

Stochastik II

Martingal- und Copulatheorie

Vorlesungsskript

Thorsten Dickhaus
Universität Bremen
Wintersemester 2019 / 2020
Version: 14. November 2019

Vorbemerkungen

Das Material zu diesem Skript habe ich im Wesentlichen den Büchern von Gaensler and Stute (1977), Luschgy (2013) und Hall and Heyde (1980), dem Übersichtsartikel von Embrechts et al. (2003) sowie Vorlesungsskripten von Gerhard Dikta und Michael Kolonko entnommen. Sollten sich in den übernommenen Teilen Fehler finden, so bin dafür natürlich ich verantwortlich. Lob und positive Kritik gebührt indes den Original-Autoren.

Für die Manuskripterstellung danke ich Konstantin Schildknecht und Yuriy Kopansky.

Übungsaufgaben und R-Programme zu diesem Kurs stelle ich auf Anfrage gerne zur Verfügung. Einige Referenzen dazu finden sich im Text an den zugehörigen Stellen.

Inhaltsverzeichnis

1	Grundlagen	1
1.1	Bedingte Verteilungen und bedingte Erwartungswerte	1
1.2	Charakteristische Funktion	7
1.3	Konvergenzarten und Folgen von Zufallsvariablen	12
2	Martingaltheorie	25
2.1	Allgemeine Definitionen und Eigenschaften	25
2.2	Stoppszeiten	28
2.3	Stoppsätze	31
2.4	Martingalkonvergenzsätze	37
2.5	Anwendungen der Martingaltheorie	43
2.5.1	Kolmogoroff'sches 0-1 Gesetz	43
2.5.2	Starkes Gesetz der großen Zahlen	44
2.5.3	Die Gleichungen von Wald	46
2.6	Optimales Stoppen (bei endlichem Zeithorizont)	48
3	Copula-Theorie	54
3.1	Das Prinzip der Quantilstransformation	54
3.2	Allgemeine Eigenschaften von Copulae	57
3.3	Abhängigkeitsmaße	65
3.3.1	Der (lineare) Korrelationskoeffizient	65
3.3.2	Konkordanzmaße	68
3.4	Tail-Abhängigkeit	72
3.5	Elliptische Copulae	75
3.6	Archimedische Copulae	84
3.6.1	Eigenschaften bivariater Archimedischer Copulae	84
3.6.2	Kendall's τ für Archimedische Copulae	90
3.6.3	Tailabhängigkeit bivariater Archimedischer Copulae	92
3.7	Marshall-Olkin Copulae	94

Abbildungsverzeichnis	100
Literaturverzeichnis	101

Kapitel 1

Grundlagen

1.1 Bedingte Verteilungen und bedingte Erwartungswerte

Erinnerung 1.1

Seien X und Y reellwertige stetige Zufallsvariablen auf dem selben Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit gemeinsamer Dichtefunktion $f_{(X,Y)} : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ bezüglich λ^2 (Lebesguemaß), so ist

(a) $f_Y(y) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dx$ eine Randdichte von Y .

(b) $f_{Y|X}(y|x) = \frac{f_{(X,Y)}(x,y)}{f_X(x)}$, $x, y \in \mathbb{R}$, eine bedingte Dichte von Y bezüglich X (mit $0/0 = 0$).

(c) Bezeichne \mathcal{B} die σ -Algebra der Borelmengen auf \mathbb{R} . Für $x \in \mathbb{R}$ mit $f_X(x) > 0$ heißt die Mengenfunktion

$$\mathcal{B} \ni B \mapsto \mathbb{P}(Y \in B|X = x) := \int_B f_{Y|X}(y|x)\lambda(dy)$$

bedingte Verteilung von Y bezüglich $X = x$.

(d) Rechenregeln:

(i) $\mathbb{P}(X \in A, Y \in B) = \int_A \mathbb{P}(Y \in B|X = x)f_X(x)\lambda(dx)$.

(ii) $\mathbb{P}(Y \in B) = \int_{-\infty}^{\infty} \mathbb{P}(Y \in B|X = x)f_X(x)dx$.

(iii) $\mathbb{P}((X, Y) \in C) = \int_{-\infty}^{\infty} \mathbb{P}(Y \in C(x)|X = x)f_X(x)dx$

für $C \in \mathcal{B}^2$ und mit $C(x) = \{y \in \mathbb{R} | (x, y) \in C\}$, dem x -Schnitt von C .

(iv) Sind $A, B \in \mathcal{B}$ mit $\mathbb{P}(X \in A) > 0$, dann ist die elementare bedingte Wahrscheinlichkeit von Y bzgl. X definiert durch $\mathbb{P}(Y \in B|X \in A) = \mathbb{P}(X \in A, Y \in B)/\mathbb{P}(X \in A)$.

Definition 1.2

Seien $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$ zwei Messräume. Eine Abbildung $q : \Omega_1 \times \mathcal{A}_2 \rightarrow [0, 1]$ heißt Übergangswahrscheinlichkeitsverteilung (Markov-Kern) von Ω_1 nach Ω_2 (bzw. \mathcal{A}_2): \Leftrightarrow

(i) $A' \mapsto q(x, A')$ ist ein Wahrscheinlichkeitsmaß auf $(\Omega_2, \mathcal{A}_2)$ für alle $x \in \Omega_1$.

(ii) $x \mapsto q(x, A')$ ist $(\mathcal{A}_1, \mathcal{B})$ -messbar für alle $A' \in \mathcal{A}_2$.

Definition und Satz 1.3

Seien $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2$ zwei Messräume. Sei μ ein Wahrscheinlichkeitsmaß auf $(\Omega_1, \mathcal{A}_1)$ und q ein Markov-Kern von Ω_1 nach Ω_2 .

a) Durch die Festlegung

$$\mu \otimes q(A_1 \times A_2) := \int_{A_1} q(x, A_2) \mu(dx), A_i \in \mathcal{A}_i, i = 1, 2$$

wird auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ das Wahrscheinlichkeitsmaß $\mu \otimes q$ definiert.

b) Für $C \in \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt

$$\mu \otimes q(C) = \int_{\Omega_1} q(x, C(x)) \mu(dx).$$

Beweis: Für $C \in \mathcal{A}_1 \otimes \mathcal{A}_2$ schreiben wir kurz $Q(C) := \int_{\Omega_1} q(x, C(x)) \mu(dx)$.

Normierungsbedingung und σ -Additivität von Q (zur Übung) $\Rightarrow Q$ ist ein Wahrscheinlichkeitsmaß auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$. Sei jetzt $A \times B \in \mathcal{A}_1 \times \mathcal{A}_2$ (Kartesisches Produkt!), so rechnen wir nach:

$$Q(A \times B) = \int_{\Omega_1} q(x, (A \times B)(x)) \mu(dx) = \int_{\Omega_1} \mathbf{1}_A(x) q(x, B) \mu(dx) = \int_A q(x, B) \mu(dx).$$

Aus dem Maßeindeutigkeitssatz folgt, da $\mathcal{A}_1 \times \mathcal{A}_2$ ein \cap -stabiles Erzeugendensystem von $\mathcal{A}_1 \otimes \mathcal{A}_2$ ist, dass $Q =: \mu \otimes q$ eindeutig definiert ist. ■

Beispiel 1.4

a) Sei $q(x, B) \equiv \nu(B)$, ν Wahrscheinlichkeitsmaß auf $(\Omega_2, \mathcal{A}_2)$. Dann ergibt sich

$$\mu \otimes q(A_1 \times A_2) = \int_{A_1} q(x, A_2) \mu(dx) = \int_{A_1} \nu(A_2) \mu(dx) = \mu \times \nu(A_1 \times A_2),$$

also das "klassische" Produktmaß (hier zur Unterscheidung mit \times notiert).

b) Seien X und Y stochastisch unabhängige Zufallsvariablen mit Werten in Ω_1 bzw. Ω_2 , dann gilt

$\mathbb{P}_{(X,Y)} = \mathbb{P}_X \times \mathbb{P}_Y$ und mit a) folgt

$$\mathbb{P}_{(X,Y)} = \mathbb{P}_X \times \mathbb{P}_Y = \mathbb{P}_X \otimes \mathbb{P}_Y, \text{ d.h. } q(t, A_2) = \mathbb{P}(Y \in A_2)$$

ist eine Version der bedingten Verteilung $\mathbb{P}_{Y|X=t}$ für alle $t \in \Omega_1$.

Satz 1.5 (Satz von Fubini für Markov-Kerne)

Unter den Bezeichnungen aus Satz 1.3 sei $f : \Omega_1 \times \Omega_2 \rightarrow \overline{\mathbb{R}}$ eine messbare Abbildung. Dann gilt

$$\int_{\Omega_1 \times \Omega_2} f d(\mu \otimes q) = \int_{\Omega_1} \left[\int_{\Omega_2} f(x, y) q(x, dy) \right] \mu(dx),$$

falls eine der folgenden Voraussetzungen gilt:

- (i) $f \geq 0$.
- (ii) f ist $(\mu \otimes q)$ -quasiintegrierbar.

Beweis: Satz 14.29 in Klenke (2008). ■

Anmerkung:

Ist ν ein Wahrscheinlichkeitsmaß auf Ω_2 und $q(x, B) := \nu(B)$, so ergibt sich der klassische Satz von Fubini über Produktmaße:

$$\int_{\Omega_1 \times \Omega_2} f d(\mu \times \nu) = \int_{\Omega_1} \left[\int_{\Omega_2} f d\nu \right] d\mu = \int_{\Omega_2} \left[\int_{\Omega_1} f d\mu \right] d\nu$$

unter den Voraussetzungen von Satz 1.5.

Bemerkung 1.6

Für den Beweis von Satz 1.5 ist es wichtig, dass die Funktion

$$h : \Omega_1 \rightarrow \overline{\mathbb{R}}, x \mapsto h(x) := \int_{\Omega_2} f(x, y) q(x, dy)$$

messbar ist. Dies zeigt man mit algebraischer Induktion (zur Übung).

Definition 1.7

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und seien X, Y Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit Werten in $(\Omega_1, \mathcal{A}_1)$ bzw. $(\Omega_2, \mathcal{A}_2)$. Dann heißt ein Markov-Kern q von Ω_1 nach Ω_2 mit der Eigenschaft

$$\mathbb{P}(X \in A_1, Y \in A_2) = \int_{A_1} q(x, A_2) \mathbb{P}_X(dx)$$

für alle $A_i \in \mathcal{A}_i, i = 1, 2$, eine reguläre Version der bedingten Verteilung von Y bezüglich X .

Kurzform: $\mathbb{P}_{(X, Y)} = \mathbb{P}_X \otimes q$.

Ist $(\Omega_2, \mathcal{A}_2) = (\mathbb{R}^d, \mathcal{B}^d)$ mit $d \in \mathbb{N}$, so existiert stets eine reguläre Version von $\mathbb{P}_{Y|X}$.

Definition 1.8

Unter den Voraussetzungen von Definition 1.7 sei $T : (\Omega_2, \mathcal{A}_2) \rightarrow (\mathbb{R}, \mathcal{B})$ eine messbare Funktion derart, dass $T(Y) \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$ ist.

Dann heißt

$$\mathbb{E}[T(Y)|X = x] := \int T(y) q(x, dy) =: g(x)$$

eine Version des bedingten Erwartungswertes von $T(Y)$ unter der Hypothese $X = x$.

Bemerkung 1.9

Unter den Voraussetzungen von Definition 1.8 gilt:

- (i) Es existiert stets eine Version von $x \mapsto \mathbb{E}[T(Y)|X = x]$.
- (ii) Alle Versionen von $x \mapsto \mathbb{E}[T(Y)|X = x]$ sind messbare und \mathbb{P}_X -integrierbare Abbildungen $g : \Omega_1 \rightarrow \mathbb{R}$.

Definition und Satz 1.10

Es seien die Voraussetzungen von Definition 1.8 mit $T = id$. gegeben.

a) Die Zufallsvariable $\mathbb{E}[Y|X] := g(X) = g \circ X$, die für $X(\omega) = x$ den Wert $g(x)$ mit $g(x) = \mathbb{E}[Y|X = x] = \int yq(x, dy)$ annimmt, heißt (eine) bedingte Erwartung von Y bezüglich X .

b) Bezeichne

$$\sigma(X) = X^{-1}(\mathcal{A}_1) = \{X^{-1}(B) | B \in \mathcal{A}_1\} = \{A \in \mathcal{F} | \exists B \in \mathcal{A}_1 : X^{-1}(B) = A\}$$

die von $X : (\Omega, \mathcal{F}) \rightarrow (\Omega_1, \mathcal{A}_1)$ erzeugte Unter- σ -Algebra von \mathcal{F} .

Dann gilt für $A \in \sigma(X)$ und mit $B \in \mathcal{A}_1$ so, dass $X^{-1}(B) = A$ ist, dass

$$\begin{aligned} \int_A Y d\mathbb{P} &= \int_{\Omega} \mathbf{1}_B(X) Y d\mathbb{P} = \int_{\Omega_1 \times \mathbb{R}} \mathbf{1}_B(x) y d\mathbb{P}_{(X,Y)}(x, y) \\ &= \int_{\Omega_1} \mathbf{1}_B(x) \left[\int_{\mathbb{R}} yq(x, dy) \right] \mathbb{P}_X(dx) = \int_{\Omega_1} \mathbf{1}_B(x) g(x) \mathbb{P}_X(dx) \\ &= \int_B g(x) \mathbb{P}_X(dx) = \int_{\Omega} \mathbf{1}_B(X) g(X) d\mathbb{P} = \int_A g \circ X d\mathbb{P} \\ &= \int_A \mathbb{E}[Y|X] d\mathbb{P}. \end{aligned}$$

c) Sei allgemein \mathcal{C} eine Sub- σ -Algebra von \mathcal{F} . Dann ist eine bedingte Erwartung $Z \in \mathbb{E}[Y|\mathcal{C}]$ (Schreibweise $Z = \mathbb{E}[Y|\mathcal{C}]$) charakterisiert durch

- (i) Z ist $(\mathcal{C}, \mathcal{B})$ -messbar.
- (ii) $\forall C \in \mathcal{C} : \int_C Z d\mathbb{P} = \int_C Y d\mathbb{P}$.

Formal kann jedes solche $\mathcal{C} \subseteq \mathcal{F}$ als $\sigma(X)$ für ein geeignetes X geschrieben werden.

Beispiel 1.11

a) Zeichen werden in einem Übertragungskanal mit einer unbekanntem Wahrscheinlichkeit gestört. Die unbekanntem Störwahrscheinlichkeit wird als Zufallsvariable X mit Werten in $(0, 1)$ modelliert. Bei gegebenem $X = p$ sollen die Störungen laut Modell iid auftreten.

Es sei $Y :=$ "Wartezeit bis zur ersten Störung", gemessen in Anzahl gesendeter Zeichen. Gesucht ist nun die mittlere "Zeit" bis zur ersten Störung, falls $X = p$ bekannt ist.

Lösung: Eine Version von $\mathbb{P}_{Y|X=p}$ ist die geometrische Verteilung mit Parameter p , also $\mathbb{P}(Y = k|X = p) = p(1 - p)^k, k \geq 0$.

$$\Rightarrow \mathbb{E}[Y|X = p] = \sum_{k=0}^{\infty} kp(1 - p)^k = \frac{1 - p}{p} = g(p).$$

Gelte nun für die Störwahrscheinlichkeit X , dass $\mathbb{P}(X = \frac{1}{2}) =: a$ und $\mathbb{P}(X = \frac{3}{4}) = 1 - a$, so folgt

$$\mathbb{E}[Y|X] = \frac{1 - X}{X} =: Z$$

mit $\mathbb{P}(Z = 1) = a = 1 - \mathbb{P}(Z = \frac{1}{3})$.

b) Sei Y reellwertige, integrierbare Zufallsvariable und X diskret mit Werten in \mathbb{N}_0 . Dann kann $g(i) := \mathbb{E}[Y|X = i], i \in \mathbb{N}_0$, wie folgt bestimmt werden. Nach elementarer bedingter Wahrscheinlichkeitsformel gilt:

$$\begin{aligned} \mathbb{P}(Y \in B|X = i) &= \frac{\mathbb{P}(Y \in B, X = i)}{\mathbb{P}(X = i)} \\ &= [\mathbb{P}(X = i)]^{-1} \int \mathbf{1}_{\{Y \in B\}} \mathbf{1}_{\{X=i\}} d\mathbb{P} \\ \Rightarrow g(i) &= \frac{\mathbb{E}[Y \mathbf{1}_{\{X=i\}}]}{\mathbb{P}(X = i)}, i \in \mathbb{N}_0. \end{aligned}$$

Ausführliche Verifikation mittels charakterisierender Integralgleichung ist eine Übungsaufgabe. Beispielsweise gilt z.B. für $X := \lfloor Y \rfloor$, dass

$$\mathbb{E}[Y|X = i] = \frac{\mathbb{E}[Y \mathbf{1}_{\{i \leq Y < i+1\}}]}{\mathbb{P}(i \leq Y < i+1)} = g(i), i \in \mathbb{N}_0.$$

Bemerkung 1.12 (Anschauliche Interpretation von $\mathbb{E}[Y|X]$)

Sei $Z := \mathbb{E}[Y|X]$ (genauer sei $Z \in \mathbb{E}[Y|X]$). Dann hat Z die folgenden Eigenschaften:

- (i) Z ist auf dem selben Wahrscheinlichkeitsraum wie Y definiert.
- (ii) Der Mittelwert von Z stimmt mit dem von Y überein, wenn auf Mengen $X^{-1}(B)$ eingeschränkt wird.
- (iii) Wegen $Z = g(X)$ "variiert" Z aber nur so stark wie X . Nimmt X also z.B. nur endlich viele Werte an, so auch $Z = \mathbb{E}[Y|X]$. Die bedingte Erwartung ist also gewissermaßen eine Glättung von Y entlang X .

