Gesamtheit mehrerer Optimierungen von jeweils wenigen Parametern. Zusammengefasst kann die PML-Methode im Vergleich zu der EML-Methode im Allgemeinen als komplexitätsreduziernd und rechentechnisch weniger aufwändig angesehen werden.

Es ist jedoch damit zu rechnen, dass die Ergebnisse der PML-Methode weniger genau als die Ergebnisse der EML-Methode sind. Die Ursache dafür ist, dass ein Zweischrittverfahren ein höheres Fehlerrisiko als ein Einschrittverfahren mit sich bringt, da man im zweiten Schritt auf möglicherweise fehlerbehafteten Schätzungen des

ersten Schrittes aufbaut. Mehrere empirische Studien (siehe z.B. Bouyé et al. (2000) oder Durrleman et al. (2000)) zeigen, dass insbesondere für große Anzahl der Beobachtungen die Abweichungen in den Ergebnissen der PML- und der EML-Methode sehr gering ausfallen.

5.1.3 2-stufige semiparametrische Maximum-Likelihood-Methode

In den Abschnitten 5.1.1 und 5.1.2 wurde von einem Modell ausgegangen, in dem sowohl die Copula als auch die Randverteilungen bis auf die Parameter als bekannt vorausgesetzt waren. In diesem Abschnitt soll ein Modell betrachtet werden, bei dem die Randverteilungen gänzlich unbekannt sind und man sich anhand der beobachteten Daten nur für den Copulaparameter interessiert. In diesem Fall ist also der zu schätzende Parameter $\vartheta = \mathbf{R}$. Ähnlich wie bei der PML-Methode wendet man ein Zweischrittverfahren an.

Zunächst werden die Randverteilungen F_1, \dots, F_d empirisch geschätzt:

$$\hat{F}_i(\mathbf{x}) = \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{(x_{ji} \leq x)}, \quad i = 1, \dots, d.$$

Im nächsten Schritt werden die Beobachtungsdaten anhand der empirischen Randverteilungen $\hat{F}_1, \dots, \hat{F}_d$ transformiert und in die multivariate Log-Likelihoodfunktion eingesetzt. In Anlehnung an die Ausführungen im Abschnitt 5.1.2 kann die Schätzung für den Parameter ϑ durch das Maximieren von

$$\sum_{j=1}^n \ln c(\hat{F}_1(x_{j1}), \dots, \hat{F}_d(x_{jd}); \vartheta)$$

erreicht werden. Da das Verfahren zum Teil auf nichtparametrischen Schätzungen der Randverteilungsfunktionen basiert, wird in diesem Fall von einem semiparametrischen Verfahren gesprochen. Aus diesem Grund soll dieses Verfahren hier als *2-stufige semiparametrische Maximum-Likelihood-Methode* (SML-Methode) bezeichnet werden, sowie der entsprechende Parameterschätzer mit ϑ^{SML} . Bei Bouyé et al. (2000) wird dieses Verfahren unter der Bezeichnung "Canonical Maximum Likelihood Method" eingeführt. Es mag unter den gegebenen Umständen als

natürlich erscheinen, den Copulaparameter R auf diese Weise zu schätzen.

Von den drei vorgestellten Anwendungen der Maximum-Likelihood-Methode werden allein bei der SML-Methode keinerlei Annahmen über Randverteilungen getroffen. Damit scheint sie für Probleme, die zum Ziel haben, allein die Abhängigkeitsstruktur der Komponenten zu untersuchen, besonders gut geeignet zu sein. Außerdem kann ein Vergleich zwischen den Ergebnissen der SML-Methode und der EML- bzw. der PML-Methode Hinweise auf die Richtigkeit der bereits getroffenen Annahmen über die Randverteilungen geben, siehe (Glauser, 2003, S. 96).

5.2 Momentenmethode

Als eine Alternative zur Maximum-Likelihood-Methode kann die Momentenmethode betrachtet werden. Diese wurde von Karl Pearson bereits gegen Ende des 19. Jahrhunderts eingeführt und ist somit eines der ältesten Verfahren zur Gewinnung von Parameterschätzern. Die Grundidee der Methode ist, die Momente als invertierbare Funktionen des unbekanntes Parameters darzustellen. Auf dieser Basis kann ein Gleichungssystem aufgestellt werden, dessen Lösung zur Schätzung des Parameters führt.

Sei X eine eindimensionale Zufallsvariable mit dem statistischen Raum $(\mathcal{X}, \mathfrak{B}, \mathbb{P}_\vartheta \mid \vartheta \in \Theta)$ und sei ϑ der zu schätzende Parameter. Die Momentenbedingungen wird folgendermaßen definiert

$$\mathbb{E}(g(X; \vartheta)) = \mathbf{0},$$

wobei die Vektorfunktion g den funktionellen Zusammenhang zwischen den Momentenfunktionen und dem Parameter ϑ beschreibt und somit mindestens die Dimension von ϑ haben muss. Sei $\mathbf{x} = (x_1, \dots, x_n)$ eine Stichprobe mit unabhängigen

5.2 Momentenmethode

Realisierungen $x_j, j = 1, \dots, n$, von \mathbf{X} . Für die Schätzung geht man zu den empirischen Momentenbedingungen über

$$\frac{1}{n} \sum_{i=1}^n g(\mathbf{x}_i; \boldsymbol{\vartheta}) = \mathbf{0}.$$

Entspricht die Dimension von g der Dimension von $\boldsymbol{\vartheta}$, so kann nach $\boldsymbol{\vartheta}$ aufgelöst werden und liefert die Schätzung

$$\hat{\boldsymbol{\vartheta}} = \arg_{\boldsymbol{\vartheta} \in \Theta} \left[\frac{1}{n} \sum_{i=1}^n g(\mathbf{x}_i; \boldsymbol{\vartheta}) = \mathbf{0} \right].$$

Ist die Dimension von g größer als die Dimension von $\boldsymbol{\vartheta}$, wird $\hat{\boldsymbol{\vartheta}}$ so gewählt, dass eine (ggf. gewichtete) quadratische Distanz zwischen $\frac{1}{n} \sum_{i=1}^n g(\mathbf{x}_i; \boldsymbol{\vartheta})$ und Nullvektor minimiert wird:

$$\hat{\boldsymbol{\vartheta}} = \arg \min_{\boldsymbol{\vartheta} \in \Theta} \left[\frac{1}{n} \sum_{i=1}^n g(\mathbf{x}_i; \boldsymbol{\vartheta}) \right]^t W \left[\frac{1}{n} \sum_{i=1}^n g(\mathbf{x}_i; \boldsymbol{\vartheta}) \right],$$

wobei W eine positiv definite und symmetrische Gewichtematrix ist. Auf diese Weise wird garantiert, dass $\hat{\boldsymbol{\vartheta}}$ die empirischen Momentenbedingungen möglichst gut erfüllt.

Üblicherweise werden als Momentenfunktionen die ersten theoretischen Momente der Zufallsvariable gewählt. Die Momentenbedingung hat dann die Form

$$\mathbb{E}(g(X; \boldsymbol{\vartheta})) = \begin{bmatrix} \mathbb{E}(X) - \varphi_1(\boldsymbol{\vartheta}) \\ \mathbb{E}(X^2) - \varphi_2(\boldsymbol{\vartheta}) \\ \vdots \\ \mathbb{E}(X^r) - \varphi_r(\boldsymbol{\vartheta}) \end{bmatrix} = \mathbf{0}.$$

Folgende Modellannahmen müssen getroffen werden:

- Für $r \in \mathbb{N}$ und für jedes $\boldsymbol{\vartheta} \in \Theta$ gelte: $\mathbb{E}_{\boldsymbol{\vartheta}}(|X|^r) < \infty$.
- Die Momente $m_i = \mathbb{E}_{\boldsymbol{\vartheta}}(X^i)$ von X seien für jedes $i = 1, \dots, r$ über bekannte Funktionen des Parameters $\boldsymbol{\vartheta}$ darstellbar, d.h. es gebe eine Borel-messbare Funktion $\varphi : \Theta \rightarrow \mathbb{R}^r$, so dass

$$(m_1, \dots, m_r)^t = \varphi(\boldsymbol{\vartheta}) = (\varphi_1(\boldsymbol{\vartheta}), \dots, \varphi_r(\boldsymbol{\vartheta}))^t.$$

Der Parameter ϑ kann also als eine Funktion der Momente $\vartheta = \varphi^{-1}(m_1, \dots, m_r)$ dargestellt werden. Für die empirische Momentenbedingungen werden die theoretischen Momente durch die empirischen Momente $\hat{m}_i = \frac{1}{n} \sum_{j=1}^n x_j^i$, $1 \leq i \leq r$, substituiert. Die Schätzung $\hat{\vartheta} = \varphi^{-1}(\hat{m}_1, \dots, \hat{m}_r)$ heißt dann die *Momentenschätzung* für ϑ . Falls φ^{-1} eine messbare Funktion ist, kann man von einer *Momentenschätzfunktion* (M-Schätzfunktion) bzw. einem *Momentenschätzer* (M-Schätzer) $\hat{\vartheta} = \hat{\vartheta}(x)$ für den Parameter ϑ sprechen.

Es ist bekannt, dass die empirischen Momente (stark konsistente) Schätzer sind. Ist außerdem φ^{-1} stetig, so lässt sich nachweisen, dass der Momentenschätzer $\hat{\vartheta}$ konsistent ist.

Satz 5.2.1 (Konsistenz)

Es sei der statistische Raum wie oben beschrieben gegeben. $\hat{\vartheta} = \varphi^{-1}(\hat{m}_1, \dots, \hat{m}_r)$ sei der M-Schätzer für ϑ , so dass $\varphi^{-1}(m_1, \dots, m_r)$ für alle $(m_1, \dots, m_r) \in \Gamma = \{(m_1, \dots, m_r) \mid m_i = \varphi_i(\theta); i = 1, \dots, r; \vartheta \in \Theta\}$ stetig ist. Dann gilt:

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\hat{\vartheta} - \vartheta| > \varepsilon) = 0 \quad \text{oder} \quad \hat{\vartheta} \xrightarrow{\mathbb{P}} \vartheta.$$

Beweis.

Siehe (Mittelhammer, 1996, S. 493).

Dem nachfolgenden Satz kann man entnehmen, dass die Momentenschätzer unter geeigneten Regularitätsbedingungen asymptotisch normalverteilt sind.

Satz 5.2.2 (Asymptotische Normalität)

Es sei der statistische Raum wie oben beschrieben gegeben. $\hat{\vartheta} = \varphi^{-1}(\hat{m}_1, \dots, \hat{m}_r)$ sei der M-Schätzer für ϑ , so dass $\varphi^{-1}(m_1, \dots, m_r)$ für alle $(m_1, \dots, m_r) \in \Gamma$ differenzierbar ist und alle Elemente der Jacobi-Matrix

$$J = \begin{pmatrix} \frac{\partial \varphi_1^{-1}(m_1, \dots, m_r)}{\partial m_1} & \dots & \frac{\partial \varphi_1^{-1}(m_1, \dots, m_r)}{\partial m_r} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_r^{-1}(m_1, \dots, m_r)}{\partial m_1} & \dots & \frac{\partial \varphi_r^{-1}(m_1, \dots, m_r)}{\partial m_r} \end{pmatrix}$$

stetig sind. Außerdem habe J für alle $(m_1, \dots, m_r) \in \Gamma$ einen vollen Rang. Dann gilt:

$$\sqrt{n}(\hat{\vartheta} - \vartheta) \xrightarrow{\mathcal{D}} N_r(\mathbf{0}, J \Sigma J^T),$$

wobei $\Sigma = \text{Cov}(m_1, \dots, m_r)$ die Kovarianzmatrix der Momente ist.

Beweis.

Siehe (Mittelhammer, 1996, S. 494).

Im Allgemeinen kann die Einschränkung auf Unabhängigkeit der Realisierungen und Eindimensionalität von X fallen gelassen werden, man spricht dann von der *verallgemeinerten Momentenmethode* ("Generalized Method of Moments"), siehe z.B. Mittelhammer (1996) oder Lee (1996). Außerdem können anstelle der theoretischen Momente m_i andere Funktionen eingesetzt werden. Wichtig ist, dass die gewählten Momentenfunktionen sich in Abhängigkeit von dem Parameter ϑ darstellen lassen. Dieser Punkt soll im nächsten Abschnitt aufgegriffen werden, indem man die theoretischen Momente durch copulabasierte Abhängigkeitsmaße ersetzt. Ein positiver Aspekt der Momentenmethode ist, dass sie fast ohne Einschränkung für beliebige Zufallsvariablen einsetzbar ist. Jedoch ist ihre Berechnung, insbesondere bei Funktionen mit mehreren Variablen wie Copulas, auch mit rechen-technischen Unterstützung oft nicht trivial. Dieses Problem kann unter anderem durch Austauschen der Momentenfunktionen gelöst werden, siehe (Glauser, 2003, S. 94).

5.3 Kendall's τ -Momentenmethode

Im Abschnitt 5.2 wurde die allgemeine Funktionsweise der Momentenmethode erläutert und bereits darauf hingewiesen, dass es je nach Modell sinnvoll sein kann, andere Funktionen als theoretische Momente zum Schätzen des Parameters einzusetzen. Im Kapitel 3 wurden Abhängigkeitsmaße betrachtet, die als Funktionale der Copulafunktionen dargestellt werden können. Daher liegt es nahe, zur Schätzung der Copulaparameter die Momentenmethode mit den copulabasierten Abhängigkeitsmaßen entsprechend zu modifizieren. Insbesondere beim Schätzen der t -Copula wird die Momentenmethode basierend auf Kendall's τ angewandt. Damit wird die Korrelationsmatrix, die die t -Copula teilweise parametrisiert, geschätzt. Die genaue Vorgehensweise wird im Kapitel 6 erläutert.

Sei \mathbf{X} ein d -dimensionaler Zufallsvektor zum statistischen Raum $(\mathcal{X}, \mathfrak{B}, \mathbb{P}_\vartheta \mid \vartheta \in \Theta)$, wobei ϑ eine Parametermatrix ist. Für jede bivariate Randverteilung wird Kendall's τ als eine (Borel-messbare) Funktion des Parameters $\varphi_{ij} : \Theta \rightarrow \mathbb{R}$ dargestellt, so dass

$$\tau_{ij} := \tau(X_i, X_j) = \varphi_{ij}(\vartheta) \quad 1 \leq i < j \leq d,$$

gilt. Sei τ eine Matrix, die aus den Koeffizienten τ_{ij} besteht,

$$\tau = \begin{pmatrix} \tau_{11} & \tau_{12} & \cdots & \tau_{1d} \\ \tau_{21} & \tau_{22} & \cdots & \tau_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{d1} & \tau_{dd} & \cdots & \tau_{dd} \end{pmatrix}.$$

Berücksichtigt man, dass $\tau_{ii} = 1$ und $\tau_{ij} = \tau_{ji}$ für alle $1 \leq i < j \leq d$ gilt, so kann die Momentenbedingung folgendermaßen definiert werden:

$$\mathbb{E}(g(\mathbf{X}; \vartheta)) = \begin{bmatrix} 1 - \varphi_{11}(\vartheta) & \tau_{12} - \varphi_{12}(\vartheta) & \cdots & \tau_{1d} - \varphi_{1d}(\vartheta) \\ \tau_{12} - \varphi_{12}(\vartheta) & 1 - \varphi_{11}(\vartheta) & \cdots & \tau_{2d} - \varphi_{2d}(\vartheta) \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{1d} - \varphi_{1d}(\vartheta) & \tau_{2d} - \varphi_{2d}(\vartheta) & \cdots & 1 - \varphi_{dd}(\vartheta) \end{bmatrix} = [\tau - \varphi(\vartheta)] = \mathbf{0}.$$

Die Werte $\tau_{ij} \in [-1, 1]$, $1 \leq i < j \leq d$, können dann über das empirische Kendall's τ geschätzt werden.

Definition 5.3.1

Sei (Y_1, Y_2) ein bivariater Zufallsvektor und $(y_{11}, y_{12}), \dots, (y_{n1}, y_{n2})$ unabhängige Realisierungen von (Y_1, Y_2) . Dann definiert ein nichtparametrischer Schätzer

$\hat{\tau}_n : \mathbb{R}^{2n} \rightarrow [-1, 1]$ für $\tau(Y_1, Y_2)$, gegeben durch

$$\hat{\tau}_n := \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \text{sgn}((y_{i1} - y_{j1})(y_{i2} - y_{j2})),$$

den empirischen Kendall's τ .

□

Folgende Eigenschaften können für das empirische Kendall's τ nachgewiesen werden.

Bemerkung 1

Der Schätzer $\hat{\tau}_n$ ist *marginal-verteilungsfrei*. Das bedeutet, wenn F und G Verteilungsfunktionen mit $U_1 := F(Y_1)$ und $U_2 := G(Y_2)$ sind, so gilt mit Wahrscheinlichkeit gleich 1 die Beziehung:

$$\begin{aligned} \hat{\tau}_n &= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \text{sgn}((y_{i1} - y_{j1})(y_{i2} - y_{j2})) \\ &= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \text{sgn}((F(y_{i1}) - F(y_{j1}))(G(y_{i2}) - G(y_{j2}))) \\ &= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \text{sgn}((u_{i1} - u_{j1})(u_{i2} - u_{j2})). \end{aligned}$$

Anderes formuliert, heißt es, dass der Schätzer $\hat{\tau}_n$ von den Randverteilungen des Zufallsvektors unabhängig definiert ist, siehe (Ender, 2005, S. 16).

Bemerkung 2

Seien die Voraussetzungen aus der Definition 5.3.1 erfüllt. Dann ist $\hat{\tau}_n$ ein erwartungstreuer und konsistenter Schätzer für τ , siehe Gibbons (2003).

Satz 5.3.1

Seien die Voraussetzungen aus der Definition 5.3.1 erfüllt und außerdem sei $\sigma^2 := \text{Var}(\tau(Y_1, Y_2)) > 0$. Dann ist $\hat{\tau}_n$ ein asymptotisch normalverteilter Schätzer für τ mit

$$\hat{\tau}_n \sim \text{N}\left(0, \frac{4}{n\sigma^2}\right).$$

Beweis.

Siehe Lee (1990).

Sei $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ eine Stichprobe mit unabhängigen Realisierungen von \mathbf{X} und bezeichne

$$\hat{\tau}_{ij}(\mathbf{x}) = \frac{2}{n(n-1)} \sum_{1 \leq k < l \leq n} \text{sgn}[(x_{ki} - x_{li})(x_{kj} - x_{lj})] \quad (5.4)$$

den empirischen Kendall's τ für $\tau(X_i, X_j)$, $1 \leq i < j \leq d$. Die Elemente der Parametermatrix lassen sich durch die Vorschrift

$$\hat{\vartheta}_{ij} = \varphi_{ij}^{-1}(\hat{\tau}_{ij})$$

schätzen. Alle auf bivariaten Randverteilungen basierte Schätzungen zusammengefasst, kann ein Schätzer für ϑ konstruiert werden, der kurz mit

$$\hat{\vartheta} = \varphi^{-1}(\hat{\tau})$$

bezeichnet wird. Diese Schätzmethode wird im Folgenden mit *Kendall's τ -Momentenmethode*, (KM-Methode), bezeichnet. Die KM-Methode kann auf ähnliche Weise für Spearman's ρ oder andere copulabasierte Abhängigkeitsmaße angepasst werden. Im diesem Fall kann allgemein von der *Methode der copulabasierten Abhängigkeitsmaße* gesprochen werden.

Es kann nicht garantiert werden, dass die konstruierte Schätzmatrix nach der Transformation durch φ^{-1} den Anforderungen der Parametermatrix entspricht. Beispielsweise kann im Falle der Korrelationsmatrix passieren, dass die Schätzmethode eine singuläre Matrix ergibt. In solcher Situation muss die Schätzung zum Beispiel mit Hilfe der Eigenwertmethode von Rousseeuw and Molenberghs (1993) in eine positiv definite Matrix umgewandelt werden, siehe (Demarta and McNeil, 2004, S. 9). Dieses Verfahren wird im Kapitel 6 näher erläutert.

Da der Schätzer $\hat{\vartheta}$ über die Funktion φ^{-1} nur von $\hat{\tau}$ abhängt, ist er ebenfalls marginal-verteilungsfrei. Das ist insbesondere beim Schätzen der Copulaparameter ein wesentlicher Vorteil, denn die Daten müssen nicht wie beispielsweise bei der Maximum-Likelihood-Methoden vortransformiert werden. So wird das Fehlerrisiko der Schätzung nicht zusätzlich erhöht.

Kapitel 6

Parameterschätzungen für Normal- und t -Copula

Die klassischen Schätzmethoden, die im Kapitel 5 angeführt wurden, sind im Allgemeinen gut analysiert und weisen wünschenswerte statistische Eigenschaften auf. Bei einer Umsetzung im konkreten Fall treten die Probleme häufig dann auf, wenn die Berechnung der exakten Schätzungen zu großen Zeit- und Rechenaufwand erfordert oder erst gar nicht möglich ist. In solchen Fällen muss nach alternativen Lösungswegen gesucht werden. Diese werden meist stark an die Ausgangsproblemstellung angepasst. So ist beispielsweise im Falle der Normal-Copula stets eine gute Herleitung der klassischen Schätzer möglich. Dagegen müssen für die t -Copula neue Verfahren entwickelt werden, da eine exakte Darstellung der ML-Schätzer nicht möglich ist. Um trotzdem ein gutes Approximationsergebnis zu erzielen, werden verschiedene Kombinationen aus einem oder mehr klassischen Verfahren und numerischer Optimierung angewandt.

6.1 Normal-Copula

6.1.1 Maximum-Likelihood-Methode

Sei \mathbf{X} ein d -dimensionaler gemäß $N_d(\mathbf{0}, \mathbf{R})$ verteilter Zufallsvektor, wobei \mathbf{R} die Korrelationsmatrix ist. $C_{\mathbf{R}}^N$ bezeichne die zugehörige Copulafunktion mit der Parametermatrix \mathbf{R} . Um \mathbf{R} mittels der ML-Methode zu schätzen wird die Log-Likelihoodfunktion benötigt, die auf der Dichtefunktion von $C_{\mathbf{R}}^N$ basiert. Die Copuladichte wird durch das Umstellen der Gleichung (5.3) nach $c_{\mathbf{R}}^N$ ermittelt:

$$\begin{aligned} c_{\mathbf{R}}^N(\Phi(x_1), \dots, \Phi(x_d); \mathbf{R}) &= \frac{\frac{1}{(2\pi)^{d/2} |\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right)}{\prod_{i=1}^d \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{1}{2} x_i^2\right)} \\ &= \frac{1}{|\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right) \exp\left(\frac{1}{2} \mathbf{x}^T \mathbf{x}\right) \\ &= \frac{1}{|\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{x}^T (\mathbf{R}^{-1} - \mathbf{I}_d) \mathbf{x}\right), \end{aligned}$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung ist. Setzt man außerdem

$$\Phi(\mathbf{x}) = (\Phi(x_1), \dots, \Phi(x_d))^T = (u_1, \dots, u_d)^T = \mathbf{u},$$

so erhält man die übliche Form der Dichtefunktion

$$c_{\mathbf{R}}^N(\mathbf{u}; \mathbf{R}) = \frac{1}{|\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2} \Phi^{-1}(\mathbf{u})^T (\mathbf{R}^{-1} - \mathbf{I}_d) \Phi^{-1}(\mathbf{u})\right).$$

Seien $\mathbf{u}_1, \dots, \mathbf{u}_n$ unabhängige Realisierungen der auf $[0, 1]^d$ gleichverteilten Zufallsvariablen \mathbf{U} . Dann lässt sich anhand dieser Beobachtungen folgende Darstellung der Log-Likelihoodfunktion für \mathbf{R} herleiten:

$$l(\mathbf{R}; \mathbf{u}_1, \dots, \mathbf{u}_n) = -\frac{n}{2} \ln |\mathbf{R}| - \frac{1}{2} \sum_{i=1}^n \Phi^{-1}(\mathbf{u}_i)^T \mathbf{R}^{-1} \Phi^{-1}(\mathbf{u}_i) + \frac{1}{2} \sum_{i=1}^n \Phi^{-1}(\mathbf{u}_i)^T \Phi^{-1}(\mathbf{u}_i). \quad (6.1)$$

Man berechne die erste Ableitung von $l(\mathbf{R}; \mathbf{u}_1, \dots, \mathbf{u}_n)$

$$\frac{\partial l(\mathbf{R}; \mathbf{u}_1, \dots, \mathbf{u}_n)}{\partial \mathbf{R}^{-1}} = \frac{n}{2} \mathbf{R} - \frac{1}{2} \sum_{i=1}^n \Phi^{-1}(\mathbf{u}_i) \Phi^{-1}(\mathbf{u}_i)^T$$

6.1 Normal-Copula

und setze diese gleich 0. Dann ist

$$\hat{\mathbf{R}}^{ML} = \frac{1}{n} \sum_{i=1}^n \Phi^{-1}(\mathbf{u}_i) \Phi^{-1}(\mathbf{u}_i)^T. \quad (6.2)$$

die gesuchte Maximum-Likelihood-Schätzung für den Copulaparameter \mathbf{R} .

Es wurde am Anfang angenommen, dass alle Komponenten des Vektors X standardnormalverteilt sind. In diesem Fall vereinfacht sich die Formel (6.2) zu

$$\hat{\mathbf{R}}^{ML} = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T,$$

wobei $\mathbf{x}_1, \dots, \mathbf{x}_n$ die beobachteten Daten sind.