(iv) Bild zur Veranschaulichung:

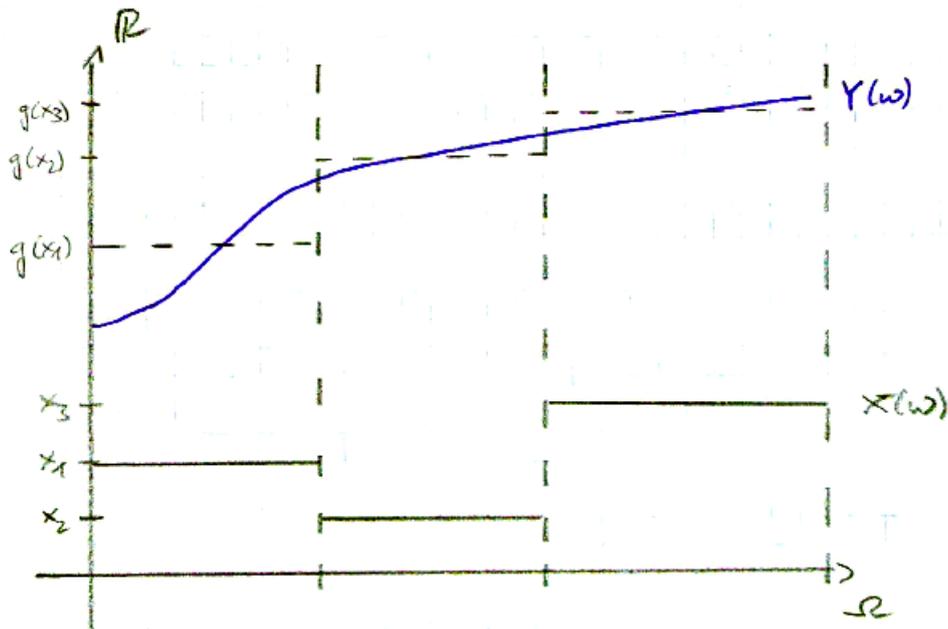


Abbildung 1.1: Skizze zur Veranschaulichung der bedingten Erwartung

(v) Liegt Y in $\mathcal{L}_2(\Omega, \mathcal{F}, \mathbb{P})$, so stellt $\mathbb{E}[Y|X]$ die beste L_2 -Approximation von Y unter allen Funktionen der Gestalt $h(X)$, $h: \Omega_1 \rightarrow \mathbb{R}$, dar, d.h. der L_2 -Abstand zwischen Y und einer (deterministischen) L_2 -Transformation von X ist am kleinsten für $\mathbb{E}[Y|X]$. Anders ausgedrückt ist $\mathbb{E}[Y|X]$ die Projektion von Y auf $\mathcal{L}_2(\Omega, \sigma(X), \mathbb{P})$.

Wir beschließen diesen Abschnitt 1.1 mit wichtigen Rechenregeln für bedingte Erwartungen.

Satz 1.13 (Rechenregeln für bedingte Erwartungen, alle Aussagen \mathbb{P} -f.s.)

Unter den Voraussetzungen von Definition 1.8 gelten die folgenden Rechenregeln.

a) Linearität der bedingten Erwartung:

$$\mathbb{E}[\alpha Y_1 + \beta Y_2 | X] = \alpha \mathbb{E}[Y_1 | X] + \beta \mathbb{E}[Y_2 | X].$$

b) Satz von der iterierten Erwartungswertbildung:

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]] = \int_{\Omega_1} \mathbb{E}[Y|X=x] \mathbb{P}_X(dx).$$

c) Sei $h: \Omega_1 \times \mathbb{R} \rightarrow \mathbb{R}$, so dass $h(X, Y)$ integrierbar ist, so folgt:

$$(i) \mathbb{E}[h(X, Y) | X = x] = \mathbb{E}[h(x, Y) | X = x] = \int h(x, y) \mathbb{P}_{Y|X=x}(dy).$$

$$(ii) X \perp Y \Rightarrow \mathbb{E}[h(X, Y) | X = x] = \mathbb{E}[h(x, Y)] = \int h(x, y) \mathbb{P}_Y(dy).$$

d) Sei $h : \Omega_1 \rightarrow \mathbb{R}$ messbar, so dass $Y \cdot h(X)$ integrierbar ist, so folgt:

$$\mathbb{E}[Y \cdot h(X)|X] = h(X) \cdot \mathbb{E}[Y|X].$$

e) Sei $g : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega', \mathcal{A}')$, so folgt:

$$\mathbb{E}[\mathbb{E}[Y|X] | g(X)] = \mathbb{E}[Y|g(X)] = \mathbb{E}[\mathbb{E}[Y|g(X)] | X].$$

f) Tower equation: Sind $\mathcal{B}_1 \subset \mathcal{B}_2$ Sub- σ -Algebren von \mathcal{F} und ist $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbb{P})$, so gilt \mathbb{P} -f.s.

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{B}_1] | \mathcal{B}_2] = \mathbb{E}[Y|\mathcal{B}_1] = \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_2] | \mathcal{B}_1].$$

Beachte: σ -Algebren können als Informationsstände interpretiert werden!

Beweis: Alle Aussagen folgen direkt aus Eigenschaften des Lebesgue-Integrals (vgl. Maß- und Integrationstheorie) oder können mit algebraischer Induktion nachgewiesen werden (man verifiziere z.B. Teil c) für Indikatorfunktionen). ■

1.2 Charakteristische Funktion

Geht es in der Wahrscheinlichkeitstheorie um die Berechnung von Momenten von Zufallsgrößen und / oder von Faltungen von Verteilungen, so sind charakteristische Funktionen zentrale Objekte, mit deren Hilfe diese Berechnung sehr einfach erfolgen kann.

Bezeichne dazu in diesem Abschnitt $i = \sqrt{-1}$ die imaginäre Einheit.

Definition 1.14

a) Sei μ ein endliches Maß auf \mathbb{R}^d für $d \in \mathbb{N}$. Die Abbildung $\varphi_\mu : \mathbb{R}^d \rightarrow \mathbb{C}$, definiert durch

$$\varphi_\mu(\mathbf{t}) := \int \exp(i\langle \mathbf{t}, \mathbf{x} \rangle) \mu(d\mathbf{x})$$

heißt Fourier - Transformierte von μ .

b) Sei $\mathbf{X} = (X_1, \dots, X_d)^\top$ ein Zufallsvektor mit (gemeinsamer) Verteilung $\mathbb{P}_{\mathbf{X}}$. Dann heißt $\varphi_{\mathbf{X}} := \varphi_{\mathbb{P}_{\mathbf{X}}}$ die charakteristische Funktion von \mathbf{X} .

c) Für eine komplexwertige Zufallsvariable Z mit Real- und Imaginärteilen $\operatorname{Re}(Z)$ und $\operatorname{Im}(Z)$ sei $\mathbb{E}[Z] := \mathbb{E}[\operatorname{Re}(Z)] + i\mathbb{E}[\operatorname{Im}(Z)]$, falls die Erwartungswerte von $\operatorname{Re}(Z)$ und $\operatorname{Im}(Z)$ (jeweils) existieren. Damit ist

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}[\exp(i\langle \mathbf{t}, \mathbf{X} \rangle)], \mathbf{t} \in \mathbb{R}^d.$$

Man beachte dabei die Euler'sche Formel $\exp(i\vartheta) = \cos(\vartheta) + i \sin(\vartheta)$.

Wegen $|\exp(i\langle \mathbf{t}, \mathbf{x} \rangle)| = 1$ für alle $\mathbf{t}, \mathbf{x} \in \mathbb{R}^d$ existiert die charakteristische Funktion für alle $\mathbf{t} \in \mathbb{R}^d$.

Satz 1.15 (Eigenschaften der charakteristischen Funktion)

a) $\forall \mathbf{t} \in \mathbb{R}^d : |\varphi_{\mathbf{X}}(\mathbf{t})| \leq 1 = \varphi_{\mathbf{X}}(\mathbf{0})$.

b) Affine Transformationen: Sei \mathbf{X} eine Zufallsgröße mit Werten in \mathbb{R}^d und $\mathbf{Y} := A\mathbf{X} + \mathbf{b}$ mit $A \in \mathbb{R}^{m \times d}$ und $\mathbf{b} \in \mathbb{R}^m$, wobei $d, m \in \mathbb{N}$. Dann gilt $\varphi_{\mathbf{Y}}(\mathbf{u}) = \exp(i\langle \mathbf{u}, \mathbf{b} \rangle) \varphi_{\mathbf{X}}(A^\top \mathbf{u})$, $\mathbf{u} \in \mathbb{R}^m$. Ist speziell $d = m = 1$ und $A = a = -1$, $b = 0$, so ergibt sich z. B.

$$\varphi_{-X}(u) = \varphi_X(-u) = \overline{\varphi_X(u)}$$

aufgrund der Symmetrieeigenschaften von Sinus und Cosinus.

c) $\mathbb{P}_X = \mathbb{P}_{-X}$ genau dann, wenn φ_X (rein) reellwertig ist.

d) Die Zufallsvariablen X_1, \dots, X_d sind genau dann stochastisch unabhängig, wenn $\forall \mathbf{u} = (u_1, \dots, u_d)^\top \in \mathbb{R}^d : \varphi_{\mathbf{X}}(\mathbf{u}) = \prod_{k=1}^d \varphi_{X_k}(u_k)$ gilt, $\mathbf{X} = (X_1, \dots, X_d)^\top$.

e) Faltungsformel: Sind \mathbf{X} und \mathbf{Y} stochastisch unabhängige Zufallsvektoren mit Werten in \mathbb{R}^d , so ist $\varphi_{\mathbf{X}+\mathbf{Y}} = \varphi_{\mathbf{X}} \cdot \varphi_{\mathbf{Y}}$.

Beweis:

zu a): $\mathbb{P}_{\mathbf{X}}(\mathbb{R}^d) = 1$.

zu b): Zur Übung (Lineare Algebra).

zu c): Folgt aus den Symmetrieeigenschaften von Sinus und Cosinus.

zu d): Folgt aus der Charakterisierung der stochastischen Unabhängigkeit über

$$\mathbb{E}[f(X_i)g(X_j)] = \mathbb{E}[f(X_i)] \mathbb{E}[g(X_j)]$$

für alle komplexwertigen, messbaren Funktionen f und g , Details z.B. in Kapitel 8 von Breiman (1992).

zu e):

$$\begin{aligned} \varphi_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) &= \mathbb{E}[\exp(i\langle \mathbf{t}, \mathbf{X} + \mathbf{Y} \rangle)] \\ &= \mathbb{E}[\exp(i\langle \mathbf{t}, \mathbf{X} \rangle + i\langle \mathbf{t}, \mathbf{Y} \rangle)] \\ &= \mathbb{E}[\exp(i\langle \mathbf{t}, \mathbf{X} \rangle) \exp(i\langle \mathbf{t}, \mathbf{Y} \rangle)] \\ &\stackrel{\text{stoch. Unabhängigkeit}}{=} \mathbb{E}[\exp(i\langle \mathbf{t}, \mathbf{X} \rangle)] \mathbb{E}[\exp(i\langle \mathbf{t}, \mathbf{Y} \rangle)]. \end{aligned}$$



Es existieren eine ganze Reihe von “Umkehrformeln”, die es erlauben, Verteilungsfunktionen, Dichtefunktionen oder Wahrscheinlichkeitsfunktionen aus charakteristischen Funktionen zurückzugewinnen.

Satz 1.16

a) Diskrete Fourier-Inversionsformel:

Sei μ endliches Maß auf $\mathbb{Z}^d \Rightarrow \forall \mathbf{x} \in \mathbb{Z}^d$ gilt:

(i)

$$\mu(\{\mathbf{x}\}) = (2\pi)^{-d} \int_{[-\pi, \pi]^d} \exp(-i\langle \mathbf{t}, \mathbf{x} \rangle) \varphi_\mu(\mathbf{t}) d\mathbf{t},$$

(ii)

$$\sum_{\mathbf{x} \in \mathbb{Z}^d} \mu(\{\mathbf{x}\})^2 = (2\pi)^{-d} \int_{[-\pi, \pi]^d} |\varphi_\mu(\mathbf{t})|^2 d\mathbf{t} \quad (\text{Plancherel}).$$

b) Besitzt μ eine λ^d -Dichte f , so gilt

$$f(\mathbf{x}) = (2\pi)^{-d} \int_{\mathbb{R}^d} \exp(-i\langle \mathbf{t}, \mathbf{x} \rangle) \varphi_\mu(\mathbf{t}) \lambda^d(d\mathbf{t}), \quad \mathbf{x} \in \mathbb{R}^d.$$

c) In Dimension $d = 1$ gilt

$$F_X(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\text{Im}(e^{-itx} \varphi_X(t))}{t} dt$$

für alle Stetigkeitspunkte x von F_X .

d) Chungs Inversionsformel (hier nur $d = 1$):

Falls $a < b$ und $\mathbb{P}(X = a) = \mathbb{P}(X = b) = 0$, so folgt

$$F_X(b) - F_X(a) = \lim_{T \rightarrow \infty} \left\{ \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t) dt \right\}.$$

Beweis:

zu a): Klenke (2008), Seiten 300-301.

zu b): Klenke (2008), Seiten 300-301.

zu c): Gil-Pelaez (1951).

zu d): Chung (2000).



Korollar 1.17 (Eindeutigkeitsatz)

Ein endliches Maß μ auf \mathbb{R}^d ist durch Angabe der charakteristischen Funktion φ_μ eindeutig festgelegt.

Satz 1.18 (Momentenberechnung)

Sei $\mathbf{X} = (X_1, \dots, X_d)^\top$ ein Zufallsvektor mit Werten in \mathbb{R}^d . Falls $\mathbb{E}[|\mathbf{X}|^m]$ für $m \in \mathbb{N}$ endlich ist, dann ist $\varphi_{\mathbf{X}}$ m -mal stetig partiell differenzierbar und es gilt für alle $\mathbf{t} \in \mathbb{R}^d$, dass

$$\frac{\partial^m}{\partial t_{j_1} \partial t_{j_2} \dots \partial t_{j_m}} \varphi_{\mathbf{X}}(\mathbf{t}) = i^m \mathbb{E}[X_{j_1} X_{j_2} \dots X_{j_m} \exp(i\langle \mathbf{t}, \mathbf{X} \rangle)].$$

Beweisskizze:

Nach Definition ist

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \int_{\mathbb{R}^d} \exp(i\langle \mathbf{t}, \mathbf{x} \rangle) \mathbb{P}_{\mathbf{X}}(d\mathbf{x}).$$

Es ist nun

$$\frac{\partial}{\partial t_{j_1}} \exp(i\langle \mathbf{t}, \mathbf{x} \rangle) = i x_{j_1} \exp(i\langle \mathbf{t}, \mathbf{x} \rangle) \quad (1.1)$$

sowie

$$\int_{\mathbb{R}^d} i x_{j_1} \exp(i\langle \mathbf{t}, \mathbf{x} \rangle) \mathbb{P}_{\mathbf{X}}(d\mathbf{x}) = i \mathbb{E}[X_{j_1} \exp(i\langle \mathbf{t}, \mathbf{X} \rangle)]. \quad (1.2)$$

Man zeigt nun (siehe, z. B., Jacod and Protter (2000), Theorem 13.2), dass Differenzieren unter dem Integral unter den gemachten Annahmen zulässig ist, und wendet (1.1) und (1.2) m -mal hintereinander an. ■

Beispiel 1.19 (Normalverteilungen)

a) Sei $X \sim \mathcal{N}(0, 1)$ im \mathbb{R}^1 . Dann

$$\begin{aligned} \varphi_X(t) = \mathbb{E}[\exp(itX)] &= \int_{\mathbb{R}} \cos(tx) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\ &\quad + i \underbrace{\int_{\mathbb{R}} \sin(tx) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx}_{=0, \text{ da Integrand ungerade Funktion}}. \end{aligned}$$

Folglich ergibt sich für die Ableitung, dass

$$\begin{aligned} \Rightarrow \varphi'_X(t) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} -x \sin(tx) \exp\left(-\frac{x^2}{2}\right) dx \\ &= -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} t \cos(tx) \exp\left(-\frac{x^2}{2}\right) dx \\ &= -t \varphi_X(t), \end{aligned}$$

wobei wir im mittleren Schritt partiell integriert haben mit $u'(x) = -x \exp(-x^2/2)$ und $v(x) = \sin(tx)$. Also ergibt sich insgesamt

$$\frac{\varphi'_X(t)}{\varphi_X(t)} = -t \quad \Rightarrow \quad \ln(\varphi_X(t)) = -\frac{t^2}{2} + C \quad \Rightarrow \quad \varphi_X(t) = \exp\left(-\frac{t^2}{2}\right) \exp(C).$$

Wegen $\varphi_X(0) = 1$ ist $C = 0$, also folgt schließlich $\varphi_X(t) = \exp(-t^2/2)$.

b) Sei $Y \sim \mathcal{N}(\mu, \sigma^2)$ im \mathbb{R}^1 . Dann ist $Y \stackrel{\mathcal{D}}{=} \sigma X + \mu$ mit $X \sim \mathcal{N}(0, 1)$. Damit gilt nach Satz 1.15.b), dass

$$\varphi_Y(t) = \exp(it\mu) \exp\left(-\frac{\sigma^2 t^2}{2}\right) = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right).$$

c) Sei $\mathbf{X} = (X_1, \dots, X_d)^\top$ standardnormalverteilt im \mathbb{R}^d . Dann liefert Satz 1.15.d), dass

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \prod_{k=1}^d \exp\left(-\frac{t_k^2}{2}\right) = \exp\left(-\frac{1}{2}|\mathbf{t}|^2\right).$$

d) Sei $\mathbf{Y} = (Y_1, \dots, Y_m)^\top$ allgemein normalverteilt, $\mathbf{Y} \sim \mathcal{N}_m(\mu, \Sigma)$.