Sei nun im Allgemeinen die Verteilung von X über die Normal-Copula $C_{\mathbf{R}}^{\mathbf{N}}$ und die Randverteilungen F_1, \dots, F_d beschrieben. In diesem Fall müssen die ursprünglichen Daten vor der Anwendung der ML-Methode entsprechend der angenommenen oder geschätzten Randverteilungen F_1, \dots, F_d in

$$\mathbf{u}_i = (F_1(x_{i1}), \dots, F_d(x_{id}))^T, \quad i = 1, \dots, n,$$

umgewandelt werden. Danach kann $\hat{\mathbf{R}}^{ML}$ nach der Formel (6.2) berechnet werden.

6.1.2 Momentenmethode

Sei X ein d -dimensionaler Vektor und die Verteilung von X über die Normal-Copula $C_{\mathbf{R}}^{\mathbf{N}}$ und die Randverteilungen F_1, \dots, F_d beschrieben.

Die Elemente der Korrelationsmatrix \mathbf{R} können in folgender Weise über die theoretischen Momentenfunktionen dargestellt werden:

$$\begin{aligned} R_{ij} &= \frac{\mathbb{E}(X_i X_j) - \mathbb{E}(X_i) \mathbb{E}(X_j)}{\sqrt{(\mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2)(\mathbb{E}(X_j^2) - \mathbb{E}(X_j)^2)}} \\ &= \frac{m_2(X_i, X_j) - m_1(X_i) m_1(X_j)}{\sqrt{m_2(X_i, X_i) - m_1(X_i)^2} \sqrt{m_2(X_j, X_j) - m_1(X_j)^2}}, \quad i, j = 1, \dots, d, \end{aligned}$$

wobei m_1 und m_2 die ersten beiden multivariaten Momente sind. Dabei muss beachtet werden, dass der Copulaparameter \mathbf{R} nur im Falle von nichtentarteten multivariaten elliptischen Verteilungen die Korrelationmatrix im üblichen Sinne ist. Daher kann die M-Schätzung für R nicht in jedem Fall direkt über die Daten, wie in der Darstellung oben angeführt, bestimmt werden. Andererseits können die ursprünglichen Daten so transformiert werden, dass sie für die Anwendung der Momentenmethode geeignet sind:

$$\Phi^{-1}(\mathbf{u}_i) = (\Phi^{-1}(F_1(x_{i1})), \dots, \Phi^{-1}(F_d(x_{id})))^T, \quad i = 1, \dots, n,$$

wobei $\mathbf{x}_1, \dots, \mathbf{x}_n$ die Beobachtungen von \mathbf{X} sind. Entscheidend dabei ist, dass für die Konstruktion der Normal-Copula die Randverteilungen standardnormalverteilt vorausgesetzt werden. Die Momentenschätzung, die man dann erhält, stimmt mit der ML-Schätzung, siehe (6.2), überein. Somit liefert die klassische Momentenmethode für die Normal-Copula keine zusätzliche Schätzung für den Parameter R .

Eine bessere Alternative ist, den bestehenden direkten Zusammenhang zwischen der Parametermatrix \mathbf{R} und der Rangkorrelationsmatrix von Kendall

$$\tau_{ij} = \tau(X_i, X_j) = \frac{2}{\pi} \arcsin(R_{ij}), \quad i = 1 \dots n, j = 1 \dots d, \quad (6.3)$$

für die Schätzung zu nutzen. Diese Beziehung ist von den Randverteilungen unabhängig und kann somit zur Approximation des Parameters direkt über die beobachteten Daten verwendet werden. Ähnliches gilt für die Rangkorrelationsmatrix von Spearman

$$\rho_{ij} = \rho(X_i, X_j) = \frac{6}{\pi} \arcsin(R_{ij}), \quad i = 1 \dots n, j = 1 \dots d. \quad (6.4)$$

6.1.3 Kendall's τ -Momentenmethode

Seien $\mathbf{x}_1, \dots, \mathbf{x}_n$ die Beobachtungen von \mathbf{X} . Gemäß (6.3) können die Elemente von \mathbf{R} durch

$$\hat{R}_{ij} = \sin\left(\frac{\pi}{2} \hat{\tau}_{ij}\right), \quad i, j = 1, \dots, d,$$

geschätzt werden, wobei $\hat{\tau}_{ij}$ wie in (5.4) definiert ist. Sei $\hat{\tau}$ die Kendall's τ -Schätzmatrix, die aus den Elementen $\hat{\tau}_{ij}$ besteht. Dann kann die Schätzung für \mathbf{R} nach der Kendall's τ -Momentenmethode allgemein geschrieben werden als

$$\hat{\mathbf{R}}^{KM} = \sin\left(\frac{\pi}{2}\hat{\tau}\right).$$

6.1.4 Spearman's ρ -Momentenmethode

Seien $\mathbf{x}_1, \dots, \mathbf{x}_n$ die Beobachtungen von \mathbf{X} . Um mit Hilfe von (6.4) eine Schätzung für \mathbf{R} zu konstruieren, muss man sich erinnern, dass

$$\rho_{ij} = \rho(X_i, X_j) = \text{Corr}(F_i(X_i), F_j(X_j))$$

gilt. Ähnlich der ML-Methode müssen die beobachteten Daten im ersten Schritt anhand der angenommenen oder geschätzten Randverteilungen transformiert werden:

$$\mathbf{u}_i = (F_1(x_{i1}), \dots, F_d(x_{id}))^T, \quad i = 1, \dots, n.$$

Anschließend kann Spearman's ρ mittels der Momentenmethode durch

$$\hat{\rho}_{ij} = \frac{\hat{m}_{2ij} - \hat{m}_{1i}\hat{m}_{1j}}{\sqrt{\hat{m}_{2ii} - \hat{m}_{1i}^2}\sqrt{\hat{m}_{2jj} - \hat{m}_{1j}^2}}, \quad i, j = 1, \dots, d,$$

geschätzt werden. Dabei sind

$$\hat{m}_{1i} = \frac{1}{n} \sum_{k=1}^n u_{ki}, \quad i = 1, \dots, d,$$

und

$$\hat{m}_{2ij} = \frac{1}{n} \sum_{k=1}^n u_{ki}u_{kj}, \quad i, j = 1, \dots, d.$$

Ein alternativer Schätzer für Spearman's ρ , der direkt aus den Ausgangsdaten berechnet wird, wird in der nachfolgenden Definition vorgestellt, siehe (Niehof, 2009, S. 44).

Definition 6.1.1

Sei (Y_1, Y_2) ein bivariater Zufallsvektor und $(y_{11}, y_{12}), \dots, (y_{n1}, y_{n2})$ unabhängige Realisierungen von (Y_1, Y_2) . Sei $\text{rg}(y_{i1})$ der Rang der Beobachtung y_{i1} in der aufsteigend geordneten Folge y_{11}, \dots, y_{n1} . Weiterhin seien $\bar{\text{rg}}_1 = \frac{1}{n} \sum_{i=1}^n \text{rg}(y_{i1})$ bzw.

$\bar{\text{rg}}_2 = \frac{1}{n} \sum_{i=1}^n \text{rg}(y_{i2})$ die arithmetischen Mittel der Ränge von Y_1 bzw. Y_2 . Es gilt $\bar{\text{rg}}_1 = \bar{\text{rg}}_2 = \frac{n+1}{2}$. Dann definiert ein nichtparametrischer Schätzer $\hat{\rho} : \mathbb{R}^{2n} \rightarrow [-1, 1]$ für $\rho(Y_1, Y_2)$, gegeben durch

$$\begin{aligned} \hat{\rho} &= \frac{\sum_{i=1}^n (\text{rg}(y_{i1}) - \bar{\text{rg}}_1)(\text{rg}(y_{i2}) - \bar{\text{rg}}_2)}{\sqrt{\sum_{i=1}^n (\text{rg}(y_{i1}) - \bar{\text{rg}}_1)^2 \sum_{i=1}^n (\text{rg}(y_{i2}) - \bar{\text{rg}}_2)^2}} \\ &= \frac{12}{n(n^2 - 1)} \sum_{i=1}^n \left(\text{rg}(y_{i1}) - \frac{n+1}{2} \right) \left(\text{rg}(y_{i2}) - \frac{n+1}{2} \right) \\ &= \frac{12}{n(n^2 - 1)} \sum_{i=1}^n \text{rg}(y_{i1}) \text{rg}(y_{i2}) - 3 \frac{n+1}{n-1}, \end{aligned}$$

den empirischen Spearman's ρ .

□

Bemerkung 3

Der empirische Spearman's ρ ist ein erwartungstreuer und schwach konsistenter Schätzer. Der Beweis ist bei Gibbons (1993) zu finden.

Sei $\hat{\rho}$ die Schätzmatrix für Spearman's ρ , die aus den Elementen $\hat{\rho}_{ij}$ besteht. Dann kann die Schätzung für \mathbf{R} nach der Spearman's ρ -Momentenmethode (SM-Methode) folgendermaßen geschrieben werden

$$\hat{\mathbf{R}}^{SM} = 2 \sin \left(\frac{\pi}{6} \hat{\rho} \right).$$

6.2 t -Copula

6.2.1 Verfahren 1 (exakte ML-Methode)

Sei \mathbf{X} ein d -dimensionaler gemäß $t_\nu^d(\mathbf{0}, \mathbf{R})$ verteilter Zufallsvektor, wobei \mathbf{R} die Korrelationsmatrix und ν die Anzahl der Freiheitsgrade sind. $C_{\mathbf{R}, \nu}^t$ bezeichne die zugehörige Copulafunktion mit den Parametern \mathbf{R} und ν .

6.2 t -Copula

Um die Log-Likelihoodfunktion für die t -Copula aufzustellen, wird wiederum die Dichtefunktion benötigt. Diese kann wie bereits gesehen durch das Umstellen der Gleichung (5.3) nach $c_{\mathbf{R},\nu}^t$ ermittelt werden:

$$\begin{aligned} c_{\mathbf{R},\nu}^t(t_\nu(x_1), \dots, t_\nu(x_d); \mathbf{R}, \nu) &= \frac{\frac{\Gamma(\frac{\nu+d}{2})}{(\pi\nu)^{d/2} \Gamma(\frac{\nu}{2}) |\mathbf{R}|^{1/2}} \left(1 + \frac{1}{\nu} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right)^{-\frac{\nu+d}{2}}}{\frac{[\Gamma(\frac{\nu+1}{2})]^d}{(\pi\nu)^{d/2} [\Gamma(\frac{\nu}{2})]^d} \prod_{i=1}^d \left(1 + \frac{x_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \\ &= \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})} \left[\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})} \right]^d \frac{\left(1 + \frac{1}{\nu} \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}\right)^{-\frac{\nu+d}{2}}}{|\mathbf{R}|^{1/2} \prod_{i=1}^d \left(1 + \frac{x_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}}, \end{aligned}$$

wobei t_ν die Verteilungsfunktion der univariaten t -Verteilung ist. Setzt man außerdem

$$\mathbf{t}_\nu(\mathbf{x}) = (t_\nu(x_1), \dots, t_\nu(x_d))^T = (u_1, \dots, u_d)^T = \mathbf{u},$$

so erhält man die übliche Form der Dichtefunktion

$$c_{\mathbf{R},\nu}^t(u_1, \dots, u_d; \mathbf{R}, \nu) = \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})} \left[\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})} \right]^d \frac{\left(1 + \frac{1}{\nu} (\mathbf{t}_\nu^{-1}(\mathbf{u}))^T \mathbf{R}^{-1} (\mathbf{t}_\nu^{-1}(\mathbf{u}))\right)^{-\frac{\nu+d}{2}}}{|\mathbf{R}|^{1/2} \prod_{i=1}^d \left(1 + \frac{(t_\nu^{-1}(u_i))^2}{\nu}\right)^{-\frac{\nu+1}{2}}}.$$

Seien $\mathbf{u}_1, \dots, \mathbf{u}_n$ unabhängige Realisierungen der auf $[0, 1]^d$ gleichverteilten Zufallsvariablen \mathbf{U} . Auf der Grundlage dieser Beobachtungen lässt sich folgende Darstellung der Log-Likelihoodfunktion für \mathbf{R} herleiten:

$$\begin{aligned} l(\mathbf{R}, \nu; \mathbf{u}_1, \dots, \mathbf{u}_n) &= n \ln \left[\frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})} \right] - dn \ln \left[\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \right] - \frac{n}{2} \ln |\mathbf{R}| \\ &\quad - \frac{\nu+d}{2} \sum_{i=1}^n \ln \left(1 + \frac{1}{\nu} (\mathbf{t}_\nu^{-1}(\mathbf{u}_i))^T \mathbf{R}^{-1} (\mathbf{t}_\nu^{-1}(\mathbf{u}_i)) \right) \\ &\quad + \frac{\nu+1}{2} \sum_{j=1}^d \sum_{i=1}^n \ln \left(1 + \frac{(t_\nu^{-1}(u_{ij}))^2}{\nu} \right). \end{aligned} \quad (6.5)$$

Leider ist es nicht möglich, für die t -Copula eine explizite Darstellung der ML-Parameterschätzer anzugeben. Eine numerische Approximation, wie beispielsweise durch das Newton-Verfahren, ist in jedem Falle erforderlich. Jedoch können die

Berechnungen schon bei kleinen Dimensionen wegen der großen Anzahl von Operationschritten zu umständlich und zu zeitintensiv sein. In solchem Fall empfiehlt es sich, eine Kombination von verschiedenen Schätzmethoden anzuwenden. Dazu werden in den nächsten Teilabschnitten drei gängige Verfahren vorgestellt, vergleiche Fantazzini et al. (2006). Bei jedem der drei Verfahren wird vorausgesetzt, dass die ursprünglich beobachteten Daten $\mathbf{x}_1, \dots, \mathbf{x}_n$ zu den auf $[0, 1]^d$ gleichverteilten Stichproben $\mathbf{u}_1, \dots, \mathbf{u}_n$ transformiert werden. Wie die Transformation zu erfolgen hat, hängt davon ab, wie viel über die Randverteilungen der Copula bekannt ist. Gegebenfalls müssen dazu parametrische oder nichtparametrische Schätzungen der Randverteilungsfunktionen eingesetzt werden.

6.2.2 Verfahren 2 (2-stufige rekursive ML-Methode)

Dieses Verfahren basiert auf einer rekursiven Berechnung der Korrelationsmatrix \mathbf{R} und besteht aus zwei Stufen.

1. Stufe (rekursive Berechnung der Korrelationsmatrix \mathbf{R}):

Die $d(d-1)/2$ unbekanntes Elemente der Korrelationsmatrix \mathbf{R} werden einzeln geschätzt, indem für jedes Variablenpaar (X_i, X_j) , $i, j = 1, \dots, d$, die Log-Likelihoodfunktion nach (6.5) aufgestellt wird. Diese Vorgehensweise ist berechtigt, da die Korrelation zweier Zufallsvariablen von den weiteren Variablen des Zufallsvektors unabhängig ist. Eine Schätzung für das Element R_{ij} wird mittels einer Rekursion ermittelt, für deren Umsetzung die Anzahl der Freiheitsgrade ν aus dem Bereich der natürlichen Zahlen angenommen wird. Es sei bemerkt, dass diese Voraussetzung im allgemeinen nicht zwingend ist. Außerdem muss berücksichtigt werden, dass $\nu > 2$ für die Existenz der Varianz vorausgesetzt werden muss. Aufgrund dieser Ausführungen wird ν zunächst gleich 3 gesetzt und für diesen fixierten Wert mittels numerischer Optimierung eine ML-Schätzung für R_{ij} berechnet:

$$\hat{R}_{ij,\nu} = \arg \max_{R_{ij}} \left[-\frac{\nu+2}{2} \sum_{k=1}^n \ln \left(1 + \frac{1}{\nu} \boldsymbol{\xi}_k^T \mathbf{R}_{ij}^{-1} \boldsymbol{\xi}_k \right) - \frac{n}{2} \ln |\mathbf{R}_{ij}| \right],$$

wobei

$$\mathbf{R}_{ij} = \begin{pmatrix} 1 & R_{ij} \\ R_{ij} & 1 \end{pmatrix}$$

und

$$\boldsymbol{\xi}_k^T = (t_\nu^{-1}(u_{ki}), t_\nu^{-1}(u_{kj})), \quad k = 1, \dots, n,$$

sind. Anschließend wird der Wert von ν um 1 erhöht und $\hat{R}_{ij,\nu}$ erneut berechnet. Der Vorgang wird so lange wiederholt, bis das Maximum der bivariaten Log-Likelihoodfunktion gegenüber dem vorherigen Schritt zum ersten Mal abfällt. Als Schätzung für R_{ij} wird dann der Wert aus dem vorletzten Schritt der Schleife herangezogen, d.h. das Maximum aller ermittelten Maxima der Log-Likelihoodfunktion:

$$\hat{R}_{ij} = \hat{R}_{ij,\tilde{\nu}} \quad \text{mit} \quad \tilde{\nu} = \arg \max_{\nu \geq 3} (l(\hat{R}_{ij,\nu}, \nu; \mathbf{u}_1, \dots, \mathbf{u}_n)).$$

2. Stufe (Berechnung der Freiheitsgrade):

Im idealen Fall sollten alle $d(d-1)/2$ Elementenschätzungen für \mathbf{R} bei gleichem Wert von ν erzeugt worden sein. Das ist aber in der Praxis kaum der Fall. Aus diesem Grund wird ν unter Einsatz von $\hat{\mathbf{R}} = (\hat{R}_{ij})_{1 \leq i, j \leq d}$ nach der ML-Methode geschätzt:

$$\hat{\nu} = \arg \max_{\nu > 2} (l(\hat{\mathbf{R}}, \nu; \mathbf{u}_1, \dots, \mathbf{u}_n)).$$

Aufgrund der vielen Schritte, die für die Berechnung der Parameter notwendig sind, kann bereits vermutet werden, dass auch dieses Verfahren zeit- und rechenintensiv ist.

6.2.3 Verfahren 3 (2-stufige KM-ML-Methode)

Das zweite Verfahren stellt eine Kombination aus der Kendall's τ - und der Maximum-Likelihood-Methoden dar.

1. Stufe (KM-Methode für R):

Analog zu der Normal-Copula kann die Korrelationsmatrix im ersten Schritt mittels der KM-Methode geschätzt werden. Dieses Vorgehen ist legitim, da die Relation zwischen der Korrelationsmatrix und der Kendall's τ -Matrix von der Anzahl der Freiheitsgrade unabhängig ist.

$$\hat{\mathbf{R}} = \hat{\mathbf{R}}^{KM} = \sin\left(\frac{\pi}{2}\hat{\tau}\right), \quad (6.6)$$

wobei $\hat{\tau}$ die, anhand von transformierten oder untransformierten Beobachtungsdaten, geschätzte Kendall's τ -Matrix ist. Andererseits gibt es keine Garantie, dass die in (6.6) erzeugte Matrix tatsächlich positiv definit ist. Gegebenenfalls muss $\hat{\mathbf{R}}^{KM}$ in eine positiv definite Matrix umgewandelt werden, siehe (Demarta and McNeil, 2004, S. 9). Dazu wird häufig die Eigenwertmethode von Rousseeuw and Molenberghs (1993) verwendet. Diese Methode besteht aus folgenden Schritten:

- Eine positiv semi-definite Korrelationsmatrix $\mathbf{R} \in \mathbb{R}^{d \times d}$ kann zerlegt werden in

$$\mathbf{R} = \mathbf{A}\mathbf{D}\mathbf{A}^T. \quad (6.7)$$

\mathbf{D} ist eine Diagonalmatrix, deren Elemente die nichtnegativen Eigenwerte von \mathbf{R} sind. \mathbf{A} ist eine Orthogonalmatrix (d.h. $\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}_d$), deren Spalten die zugehörigen Eigenvektoren beinhalten. Falls \mathbf{R} positiv definit ist, sind alle ihre Eigenwerte positiv. Wenn \mathbf{R} symmetrisch, aber nicht positiv semi-definit ist, gilt (6.7) weiterhin, jedoch sind dann einige Eigenwerte negativ. In der Regel haben aber die negativen Eigenwerte nur kleine Absolutbeträge.

- Wenn die Ergebnismatrix positiv semi-definit sein soll, müssen alle negativen Eigenwerte durch 0 ersetzt werden. Soll die Ergebnismatrix sogar positiv definit sein, müssen die negativen Eigenwerte durch kleine positive Zahlen ersetzt werden. In jedem Fall bezeichnet \mathbf{D}' die neue Diagonalmatrix, dann kann $\mathbf{R}' = \mathbf{A}\mathbf{D}'\mathbf{A}^T$ berechnet werden.
- Durch die Umformung ist es möglich, dass die Diagonalelemente von \mathbf{R}' nicht mehr 1 sind. Dieses Problem wird durch eine Transformation gelöst.

Dabei wird \mathbf{R}' in $\mathbf{R}'' = \mathbf{D}''\mathbf{R}'\mathbf{D}''^T$ umgewandelt. \mathbf{D}'' ist eine Diagonalmatrix, deren Elemente $1/\sqrt{R''_{ii}}$, $i = 1, \dots, d$, sind, siehe (Rousseeuw and Molenberghs, 1993, S. 971-972). Damit erfüllt \mathbf{R}'' die notwendigen Kriterien und ist die gesuchte Approximation für \mathbf{R} .

2. Stufe (ML-Methode für ν):

Der plausibelste Ansatz, den noch verbleibenden Parameter ν zu schätzen, ist das Maximieren der Log-Likelihoodfunktion in (6.5) nach ν , wobei \mathbf{R} durch $\hat{\mathbf{R}}$ ersetzt wird:

$$\hat{\nu} = \arg \max_{\nu \geq 2} (l(\hat{\mathbf{R}}, \nu; \mathbf{u}_1, \dots, \mathbf{u}_n)).$$

Bei diesem Verfahren liegt die Vermutung nahe, dass es im Schnitt weniger Operationsschritte als das zuvor beschriebene benötigt. Einer der Gründe dafür ist, dass für die Berechnung von \mathbf{R} keine numerische Optimierung gebraucht wird.

6.2.4 Verfahren 4 (3-stufige M-ML-Methode)

Dieses Verfahren basiert auf der Momenten- und der Maximum-Likelihood-Methode.

1. Stufe (erste M-Methode für R):

Die Korrelationsmatrix \mathbf{R} wird das erste Mal unter der Annahme, dass die Daten gemäß einer Normal-Copula verteilt sind, geschätzt. Dazu wird die in (6.2) bereits hergeleitete ML-Schätzung (die der Momentenschätzung entspricht) verwendet:

$$\tilde{\mathbf{R}} = \frac{1}{n} \sum_{i=1}^n \Phi^{-1}(\mathbf{u}_i) \Phi^{-1}(\mathbf{u}_i)^T. \quad (6.8)$$

2. Stufe (ML-Methode für ν):

Durch Einsetzen der vorläufigen Korrelationsschätzung in (6.5) wird die ML-Schätzung für ν ermittelt:

$$\hat{\nu} = \arg \max_{\nu \geq 2} (l(\tilde{\mathbf{R}}, \nu; \mathbf{u}_1, \dots, \mathbf{u}_n)).$$

3. Stufe (zweite M-Methode für R):

Im letzten Schritt wird \mathbf{R} wiederholt mittels der Momenten-Methode und unter der Annahme, dass die Daten $t_{\hat{\nu}}$ -verteilt sind, geschätzt:

$$\hat{\mathbf{R}} = \left(\frac{\sum_{k=1}^n t_{\hat{\nu}}^{-1}(u_{ki}) t_{\hat{\nu}}^{-1}(u_{kj})}{\sqrt{\sum_{k=1}^n (t_{\hat{\nu}}^{-1}(u_{ki}))^2 \sum_{k=1}^n (t_{\hat{\nu}}^{-1}(u_{kj}))^2}} \right)_{i,j=1,\dots,d}.$$

Diese Methode ist in ähnlicher Weise bei Fantazzini et al. (2006) zu finden.

Alternativ kann \mathbf{R} in der letzten Etappe nach einem von Bouyé et al. (2000) vorgeschlagenem Verfahren geschätzt werden:

• **Alternative zur 3. Stufe:**

Die erste partielle Ableitung von (6.5) nach \mathbf{R}^{-1} ergibt:

$$\frac{\partial l(\mathbf{R}, \nu; \mathbf{u}_1, \dots, \mathbf{u}_n)}{\partial \mathbf{R}^{-1}} = \frac{n}{2} \mathbf{R} - \frac{\nu + d}{2} \sum_{i=1}^n \frac{\frac{1}{\nu} \mathbf{t}_{\nu}^{-1}(\mathbf{u}_i) (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))^T}{1 + \frac{1}{\nu} (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))^T \mathbf{R}^{-1} (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))}.$$

Durch das Setzen gleich 0, erhält man folgenden Ausdruck für \mathbf{R} :

$$\mathbf{R}^* = \frac{1}{n} \frac{\nu + d}{\nu} \sum_{i=1}^n \frac{\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i) (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))^T}{1 + \frac{1}{\nu} (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))^T (\mathbf{R}^*)^{-1} (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))}.$$

Das liefert die Grundlage für eine rekursive Berechnung. Sei $\hat{\mathbf{R}}_0 = \tilde{\mathbf{R}}$ wie in (6.8) der Startwert. Dann kann eine Schätzung für \mathbf{R} durch folgende Formel ermittelt werden:

$$\hat{\mathbf{R}}_{i+1} = \frac{1}{n} \frac{\nu + d}{\nu} \sum_{i=1}^n \frac{\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i) (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))^T}{1 + \frac{1}{\nu} (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))^T \hat{\mathbf{R}}_i^{-1} (\mathbf{t}_{\nu}^{-1}(\mathbf{u}_i))}, \quad i = 0, 1, 2, \dots$$

Sobald $\hat{\mathbf{R}}_{i+1} = \hat{\mathbf{R}}_i (= \hat{\mathbf{R}}_{\infty})$ eintritt, kann die Rekursion abgebrochen werden. Problematisch bei diesem Verfahren ist, dass keine Garantie gibt, ob die Rekursionen tatsächlich konvergieren. Das kann im Allgemeinen auch stark von dem Startwert abhängen, der über die Daten berechnet wird und somit nicht frei wählbar ist. Außerdem kann die Berechnung für große Dimensionen aufwändig werden und bei der Bestimmung des Matrizeninversen sogar zur numerischen Instabilität führen, siehe Fantazzini et al. (2006).