Dann lässt sich $\Sigma = QQ^\top$ zerlegen und $\mathbf{Y} = Q\mathbf{X} + \mu$ schreiben, wobei \mathbf{X} standardnormalverteilt ist. Somit gilt

$$\begin{aligned} \varphi_{\mathbf{Y}}(\mathbf{u}) &= \exp(i\langle \mathbf{u}, \mu \rangle) \exp\left(-\frac{1}{2}|Q^\top \mathbf{u}|^2\right) = \exp(i\langle \mathbf{u}, \mu \rangle) \exp\left(-\frac{1}{2}\langle Q^\top \mathbf{u}, Q^\top \mathbf{u} \rangle\right) \\ &= \exp(i\langle \mathbf{u}, \mu \rangle) \exp\left(-\frac{1}{2}(Q^\top \mathbf{u})^\top Q^\top \mathbf{u}\right) = \exp(i\langle \mathbf{u}, \mu \rangle) \exp\left(-\frac{1}{2}\mathbf{u}^\top Q Q^\top \mathbf{u}\right) \\ &= \exp(i\langle \mathbf{u}, \mu \rangle) \exp\left(-\frac{1}{2}\mathbf{u}^\top \Sigma \mathbf{u}\right) = \exp\left(i\langle \mathbf{u}, \mu \rangle - \frac{1}{2}\mathbf{u}^\top \Sigma \mathbf{u}\right) \\ &= \exp\left(i\langle \mathbf{u}, \mu \rangle - \frac{1}{2}\langle \mathbf{u}, \Sigma \mathbf{u} \rangle\right). \end{aligned}$$

Beispiel 1.20 (Weitere Beispiele (in $d = 1$))

a) Binomialverteilung: Sei $X \sim \text{Bin}(n, p)$, so gilt:

$$\begin{aligned} \varphi_X(t) &= \sum_{k=0}^n \exp(itk) p^k (1-p)^{n-k} \binom{n}{k} \\ &= \sum_{k=0}^n [\exp(it)p]^k (1-p)^{n-k} \binom{n}{k} \\ &\stackrel{\text{bin. Lehrsatz}}{=} [p \exp(it) + (1-p)]^n. \end{aligned}$$

b) *Gammaverteilung: Sei $Y \sim \text{Gamma}(1, r)$, so gilt:*

$$\begin{aligned}\varphi_Y(t) &= \int_0^\infty \exp(it y) \frac{y^{r-1}}{\Gamma(r)} e^{-y} dy \\ &= \int_0^\infty \frac{y^{r-1}}{\Gamma(r)} \exp(-y(1-it)) dy \\ &= (1-it)^{-r} \int_0^\infty \frac{(1-it)^r}{\Gamma(r)} y^{r-1} \exp(-y(1-it)) dy \\ &= (1-it)^{-r} \text{ wegen Normierungsbedingung von "Gamma}(1-it, r)\text{".}\end{aligned}$$

Sei nun $X \sim \text{Gamma}(\alpha, r)$, so gilt $X \stackrel{D}{=} Y/\alpha$ und damit

$$\varphi_X(t) = \varphi_Y\left(\frac{t}{\alpha}\right) = \left(1 - \frac{it}{\alpha}\right)^{-r} = \left(\frac{\alpha - it}{\alpha}\right)^r.$$

c) *Sei $X \sim \text{UNI}[a, b]$ (stetige Gleichverteilung auf dem Intervall $[a, b]$). Dann ist*

$$\begin{aligned}\varphi_X(t) &= \int_a^b \frac{\exp(itx)}{(b-a)} dx = \left[(it(b-a))^{-1} \exp(itx) \right]_a^b \\ &= \frac{\exp(itb) - \exp(ita)}{it(b-a)} \quad (\varphi_X(0) = 1). \\ a = -b \Rightarrow \varphi_X(t) &= \frac{\exp(itb) - \exp(-itb)}{2itb} \\ &= \frac{\cos(tb) + i \sin(tb) - \cos(-tb) - i \sin(-tb)}{2itb} \\ &= \frac{\sin(tb)}{tb}.\end{aligned}$$

d) *Seien $(X_i)_{i \in \mathbb{N}}$ stochastisch unabhängig und identisch verteilt. Sei N eine weitere Zufallsvariable, stochastisch unabhängig von den X_i , mit Werten in \mathbb{N} . Sei $S := \sum_{i=1}^N X_i$.*

$$\begin{aligned}\Rightarrow \varphi_S(t) &= \mathbb{E} \left[\exp \left(it \sum_{j=1}^N X_j \right) \right] \\ &= \sum_{n \in \mathbb{N}} \mathbb{P}(N = n) \varphi_{X_1}^n(t) = \sum_n \mathbb{P}(N = n) \exp(n \ln \varphi_{X_1}(t)) \\ &= \mathbb{E} [\exp(N \ln \varphi_{X_1}(t))] = \mathbb{E} [\exp(iN(-i) \ln \varphi_{X_1}(t))] \\ &= \varphi_N(-i \ln \varphi_{X_1}(t)) \text{ bei entsprechendem Konvergenzradius in } \mathbb{C}.\end{aligned}$$

1.3 Konvergenzarten und Folgen von Zufallsvariablen

In diesem Kapitel betrachten wir Folgen $(X_n)_{n \geq 1}$ von (reellwertigen) Zufallsvariablen mit $X_n : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $n \geq 1$, und beschreiben, in welchen Weisen die Folge $(X_n)_{n \geq 1}$ gegen einen Grenzwert, also eine Grenz-Zufallsvariable $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ konvergieren

kann (für $n \rightarrow \infty$). Da X_n , $n \geq 1$, und X Funktionen sind, lassen sich (wie in der Funktionalanalysis) verschiedene Konvergenzarten unterscheiden, die in der Wahrscheinlichkeitstheorie mit besonderen Begriffen belegt werden. Es bestehen ferner Implikationsbeziehungen zwischen den Konvergenzarten, d.h., die ‘‘Stärke’’ der Konvergenz lässt sich unterscheiden.

Definition 1.21 (Konvergenzarten)

Sei $(X_n)_{n \geq 1}$ eine Folge von Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum, d. h., für alle $n \in \mathbb{N}$ ist $X_n : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ eine messbare Abbildung. Ferner sei $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ eine weitere (reellwertige) Zufallsvariable auf dem gleichen Wahrscheinlichkeitsraum wie $(X_n)_{n \geq 1}$.

a) Die Folge $(X_n)_{n \geq 1}$ konvergiert \mathbb{P} -fast sicher (mit Wahrscheinlichkeit 1) gegen X für $n \rightarrow \infty$

$$\begin{aligned} &:\Leftrightarrow \mathbb{P} \left(\left\{ \omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) \right\} \right) = 1 \\ &\Leftrightarrow \mathbb{P} \left(\lim_{n \rightarrow \infty} X_n = X \right) = 1. \end{aligned}$$

In Zeichen: $X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X$

b) Die Folge $(X_n)_{n \geq 1}$ konvergiert \mathbb{P} -stochastisch (in Wahrscheinlichkeit) gegen X für $n \rightarrow \infty$

$$:\Leftrightarrow \forall \varepsilon > 0 : \lim_{n \rightarrow \infty} \mathbb{P} (|X_n - X| > \varepsilon) = 0.$$

In Zeichen: $X_n \xrightarrow{\mathbb{P}} X$

c) Die Folge $(X_n)_{n \geq 1}$ konvergiert in Verteilung (schwach) gegen X für $n \rightarrow \infty$

$$:\Leftrightarrow \forall x \in C(F_X) : \lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x).$$

In Zeichen: $X_n \xrightarrow{\mathcal{D}} X$ bzw. $\mathcal{L}(X_n) \xrightarrow{w} \mathcal{L}(X)$.

Beachte: Das Maß \mathbb{P} wird für die Definition der Verteilungskonvergenz nicht benötigt. Daher können die X_n und/oder X in dieser Definition sogar auf unterschiedlichen Wahrscheinlichkeitsräumen definiert sein. Eine exaktere Definition lautet daher:

Sei (Ω', d) ein metrischer Raum und \mathcal{A}' die von den offenen Kugeln in der Metrik d erzeugte σ -Algebra. Seien \mathbb{P} und $(\mathbb{P}_n)_{n \geq 1}$ Wahrscheinlichkeitsmaße auf dem Messraum (Ω', \mathcal{A}') . Dann konvergiert die Folge $(\mathbb{P}_n)_{n \geq 1}$ schwach gegen \mathbb{P} für $n \rightarrow \infty$

$$:\Leftrightarrow \forall f \in \mathcal{C}_b(\Omega') : \lim_{n \rightarrow \infty} \int f d\mathbb{P}_n = \int f d\mathbb{P},$$

wobei $\mathcal{C}_b(\Omega')$ die Menge aller stetigen und beschränkten Abbildungen $f : \Omega' \rightarrow \mathbb{R}$ bezeichnet.

d) Sei $p \geq 1$ und seien X, X_1, X_2, \dots reellwertige Zufallsvariablen mit in \mathbb{R} existierendem p -ten Moment. Dann konvergiert die Folge $(X_n)_{n \geq 1}$ im p -ten Mittel gegen X für $n \rightarrow \infty$

$$:\Leftrightarrow \lim_{n \rightarrow \infty} \mathbb{E}[|X_n - X|^p] = 0.$$

In Zeichen: $X_n \xrightarrow{L_p} X$

Spezialfälle:

$p = 1$: Konvergenz im Mittel

$p = 2$: Konvergenz im quadratischen Mittel

Aus der Diskussion in Definition 1.21.c) über die Verteilungskonvergenz (schwache Konvergenz der Verteilungsgesetze) hat sich bereits ergeben, dass es unterschiedliche, äquivalente Charakterisierungen der vier in Definition 1.21 beschriebenen Konvergenzarten gibt. Dazu nun mehr.

Satz 1.22 (Alternative Charakterisierungen)

a)

$$\begin{aligned} X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X &\Leftrightarrow \mathbb{P}\left(\liminf_{n \rightarrow \infty} (X_n - X) = \limsup_{n \rightarrow \infty} (X_n - X) = 0\right) = 1 \\ &\Leftrightarrow \forall \omega \in \Omega \setminus N : \lim_{n \rightarrow \infty} (X_n(\omega) - X(\omega)) = 0, \\ &\text{wobei } N \text{ eine } \mathbb{P}\text{-Nullmenge bezeichnet, d. h., } \mathbb{P}(N) = 0. \end{aligned}$$

Beachte: $Y_n := X_n - X \Rightarrow$

$$\left\{ \lim_{n \rightarrow \infty} Y_n = 0 \right\} = \bigcap_{m=1}^{\infty} \bigcup_{k=0}^{\infty} \bigcap_{n=k}^{\infty} \left\{ |Y_n| < \frac{1}{m} \right\}$$

und damit messbar!

b)

$$X_n \xrightarrow{\mathcal{D}} X \Leftrightarrow \forall f \in \mathcal{C}_b(\mathbb{R}) : \mathbb{E}[f(X_n)] = \int f d\mathcal{L}(X_n) \xrightarrow{n \rightarrow \infty} \int f d\mathcal{L}(X) = \mathbb{E}[f(X)].$$

Beweis:

zu a): Die Aussage ist unmittelbar klar.

zu b): Der Beweis macht von dem folgendem Hilfssatz Gebrauch, der Bezüge zwischen der Topologie und der Integrationstheorie auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ herstellt. Er ist Teil des sogenannten "Portmanteau Theorem" und findet sich z.B. in Ash (1972), Theorem 5.4.1 d) + e).

Hilfssatz 1.23 (ohne Beweis)

$$\begin{aligned} & \forall f \in \mathcal{C}_b(\mathbb{R}) : \mathbb{E}[f(X_n)] \xrightarrow{(n \rightarrow \infty)} \mathbb{E}[f(X)] \\ \Leftrightarrow & \liminf_{n \rightarrow \infty} \mathbb{P}_{X_n}(A) \geq \mathbb{P}_X(A) \text{ f\"ur alle offenen Teilmengen } A \text{ von } \mathbb{R} \\ \Leftrightarrow & \mathbb{P}_{X_n}(A) \rightarrow \mathbb{P}_X(A) \forall A \in \mathcal{B}(\mathbb{R}) \text{ mit } \mathbb{P}_X(\partial A) = 0 \text{ ("randlose Mengen").} \end{aligned}$$

Da $(-\infty, x]$ für $x \in C(F_X)$ eine randlose Menge ist, liefert die zweite Äquivalenz im Hilfsatz unmittelbar die " \Leftarrow "-Richtung der Aussage unter b).

Zum Beweis der " \Rightarrow "-Richtung zeigen wir:

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x) \forall x \in C(F_X) \Rightarrow \forall A \subseteq \mathbb{R} \text{ offen: } \liminf_{n \rightarrow \infty} \mathbb{P}_{X_n}(A) \geq \mathbb{P}_X(A).$$

Sei dazu $A \subseteq \mathbb{R}$ offen beliebig ausgewählt. Wir schreiben A als disjunkte Vereinigung offener Intervalle I_1, I_2, \dots . Dies ist eine Charakterisierung der offenen Teilmengen von \mathbb{R} , vgl. Satz 1.6 in Boto von Querenburg (2001). Damit ergibt sich nach dem Lemma von Fatou:

$$\liminf_{n \rightarrow \infty} \mathbb{P}_{X_n}(A) = \liminf_{n \rightarrow \infty} \sum_k \mathbb{P}_{X_n}(I_k) \geq \sum_k \liminf_{n \rightarrow \infty} \mathbb{P}_{X_n}(I_k). \quad (1.3)$$

Da F_X nur abzählbar viele Unstetigkeitsstellen besitzen kann, lässt sich für jede Konstante $\varepsilon > 0$ die folgende Konstruktion durchführen:

Für jedes k sei I'_k ein rechtsseitig abgeschlossenes Teilintervall von I_k , so dass

(1) alle Endpunkte der I'_k in $C(F_X)$ enthalten sind und

(2) $\forall k : \mathbb{P}_X(I'_k) \geq \mathbb{P}_X(I_k) - \varepsilon 2^{-k}$.

Da $X_n \xrightarrow{\mathcal{D}} X$, gilt nun

$$\liminf_{n \rightarrow \infty} \mathbb{P}_{X_n}(I_k) \geq \liminf_{n \rightarrow \infty} \mathbb{P}_{X_n}(I'_k) = \mathbb{P}_X(I'_k).$$

Folglich gilt für (1.3), dass

$$\liminf_{n \rightarrow \infty} \mathbb{P}_{X_n}(A) \geq \sum_k \mathbb{P}_X(I'_k) \geq \sum_k \mathbb{P}_X(I_k) - \varepsilon = \mathbb{P}_X(A) - \varepsilon,$$

da $\sum_{k=1}^{\infty} 2^{-k} = 1$ ist. Da ε beliebig klein gewählt werden kann, ist hiermit alles gezeigt. ■

Satz 1.24 (Lévy'scher Stetigkeitssatz)

Es sei $(X_n)_{n \geq 1}$ eine Folge von Zufallsvariablen mit zugehörigen charakteristischen Funktionen $(\varphi_n)_{n \geq 1}$.

a) Falls (X_n) gegen eine Zufallsvariable X in Verteilung konvergiert, dann konvergiert (φ_n) gegen die charakteristische Funktion von X , und zwar gleichmäßig auf jedem endlichen Intervall.

b) Falls (φ_n) punktweise gegen eine Funktion φ konvergiert, deren Realteil im Punkte $(0, 1)$ stetig ist, dann gilt:

(i) φ ist eine charakteristische Funktion, und damit existiert (genau) eine Wahrscheinlichkeitsverteilung μ , deren charakteristische Funktion gerade φ ist.

(ii) $\mathcal{L}(X_n) \xrightarrow{w} \mu$ für $n \rightarrow \infty$.

Beweis: Satz 15.23 in Klenke (2008). ■

Anmerkung: Analoge Stetigkeitssätze gelten auch für erzeugende Funktionen und Laplace-Transformierte.

Satz 1.25 (Implikationsbeziehungen zwischen Konvergenzarten)

Es sei $(X_n)_{n \geq 1}$ eine Folge von Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und X eine weitere Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Ferner sei $p \geq 1$ eine reelle Konstante.

(a) $X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X \Rightarrow X_n \xrightarrow{\mathbb{P}} X$.

(b) $X_n \xrightarrow{\mathbb{P}\text{-f.s.}} X$ impliziert die Existenz des p -ten Moments von X sowie $X_n \xrightarrow{L_p} X$ genau dann, wenn $\mathcal{H} := \{|X_n|^p : n \geq 1\}$ gleichgradig integrierbar ist, d.h., falls

$$\lim_{c \rightarrow \infty} \sup_{f \in \mathcal{H}} \int_{\{|f| \geq c\}} |f| d\mathbb{P} = 0.$$

(c) $X_n \xrightarrow{L_p} X \Rightarrow X_n \xrightarrow{L_q} X, \forall 1 \leq q \leq p$.

(d) $X_n \xrightarrow{L_p} X \Rightarrow X_n \xrightarrow{\mathbb{P}} X$.

(e) $X_n \xrightarrow{\mathbb{P}} X \Rightarrow X_n \xrightarrow{\mathcal{D}} X$.

(f) Es ergibt sich damit die in Abbildung 1.2 dargestellte Grafik.