Kapitel 7

Simulationen

In diesem Kapitel wird die Arbeitsweise der im Kapitel 6 entwickelten Verfahren untersucht. Dazu wurden auf Basis der *R* Statistik-Software Simulationen durchgeführt. Eine Anleitung zur Installation und Anwendung von *R* ist unter <http://www.r-project.org/> zu finden. Ein Ausschnitt des erstellten Simulationsprogramms ist im Anhang A aufgeführt.

Die Umsetzung der Methoden in *R* ist für den Fall der Normal-Copula bei Yan (2006) ausreichend dokumentiert und muss hier nicht erneut erläutert werden. Es wird nur der Fall der bivariaten *t*-Copula mit *t*-verteilten Randverteilungen betrachtet. Mit dem Paket *copula* können verschiedene Copulafunktionen erstellt werden. Ausserdem beinhaltet es Prozeduren, die sowohl das Erzeugen von Realisierungen einer vordefinierten Copula als auch das Schätzen ihrer Parameter ermöglichen. Für diese Simulationsstudie wurde das Paket *copula* zur Definition der *t*-Copulas und zum Erzeugen von Stichproben herangezogen, bei der Implementierung der Schätzverfahren wurde darauf verzichtet, siehe Anhang A.

Folgende Verfahren kommen bei der Simulation zum Einsatz: Exakte Maximum-Likelihood-Methode (EMLM), rekursive Maximum-Likelihood-Methode (RMLM), Kendall's τ -Momenten-Maximum-Likelihood-Methode (KM-MLM) und Momenten-Maximum-Likelihood-Methode (M-MLM). Um die Übersicht beim Methodenvergleich anschaulicher zu gestalten, wurde die Methode nach Bouyé et al. (2000) bei dieser Simulation nicht umgesetzt.

7.1 Grundlagen

Einer der Schwerpunkte, welche hier untersucht werden, betrifft den Einfluss einer möglichen Störung, die in Form eines Nichtzentralitätsparameters der Randverteilungen ausgedrückt wird, auf das Schätzen der Copulaparameter. Bei (Witting, 1985, S. 221) wird die nichtzentrale t -Verteilung wie folgt definiert.

Definition 7.1.1

Y_0, Y_1, \dots, Y_ν seien unabhängige Zufallsvariablen mit $Y_0 \sim N(\delta, 1)$ und $Y_i \sim N(0, 1)$ für $i = 1, \dots, \nu$. Dann heißt die Verteilung von

$$Z := \frac{Y_0}{\sqrt{\frac{1}{\nu} \sum_{i=1}^{\nu} Y_i^2}}$$

eine nichtzentrale t_ν -Verteilung mit dem Nichtzentralitätsparameter δ ; kurz: $t_\nu(\delta)$.

□

Für die Simulation wurde eine t -Copula mit 7 Freiheitsgraden angenommen, abgekürzt $C_{R;\nu}^t$, wobei R die Korrelation zwischen den beiden Randverteilungen und ν die Anzahl der Freiheitsgrade beschreiben. Für $R = 0, 2; 0, 5; 0, 7; 0, 8$ wurden jeweils 40 Stichproben der Copula erzeugt und anschließend anhand von t -verteilten Randverteilungen mit verschiedenen Parametern transformiert. Auf diese Weise soll erreicht werden, dass die gleichen Daten mit unterschiedlichen Störungseinflüssen simuliert werden. Dabei wird nur der Nichtzentralitätsparameter der Randverteilungen verändert, die Anzahl der Freiheitsgrade beträgt immer 10. So transformierte Daten bilden die Grundlage für die Schätzungen. Die Anzahl der Stichproben wurde bewusst moderat gewählt, um die Simulation nah an den Realbedingungen zu gestalten.

7.2 Parametrischer Fall

Im ersten Fall werden die Ausgangsdaten anhand von parametrisch geschätzten Randverteilungen umgewandelt, bevor sie bei den Schätzmethoden zum Einsatz

kommen. Die Ergebnisse von 30 Wiederholungen werden in Form von Streudiagrammen in den Abbildungen 7.1 bis 7.4 grafisch dargestellt. Zwischen den einzelnen Bildern in den Abbildungen sind keine großen Unterschiede zu erkennen. Diese Tatsache spricht dafür, dass mögliche (geringe) Fehler, die beim Schätzen der Randverteilungsparameter entstehen, keinen großen Einfluss auf die Schätzung der Copulaparameter haben. Das kann wiederum als Bestätigung für die Trennung von Copulafunktion und den Randverteilungen aufgefasst werden.

Deutlich erkennbar ist, dass bei der Schätzung der Korrelation alle angewandten Verfahren ungefähr gleich stark um den wahren Parameterwert streuen. Für die Anzahl der Freiheitsgrade liefert das Einschrittverfahren, die exakte Maximum-Likelihood-Methode, deutlich bessere Schätzergebnisse als die Mehrschrittverfahren. Diese Aussage wird durch die in den Tabellen 7.1 bis 7.4 angegebenen Durchschnittswerte der Parameterschätzungen bestätigt. Hierfür wurden exemplarisch die auf $t_{10}(0)$ -verteilten Randverteilungen basierte Schätzdaten genommen.

Außerdem ist gut zu erkennen, dass alle Methoden den Korrelationsparameter umso besser approximieren, je näher sein Wert an der 1 liegt. Bei der exakten MLM verbessert sich sogar bei einer starken Korrelation die Schätzung für die Anzahl der Freiheitsgrade. Bei den Mehrschrittverfahren ist dies nicht zu beobachten. Generell wird bei den Mehrschrittverfahren die Anzahl der Freiheitsgrade für die gewählte Anzahl von Stichproben im Schnitt etwas zu hoch geschätzt.

Wie sich eine höhere Anzahl von Stichproben auf das Schätzergebnis auswirkt, zeigt die Abbildung 7.5. Hierfür wurden die Schätzungen bei 12 Wiederholungen für jeweils 100 Stichproben berechnet. Für den Korrelationsparameter erreichen alle Methoden sehr gute Werte, die im Durchschnitt sehr nah an dem exakten Parameterwert liegen. Auch die Anzahl der Freiheitsgrade wird von allen Methoden gut approximiert. Jedoch ist auch hier zu beobachten, dass die Streuung der Schätzung um den wahren Wert bei dem Einschrittverfahren wesentlich geringer als bei den Mehrschrittverfahren ist.

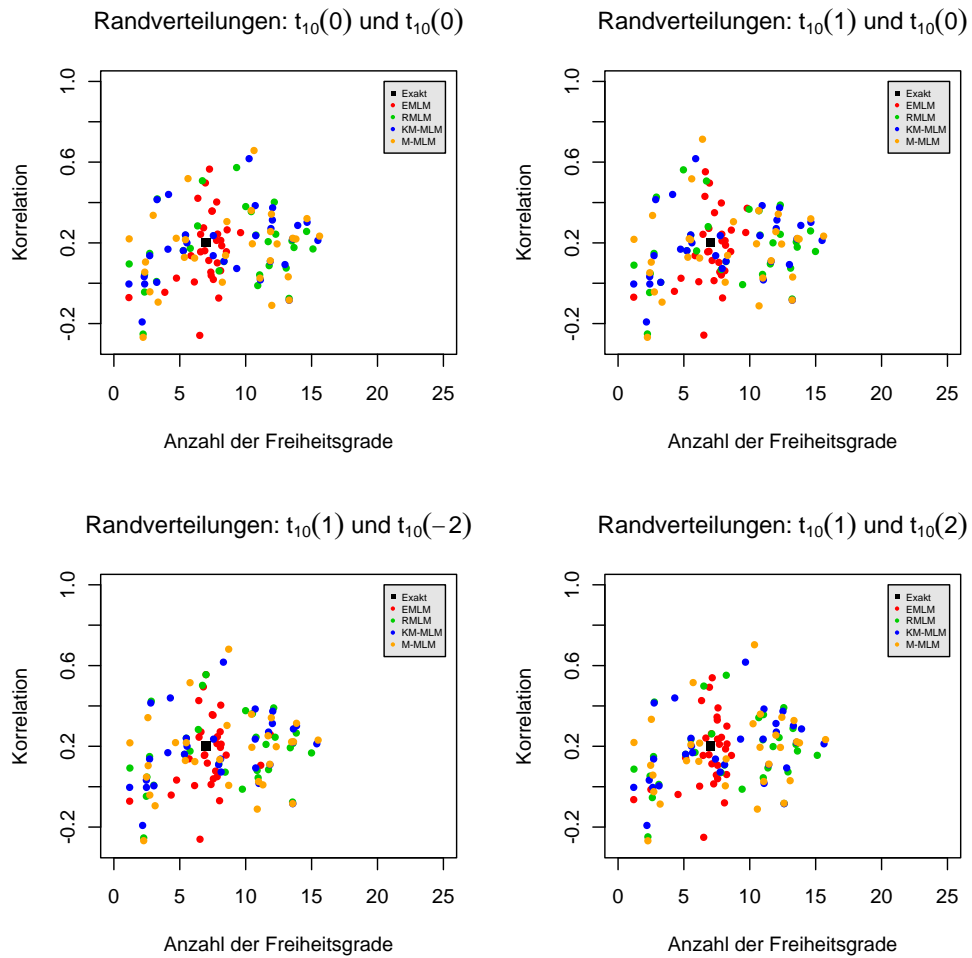


Abbildung 7.1: Parameterschätzungen für $C_{0,2;7}^t$ bei 40 Stichproben und 30 Wiederholungen.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.1687	0.1820	7.020	1.583
RMLM	0.1757	0.1781	8.568	4.326
KMMLM	0.1831	0.1708	8.308	4.474
MMLM	0.1646	0.1911	8.559	4.451

Tabelle 7.1: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,2;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.

7.2 Parametrischer Fall

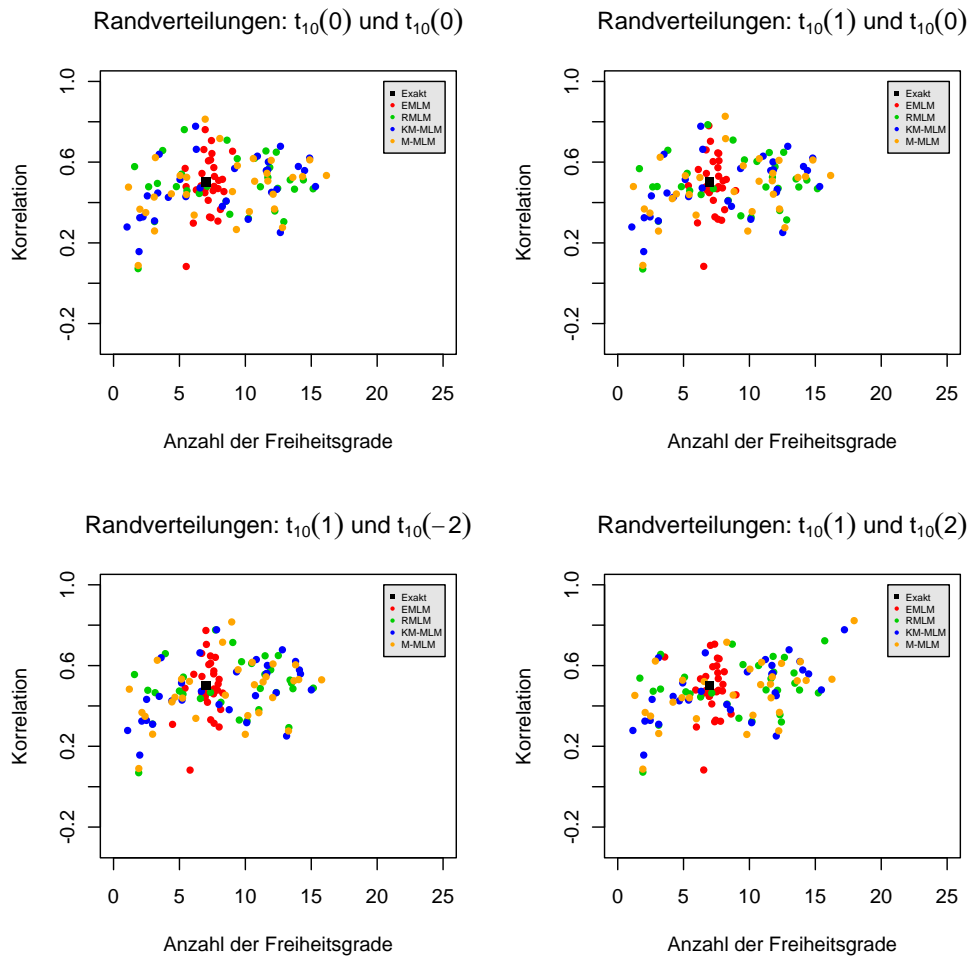


Abbildung 7.2: Parameterschätzungen für $C_{0,5;7}^t$ bei 40 Stichproben und 30 Wiederholungen.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.489	0.1395	7.205	0.8568
RMLM	0.4908	0.1411	8.669	4.1266
KMMLM	0.4765	0.1446	8.138	4.5550
MMLM	0.4704	0.1487	8.312	4.4284

Tabelle 7.2: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,5;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.

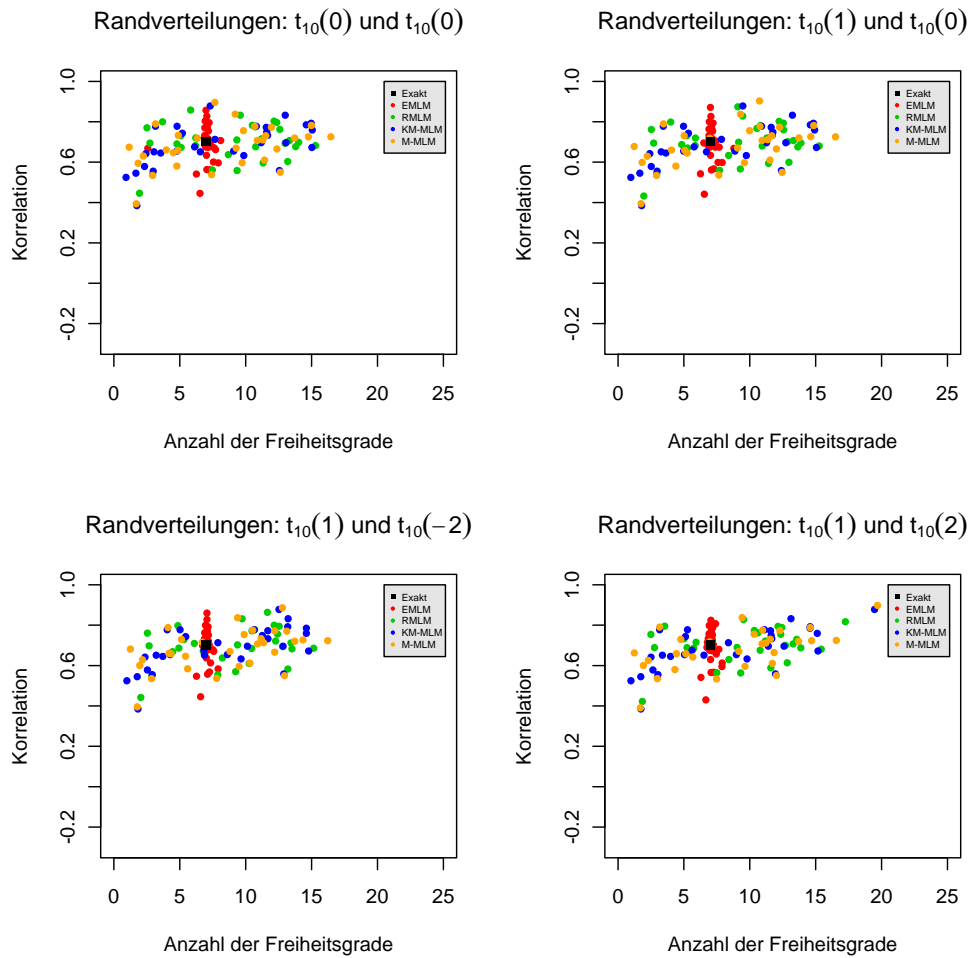


Abbildung 7.3: Parameterschätzungen für $C_{0,7;7}^t$ bei 40 Stichproben und 30 Wiederholungen.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.6919	0.09213	6.999	0.9122
RMLM	0.6953	0.09312	8.862	4.0036
KMMLM	0.6833	0.10511	7.947	4.7126
MMLM	0.6791	0.10272	8.280	4.4763

Tabelle 7.3: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.

7.2 Parametrischer Fall

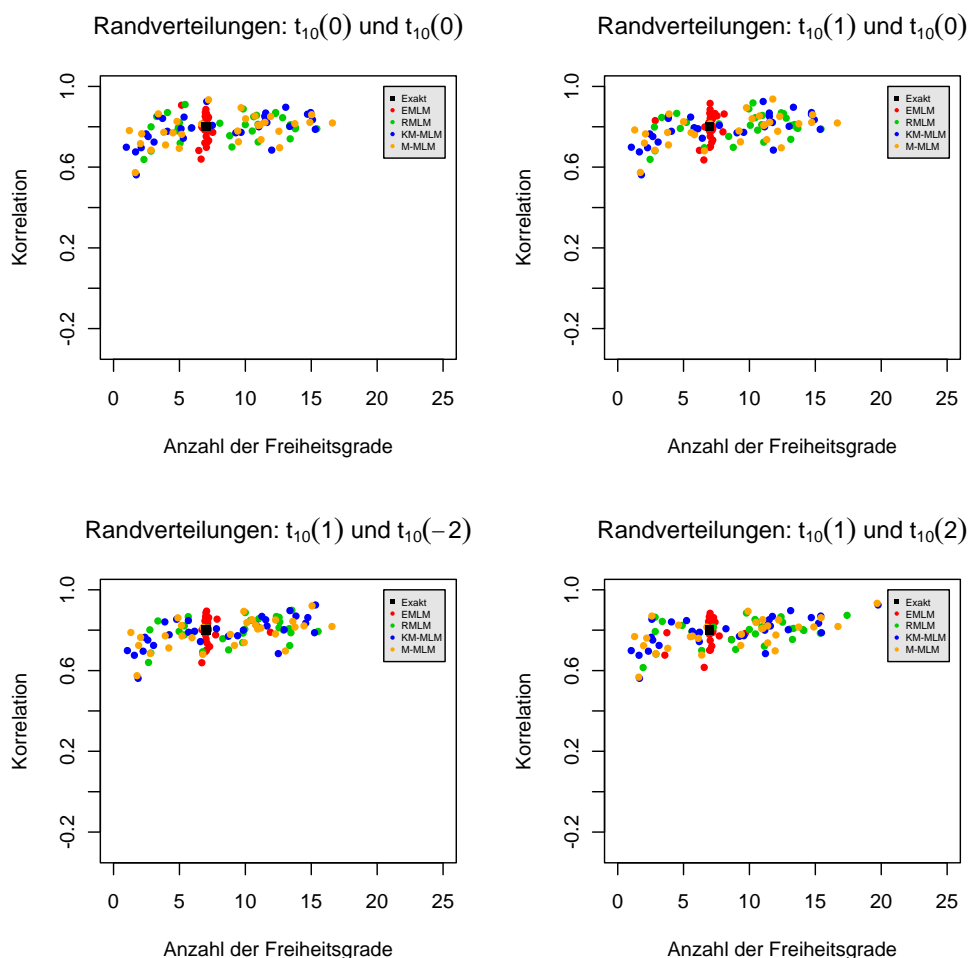


Abbildung 7.4: Parameterschätzungen für $C_{0,8;7}^t$ bei 40 Stichproben und 30 Wiederholungen.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.7956	0.06243	6.969	0.3948
RMLM	0.7984	0.06304	8.895	4.0057
KMMLM	0.7897	0.07641	7.896	4.7759
MMLM	0.7854	0.07315	8.200	4.5468

Tabelle 7.4: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,8;7}^t$ bei 40 Stichproben und 30 Wiederholungen, parametrischer Fall.

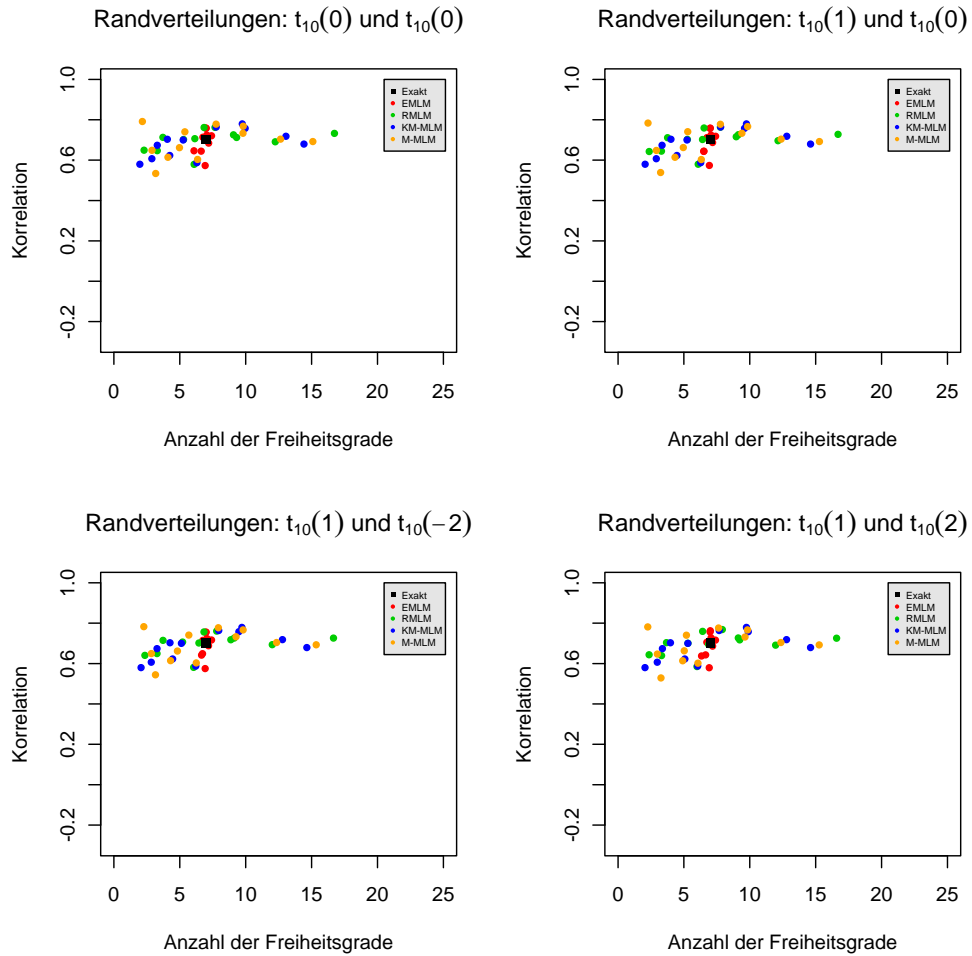


Abbildung 7.5: Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.6951	0.05250	6.920	0.3303
RMLM	0.6981	0.05213	7.398	4.0652
KMMLM	0.6807	0.06910	6.917	4.0930
MMLM	0.6887	0.07908	7.019	4.0924

Tabelle 7.5: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen, parametrischer Fall.

7.3 Nichtparametrischer Fall

Im zweiten Fall werden die Ausgangsdaten erst anhand von nichtparametrisch geschätzten Randverteilungen umgewandelt und dann bei den Schätzmethoden eingesetzt. Die Schätzergebnisse sind in den Abbildungen 7.6 und 7.7 zusammen mit dem parametrischen Fall für $t_{10}(0)$ -verteilte Randverteilungen grafisch dargestellt. Die Durchschnittswerte der Copulaparameterschätzungen werden in den Tabellen 7.6 bis 7.9 präsentiert. Außer der Kendall's τ -Momenten-Maximum-Likelihood-Methode, bei der die Schätzwerte für die Korrelation im parametrischen und nichtparametrischen Fall identisch sind, liefern die anderen Verfahren bei der Schätzung der Korrelation im parametrischen Fall im Durchschnitt minimal bessere Ergebnisse. Auf der anderen Seite ist zu beobachten, dass bei der Schätzung der Anzahl der Freiheitsgrade insbesondere die Mehrschrittverfahren im nichtparametrischen Fall im Durchschnitt minimal besser abschneiden. Auch bei der für die Verfahren berechneten Standardabweichung ist bei keinem der Verfahren und in keinem der beiden Fälle eine deutliche Verbesserung zu erkennen.

In der Abbildung 7.8 und in der Tabelle 7.10 sind die entsprechenden Schätzergebnisse für eine höhere Anzahl von Stichproben dargestellt. Bei den approximativen Werten für die Korrelation wurden zwischen dem parametrischen und nichtparametrischen Fall sowohl in den Durchschnittswerten als auch in den Streuwerten keine großen Unterschiede festgestellt. Die Schätzungen für die Anzahl der Freiheitsgrade sind im Durchschnitt bei dem parametrischen Fall besser, auch die Streuwerte fallen insbesondere bei den Mehrschrittverfahren minimal besser aus.