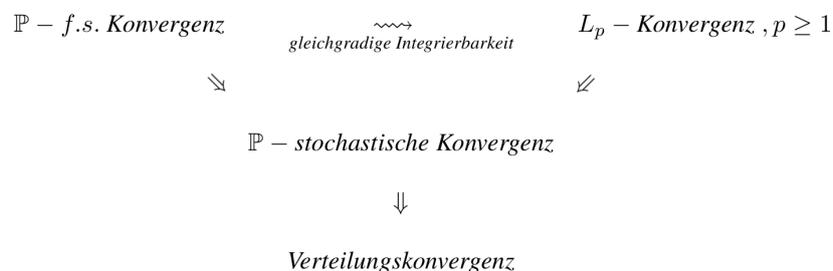


Abbildung 1.2: Zusammenhang von Konvergenzarten

Beweis:

zu a): Sei $\varepsilon > 0$ beliebig, aber fest. Definiere $A_n(\varepsilon) := \{\omega \in \Omega : |X_n(\omega) - X(\omega)| \leq \varepsilon\}$. Wir müssen zeigen, dass

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n(\varepsilon)) = 1.$$

Sei dazu $A := \{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$. Nach Voraussetzung dürfen wir annehmen, dass $\mathbb{P}(A) = 1$ ist. Nun ist aber für hinreichend großes n die Menge A eine Teilmenge von $A_n(\varepsilon)$, denn für alle $\omega \in A$ liegen alle bis auf endlich viele Folgenglieder $X_n(\omega)$ in jeder ε -Umgebung von $X(\omega)$. Dies impliziert die Aussage wegen der Monotonie von \mathbb{P} .

zu b): vgl. Abschnitt 6.2 in Klenke (2008); mehr zur gleichgradigen Integrierbarkeit in Abschnitt 2.4.

zu c): Die Funktion g , definiert durch $g(t) := t^{\frac{p}{q}}$, ist konvex auf $\mathbb{R}_{\geq 0} \ni t$. Nach der Jensen'schen Ungleichung gilt daher für alle $n \in \mathbb{N}$, dass

$$\mathbb{E}[|X_n - X|^p] = \mathbb{E}\left[|X_n - X|^{\frac{p}{q}}\right]^q \geq (\mathbb{E}[|X_n - X|^q])^{\frac{p}{q}}$$

und damit

$$(\mathbb{E}[|X_n - X|^p])^{\frac{1}{p}} \geq (\mathbb{E}[|X_n - X|^q])^{\frac{1}{q}}.$$

zu d): Wir wenden die Markov-Ungleichung auf $Y_n := |X_n - X|$ mit $h(t) := t^p$ an und erhalten für jedes $\varepsilon > 0$, dass

$$\mathbb{P}(|X_n - X| > \varepsilon) \leq \varepsilon^{-p} \mathbb{E}[|X_n - X|^p].$$

zu e): Sei f eine gleichmäßig stetige, beschränkte Funktion auf \mathbb{R} und $\varepsilon > 0$ beliebig vorgegeben. Dann gibt es ein $\delta > 0$ mit der Eigenschaft

$$|x - y| \leq \delta \Rightarrow |f(x) - f(y)| < \varepsilon, \quad x, y \in \mathbb{R}.$$

Wir rechnen:

$$\begin{aligned} & \left| \int f(X_n) d\mathbb{P} - \int f(X) d\mathbb{P} \right| \leq \int |f(X_n) - f(X)| d\mathbb{P} \\ &= \int_{\{|X_n - X| \leq \delta\}} |f(X_n) - f(X)| d\mathbb{P} + \int_{\{|X_n - X| > \delta\}} |f(X_n) - f(X)| d\mathbb{P} \\ &\leq \varepsilon \mathbb{P}(|X_n - X| \leq \delta) + 2 \sup_{x \in \mathbb{R}} |f(x)| \cdot \mathbb{P}(|X_n - X| > \delta). \end{aligned}$$

Also gilt wegen $X_n \xrightarrow{\mathbb{P}} X$, dass

$$\limsup_{n \rightarrow \infty} \left| \int f(X_n) d\mathbb{P} - \int f(X) d\mathbb{P} \right| \leq \varepsilon$$

und damit

$$\int f(X_n) d\mathbb{P} \xrightarrow{n \rightarrow \infty} \int f(X) d\mathbb{P},$$

da ε beliebig gewählt wurde. Da aber nach Transformationssatz

$$\int f(X_n) d\mathbb{P} \xrightarrow{n \rightarrow \infty} \int f(X) d\mathbb{P} \iff \int f d\mathbb{P}_{X_n} \xrightarrow{n \rightarrow \infty} \int f d\mathbb{P}_X$$

gilt, ist hiermit alles gezeigt. ■

Bemerkung 1.26

Die Implikationen aus Satz 1.25 sind im Allgemeinen strikt, d.h., die Umkehrungen gelten allgemein nicht. Ein Beispiel für $X_n \xrightarrow{\mathcal{D}} X$, aber $X_n \not\xrightarrow{\mathbb{P}} X$ ist gegeben durch

$$X_n(\omega) = \mathbf{1}_{[0, \frac{1}{2}]}(\omega), n \geq 1, \text{ und } X(\omega) = \mathbf{1}_{(\frac{1}{2}, 1]}(\omega)$$

auf $([0, 1], \mathcal{B}([0, 1]), \text{UNI}[0, 1])$.

In dem Spezialfall, dass $X \equiv x_0$ \mathbb{P} -fast sicher konstant ist, gilt jedoch:

$$X_n \xrightarrow{\mathbb{P}} x_0 \iff X_n \xrightarrow{\mathcal{D}} X = x_0.$$

Beweis: siehe Bauer (1991), Beweis von Satz 5.1. ■

Ein für die mathematische Statistik ungemein wichtiger Satz beschließt den technischen Teil dieses Kapitels.

Satz 1.27 (Satz von Cramér-Slutsky (Slutzky))

Seien $(X_n)_{n \geq 1}$ und $(Y_n)_{n \geq 1}$ zwei Folgen von Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit Werten in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Sei $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ eine weitere Zufallsvariable.

a) $X_n \xrightarrow{\mathcal{D}} X$ und $|X_n - Y_n| \xrightarrow{\mathbb{P}} 0 \Rightarrow Y_n \xrightarrow{\mathcal{D}} X$.

b) Sei $c \in \mathbb{R}$. $X_n \xrightarrow{\mathcal{D}} X$ und $Y_n \xrightarrow{\mathcal{D}} c \Rightarrow \left\{ \begin{array}{l} (i) \quad X_n + Y_n \xrightarrow{\mathcal{D}} X + c. \\ (ii) \quad X_n Y_n \xrightarrow{\mathcal{D}} cX. \end{array} \right\}$

Beweis:

zu a): Sei $f \in \mathcal{C}_b(\mathbb{R})$ mit Lipschitz-Konstante K . Dann ist

$$|f(x) - f(y)| \leq K|x - y| \wedge 2 \sup_{u \in \mathbb{R}} |f(u)|, \forall x, y \in \mathbb{R}.$$

Der Satz von der majorisierten Konvergenz liefert, dass

$$\limsup_{n \rightarrow \infty} \mathbb{E} [|f(X_n) - f(Y_n)|] = 0.$$

Also ergibt sich

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} |\mathbb{E}[f(Y_n)] - \mathbb{E}[f(X)]| \\
& \leq \limsup_{n \rightarrow \infty} |\mathbb{E}[f(X)] - \mathbb{E}[f(X_n)]| + \limsup_{n \rightarrow \infty} |\mathbb{E}[f(X_n) - f(Y_n)]| \\
& = 0.
\end{aligned}$$

zu b): (i) Definiere $Z_n := X_n + c$ und $\tilde{Z}_n := X_n + Y_n$. Dann gilt $Z_n \xrightarrow{\mathcal{D}} X + c$ und $|Z_n - \tilde{Z}_n| \xrightarrow{\mathbb{P}} 0$. Also kann Teil a) angewendet werden.

(ii) Siehe Theorem 2.3.3 in Lehmann (1999); Beweis in Bickel and Doksum (1977) bzw. Cramér (1946). ■

“Stillschweigend” haben wir den folgenden Satz benutzt.

Satz 1.28 (Continuous Mapping Theorem, siehe Abschnitt 1.7 in Serfling (1980))

Sei $h : \mathbb{R} \rightarrow \mathbb{R}$ messbar und stetig $\Rightarrow [X_n \xrightarrow{\mathcal{D}} X \Rightarrow h(X_n) \xrightarrow{\mathcal{D}} h(X)]$.

Wir kommen nun zu Anwendungen der Konvergenztheorie für Folgen von Zufallsvariablen.

Satz 1.29 (Kolmogoroffsches 0 – 1 Gesetz)

Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge stochastisch unabhängiger Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit beliebigen Wertebereichen. Dann gilt für jedes terminale (bzw. asymptotische) Ereignis, d.h., für jedes Ereignis $A \in \mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(\{X_m\} : m > n)$, entweder $\mathbb{P}(A) = 0$ oder $\mathbb{P}(A) = 1$. Wir nennen \mathcal{T} die terminale σ -Algebra zu $(X_n)_{n \in \mathbb{N}}$.

Beweis: Sei $(\Omega'_k, \mathcal{A}'_k)$ der Wertebereich von $X_k, k \in \mathbb{N}$, und seien $n \in \mathbb{N}$ sowie $C_k \in \mathcal{A}'_k, k = 1, \dots, n$, beliebig ausgewählt. Definiere $C := \{X_1 \in C_1, \dots, X_n \in C_n\}$. Dann ist

$$\mathbf{1}_{\{(X_k)_{k \geq 1} \in C\}} = \prod_{k=1}^n \mathbf{1}_{C_k}(X_k)$$

stochastisch unabhängig von $\mathbf{1}_A$.

Ferner erzeugt das System aller Mengen C (mit allen möglichen Werten $n \geq 1$) die Produkt- σ -Algebra $\bigotimes_{k \geq 1} \mathcal{A}'_k$ und deswegen ist $(X_k)_{k \geq 1} \perp\!\!\!\perp \mathbf{1}_A$. Insbesondere ist A als Element von \mathcal{T} damit stochastisch unabhängig von $A = \{\mathbf{1}_A = 1\}$, d.h., $\mathbb{P}(A \cap A) = \mathbb{P}(A)\mathbb{P}(A) \Rightarrow \mathbb{P}(A) = [\mathbb{P}(A)]^2$.

Die Gleichung $x = x^2$ hat aber nur die Lösungen 0 und 1. ■

Korollar 1.30

Es sei $(X_n)_{n \geq 1}$ eine Folge stochastisch unabhängiger, reellwertiger Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann sind $\liminf_{n \rightarrow \infty} X_n, \limsup_{n \rightarrow \infty} X_n$, sowie die Cesàro-Limiten $\liminf_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n X_i$ und $\limsup_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n X_i$ allesamt \mathbb{P} -fast sicher konstant.

Beweis: Korollar 2.39 in Klenke (2008). ■

Satz 1.31 (Lemma von Borel-Cantelli)

Sei $(A_k)_{k \geq 1}$ eine Folge von Ereignissen in einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und $A := \limsup_{k \rightarrow \infty} A_k = \{\omega \in \Omega : \omega \in A_k \text{ für unendlich viele } k\}$.

(a) Ist $\sum_{k \geq 1} \mathbb{P}(A_k) < \infty$, so ist $\mathbb{P}(A) = 0$.

(b) Ist $\sum_{k \geq 1} \mathbb{P}(A_k) = \infty$ und sind alle $(A_k)_{k \geq 1}$ stochastisch unabhängig, so ist $\mathbb{P}(A) = 1$.

Beweis:

zu (a): Für alle $m \in \mathbb{N}$ ist $A \subseteq \bigcup_{k \geq m} A_k$ und daher $\mathbb{P}(A) \leq \sum_{k \geq m} \mathbb{P}(A_k)$. Falls nun $\sum_{k \geq 1} \mathbb{P}(A_k) < \infty$, so folgt $\lim_{m \rightarrow \infty} \sum_{k \geq m} \mathbb{P}(A_k) = 0$ und damit $\mathbb{P}(A) = 0$.

zu (b): Wir beachten, dass $A^c = \bigcup_{m \geq 1} \bigcap_{k \geq m} A_k^c$ ist. Es folgt

$$\begin{aligned} \mathbb{P}(A^c) &\leq \sum_{m \geq 1} \mathbb{P} \left(\bigcap_{k \geq m} A_k^c \right) = \sum_{m \geq 1} \lim_{n \rightarrow \infty} \mathbb{P} \left(\bigcap_{k=m}^n A_k^c \right) \\ &= \sum_{m \geq 1} \lim_{n \rightarrow \infty} \prod_{k=m}^n (1 - \mathbb{P}(A_k)) \end{aligned}$$

wegen der vorausgesetzten stochastischen Unabhängigkeit der $(A_k)_{k \geq 1}$. Anwendung der bekannten Abschätzung $1 - x \leq \exp(-x) \forall x \in [0, 1]$ ergibt, dass

$$\begin{aligned} \mathbb{P}(A^c) &\leq \sum_{m \geq 1} \lim_{n \rightarrow \infty} \exp \left(- \sum_{k=m}^n \mathbb{P}(A_k) \right) \\ &= \sum_{m \geq 1} 0 = 0. \end{aligned}$$

■

Eine zentrale Fragestellung in der (mathematischen) Statistik lautet: „Unter welchen Voraussetzungen konzentriert sich der arithmetische Mittelwert (das empirische Mittel) einer Folge $(X_n)_{n \geq 1}$ von Zufallsvariablen ‘hinreichend gut’ um die theoretischen Mittelwerte $\mathbb{E}[X_n]$ für $n \rightarrow \infty$?“

Die Beantwortung dieser Frage ist zentral zur Beurteilung der Qualität von Schätz- und Testverfahren. Das einfachste Beispiel ist vermutlich ein Bernoulli’sches Versuchsschema. Kann die Trefferwahrscheinlichkeit p aus einer ”langen“ Messreihe ”gut“ inferiert werden?

Wahrscheinlichkeitstheoretisch wird dieser Problemkreis mit den Gesetzen der großen Zahlen bearbeitet.

Satz 1.32 (Gesetze der großen Zahlen)

Es sei $(X_n)_{n \geq 1}$ eine Folge von integrierbaren, reellwertigen Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Sei

$$S_n := \sum_{i=1}^n (X_i - \mathbb{E}[X_i]).$$

Wir sagen, dass $(X_n)_{n \geq 1}$ dem schwachen bzw. starken Gesetz der großen Zahlen genügt, falls

$$n^{-1}S_n \xrightarrow{\mathbb{P}} 0 \text{ bzw. } n^{-1}S_n \xrightarrow{\mathbb{P}\text{-f.s.}} 0.$$

(a) $(X_n)_{n \geq 1}$ genügt dem schwachen Gesetz der großen Zahlen, falls die $(X_n)_{n \in \mathbb{N}}$ paarweise unkorreliert sind und

$$\lim_{n \rightarrow \infty} n^{-2} \sum_{i=1}^n \text{Var}(X_i) = 0$$

gilt.

(b) $(X_n)_{n \geq 1}$ genügt dem starken Gesetz der großen Zahlen, falls die $(X_n)_{n \in \mathbb{N}}$ identisch verteilt und paarweise stochastisch unabhängig sind.

Beweis:

zu (a): Offenbar besitzt X_n eine endliche Varianz für alle $n \in \mathbb{N}$. Ferner ist $\mathbb{E}[S_n] = 0$ und $\text{Var}(S_n) = \sum_{i=1}^n \text{Var}(X_i)$ (nach Bienaymé) für alle $n \in \mathbb{N}$.

$$\text{Also ist } \text{Var}(n^{-1}S_n) = n^{-2} \sum_{i=1}^n \text{Var}(X_i) =: \sigma_n^2.$$

Nach der Chebyshev-Ungleichung folgt, dass

$$\forall \varepsilon > 0 : \mathbb{P}(|n^{-1}S_n| \geq \varepsilon) \leq \varepsilon^{-2} \sigma_n^2.$$

Die Bedingung $\sigma_n^2 \xrightarrow{n \rightarrow \infty} 0$ impliziert die \mathbb{P} -stochastische Konvergenz von $n^{-1}S_n$.

zu (b): Etemadi (1981) benutzt das Lemma von Borel Cantelli (Satz 1.31), den Satz von der monotonen Konvergenz und eine Abschneidetechnik, die ähnlich auch beim Zentralen Grenzwertsatz in der Version von Lindeberg/Feller (siehe unten) gebraucht wird. ■

Satz 1.33 (Zentraler Grenzwertsatz)

Sei $(X_n)_{n \geq 1}$ eine Folge (reellwertiger) stochastisch unabhängiger Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit endlichen zweiten Momenten und nicht-trivialer Varianz. O.B.d.A. sei $\mathbb{E}[X_k] = 0$ für alle $k \in \mathbb{N}$. Wir bezeichnen ferner für $k \in \mathbb{N}$ mit $\sigma_k^2 := \text{Var}(X_k) = \mathbb{E}[X_k^2] > 0$ die Varianz von X_k .

Sei $S_n := \sum_{j=1}^n X_j$. Beachte, dass $\text{Var}(S_n) = \sum_{k=1}^n \sigma_k^2$.

Wir sagen, dass für die Folge $(X_n)_{n \geq 1}$ ein Zentraler Grenzwertsatz gilt, falls

$$\mathcal{L} \left(\frac{S_n}{\sqrt{\text{Var}(S_n)}} \right) \xrightarrow[n \rightarrow \infty]{w} \mathcal{N}(0, 1).$$

Die folgenden drei Bedingungen sind jeweils hinreichend dafür, dass ein Zentraler Grenzwertsatz für $(X_n)_{n \geq 1}$ gilt:

(i) Alle X_k , $k \in \mathbb{N}$, haben die selbe Verteilung.

(ii) Ljapunov-Bedingung:

$$\begin{aligned} \exists \delta > 0 : \alpha_k := \mathbb{E} \left[|X_k^{2+\delta}| \right] < \infty \quad \forall k \in \mathbb{N} \quad \text{und} \quad \sum_{i=1}^n \alpha_i = o \left((\text{Var}(S_n))^{\frac{2+\delta}{2}} \right) \\ \Leftrightarrow \lim_{n \rightarrow \infty} (\text{Var}(S_n))^{-\frac{2+\delta}{2}} \sum_{j=1}^n \mathbb{E} \left[|X_j|^{2+\delta} \right] = 0. \end{aligned}$$

(iii) Lindeberg-Bedingung:

$$\forall \varepsilon > 0 : [\text{Var}(S_n)]^{-1} \sum_{j=1}^n \int_{\{|y| \geq \varepsilon \sqrt{\text{Var}(S_n)}\}} y^2 F_j(dy) \xrightarrow[(n \rightarrow \infty)]{} 0,$$

wobei $F_j(x) = \mathbb{P}(X_j \leq x)$, $j \in \mathbb{N}$.

Bemerkung 1.34

a) (i) \Rightarrow (ii) \Rightarrow (iii).

b) Die Lindeberg-Bedingung stellt sicher, dass die individuellen Varianzen der X_k klein sind im Vergleich zu ihrer Summe, denn (iii) impliziert, dass für gegebenes $\delta > 0$ ein $N(\delta)$ existiert mit der Eigenschaft

$$\forall n > N(\delta) : \frac{\sigma_k}{\sqrt{\text{Var}(S_n)}} < \delta \quad \forall k = 1, \dots, n.$$

c) (i) \Rightarrow (iii) ist leicht einzusehen. Sind $(X_n)_{n \geq 1}$ stochastisch unabhängig und identisch verteilt, so ist $\text{Var}(S_n) = n\sigma^2$ (mit $\sigma^2 = \text{Var}(X_1)$) und die linke Seite der Lindeberg-Bedingung wird zu $\sigma^{-2} \int_{\{|y| \geq \varepsilon \sqrt{n}\sigma\}} y^2 F(dy)$ mit $F(x) = \mathbb{P}(X_1 \leq x)$.

Da X_1 ein endliches zweites Moment besitzt und der Integrationsweg für $n \rightarrow \infty$ verschwindet, folgt die Gültigkeit der Lindeberg-Bedingung.

Beweis:

Beweis unter (i):

Sei φ die charakteristische Funktion von X_1/σ , wobei $\sigma^2 = \text{Var}(X_1)$. Wir müssen zeigen, dass

$$\mathcal{L} \left(\sum_{j=1}^n X_j / (\sqrt{n}\sigma) \right) \xrightarrow[n \rightarrow \infty]{w} \mathcal{N}(0, 1).$$

Für fixes n ist die charakteristische Funktion von $\sum_{j=1}^n X_j/(\sqrt{n}\sigma)$ gegeben durch $t \mapsto \varphi^n\left(\frac{t}{\sqrt{n}}\right)$. Es bleibt nach Lévy'schem Stetigkeitssatz (Satz 1.24) zu zeigen:

$$\lim_{n \rightarrow \infty} \varphi^n\left(\frac{t}{\sqrt{n}}\right) = \exp\left(-\frac{t^2}{2}\right) \text{ punktweise für alle } t \in \mathbb{R}.$$

Da X_1 ein endliches zweites Moment besitzt, ist φ nach Satz 1.18 zweimal stetig differenzierbar. Wegen

$$\mathbb{E}[X_1] = 0 = \left. \frac{d}{dt} \varphi(t) \right|_{t=0} \quad \text{sowie} \quad \mathbb{E}\left[\left(\frac{X_1}{\sigma}\right)^2\right] = 1 = -\left. \frac{d^2}{dt^2} \varphi(t) \right|_{t=0}$$

gilt somit für die Taylorentwicklung um 0, dass

$$\varphi\left(\frac{t}{\sqrt{n}}\right) = 1 + 0 - \frac{t^2}{2n} + o(n^{-1}).$$

$$\begin{aligned} \text{Damit ist} \quad \lim_{n \rightarrow \infty} \varphi^n\left(\frac{t}{\sqrt{n}}\right) &= \lim_{n \rightarrow \infty} \left(1 - \frac{t^2}{2n}\right)^n \\ &= \exp\left(-\frac{t^2}{2}\right), \text{ da } \forall x \in \mathbb{R} : \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \exp(x). \end{aligned}$$

Der Beweis unter (iii), der die Aussage unter (ii) impliziert, wird ähnlich geführt und findet sich in Feller (1971), Theorem 1 in Abschnitt XV.6. ■

Zur Gewinnung von präzisen Aussagen in der Statistik ist es überdies nützlich, dass unter der Annahme der Existenz dritter Momente auch die (asymptotische) Größenordnung der Differenz der Verteilungsfunktion der standardisierten Summenstatistik und Φ (der Verteilungsfunktion von $\mathcal{N}(0, 1)$) angegeben werden kann.

Satz 1.35 (Satz von Berry und Esséen)

Unter den Voraussetzungen von Satz 1.33 sei F_n die Verteilungsfunktion von $S_n/\sqrt{\text{Var}(S_n)}$, $n \in \mathbb{N}$. Dann gilt

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{6}{(\text{Var}(S_n))^{\frac{3}{2}}} \sum_{j=1}^n \mathbb{E}[|X_j|^3].$$

Sind $(X_n)_{n \geq 1}$ stochastisch unabhängig und identisch verteilt, so ergibt sich

$$\sup_{x \in \mathbb{R}} |F_n(x) - \Phi(x)| \leq \frac{6}{\sqrt{n}(\text{Var}(X_1))^{\frac{3}{2}}} \mathbb{E}[|X_1|^3] \sim \frac{1}{\sqrt{n}}.$$

Beweis: Satz 4.2.10 in Gaenssler and Stute (1977). ■

Zum Abschluss dieses Kapitels nun noch der sogenannte "Hauptsatz der Statistik".

Satz 1.36 (Glivenko-Cantelli)

Sei $((X_{n1}, \dots, X_{nm}))_{n \in \mathbb{N}}$ ein Dreiecksschema von zeilenweise stochastisch unabhängigen Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Für jedes $n \in \mathbb{N}$ seien also

X_{n1}, \dots, X_{nm} stochastisch unabhängig mit zugehörigen Verteilungsfunktionen F_{n1}, \dots, F_{nm} . Bezeichne $\bar{F}_n = n^{-1} \sum_{i=1}^n F_{ni}$, $n \in \mathbb{N}$.

Für jedes $n \in \mathbb{N}$ sei $G_n : \mathbb{R} \rightarrow [0, 1]$, definiert durch $G_n(t) = n^{-1} \sum_{i=1}^n \mathbf{1}_{(-\infty, t]}(X_{ni})$ für $t \in \mathbb{R}$, die sogenannte empirische Verteilungsfunktion von $(X_{nj})_{j=1, \dots, n}$.

Dann gilt:

$$\sup_{t \in \mathbb{R}} |G_n(t) - \bar{F}_n(t)| \xrightarrow{\mathbb{P}\text{-f.s.}} 0 \text{ für } n \rightarrow \infty.$$

Ist insbesondere $(X_n)_{n \geq 1}$ eine Folge von stochastisch unabhängigen und identisch verteilten Zufallsvariablen auf $(\Omega, \mathcal{A}, \mathbb{P})$ mit Verteilungsfunktion F von X_1 , so gilt:

$$\sup_{t \in \mathbb{R}} |G_n(t) - F(t)| \xrightarrow{\mathbb{P}\text{-f.s.}} 0 \text{ für } n \rightarrow \infty.$$

Beweis: Theorem 3.2.1 in Shorack and Wellner (1986). ■

Bemerkung 1.37

Für jedes fixe $t \in \mathbb{R}$ folgt die (punktweise) \mathbb{P} -fast sichere Konvergenz bereits aus dem starken Gesetz der großen Zahlen, falls die $(X_{ni})_{i=1, \dots, n}$ stochastisch unabhängig und identisch verteilt sind. Der allgemeine Fall wird bewiesen unter Anwendung des Prinzips der Quantilstransformation und des Lemmas von Borel-Cantelli (Satz 1.31).

Kapitel 2

Martingaltheorie

2.1 Allgemeine Definitionen und Eigenschaften

Definition 2.1

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(\mathcal{F}_n)_{n \geq 0}$ eine monoton wachsende Folge von Sub- σ -Algebren von \mathcal{F} , d.h., $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_i \subset \dots \subset \mathcal{F}$. Dann heißt $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration und $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ ein gefilterter Wahrscheinlichkeitsraum. Ferner setzen wir $\mathcal{F}_\infty := \sigma(\bigcup_{i \geq 0} \mathcal{F}_i)$.

Definition 2.2

Sei $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ ein gefilterter Wahrscheinlichkeitsraum und $(S_n)_{n \geq 0}$ eine Folge von Zufallsvariablen so, dass S_n für alle $n \geq 0$ \mathcal{F}_n -messbar ist. Dann nennt man $(S_n)_{n \geq 0}$ adaptiert zu $(\mathcal{F}_n)_{n \geq 0}$ und schreibt abkürzend $(S_n, \mathcal{F}_n)_{n \geq 0}$. Wird $\forall n \geq 0: \mathcal{F}_n = \sigma(S_0, \dots, S_n)$ gesetzt, so heißt $(\mathcal{F}_n)_{n \geq 0}$ die natürliche Filtration zu $(S_n)_{n \geq 0}$.

Definition 2.3 (Martingal, Sub- und Supermartingal)

Sei $(S_n, \mathcal{F}_n)_{n \geq 0}$ eine adaptierte Folge auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}[|S_n|] < \infty$ für alle $n \geq 0$. Dann nennt man $(S_n, \mathcal{F}_n)_{n \geq 0}$ ein

- (i) Martingal: $\Leftrightarrow \mathbb{E}[S_m | \mathcal{F}_n] = S_n \quad \forall m \geq n$,
- (ii) Submartingal: $\Leftrightarrow \mathbb{E}[S_m | \mathcal{F}_n] \geq S_n \quad \forall m \geq n$,
- (iii) Supermartingal: $\Leftrightarrow \mathbb{E}[S_m | \mathcal{F}_n] \leq S_n \quad \forall m \geq n$,

wobei die Relationen in (i) bis (iii) \mathbb{P} -fast sicher gelten sollen.

Anmerkung:

- (i) Martingale bleiben im bedingten Mittel gleich.
- (ii) Submartingale wachsen im bedingten Mittel.
- (iii) Supermartingale fallen im bedingten Mittel.
- (iv) Man kann sich die Folge $(S_n)_{n \geq 0}$ etwa als Spielstand in einem Spiel mit vielen Runden vorstellen.

Beispiel 2.4

(a) Sei $(X_n)_{n \geq 1}$ eine iid-Folge von Zufallsvariablen mit $X_1 \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathbb{E}[X_1] = \mu$. Sei $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ und $S_0 \equiv 0$, $S_n = \sum_{i=1}^n X_i$, $n \geq 1$.

Wir rechnen

$$\mathbb{E}[S_m | \mathcal{F}_n] = \mathbb{E} \left[\sum_{1 \leq i \leq n} X_i | \mathcal{F}_n \right] + \mathbb{E} \left[\sum_{n+1 \leq i \leq m} X_i | \mathcal{F}_n \right].$$

Nun ist S_n messbar bzgl. \mathcal{F}_n und damit ist $\mathbb{E} \left[\sum_{1 \leq i \leq n} X_i | \mathcal{F}_n \right] = S_n$. Andererseits ist $X_i \perp \mathcal{F}_n$ für

$i > n$ und damit $\mathbb{E} \left[\sum_{n+1 \leq i \leq m} X_i | \mathcal{F}_n \right] = (m - n)\mu$, \mathbb{P} -fast sicher.

Insgesamt erhalten wir also

$$\mathbb{E}[S_m | \mathcal{F}_n] = S_n + (m - n)\mu$$

und folglich ist

$$(S_n, \mathcal{F}_n)_{n \geq 0} \begin{cases} \text{Martingal, falls } \mu = 0, \\ \text{Submartingal, falls } \mu \geq 0, \\ \text{Supermartingal, falls } \mu \leq 0. \end{cases}$$

(b) Sei $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ und $(\mathcal{F}_n)_{n \geq 0}$ (irgend) eine Filtration auf (Ω, \mathcal{F}) . Definiere $S_n = \mathbb{E}[X | \mathcal{F}_n] \forall n \geq 0$. Dann liefert die Tower equation (siehe Satz 1.13.f)), dass

$$\forall m \geq n : \mathbb{E}[S_m | \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X | \mathcal{F}_m] | \mathcal{F}_n] = \mathbb{E}[X | \mathcal{F}_n] = S_n.$$

Somit ist $(S_n, \mathcal{F}_n)_{n \geq 0}$ ein Martingal.

Lemma 2.5

Eine adaptierte Folge $(S_n, \mathcal{F}_n)_{n \geq 0} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ ist genau dann ein Martingal, wenn

$$\mathbb{E}[S_{n+1} | \mathcal{F}_n] = S_n \tag{2.1}$$

\mathbb{P} -fast sicher für alle $n \geq 0$ gilt. Analoge Bedingungen gelten für Sub- und Supermartingale.

Beweis: Es gelte (2.1) und es sei $m > n$. Wir rechnen

$$\begin{aligned} \mathbb{E}[S_m | \mathcal{F}_n] &= \mathbb{E}[\mathbb{E}[S_m | \mathcal{F}_n] | \mathcal{F}_{m-1}] = \mathbb{E}[\mathbb{E}[S_m | \mathcal{F}_{m-1}] | \mathcal{F}_n] \\ &\stackrel{(2.1)}{=} \underbrace{\mathbb{E}[S_{m-1} | \mathcal{F}_n]}_{\text{Argumentation iterativ anwenden}} = \dots \\ &= \mathbb{E}[S_{n+1} | \mathcal{F}_n] = S_n. \end{aligned}$$

■

Lemma 2.6

Sei $(S_n, \mathcal{F}_n)_{n \geq 0} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ ein Submartingal und $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ eine konvexe Funktion mit $\varphi \circ S_n \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ für alle $n \geq 0$. Dann gilt:

- (i) Falls φ isoton ist, so ist $(\varphi(S_n), \mathcal{F}_n)_{n \geq 0}$ ebenfalls ein Submartingal.
- (ii) Falls $(S_n, \mathcal{F}_n)_{n \geq 0}$ ein Martingal ist, so ist $(\varphi(S_n), \mathcal{F}_n)_{n \geq 0}$ stets ein Submartingal.

Beweis: Setze $T_n = \varphi(S_n)$, $n \geq 0$. Dann gilt sowohl unter (i) als auch unter (ii), dass

$$T_n = \varphi \circ S_n \leq \mathbb{E}[\varphi \circ S_{n+1} | \mathcal{F}_n] = \mathbb{E}[T_{n+1} | \mathcal{F}_n] \quad (2.2)$$

ist. Die Behauptung folgt daraus mit Hilfe von Lemma 2.5. Die Gültigkeit der Ungleichung in (2.2) gilt unter Teil (i) wegen der Isotonie von φ und unter Teil (ii) wegen der Jensen'schen Ungleichung. ■

Korollar 2.7

Sei $(S_n, \mathcal{F}_n)_{n \geq 0}$ eine adaptierte Folge.

- (i) Falls $(S_n, \mathcal{F}_n)_{n \geq 0} \subset L_p(\Omega, \mathcal{F}, \mathbb{P})$ für $1 \leq p \leq \infty$ ein Martingal ist, so ist $(|S_n|^p, \mathcal{F}_n)_{n \geq 0}$ ein Submartingal.
- (ii) Falls $(S_n, \mathcal{F}_n)_{n \geq 0}$ ein Submartingal und $c \in \mathbb{R}$ eine vorgegebene Konstante ist, dann ist $(\max(c, S_n), \mathcal{F}_n)_{n \geq 0}$ ebenfalls ein Submartingal. Wählen wir $c = 0$, so ist insbesondere $(S_n^+, \mathcal{F}_n)_{n \geq 0}$ ein Submartingal.
- (iii) Falls $(S_n, \mathcal{F}_n)_{n \geq 0}$ ein Supermartingal ist, so ist $(S_n^-, \mathcal{F}_n)_{n \geq 0}$ ein Submartingal, denn $(-S_n, \mathcal{F}_n)_{n \geq 0}$ ist ein Submartingal und $S_n^- = \max(-S_n, 0)$.

Beispiel 2.8

Gegeben sei eine Filtration $(\mathcal{F}_n)_{n \geq 0}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Wir betrachten ein faires Spiel, d.h. ein Martingal $(X_n, \mathcal{F}_n)_{n \geq 0}$, wobei $X_n - X_{n-1}$ den Nettogewinn pro eingesetzter Geldeinheit in den n -ten Spielrunde angibt. Aus der Martingaleigenschaft von $(X_n, \mathcal{F}_n)_{n \geq 0}$ sowie der Linearität der bedingten Erwartung folgt $\mathbb{E}[X_n - X_{n-1} | \mathcal{F}_{n-1}] = 0$, $n \geq 1$, wobei selbstverständlich $X_0 := 0$ gesetzt wird. Angenommen, es darf in der n -ten Spielrunde ein beliebiger (aber endlicher) Einsatz C_n getätigt werden, nachdem man alle vorherigen Spielrunden beobachtet hat (d.h., C_n ist \mathcal{F}_{n-1} -messbar).

Frage: Kann C_n so gewählt werden (geschickte Spielstrategie!), dass man die Bank sprengt?

Lösung: Bezeichne $Y_n = \sum_{i=1}^n C_i(X_i - X_{i-1})$ den Gesamtgewinn bis nach Spielrunde $n \geq 1$. Dann folgt unter den gemachten Annahmen, dass

$$\begin{aligned} \mathbb{E}[Y_{n+1} | \mathcal{F}_n] &= Y_n + \mathbb{E}[C_{n+1}(X_{n+1} - X_n) | \mathcal{F}_n] \\ &= Y_n + C_{n+1} \mathbb{E}[(X_{n+1} - X_n) | \mathcal{F}_n] \\ &= Y_n + C_{n+1}(X_n - X_n) = Y_n \end{aligned}$$

fast sicher, egal, welche Einsatzstrategie verfolgt wird. Man kann also keine Strategie angeben, um im jeweils nächsten Spiel den Gewinn im bedingten Mittel zu steigern.

2.2 Stoppzeiten

Definition 2.9

Sei $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ ein gefilterter Wahrscheinlichkeitsraum. Dann nennt man eine Abbildung

$$\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$$

mit der Eigenschaft

$$\{\tau = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}_0 \tag{2.3}$$

eine Stoppzeit zur Filtration $(\mathcal{F}_n)_{n \geq 0}$.

Bemerkung 2.10

Die Eigenschaft (2.3) ist äquivalent zu

$$\forall n \in \mathbb{N} : \{\tau \leq n\} \in \mathcal{F}_n, \tag{2.4}$$

siehe Übungsaufgabe.

Lemma 2.11

Sei $(\tau_k)_{k \geq 1}$ eine Folge von Stoppzeiten zur Filtration $(\mathcal{F}_n)_{n \geq 0}$, dann sind $\inf_{k \geq 1} \tau_k$, $\sup_{k \geq 1} \tau_k$, $\liminf_{k \geq 1} \tau_k$ und $\limsup_{k \geq 1} \tau_k$ ebenfalls Stoppzeiten zu $(\mathcal{F}_n)_{n \geq 0}$.

Beweis: Übungsaufgabe. ■

Beispiel 2.12

Sei $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration bezüglich (Ω, \mathcal{F}) und $(S_n)_{n \geq 0}$ eine adaptierte Folge reellwertiger Zufallsvariablen.

(a) Die konstante Zufallsvariable $\tau \equiv m$, $m \in \mathbb{N}$, ist eine Stoppzeit zu $(\mathcal{F}_n)_{n \geq 0}$.

(b) Für jede Borelmenge $B \in \mathcal{B}(\mathbb{R})$ ist die Ersteintrittszeit der Folge $(S_n)_{n \geq 0}$ in B , d.h.

$$\tau_B := \inf\{n \in \mathbb{N}_0 : S_n \in B\},$$

eine Stoppzeit zu $(\mathcal{F}_n)_{n \geq 0}$.