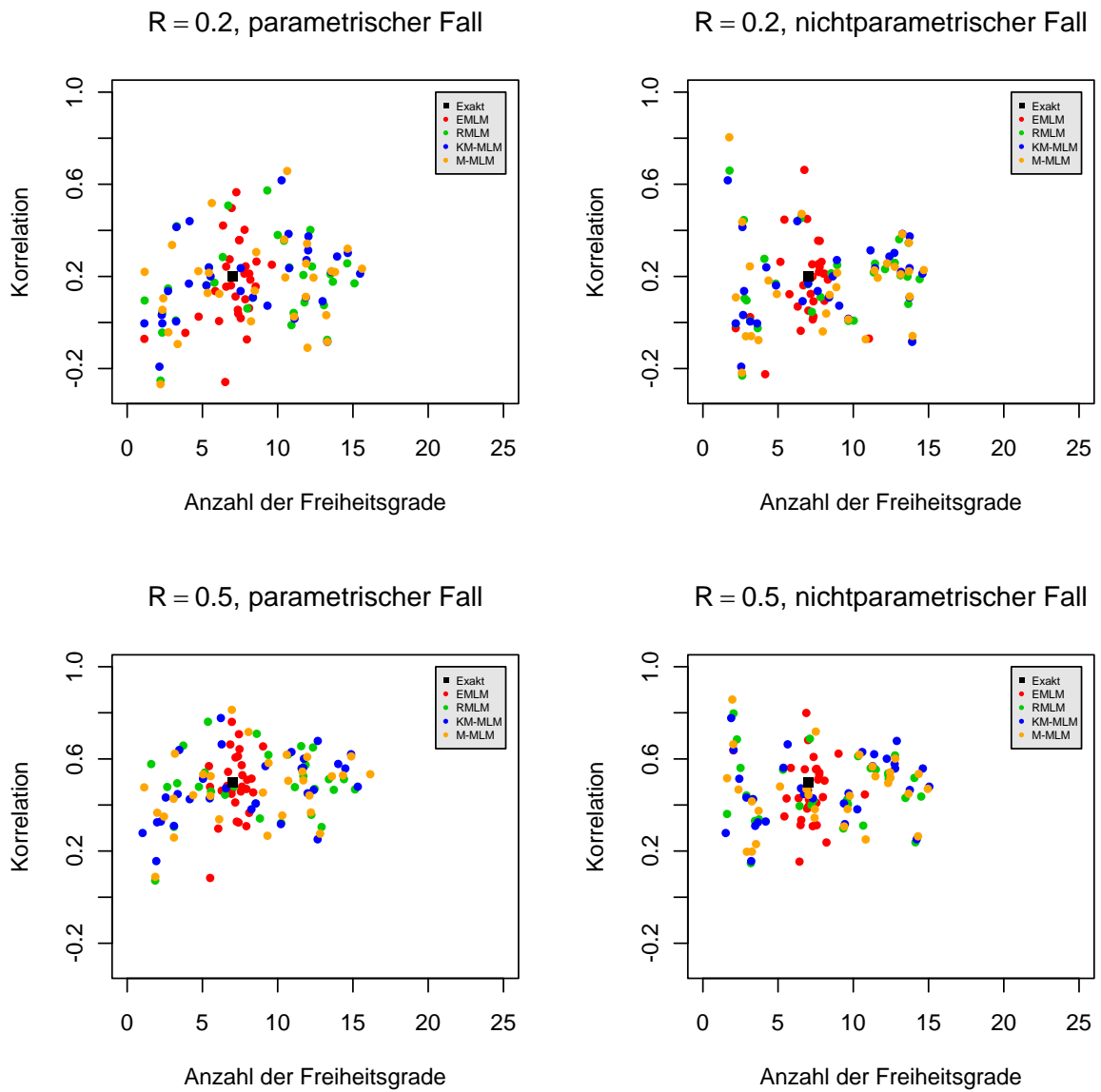


Abbildung 7.6: Parameterschätzungen für $C_{0,2;7}^t$ und $C_{0,5;7}^t$ bei 40 Stichproben und 30 Wiederholungen.

7.3 Nichtparametrischer Fall

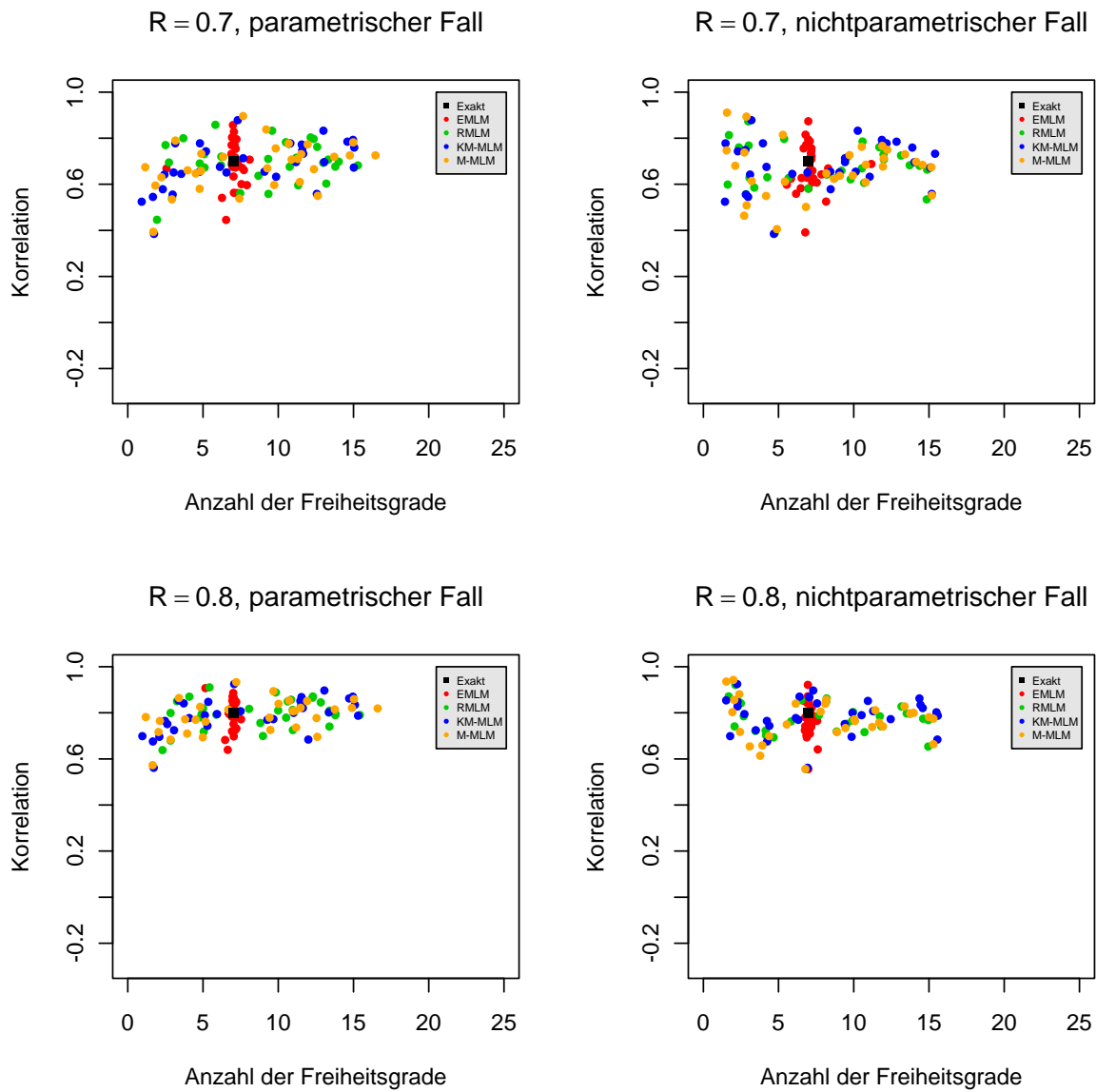


Abbildung 7.7: Parameterschätzungen für $C_{0,7;7}^t$ und $C_{0,8;7}^t$ bei 40 Stichproben und 30 Wiederholungen.

Kapitel 7 Simulationen

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.1718	0.1770	6.896	1.657
RMLM	0.1705	0.1778	8.268	4.340
KMMLM	0.1831	0.1708	8.215	4.396
MMLM	0.1649	0.2032	8.408	4.368

Tabelle 7.6: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,2;7}^t$.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.4635	0.1380	7.091	1.395
RMLM	0.4658	0.1403	8.120	4.269
KMMLM	0.4765	0.1446	8.020	4.473
MMLM	0.4571	0.1531	8.207	4.353

Tabelle 7.7: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,5;7}^t$.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.6697	0.09853	7.191	0.9655
RMLM	0.6718	0.09813	8.136	4.3769
KMMLM	0.6833	0.10511	8.151	4.6018
MMLM	0.6667	0.11510	8.154	4.5325

Tabelle 7.8: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.7719	0.07400	7.045	0.2047
RMLM	0.7743	0.07345	8.066	4.3864
KMMLM	0.7897	0.07641	8.046	4.6642
MMLM	0.7693	0.08812	8.054	4.5888

Tabelle 7.9: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,8;7}^t$.

7.3 Nichtparametrischer Fall

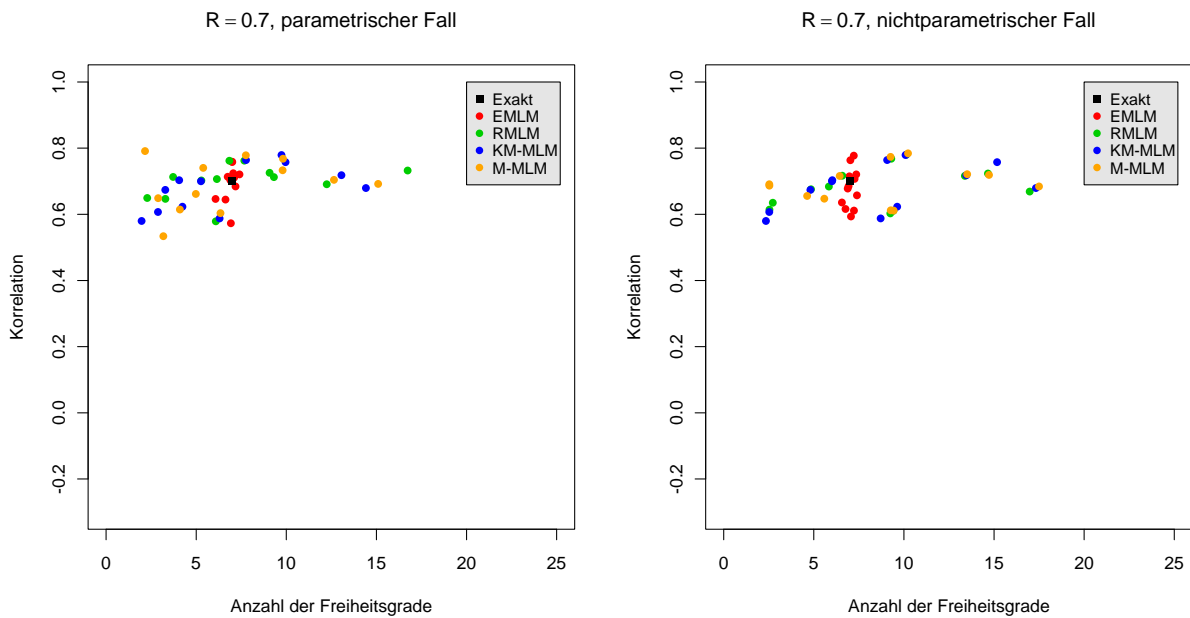


Abbildung 7.8: Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen.

Methoden	Korrelation		Anzahl der Freiheitsgrade	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
EMLM	0.6797	0.05946	7.061	0.2497
RMLM	0.6826	0.05994	8.811	4.5763
KMMLM	0.6807	0.06910	8.772	4.7663
MMLM	0.6914	0.05530	8.810	4.7456

Tabelle 7.10: Mittelwert und Standardabweichung der Parameterschätzungen für $C_{0,7;7}^t$ bei 100 Stichproben und 12 Wiederholungen, nichtparametrischer Fall.

7.4 Fazit

Für den untersuchten Fall lässt sich zusammenfassend sagen, dass, sofern keine entscheidenden Kriterien dagegen sprechen, das Einschrittverfahren den Mehrschrittverfahren vorzuziehen ist. Insbesondere beim Schätzen der Anzahl der Freiheitsgrade gehen offensichtlich durch das Einsetzen von approximativen Werten in die Maximum-Likelihood-Funktion die Informationen verloren, was mitunter zu sehr unbefriedigenden Ergebnissen führen kann. Von den Mehrschrittverfahren kann man anhand dieser Simulationstudie keines favorisieren. Alle drei untersuchten Verfahren haben im Schnitt nur geringfügig voneinander abweichende Schätzergebnisse geliefert.

Es scheint auch keine große Rolle zu spielen, ob die Randverteilungen teilweise bekannt oder vollständig unbekannt sind, denn die Unterschiede im parametrischen und nichtparametrischen Fall waren bei dieser Studie nicht besonders groß. Auch kleine Abweichungen von exakten Werten bei den geschätzten Randverteilungsparametern wirkten sich nicht bedeutend stark auf die Schätzungen der Copulaparameter aus. Hierzu wäre es interessant, die Studie auf eine höhere Dimension auszuweiten, sowie mit anderen als t -verteilten Randverteilungen durchzuführen.