Beweis:

zu (a): $\{\tau = m\} = \Omega$, $\{\tau = m'\} = \emptyset$, $\forall m' \neq m$.

zu (b): $\{\tau_B = n\} = \{S_0 \notin B\} \cap \dots \cap \{S_{n-1} \notin B\} \cap \{S_n \in B\}$. Definierende Eigenschaften von Filtrationen liefern das Ergebnis. ■

Definition 2.13

Sei τ eine Stoppzeit bzgl. der Filtration $(\mathcal{F}_n)_{n \geq 0}$ auf (Ω, \mathcal{F}) . Dann wird durch

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \mid A \cap \{\tau \leq n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}_0\}$$

die σ -Algebra der τ -Vergangenheit definiert.

Bemerkung 2.14

(a) Angenommen, $\mathcal{F}_n = \sigma(S_0, \dots, S_n)$ für eine Folge von Zufallsvariablen $(S_n)_{n \geq 0}$ auf (Ω, \mathcal{F}) .

Dann bedeutet $A \in \mathcal{F}_\tau$ heuristisch, dass man aus der Kenntnis von $S_0(\omega), \dots, S_\tau(\omega)$ bereits erschließen kann, ob $\omega \in A$ ist oder nicht.

(b) Es gilt stets $\mathcal{F}_\tau \subseteq \mathcal{F}_\infty$ und τ ist \mathcal{F}_τ -messbar.

Lemma 2.15

Seien τ_1, τ_2 zwei Stoppzeiten zur Filtration $(\mathcal{F}_n)_{n \geq 0}$ auf (Ω, \mathcal{F}) . Dann gilt:

- (i) Die Mengen $\{\tau_1 < \tau_2\}$, $\{\tau_1 = \tau_2\}$ und $\{\tau_1 \leq \tau_2\}$ sind in \mathcal{F}_{τ_1} und in \mathcal{F}_{τ_2} .
- (ii) $\forall A \in \mathcal{F}_{\tau_1}: A \cap \{\tau_1 \leq \tau_2\} \in \mathcal{F}_{\tau_2}$.
- (iii) $\tau_1 \leq \tau_2 \Rightarrow \mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

Beweis:

zu (i): Sei $n \geq 0$ beliebig. Dann gilt:

$$\begin{aligned} \{\tau_1 < \tau_2\} \cap \{\tau_1 \leq n\} &= \bigcup_{i=0}^n \{\tau_1 < \tau_2, \tau_1 = i\} \\ &= \bigcup_{i=0}^n \{i < \tau_2\} \cap \{\tau_1 = i\} \\ &= \bigcup_{i=0}^n \{\tau_2 \leq i\}^c \cap \{\tau_1 = i\} \in \mathcal{F}_n. \end{aligned}$$

Ferner gilt

$$\begin{aligned} \{\tau_1 < \tau_2\} \cap \{\tau_2 \leq n\} &= \bigcup_{i=1}^n \{\tau_1 < \tau_2, \tau_2 = i\} \\ &= \bigcup_{i=1}^n \{\tau_1 < i\} \cap \{\tau_2 = i\} \\ &= \bigcup_{i=1}^n \left(\bigcup_{j=0}^{i-1} \{\tau_1 = j\} \right) \cap \{\tau_2 = i\} \in \mathcal{F}_n. \end{aligned}$$

Hieraus folgt, dass $\{\tau_1 < \tau_2\} \in \mathcal{F}_{\tau_i}$, $i = 1, 2$. Für die anderen beiden Mengen rechnet man analog.

zu (ii): Seien $A \in \mathcal{F}_{\tau_1}$ und $n \in \mathbb{N}_0$ beliebig vorgegeben. Dann gilt:

$$A \cap \{\tau_1 \leq \tau_2\} \cap \{\tau_2 \leq n\} = \bigcup_{i=0}^n (A \cap \{\tau_1 \leq i\}) \cap \{\tau_2 = i\} \in \mathcal{F}_n.$$

zu (iii): Folgt sofort aus (ii) mit $\{\tau_1 \leq \tau_2\} = \Omega$. ■

Lemma 2.16

Seien $\tau, (\tau_k)_{k \geq 1}$ Stoppzeiten zur Filtration $(\mathcal{F}_n)_{n \geq 0}$ auf (Ω, \mathcal{F}) mit $\tau_k \downarrow \tau$. Dann ist $\mathcal{F}_\tau = \bigcap_{k \geq 1} \mathcal{F}_{\tau_k}$.

Beweis: Aufgrund von Lemma 2.15 gilt für alle $k \in \mathbb{N}$, dass $\mathcal{F}_\tau \subseteq \mathcal{F}_{\tau_k} \subseteq \mathcal{F}_{\tau_{k-1}} \subseteq \dots \subseteq \mathcal{F}_{\tau_1}$.

Also ist $\mathcal{F}_\tau \subseteq \bigcap_{k \geq 1} \mathcal{F}_{\tau_k}$.

Sei jetzt $A \in \bigcap_{k \geq 1} \mathcal{F}_{\tau_k}$ und $\ell \in \mathbb{N}$. Wir beachten zunächst, dass $\{\tau \leq \ell\} = \bigcup_{k \geq 1} \{\tau_k \leq \ell\}$. Damit folgt aber, dass $A \cap \{\tau \leq \ell\} = \bigcup_{k \geq 1} \underbrace{(A \cap \{\tau_k \leq \ell\})}_{\in \mathcal{F}_\ell} \in \mathcal{F}_\ell$. Dies impliziert, dass $A \in \mathcal{F}_\tau$. ■

Lemma 2.17

Sei $(S_n, \mathcal{F}_n)_{n \geq 0}$ eine adaptierte Folge und τ eine endliche Stoppzeit, d.h. $\mathbb{P}(\tau < \infty) = 1$. Dann ist S_τ , also der Wert des zum Zeitpunkt τ gestoppten Prozesses, \mathbb{P} -fast sicher wohldefiniert und \mathcal{F}_τ -messbar.

Beweis: Sei $x \in \mathbb{R}$ und $n \in \mathbb{N}_0$ beliebig. Dann gilt:

$$\{S_\tau \leq x\} \cap \{\tau \leq n\} = \bigcup_{k=0}^n \underbrace{\{S_k \leq x\}}_{\in \mathcal{F}_k} \cap \underbrace{\{\tau = k\}}_{\in \mathcal{F}_k}.$$

Da $\mathcal{F}_k \subseteq \mathcal{F}_n \forall k \leq n$ folgt damit die Behauptung aus der Definition von \mathcal{F}_τ . ■

Lemma 2.18

Sei $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ und τ eine Stoppzeit zu einer Filtration $(\mathcal{F}_n)_{n \geq 0}$ auf (Ω, \mathcal{F}) . Dann gilt \mathbb{P} -fast sicher für alle $n \geq 0$, dass

$$\mathbf{1}_n(\tau) \mathbb{E}[X | \mathcal{F}_\tau] = \mathbf{1}_n(\tau) \mathbb{E}[X | \mathcal{F}_n]. \tag{2.5}$$

Beweis: Wir müssen die charakterisierende Integralgleichung über \mathcal{F}_τ -Mengen verifizieren. Sei also $F \in \mathcal{F}_\tau$. Wir rechnen:

$$\begin{aligned} \int_F \mathbf{1}_n(\tau) \mathbb{E}[X | \mathcal{F}_\tau] d\mathbb{P} &= \int_F \mathbb{E}[\mathbf{1}_n(\tau) X | \mathcal{F}_\tau] d\mathbb{P} \\ &\stackrel{F \in \mathcal{F}_\tau}{=} \int_F \mathbf{1}_n(\tau) X d\mathbb{P} = \int_{F \cap \{\tau=n\}} X d\mathbb{P}. \end{aligned} \tag{2.6}$$

Wir beachten, dass $F \cap \{\tau = n\} \in \mathcal{F}_n$. Damit ist die rechte Seite von (2.6) gleich

$$\int_{F \cap \{\tau = n\}} \mathbb{E}[X|\mathcal{F}_n] d\mathbb{P} = \int_F \mathbf{1}_n(\tau) \mathbb{E}[X|\mathcal{F}_n] d\mathbb{P}.$$

■

Satz 2.19

Sei $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ und τ eine endliche Stoppzeit zu einer Filtration $(\mathcal{F}_n)_{n \geq 0}$ auf (Ω, \mathcal{F}) . Dann wird durch $S_n := \mathbb{E}[X|\mathcal{F}_n]$ ein Martingal definiert, vgl. Beispiel 2.4.(b). Nach Lemma 2.18 gilt für dieses Martingal \mathbb{P} -fast sicher, dass $S_\tau = \mathbb{E}[X|\mathcal{F}_\tau]$ ist.

Definition 2.20

Ein Martingal $(S_n, \mathcal{F}_n)_{n \geq 0}$ heißt regulär, falls ein $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ existiert mit $S_n = \mathbb{E}[X|\mathcal{F}_n]$, $n \in \mathbb{N}_0$.

2.3 Stoppsätze

Wir interpretieren Stoppzeiten (insbesondere: Ersteintrittszeiten) als Spiel- bzw. Handelsstrategien. Ziel ist es nun, verschiedene Strategien hinsichtlich ihrer (zu erwartenden) „Qualität“ zu vergleichen und damit optimale Strategien zu finden. Stoppsätze liefern Ergebnisse zum Vergleich von Stoppzeiten.

Satz 2.21

Sei $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$, $N \in \mathbb{N}$, ein Supermartingal mit endlichem Zeithorizont und τ, σ zwei Stoppzeiten bzgl. der Filtration $(\mathcal{F}_n)_{1 \leq n \leq N}$ auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann gilt \mathbb{P} -fast sicher, dass

$$\mathbf{1}_{\{\tau \geq \sigma\}} S_\sigma \geq \mathbf{1}_{\{\tau \geq \sigma\}} \mathbb{E}[S_\tau | \mathcal{F}_\sigma] \text{ ist.} \quad (2.7)$$

Bemerkung 2.22

Setzen wir $\sigma \equiv 1$, so liefert 2.7, dass

$$S_1 \geq \mathbb{E}[S_\tau | \mathcal{F}_1] \Rightarrow \mathbb{E}[S_1] \geq \mathbb{E}[\mathbb{E}[S_\tau | \mathcal{F}_1]] = \mathbb{E}[S_\tau].$$

Das bedeutet, dass jedes gestoppte Supermartingal (egal nach welcher Strategie τ !) mit endlichem Zeithorizont im Mittel kleiner wird als das Mittel von S_1 .

Beweis von Satz 2.21. Wir beachten zunächst, dass

$$\mathbb{E}[|S_\tau|] = \sum_{n=1}^N \int_{\{\tau=n\}} |S_n| d\mathbb{P} \leq \sum_{n=1}^N \mathbb{E}[|S_n|] < \infty$$

nach Definition 2.3. Also ist S_τ integrierbar. Ferner ist $\mathbf{1}_{\{\tau \geq \sigma\}} = \sum_{n=1}^N \mathbf{1}_{\{\tau \geq n, \sigma = n\}}$. Damit genügt es zu zeigen, dass

$$\mathbf{1}_{\{\tau \geq n, \sigma = n\}} S_n \geq \mathbf{1}_{\{\tau \geq n, \sigma = n\}} \mathbb{E}[S_\tau | \mathcal{F}_\sigma] = \mathbf{1}_{\{\tau \geq n, \sigma = n\}} \mathbb{E}[S_\tau | \mathcal{F}_n]$$

(siehe Lemma 2.18) für alle $1 \leq n \leq N$ zutrifft. Sei also n beliebig, aber fest und $A \in \mathcal{F}_n$. Wir müssen verifizieren, dass

$$\int_A \mathbf{1}_{\{\tau \geq n, \sigma = n\}} S_n d\mathbb{P} \geq \int_A \mathbf{1}_{\{\tau \geq n, \sigma = n\}} \mathbb{E}[S_\tau | \mathcal{F}_n] d\mathbb{P} = \int_A \mathbf{1}_{\{\tau \geq n, \sigma = n\}} S_\tau d\mathbb{P} \text{ ist.}$$

Wir rechnen

$$\begin{aligned} \int_{A \cap \{\tau \geq n, \sigma = n\}} (S_n - S_\tau) d\mathbb{P} &= \int_{A \cap \{\tau = n, \sigma = n\}} (S_n - S_n) d\mathbb{P} + \int_{A \cap \{\tau > n, \sigma = n\}} (S_n - S_\tau) d\mathbb{P} \\ &= \int_{A \cap \{\tau \geq n+1, \sigma = n\}} (S_n - S_\tau) d\mathbb{P} \\ &\geq \int_{A \cap \{\tau \geq n+1, \sigma = n\}} (\mathbb{E}[S_{n+1} | \mathcal{F}_n] - S_\tau) d\mathbb{P} \end{aligned} \quad (2.8)$$

wegen der Supermartingaleigenschaft.

Beachte nun, dass $\{\tau \geq n+1\} = \{\tau \leq n\}^c$ und somit $A \cap \{\tau \geq n+1, \sigma = n\} \in \mathcal{F}_n$ ist.

Damit ist die rechte Seite von (2.8) gleich

$$\int_{A \cap \{\tau \geq n+1, \sigma = n\}} (S_{n+1} - S_\tau) d\mathbb{P}.$$

Insgesamt ist also

$$\int_{A \cap \{\tau \geq n, \sigma = n\}} (S_n - S_\tau) d\mathbb{P} \geq \int_{A \cap \{\tau \geq n+1, \sigma = n\}} (S_{n+1} - S_\tau) d\mathbb{P}.$$

Iterative Anwendung dieser Argumentation führt zu

$$\int_{A \cap \{\tau \geq n, \sigma = n\}} (S_n - S_\tau) d\mathbb{P} \geq \int_{A \cap \{\tau = N, \sigma = n\}} (S_N - S_\tau) d\mathbb{P} = 0,$$

was den Beweis komplettiert. ■

Korollar 2.23

Sei $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ mit $N \in \mathbb{N}$ ein Supermartingal mit endlichem Zeithorizont und τ, σ zwei Stoppzeiten bzgl. der Filtration $(\mathcal{F}_n)_{1 \leq n \leq N}$ auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit der Eigenschaft $\tau \geq \sigma$. Dann gilt

$$\mathbb{E}[S_1] \geq \mathbb{E}[S_\sigma] \geq \mathbb{E}[S_\tau] \geq \mathbb{E}[S_N].$$

Bemerkung 2.24

Handelt es sich bei $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ um ein Martingal bzw. Submartingal, so bleiben die Aussagen richtig, wenn „ \geq “ durch „ $=$ “ bzw. „ \leq “ ersetzt wird.

Anwendung 2.25 (Maximalungleichung)

Sei $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$, $N \in \mathbb{N}$, ein Submartingal mit endlichem Zeithorizont auf $(\Omega, \mathcal{F}, \mathbb{P})$ und $\varepsilon > 0$. Dann gelten:

$$\varepsilon \mathbb{P}\left(\max_{1 \leq n \leq N} S_n \geq \varepsilon\right) \leq \int_{\{\max_{1 \leq n \leq N} S_n \geq \varepsilon\}} S_N d\mathbb{P} \leq \mathbb{E}[S_N^+], \quad (2.9)$$

$$\varepsilon \mathbb{P}\left(\min_{1 \leq n \leq N} S_n \leq -\varepsilon\right) \leq -\mathbb{E}[S_1] + \int_{\{\min_{1 \leq n \leq N} S_n > -\varepsilon\}} S_N d\mathbb{P}. \quad (2.10)$$

Beweis: Zum Nachweis von (2.9) definiere

$$\tau := \begin{cases} \min\{n \leq N \mid S_n \geq \varepsilon\}, & \text{falls } \{n \leq N \mid S_n \geq \varepsilon\} \neq \emptyset \\ N, & \text{sonst.} \end{cases}$$

Dann ist τ eine Stoppzeit und wir erhalten aus Korollar 2.23, dass

$$\begin{aligned} \mathbb{E}[S_N] &\geq \mathbb{E}[S_\tau] \\ &= \int_{\{\max_{1 \leq n \leq N} S_n \geq \varepsilon\}} S_\tau d\mathbb{P} + \int_{\{\max_{1 \leq n \leq N} S_n < \varepsilon\}} S_\tau d\mathbb{P} \\ &\geq \varepsilon \mathbb{P}\left(\max_{1 \leq n \leq N} S_n \geq \varepsilon\right) + \int_{\{\max_{1 \leq n \leq N} S_n < \varepsilon\}} S_N d\mathbb{P}. \end{aligned}$$

Also ist

$$\varepsilon \mathbb{P}\left(\max_{1 \leq n \leq N} S_n \geq \varepsilon\right) \leq \mathbb{E}[S_N] - \int_{\{\max_{1 \leq n \leq N} S_n < \varepsilon\}} S_N d\mathbb{P} = \int_{\{\max_{1 \leq n \leq N} S_n \geq \varepsilon\}} S_N d\mathbb{P}$$

wie gewünscht.

Die zweite Ungleichung (2.10) folgt durch analoge Rechnung mit Hilfe von

$$\tau := \begin{cases} \min\{n \leq N \mid S_n \leq -\varepsilon\}, & \text{falls } \{n \leq N \mid S_n \leq -\varepsilon\} \neq \emptyset \\ N, & \text{sonst} \end{cases}$$

unter Beachtung von $\mathbb{E}[S_1] \leq \mathbb{E}[S_\tau]$. ■

Korollar 2.26 (Kolmogoroff'sche Ungleichung)

Sei $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$, $N \in \mathbb{N}$, ein Martingal mit endlichem Zeithorizont auf $(\Omega, \mathcal{F}, \mathbb{P})$ und $\varepsilon > 0$. Ferner sei $(S_n)_{1 \leq n \leq N} \subset L_2(\Omega, \mathcal{F}, \mathbb{P})$. Dann ist nach Korollar 2.7.(i) die Folge $(S_n^2, \mathcal{F}_n)_{1 \leq n \leq N}$ ein Submartingal und es folgt aus Anwendung 2.25, dass

$$\mathbb{P}(\max_{1 \leq n \leq N} |S_n| \geq \varepsilon) \leq \varepsilon^{-2} \mathbb{E}[S_N^2].$$

Beispiel 2.27

Sei $(X_i)_{i \geq 1}$ eine iid. Folge mit $\mathbb{E}[X_1] = 0$ und $\text{Var}(X_1) = \sigma^2 < \infty$, und sei $\forall 1 \leq n \leq N$: $S_n = \sum_{i=1}^n X_i$. Nach Beispiel 2.4.(a) ist $(S_n)_{1 \leq n \leq N}$ ein Martingal bzgl. der natürlichen Filtration. Die Kolmogoroff'sche Ungleichung liefert nun

$$\mathbb{P}(\max_{1 \leq n \leq N} |S_n| \geq \varepsilon) \leq \frac{N\sigma^2}{\varepsilon^2},$$

während die Tschebyscheff'sche Ungleichung nur

$$\mathbb{P}(|S_N| \geq \varepsilon) \leq \frac{N\sigma^2}{\varepsilon^2}$$

liefert.

Satz 2.28 (Transformation durch Stoppzeiten)

Sei $(S_n, \mathcal{F}_n)_{n \geq 1} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ ein (Super-) Martingal und τ eine Stoppzeit. Dann wird durch

$$\hat{S}_n := S_{\tau \wedge n}, \quad n \geq 1$$

der bei τ gestoppte Prozess definiert und es gilt:

$$(\hat{S}_n, \mathcal{F}_n)_{n \geq 1} \text{ ist ein (Super-) Martingal.}$$

Beweis: Wir beachten, dass für jedes $n \geq 1$ die Zufallsvariable $\tau \wedge n$ eine endliche Stoppzeit ist. Nach Lemma 2.17 ist \hat{S}_n damit messbar bzgl. $\mathcal{F}_{\tau \wedge n}$. Aus der Definition von $\mathcal{F}_{\tau \wedge n}$ folgt aber direkt, dass $\mathcal{F}_{\tau \wedge n} \subseteq \mathcal{F}_n$. Also ist $(\hat{S}_n)_{n \geq 1}$ adaptiert zu $(\mathcal{F}_n)_{n \geq 1}$.

Nehmen wir nun an, dass $(S_n, \mathcal{F}_n)_{n \geq 1}$ ein Supermartingal ist, und sei $m > n$. Dann ist $\tau \wedge n \leq \tau \wedge m$ und wir erhalten aus Satz 2.21, dass

$$\hat{S}_n = S_{\tau \wedge n} \geq \mathbb{E}[S_{\tau \wedge m} \mid \mathcal{F}_{\tau \wedge n}] = \mathbb{E}[\hat{S}_m \mid \mathcal{F}_{\tau \wedge n}]. \quad (2.11)$$

Nach Lemma 2.15.(iii) folgt aus $\tau \wedge n \leq \tau \wedge m$, dass $\mathcal{F}_{\tau \wedge n} \subseteq \mathcal{F}_{\tau \wedge m}$. Also besagt (2.11), dass $(\hat{S}_n, \mathcal{F}_{\tau \wedge n})_{n \geq 1}$ wieder ein Supermartingal auf $(\Omega, \mathcal{F}, \mathbb{P})$ ist. Wir betrachten indes die Filtration $(\mathcal{F}_n)_{n \geq 1}$.

Dazu beachten wir folgende Fallunterscheidung.

1. Fall: $\tau \geq n$: In diesem Fall gilt $\tau \wedge n = n$ und Lemma 2.18 liefert

$$\mathbf{1}_n(\tau \wedge n) \mathbb{E}[\hat{S}_m | \mathcal{F}_{\tau \wedge n}] = \mathbf{1}_n(\tau \wedge n) \mathbb{E}[\hat{S}_m | \mathcal{F}_n].$$

Zusammen mit (2.11) ist in diesem Fall also

$$\mathbf{1}_{\{\tau \geq n\}} \hat{S}_n \geq \mathbf{1}_{\{\tau \geq n\}} \mathbb{E}[\hat{S}_m | \mathcal{F}_n].$$

2. Fall: $\tau = k < n$: In diesem Fall gilt $\mathbf{1}_k(\tau \wedge n) \hat{S}_n = \mathbf{1}_k(\tau) S_k$ und

$$\begin{aligned} \mathbf{1}_k(\tau \wedge n) \mathbb{E}[\hat{S}_m | \mathcal{F}_n] &= \mathbf{1}_k(\tau) \mathbb{E}[\hat{S}_m | \mathcal{F}_n] \\ &= \mathbf{1}_k(\tau) \mathbb{E}[\mathbf{1}_k(\tau) \hat{S}_m | \mathcal{F}_n] \\ &= \mathbf{1}_k(\tau) \mathbb{E}[S_k | \mathcal{F}_n] \\ &= \mathbf{1}_k(\tau) S_k \\ &= \mathbf{1}_k(\tau \wedge n) \hat{S}_n \quad \mathbb{P} - \text{fast sicher.} \end{aligned}$$

Fassen wir beide Fälle zusammen, so gilt also für $m \geq n$, dass

$$\hat{S}_n \geq \mathbb{E}[\hat{S}_m | \mathcal{F}_n].$$

Der Martingalfall wird analog bewiesen. ■

Bemerkung 2.29

Falls $(S_n, \mathcal{F}_n)_{n \geq 1}$ ein Submartingal ist, so ist $(-S_n, \mathcal{F}_n)_{n \geq 1}$ ein Supermartingal. Daher gilt Satz 2.28 auch für Submartingale.

Satz 2.30

Sei $(S_n, \mathcal{F}_n)_{n \geq 1} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ ein (Super-) Martingal und τ eine Stoppzeit. Ferner sei eine der folgenden drei Bedingungen erfüllt.

- (i) Es gibt ein $N \in \mathbb{N}$ mit $\mathbb{P}(\tau \leq N) = 1$.
- (ii) $\mathbb{P}(\tau < \infty) = 1$ und $\exists K > 0 : |S_n| < K$ für alle $n \geq 1$.
- (iii) $\mathbb{E}[\tau] < \infty$ und $\exists K > 0 : |S_n - S_{n-1}| < K$ für alle $n \geq 2$.

Dann gilt, dass

$$\mathbb{E}[S_\tau] = (\leq) \mathbb{E}[S_1]$$

ist, falls (S_n, \mathcal{F}_n) ein Martingal (Supermartingal) ist.

Beweis: Sei $(S_n, \mathcal{F}_n)_{n \geq 1}$ ein Supermartingal. Dann ist

$$\forall n \in \mathbb{N} : \mathbb{E}[S_{\tau \wedge n} - S_1] \leq 0 \quad (2.12)$$

wegen Korollar 2.23 (endlicher Zeithorizont).

Teil (i) folgt sofort aus (2.12) unter Beachtung von $\tau = \tau \wedge N$ \mathbb{P} -fast sicher.

Zum Nachweis von Teil (ii) beachten wir $|S_n| < K \forall n \in \mathbb{N} \Rightarrow \mathbb{E}[|S_{\tau \wedge n}|] \leq K$. Ferner folgt aus $\mathbb{P}(\tau < \infty) = 1$, dass $S_{\tau \wedge n} \rightarrow S_\tau$ \mathbb{P} -fast sicher für $n \rightarrow \infty$. Daraus folgt $\mathbb{E}[S_{\tau \wedge n}] \rightarrow \mathbb{E}[S_\tau]$ für $n \rightarrow \infty$. Zusammen mit (2.12) ergibt sich die Behauptung.

Teil (iii) ist eine Übungsaufgabe. ■

Beispiel 2.31 (Petersburger Paradoxon)

Wir betrachten das Roulette mit der sogenannten Verdopplungsstrategie. Angenommen, wir setzen stets auf Rot. Bei Gewinn erhalten wir dann den doppelten Einsatz. Wir starten die Spielserie mit 1 Euro Einsatz. Falls wir verlieren, verdoppeln wir den Einsatz. Gewinnen wir jetzt (in Spielrunde 2), so erhalten wir $2 \cdot 2 = 4$ Euro. Unser Gewinn ist damit $4 - (1 + 2) = 1$ Euro. Leicht rechnet man induktiv nach, dass dies für jede beliebige Spielrunde n richtig bleibt, wenn wir die Verdopplungsstrategie konsequent beibehalten. Gewinnen wir nämlich in Spielrunde n , so ist unser Gesamtgewinn gegeben durch

$$2 \cdot 2^{n-1} - \sum_{i=0}^{n-1} 2^i = 2^n - \frac{1 - 2^n}{1 - 2} = 2^n + 1 - 2^n = 1.$$

Fragen:

1. Wie viele Spielrunden braucht man im Mittel bis zum Gewinn?
2. Wie viel Einsatzkapital ist im Mittel bis zum Gewinn erforderlich?

Nehmen wir dazu vereinfachend an, dass $\mathbb{P}(\text{Rot}) = 1/2$ in jeder Spielrunde gelte. Betrachte die iid. Folge $(X_n)_{n \geq 1}$ mit $X_n = \mathbf{1}_{\{n\text{-tes Spiel verloren}\}}$. Sei $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$, $n \geq 1$, und definiere den gesamten verlorenen Einsatz bis nach der n -ten Spielrunde vermittels

$$S_n = \sum_{i=1}^n 2^{i-1} X_i,$$

sowie die Stoppzeit $\tau = \inf\{n \in \mathbb{N} : X_n = 0\}$.

Für alle $n \in \mathbb{N}$ gilt, dass wir im Falle $\tau = n$ einen Euro gewinnen. Ferner ist $(S_n, \mathcal{F}_n)_{n \geq 1}$ ein Submartingal, denn

$$\begin{aligned}
\mathbb{E}[S_{n+1}|\mathcal{F}_n] &= \mathbb{E}[S_n + 2^n X_{n+1}|\mathcal{F}_n] \\
&= S_n + 2^n \mathbb{E}[X_{n+1}] \\
&= S_n + 2^{n-1} \geq S_n.
\end{aligned}$$

Antwort auf Frage 1):

$$\begin{aligned}
\mathbb{E}[\tau] &= \sum_{n \geq 1} n 2^{-n} = \lim_{n \rightarrow \infty} \sum_{k=1}^n k \left(\frac{1}{2}\right)^k \\
&= \lim_{n \rightarrow \infty} \frac{\left(\frac{1}{2}\right)^{n+2} - (n+1)\left(\frac{1}{2}\right)^{n+1} + \frac{1}{2}}{\frac{1}{4}} \\
&= \lim_{n \rightarrow \infty} 2 + \left(\frac{1}{2}\right)^n - (n+1) \left(\frac{1}{2}\right)^{n-1} = 2.
\end{aligned}$$

Antwort auf Frage 2):

$$\begin{aligned}
\mathbb{E}[S_\tau] &= \sum_{n \geq 1} \left(\sum_{i=1}^n 2^{i-1} \right) \mathbb{P}(\tau = n) \\
&= \sum_{n \geq 1} \left(\sum_{\ell=0}^{n-1} 2^\ell \right) 2^{-n} \\
&= \sum_{n \geq 1} \frac{2^n - 1}{2 - 1} 2^{-n} = +\infty.
\end{aligned}$$

2.4 Martingalkonvergenzsätze

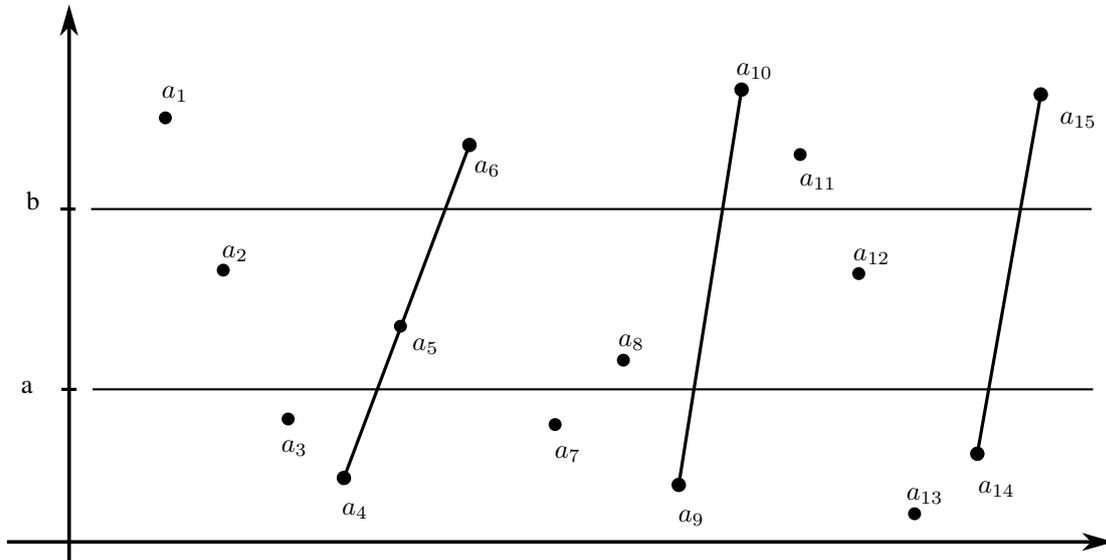
Definition 2.32

Seien a_1, \dots, a_n und $a < b$ reelle Zahlen. Dann heißt

$$U[a, b] := \sup\{m \mid \exists 1 \leq i_1 < \dots < i_{2m} \leq n : \forall 1 \leq \lambda \leq m : a_{i_{2\lambda-1}} \leq a, a_{i_{2\lambda}} \geq b\}$$

die Anzahl der aufsteigenden Überquerungen des Intervalls $[a, b]$ durch a_1, \dots, a_n .

Abbildung 2.33



Hier ist $U[a, b] = m = 3$. Wähle hierzu etwa $i_1 = 4$, $i_2 = 6$, $i_3 = 9$, $i_4 = 10$, $i_5 = 14$, $i_6 = 15$.

Lemma 2.34

Sei $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$, $N \in \mathbb{N}$, ein Submartingal mit endlichem Zeithorizont auf $(\Omega, \mathcal{F}, \mathbb{P})$. Bezeichne $U[a, b]$ die zufällige Anzahl der aufsteigenden Überquerungen eines Intervalls $[a, b]$ durch S_1, \dots, S_N . Dann gilt die Ungleichungskette

$$\begin{aligned}
 (b - a)\mathbb{E}[U[a, b]] &\leq \mathbb{E}[S_N - a] & (2.13) \\
 &\leq \mathbb{E}[(S_N - a)^+] \\
 &\leq \mathbb{E}[S_N^+] + |a|.
 \end{aligned}$$

Beweis: Für festes $\omega \in \Omega$ ermittelt man $(U[a, b])(\omega)$ nach dem folgenden Auszählverfahren: Suche den kleinsten Index $1 \leq i \leq N$ mit $S_i \leq a$, anschließend den nächst größeren Index $j > i$ mit $S_j \geq b$, usw. Über dieses Auszählverfahren definieren wir eine Folge von Stoppzeiten:

$$\tau_0 := 0,$$

$$\tau_1 := \begin{cases} \min\{1 \leq n \leq N : S_n \leq a\}, & \text{falls } \{1 \leq n \leq N : S_n \leq a\} \neq \emptyset \\ N, & \text{sonst} \end{cases}$$

$$\tau_2 := \begin{cases} \min\{\tau_1 \leq n \leq N : S_n \geq b\}, & \text{falls } \{\tau_1 \leq n \leq N : S_n \geq b\} \neq \emptyset \\ N, & \text{sonst} \end{cases}$$

⋮

$$\tau_{2m-1} := \begin{cases} \min\{\tau_{2m-2} \leq n \leq N : S_n \leq a\}, & \text{falls } \{\tau_{2m-2} \leq n \leq N : S_n \leq a\} \neq \emptyset \\ N, & \text{sonst} \end{cases}$$

$$\tau_{2m} := \begin{cases} \min\{\tau_{2m-1} \leq n \leq N : S_n \geq b\}, & \text{falls } \{\tau_{2m-1} \leq n \leq N : S_n \geq b\} \neq \emptyset \\ N, & \text{sonst} \end{cases}$$

usw.

Im Folgenden können wir o.B.d.A. annehmen, dass $S_n \geq a$ für alle $1 \leq n \leq N$ ist, denn ansonsten können wir zum Submartingal $T_n = \max(S_n, a)$ übergehen, vgl. Korollar 2.7.(ii).

Öffensichtlich ist $0 = \tau_0 \leq \tau_1 \leq \dots \leq N < \infty$ eine wachsende Folge von Stoppzeiten zu $(\mathcal{F}_n)_{1 \leq n \leq N}$. Damit ist, gemäß Bemerkung 2.24, $(S'_n, \mathcal{F}_n)_{0 \leq n \leq N}$ wieder ein Submartingal, wobei wir $S_0 := a, S'_n := S_{\tau_n}$ setzen. Ferner gilt

$$\begin{aligned} S_N - S_0 &= \sum_{k=0}^{N-1} (S'_{k+1} - S'_k) \\ &= \sum_{k:0 < 2k \leq N} (S'_{2k} - S'_{2k-1}) + \sum_{k:0 \leq 2k \leq N-1} (S'_{2k+1} - S'_{2k}) \\ &=: I_{1,N} + I_{2,N}. \end{aligned}$$

Aus der Wahl der Stoppzeiten erhalten wir, dass

$$I_{1,N} \geq (b - a)U[a, b]$$

ist. Aus der Submartingaleigenschaft von $(S'_n, \mathcal{F}_n)_{0 \leq n \leq N}$ folgt zudem $\mathbb{E}[S'_{2k+1} - S'_{2k}] \geq 0$ für alle Summanden von $I_{2,N}$ und damit $\mathbb{E}[I_{2,N}] \geq 0$. Zusammengenommen ergibt sich

$$\mathbb{E}[S_N - a] = \mathbb{E}[I_{1,N}] + \mathbb{E}[I_{2,N}] \geq (b - a)\mathbb{E}[U[a, b]], \text{ also (2.13).}$$

Der Rest der Ungleichungskette folgt aus der Definition des Positivteils einer Funktion. ■

Satz 2.35 (Martingalkonvergenzsatz von Doob)

Sei $(S_n, \mathcal{F}_n)_{n \geq 1}$ ein Submartingal auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit der Eigenschaft

$$\sup_{n \geq 1} \mathbb{E}[S_n^+] \leq K < \infty. \quad (2.14)$$

Dann existiert eine Zufallsvariable $S_\infty \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ mit $S_n \rightarrow S_\infty$ \mathbb{P} -fast sicher.

Beweis: Definiere

$$A := \{\omega \in \Omega : \liminf_{n \rightarrow \infty} S_n(\omega) = \limsup_{n \rightarrow \infty} S_n(\omega)\} = \{\liminf_{n \rightarrow \infty} S_n = \limsup_{n \rightarrow \infty} S_n\}.$$

Es genügt zu zeigen: $\mathbb{P}(A) = 1$.