In fast allen Fällen führte die erhöhte Stichprobenzahl im Durchschnitt zu besseren Schätzergebnissen. Jedoch veränderten sich die Streuungswerte nur unbedeutend, so dass man insgesamt bei dieser Erhöhung der Stichprobenzahl von keiner deutlichen Verbesserung sprechen kann. Auch hierzu wäre eine Studie mit mehr als zwei Randverteilungen interessant.

Aus Zeitgründen wurde in dieser Simulationsstudie der Einfluss der Startwerte bei der programmtechnischen Umsetzung nicht untersucht. Dieser Aspekt sollte bei einer möglichen Weiterentwicklung bzw. Verbesserung des Simulationsprogramms mit berücksichtigt werden.

Anhang A

Programmausschnitt

```
library(copula)
library(stats)
library(graphics)

# Definition der Schätzmethoden

# Log-Likelihood-Funktion
ll <- function(rho, nu, dat, n){
  l <- c(1 : n)
  for(i in 1 : n){
    ll[i] <- -dmvt(qt(dat[i, ], nu), sigma = matrix(c(1, rho, rho, 1), 2), df = nu, log = TRUE) -
    log(dt(qt(dat[i, 1], nu), nu) * dt(qt(dat[i, 2], nu), nu))}
  return(-sum(l))}

# Exakte Maximum-Likelihood-Methode
emlm <- function(param.start, dat, n){
  ll_theta <- function(theta, dat, n){
    rho <- -theta[1]
    nu <- -theta[2]
    ll(rho, nu, dat, n)}
  return(optim(param.start, ll_theta, dat = dat, n = n, method = "CG")$par)}
```


Rekursive Maximum-Likelihood-Methode

```
rmlm <- function(param.start, dat, n){
  nu <- -3
  rho.hat <- optim(param.start[1], ll, nu = nu, dat = dat, n = n, method = "BFGS")$par
  repeat{
    nu <- -nu + 1
    rho <- optim(param.start[1], ll, nu = nu, dat = dat, n = n, method = "BFGS")$par
    if(round(ll(rho, nu, dat, n), 3) < round(ll(rho.hat, nu, dat, n), 3))
      (rho.hat <- -rho) else break}
    nu.hat <- optim(param.start[2], ll, rho = rho.hat, dat = dat, n = n, method = "CG")$par
  return(cbind(rho.hat, nu.hat))}
```

Kendall's-tau-Momenten-Maximum-Likelihood-Methode

```
kmmlm <- function(param.start, dat, n){
  rho.hat <- -sin(cor(dat[, 1], dat[, 2], method = "kendall") * pi/2)
  nu.hat <- optim(param.start[2], ll, rho = rho.hat, dat = dat, n = n, method = "CG")$par
  return(cbind(rho.hat, nu.hat))}
```

Momenten-Maximum-Likelihood-Methode

```
mmlm <- function(param.start, dat, n){
  rho.hat.0 <- -cor(qnorm(dat[, 1], mean = 0, sd = 1),
  qnorm(dat[, 2], mean = 0, sd = 1), method = "pearson")
  nu.hat <- optim(param.start[2], ll, rho = rho.hat.0, dat = dat, n = n, method = "CG")$par
  rho.hat <- -cor(qt(dat[, 1], df = nu.hat), qt(dat[, 2], df = nu.hat), method = "pearson")
  return(cbind(rho.hat, nu.hat))}
```

Prozedur zum Schätzen der Parameter der Randverteilungen

```
marg.hat <- function(param.start, dat){
  l <- function(param, dat) sum(-dt(dat, df = param[1], ncp = param[2], log = TRUE))
  param.hat <- optim(param.start, l, dat = dat, method = "CG")$par
  return(param.hat)}
```

Prozedur zum Schätzen der Copulaparameter

```
do_est <- function(myTcopula, p.cop, p.marg, n, wdh, trans, method){
  erg <- matrix(nrow = wdh, ncol = 2)
  for(k in 1:wdh){
    set.seed(10081977 + k)
```

```

.Random.seed
u <- rcopula(myTcopula, n)
x <- matrix(nrow = n, ncol = 2)
for(i in 1:n){
  for(j in 1:2){
    x[i,j] <- -qt(u[i,j], df = p.marg[1,j], ncp = p.marg[2,j])}}
start.p.cop <- -p.cop
start.p.marg <- -p.marg
if(trans == "p"){
  p.marg.hat <- -matrix(c(marg.hat(start.p.marg[,1], x[,1]), marg.hat(start.p.marg[,2], x[,2])), 2)
  p.u <- matrix(nrow = n, ncol = 2)
  for(i in 1:n){
    for(j in 1:2){
      p.u[i,j] <- -pt(x[i,j], df = p.marg.hat[1,j], ncp = p.marg.hat[2,j])}}
  if(method == "emlm"){erg[k,] <- -emlm(start.p.cop, p.u, n)}
  else if(method == "rmlm"){erg[k,] <- -rmlm(start.p.cop, p.u, n)}
  else if(method == "kmmlm"){erg[k,] <- -kmmlm(start.p.cop, p.u, n)}
  else if(method == "mmlm"){erg[k,] <- -mmlm(start.p.cop, p.u, n)}
  }else if(trans == "np"){
    np.u <- -cbind((rank(x[,1]) - 0.5)/n, (rank(x[,2]) - 0.5)/n)
    if(method == "emlm"){erg[k,] <- -emlm(start.p.cop, np.u, n)}
    else if(method == "rmlm"){erg[k,] <- -rmlm(start.p.cop, np.u, n)}
    else if(method == "kmmlm"){erg[k,] <- -kmmlm(start.p.cop, np.u, n)}
    else if(method == "mmlm"){erg[k,] <- -mmlm(start.p.cop, np.u, n)}
  }}
return(erg)}

...

```


Literaturverzeichnis

- Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G., and Roncalli, T. (2000). *Copulas for Finance - A Reading Guide and Some Application*. Working paper, Financial Econometrics Research Centre, City University Business School, London., URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1032533. [Stand: 2009-07-26].
- Demarta, S. and McNeil, A. (2004). *The t Copula and Related Copulas*. Working paper, Department of Mathematics, ETH, Zürich, URL: <http://www.math.ethz.ch/~mcneil/ftp/tCopula.pdf>. [Stand: 2009-10-12].
- Durrleman, V., Nikeghbali, A., and Roncalli, T. (2000). *Which copula is the right one?* Working paper, Groupe de Recherche Opérationnelle, Crédit Lyonnais, Paris, URL: <http://www.risklatte.com/copula/copula003.pdf?func=download&id=77>. [Stand: 2009-07-26].
- Embrechts, P., Lindskog, F., and McNeil, A. (2001). *Modelling Dependence with Copulas and Applications to Risk Management*. Working paper, Department of Mathematics, ETH, Zürich, URL: <http://www.math.ethz.ch/%7Ebaltes/ftp/copchapter.pdf>. [Stand: 2009-02-10].
- Embrechts, P., McNeil, A., and Straumann, D. (1999). *Correlation and dependence in risk management: properties and pitfalls*. Working paper, Department of Mathematics, ETH, Zürich, URL: <http://www.math.ethz.ch/%7Estrauman/preprints/pitfalls.pdf>. [Stand: 2009-02-10].
- Ender, P. (2005). *Ein Anpassungstest für Kopulafunktionen*. PhD thesis, Fakultät für Mathematik, Universität Karlsruhe, URL: <http://digbib.ubka.uni-karlsruhe.de/volltexte/1000003403>. [Stand: 2009-08-05].
- Fang, K.-T., Kotz, S., and Ng, K.-W. (1989). *Symmetric Multivariate and Related Distributions*, volume 36 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

- Fantazzini, D., Giudici, P., and Valle, L. D. (2006). *Copulae and Operational Risks*. Working paper, Department of Economics and Quantitative Methods, University of Pavia, Italy, URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=944174. [Stand: 2009-11-26].
- Feller, W. (1968). *An introduction to probability theory and its applications*, volume 1. John Wiley & Sons, New York.
- Frahm, G. (2004). *Generalized Elliptical Distributions: Theory and Applications*. PhD thesis, Wirtschafts- und Sozialwissenschaftliche Fakultät, Universität zu Köln, URL: http://deposit.d-nb.de/cgi-bin/dokserv?idn=973169869&dok_var=d1&dok_ext=pdf&filename=973169869.pdf. [Stand: 2009-08-27].
- Gibbons, J. D. (1993). *Nonparametric Measures Of Associations*. Sage Publications, Newbury Park.
- Gibbons, J. D. (2003). *Nonparametric Statistical Inference*. Marcel Dekker, New York, 4. edition.
- Glauser, M. (2003). *Messung von Marktrisiken unter Verwendung von Copulafunktionen. Eine Studie für den Schweizer Aktienmarkt*. PhD thesis, Wirtschafts- und Sozialwissenschaftliche Fakultät, Universität Freiburg, Schweiz, URL: <http://ethesis.unifr.ch/theses/downloads.php?file=GlauserM.pdf>. [Stand: 2009-03-10].
- Joe, H. and Xu, J. (1996). *The Estimation Method of Inference Functions for Margins for Multivariate Models*. Working paper, Department of Statistics, University of British Columbia, URL: <http://www.stat.ubc.ca/~harry/ifm.pdf>. [Stand: 2009-07-26].
- Kaiser, J. (2009). *Einige Eigenschaften gestutzter bivariater t-Verteilungen*. PhD thesis, Fakultät für Statistik, Technische Universität Dortmund, URL: <http://dspace.hrz.uni-dortmund.de:8080/bitstream/2003/26178/1/DissertationJKaiser.pdf>. [Stand: 2009-08-27].
- Lee, A. J. (1990). *U-Statistics. Theory and Practice*. Marcel Dekker, New York.
- Lee, M. (1996). *Method of Moments and Semiparametric Econometrics for Limited Dependent Variable Models*. Springer, New York.
- Lindskog, F. (2000). *Modelling Dependence with Copulas and Applications to Risk Management*. PhD thesis, ETH, Zürich, URL:

Literaturverzeichnis

- http://www.math.mcmaster.ca/~zhangx/resources/2000_masterthesis_lindskog.pdf.
[Stand: 2009-05-12].
- Lindskog, F., McNeil, A., and Schmock, U. (2001). *Kendall's tau for elliptical distributions*. Working paper, Department of Mathematics, ETH, Zürich, URL: <http://www.ma.hw.ac.uk/~mcneil/ftp/KendallsTau.pdf>. [Stand: 2009-03-30].
- Mittelhammer, R. C. (1996). *Mathematical Statistics for Economics and Business*. Springer, New York.
- Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, New York.
- Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York, 2. edition.
- Niehof, M. (2009). *Modellierung von stochastischen Abhängigkeiten mittels Copula*. Master's thesis, Institut für mathematische Statistik, Westfälische Wilhelms-Universität Münster, URL: <http://wwwmath.uni-muenster.de/statistik/paulsen/Diplomarbeiten/Niehof.pdf>. [Stand: 2010-04-10].
- Pertsch, E. and Lange-Kowal, E. E. (1999). *Langenscheidts Schulwörterbuch Latein. Lateinisch-Deutsch/ Deutsch-Lateinisch*. Langenscheidt, Berlin, 27. edition.
- Pfanzagl, J. (1969). On the measurability and consistency of minimum contrast estimates. *Metrika*, Bd.14(1.): S.249–272.
- Rousseeuw, P. and Molenberghs, G. (1993). Transformation of non positive semi-definite correlation matrices. *Communications in Statistics - Theory and Methods*, Bd.22(4.): S.965–984.
- Schwarz, H. R. (1997). *Numerische Mathematik*. B.G. Teubner, Stuttgart, 3. edition.
- Schweizer, B. (1991). Thirty years of copulas. In Dall'Aglio, G., Kotz, S., and Salinetti, G., editors, *Advances in Probability Distributions with Given Marginals*, volume 67 of *Mathematics and Its Applications*, pages S.13–50. Kluwer Academic Publishers.
- Schweizer, B. and Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics*, Bd.9(4.): S.879–885.
- Witting, H. (1985). *Mathematische Statistik I. Parametrische Verfahren bei festem Stichprobenumfang*. B.G. Teubner, Stuttgart.

Witting, H. and Müller-Funk, U. (1995). *Mathematische Statistik II. Asymptotische Statistik: Parametrische Modelle und nichtparametrische Funktionale*. B.G. Teubner, Stuttgart.

Yan, J. (2006). *Enjoy the Joy of Copulas*. Working paper, Department of Statistics and Actuarial Science, University of Iowa, USA, URL: <http://www.stat.uiowa.edu/techrep/tr365.pdf>. [Stand: 2009-02-10].

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die Diplomarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Potsdam, den 23. April 2010

Olga Kuxhaus