Seien dazu $a, b \in \mathbb{R}$ mit $a < b$ beliebig ausgewählt und sei

$$A(a, b) := \{\liminf_{n \rightarrow \infty} S_n < a < b < \limsup_{n \rightarrow \infty} S_n\}.$$

Offenbar gilt für jedes $\omega \in A(a, b)$, dass $U[a, b](\omega) = \infty$. Daraus folgt aber, dass

$$\mathbb{E}[\mathbf{1}\{A(a, b)\}U[a, b]] = \begin{cases} \infty, & \text{falls } \mathbb{P}(A(a, b)) > 0, \\ 0, & \text{falls } \mathbb{P}(A(a, b)) = 0. \end{cases} \quad (2.15)$$

Nun liefert aber Lemma 2.34, dass

$$\mathbb{E}[\mathbf{1}\{A(a, b)\}U[a, b]] \leq \mathbb{E}[U[a, b]] \leq \sup_{n \geq 1} \frac{\mathbb{E}[S_n^+ + |a|]}{b - a} \stackrel{(2.14)}{\leq} \frac{K + |a|}{b - a} < \infty.$$

Zusammen mit (2.15) folgen wir, dass $\mathbb{P}(A(a, b)) = 0$ für alle $a < b$. Also ist

$$\begin{aligned} \mathbb{P}(A^c) &= \mathbb{P}(\liminf_{n \rightarrow \infty} S_n < \limsup_{n \rightarrow \infty} S_n) \\ &= \mathbb{P}\left(\bigcup_{a, b \in \mathbb{Q}: a < b} A(a, b)\right) \\ &\leq \sum_{a, b \in \mathbb{Q}: a < b} \mathbb{P}(A(a, b)) = 0 \end{aligned}$$

und damit $\mathbb{P}(A) = 1$. Folglich existiert eine Zufallsvariable S_∞ mit $S_n \rightarrow S_\infty$ \mathbb{P} -fast sicher für $n \rightarrow \infty$.

Zum Nachweis, dass $S_\infty \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ ist, rechnen wir

$$\mathbb{E}[|S_\infty|] = \mathbb{E}[\liminf_{n \rightarrow \infty} |S_n|] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[|S_n|]$$

nach dem Lemma von Fatou aus der Maßtheorie.

Beachte schließlich, dass $|S_n| = 2S_n^+ - S_n$ und damit $\mathbb{E}[|S_n|] = 2\mathbb{E}[S_n^+] - \mathbb{E}[S_n] \leq 2K - \mathbb{E}[S_1]$ für alle $n \in \mathbb{N}$. ■

Abschließend charakterisieren wir im Martingal-Fall noch die Zufallsvariable S_∞ näher.

Ziel dabei:

$\forall n \geq 1 : S_n = \mathbb{E}[S_\infty | \mathcal{F}_n] \mathbb{P}$ - fast sicher.

Dies führt aber nach Definition 2.20 auf die Untersuchung der Regularität von $(S_n, \mathcal{F}_n)_{n \geq 1}$.

Wir werden sehen, dass dazu die Eigenschaft der gleichgradigen Integrierbarkeit entscheidend ist.

Erinnerung 2.36

Sei $\mathcal{H} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ eine Menge integrierbarer Funktionen. Dann nennt man \mathcal{H} gleichgradig integrierbar, falls gilt:

$$\forall \varepsilon > 0 : \exists K \equiv K(\varepsilon) > 0 : \sup_{X \in \mathcal{H}} \int_{\{|X| > K\}} |X| d\mathbb{P} < \varepsilon.$$

Beispiel 2.37

a) Sei $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$. Dann ist $\mathcal{H} = \{X\}$ gleichgradig integrierbar. Betrachte dazu die Folge $X_n := |X| \mathbf{1}_{\{|X| \geq n\}}$, $n \geq 1$. Offenbar konvergiert $X_n \rightarrow 0 \mathbb{P}$ - fast sicher für $n \rightarrow \infty$. Ferner besitzt $\{X_n\}_{n \in \mathbb{N}}$ in $|X|$ eine integrierbare Majorante. Daher folgt mit majorisierter Konvergenz, dass

$$\lim_{n \rightarrow \infty} \int_{\{|X| > K\}} |X| d\mathbb{P} = \lim_{n \rightarrow \infty} \int X_n d\mathbb{P} = \int \lim_{n \rightarrow \infty} X_n d\mathbb{P} = 0.$$

Demnach gilt nach Definition des Grenzwertes, dass

$$\forall \varepsilon > 0 \exists n(\varepsilon) : \int_{\{|X| > n(\varepsilon)\}} |X| d\mathbb{P} < \varepsilon,$$

also die gleichgradige Integrierbarkeit von \mathcal{H} .

b) Sei $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, 1], \mathcal{B}([0, 1]), \lambda)$ mit dem Lebesguemaß (stetige Gleichverteilung) λ , und $\mathcal{H} = \{X_n : n \geq 1\}$ mit $X_n = n \mathbf{1}_{[0, 1/n]}$.

Dann ist $\mathcal{H} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$, denn

$$\forall n \in \mathbb{N} : \mathbb{E}[X_n] = n \int \mathbf{1}_{[0, 1/n]} d\lambda = n\lambda([0, 1/n]) = 1.$$

Allerdings ist \mathcal{H} nicht gleichgradig integrierbar. Sei nämlich $K > 0$ eine beliebige Konstante. Wir rechnen:

$$\begin{aligned} \int_{\{|X_n|>K\}} |X_n| d\mathbb{P} &= n \int_{\{\mathbf{1}_{[0,1/n]}>K/n\}} \mathbf{1}_{[0,1/n]} d\lambda = n \int_{\{n>K\}} \mathbf{1}_{[0,1/n]} d\lambda \\ &= \int_{\{n>K\} \cap [0,1/n]} n d\lambda = \int_{[0,1/n]} n d\lambda - \int_{\{n \leq K\} \cap [0,1/n]} n d\lambda \\ &\geq 1 - \int_{[0,1/n]} K d\lambda = 1 - \frac{K}{n}. \end{aligned}$$

Also ist

$$\sup_{n \geq 1} \int_{\{|X_n|>K\}} |X_n| d\mathbb{P} \equiv 1$$

für alle $K > 0$ und somit \mathcal{H} nicht gleichgradig integrierbar.

Lemma 2.38

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $\mathcal{H} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ eine Menge integrierbarer Zufallsvariablen. Dann ist \mathcal{H} gleichgradig integrierbar, falls eine der beiden folgenden Bedingungen erfüllt ist.

(i) \mathcal{H} ist beschränkt in $L_p(\Omega, \mathcal{F}, \mathbb{P})$ für ein $p > 1$. Das heißt, $\exists A > 0$ mit $\sup_{X \in \mathcal{H}} \mathbb{E}[|X|^p] < A$.

(ii) Es gibt eine Zufallsvariable $Y \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ mit $\sup_{X \in \mathcal{H}} |X(\omega)| \leq Y(\omega) \forall \omega \in \Omega$.

Beweis:

zu (i):

Sei $K > 0$ beliebig. Dann ist

$$\begin{aligned} \sup_{X \in \mathcal{H}} \int_{\{|X|>K\}} |X| d\mathbb{P} &\leq \sup_{X \in \mathcal{H}} \int_{\{|X|>K\}} K^{1-p} |X|^{p-1} |X| d\mathbb{P} \\ &= \sup_{X \in \mathcal{H}} \int_{\{|X|>K\}} K^{1-p} |X|^p d\mathbb{P} \\ &\leq K^{1-p} A \rightarrow 0, \quad K \rightarrow \infty. \end{aligned}$$

zu (ii):

Folgt aus Beispiel 2.37.a) unter Beachtung von

$$\sup_{X \in \mathcal{H}} \int_{\{|X|>K\}} |X| d\mathbb{P} \leq \int_{\{|Y|>K\}} |Y| d\mathbb{P}.$$

■

Satz 2.39 (Satz 6.4.13 in Gaenssler and Stute (1977))

Sei $(S_n, \mathcal{F}_n)_{n \geq 1}$ ein Martingal auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann sind die folgenden Aussagen äquivalent.

(i) $\mathcal{H} = \{S_n : n \geq 1\}$ ist gleichgradig integrierbar.

(ii) $\exists S_\infty \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ mit $S_n \rightarrow S_\infty$ \mathbb{P} -fast sicher und im Mittel, $n \rightarrow \infty$.

(iii) $\mathcal{H} = \{S_n : n \geq 1\}$ ist regulär.

Korollar 2.40

Sei $(\mathcal{F}_n)_{n \geq 1}$ eine Filtration auf $(\Omega, \mathcal{F}, \mathbb{P})$ und $X \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ eine Zufallsvariable. Dann ist $(S_n, \mathcal{F}_n)_{n \geq 1}$ mit

$$\forall n \in \mathbb{N} : S_n := \mathbb{E}[X | \mathcal{F}_n]$$

ein reguläres Martingal mit $S_\infty = \mathbb{E}[X | \mathcal{F}_\infty]$.

2.5 Anwendungen der Martingaltheorie

2.5.1 Kolmogoroff'sches 0-1 Gesetz

Erinnerung 2.41

Sei $(X_n)_{n \geq 1}$ eine Folge von stochastisch unabhängigen Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$. Dann heißt

$$\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(\{X_m\} : m > n)$$

die terminale σ -Algebra zu $(X_n)_{n \geq 1}$ und jedes $A \in \mathcal{T}$ ein terminales Ereignis.

Beispielweise ist

$$A := \{\omega \in \Omega : \liminf_{n \rightarrow \infty} X_n(\omega) = \limsup_{n \rightarrow \infty} X_n(\omega)\}$$

ein terminales Ereignis.

Satz 2.42 (siehe Satz 1.29)

Unter der Voraussetzungen von Erinnerung 2.41 gilt $\forall A \in \mathcal{T} : \mathbb{P}(A) \in \{0, 1\}$.

Beweis: (mit Martingalargumenten)

Sei $A \in \mathcal{T}$ beliebig vorgegeben. Damit liegt A in $\sigma(X_{n+1}, X_{n+2}, \dots)$. Diese σ -Algebra ist aber unabhängig von $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ wegen der stochastischen Unabhängigkeit der $(X_n)_{n \geq 1}$. Also ist \mathbb{P} -fast sicher

$$\mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A] = \mathbb{E}[\mathbf{1}_A | \mathcal{F}_n]. \quad (2.16)$$

Nach Korollar 2.40 wissen wir, dass

$$S_n := \mathbb{E}[\mathbf{1}_A | \mathcal{F}_n], \quad n \geq 1,$$

ein reguläres Martingal ist, das gegen

$$S_\infty = \mathbb{E}[\mathbf{1}_A | \mathcal{F}_\infty]$$

konvergiert. Wegen (2.16) ist aber $S_n \equiv \mathbb{P}(A)$ und damit folgt $\mathbb{P}(A) = \mathbb{E}[\mathbf{1}_A | \mathcal{F}_\infty]$. Nach Definition von \mathcal{F}_∞ ist jedes X_n messbar bezüglich \mathcal{F}_∞ . Also gilt $\forall n \geq 1 : \sigma(X_{n+1}, X_{n+2}, \dots) \subseteq \mathcal{F}_\infty$ und damit $\mathcal{T} \subseteq \mathcal{F}_\infty$.

Demnach ist $\mathbb{E}[\mathbf{1}_A | \mathcal{F}_\infty] = \mathbf{1}_A$ \mathbb{P} -fast sicher nach Satz 1.13.d). Insgesamt erhalten wir $\mathbb{P}(A) = \mathbf{1}_A$. Indikatorvariablen können aber nur die Werte Null oder Eins annehmen. ■

2.5.2 Starkes Gesetz der großen Zahlen

Definition 2.43

Sei $(S_n)_{n \geq 1}$ eine Folge von integrierbaren Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbb{P})$. Ferner sei $(\mathcal{F}_n)_{n \geq 1}$ eine monoton fallende Folge von σ -Algebren auf Ω mit $\mathcal{F}_n \subseteq \mathcal{F} \forall n \geq 1$ derart, dass $S_n \forall n \in \mathbb{N}$ \mathcal{F}_n -messbar ist. Falls dann $\forall n \geq 1 : \mathbb{E}[S_{n-1} | \mathcal{F}_n] = S_n$ ist, so nennt man (S_n, \mathcal{F}_n) ein reverses Martingal bzw. ein Rückwärtsmartingal.

Bemerkung 2.44

Wählt man als Indexmenge eines Martingals (S_n, \mathcal{F}_n) nicht \mathbb{N} sondern $-\mathbb{N}$, und setzt man sodann $\tilde{S}_n := S_{-n}$ und $\tilde{\mathcal{F}}_n := \mathcal{F}_{-n}$, so ist $(\tilde{S}_n, \tilde{\mathcal{F}}_n)_{n \geq 1}$ ein reverses Martingal.

Deswegen gelten der Martingalkonvergenzsatz 2.35 sowie Satz 2.39 entsprechend für reverse Martingale. Analoges gilt für Sub- und Supermartingale.

Korollar 2.45 (zu Satz 2.39)

Jedes Rückwärtsmartingal $(S_n, \mathcal{F}_n)_{n \geq 1}$ auf $(\Omega, \mathcal{F}, \mathbb{P})$ konvergiert \mathbb{P} -fast sicher und im Mittel gegen $S_\infty = \mathbb{E}[S_1 | \bigcap_{n \geq 1} \mathcal{F}_n]$.

Satz 2.46 (Starkes Gesetz der großen Zahlen)

Sei $(X_n)_{n \geq 1} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ eine Folge stochastisch unabhängiger und identisch verteilter Zufallsvariablen. Dann gilt

$$\bar{X}_n := n^{-1} \sum_{i=1}^n X_i \rightarrow \mu := \mathbb{E}[X_1] \quad \mathbb{P}\text{-f.s.}, n \rightarrow \infty.$$

Beweis: (mit Hilfe des Martingalkonvergenzsatzes)

Sei $S_n := \sum_{i=1}^n X_i$ und $\mathcal{F}_n := \sigma(S_n, X_{n+1}, X_{n+2}, \dots)$ für $n \in \mathbb{N}$. Offenbar ist S_n messbar bzgl. \mathcal{F}_n , und $(\mathcal{F}_n)_{n \geq 1}$ ist fallend, d.h. $\mathcal{F}_{n+1} \subseteq \mathcal{F}_n$. Letzteres sieht man dadurch, dass $S_{n+1} = S_n + X_{n+1}$ ist und die rechte Seite \mathcal{F}_n -messbar ist. Damit sind alle Zufallsvariablen, die \mathcal{F}_{n+1} erzeugen, \mathcal{F}_n -messbar und somit $\mathcal{F}_{n+1} \subseteq \mathcal{F}_n$. Da $(X_i)_{1 \leq i \leq n} \perp (X_k)_{k > n}$ ist, gilt $\forall 1 \leq i \leq n : \mathbb{E}[X_i | \mathcal{F}_n] = \mathbb{E}[X_i | \mathcal{F}_i]$.

Ferner ist aus Symmetriegründen

$$\forall 1 \leq i \leq n : \mathbb{E}[X_i | S_n] = \mathbb{E}[X_1 | S_n]. \quad (2.17)$$

Folglich ist

$$n\mathbb{E}[X_1 | \mathcal{F}_n] = n\mathbb{E}[X_1 | S_n] \stackrel{(2.17)}{=} \sum_{i=1}^n \mathbb{E}[X_i | S_n] = \mathbb{E} \left[\sum_{i=1}^n X_i | S_n \right] = S_n$$

und damit

$$\forall n \in \mathbb{N} : \forall 1 \leq i \leq n : \mathbb{E}[X_i | \mathcal{F}_n] = \bar{X}_n. \quad (2.18)$$

Betrachte nun die adaptierte Folge $(\bar{X}_n, \mathcal{F}_n)_{n \geq 1}$. Wie bereits gesehen ist $(\mathcal{F}_n)_{n \geq 1}$ monoton fallend. Ferner gilt für alle $m < n$:

$$\mathbb{E}[\bar{X}_m | \mathcal{F}_n] = \frac{1}{m} \sum_{i=1}^m \mathbb{E}[X_i | \mathcal{F}_n] \stackrel{(2.18)}{=} \frac{1}{m} \sum_{i=1}^m \bar{X}_n = \bar{X}_n.$$

Also ist $(\bar{X}_n, \mathcal{F}_n)_{n \geq 1}$ ein reverses Martingal. Damit gilt nach Korollar 2.45, dass

$$\bar{X}_n \rightarrow \bar{X}_\infty \quad \mathbb{P}\text{-f.s. für } n \rightarrow \infty.$$

Bleibt zu zeigen: $\bar{X}_\infty = \mu$ \mathbb{P} -fast sicher.

Betrachten wir dazu die Verteilungsfunktion von \bar{X}_∞ . Es gilt

$$\mathbb{P}(\bar{X}_\infty \leq x) = \mathbb{P}(\limsup_{n \rightarrow \infty} \bar{X}_n \leq x), \quad x \in \mathbb{R}.$$

Nun ist aber nach Korollar 1.30 der Cesàro-Limes $\limsup_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n X_i$ \mathbb{P} -fast sicher konstant, sagen wir gleich $x_0 \in \mathbb{R}$. Da x_0 eine Konstante ist, muss $x_0 = \mathbb{E}[\bar{X}_\infty]$ sein.

Schließlich folgern wir

$$x_0 = \mathbb{E}[\bar{X}_\infty] = \mathbb{E} \left[\mathbb{E}[X_1 | \bigcap_{n \geq 1} \mathcal{F}_n] \right] = \mathbb{E}[X_1] = \mu.$$

■

Anmerkung:

Korollar 2.45 liefert darüber hinaus, dass

$$\int |\bar{X}_n - \mu| d\mathbb{P} \rightarrow 0, \quad n \rightarrow \infty.$$

2.5.3 Die Gleichungen von Wald

Erinnerung 2.47

Seien $(X_i)_{1 \leq i \leq n}$ iid. Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}[X_1] = \mu$ und $\text{Var}(X_1) = \sigma^2 < \infty$.
Definiere $S_n := \sum_{i=1}^n X_i$. Dann gilt

$$\mathbb{E}[S_n] = n\mu \quad \text{und} \quad \text{Var}(S_n) = n\sigma^2.$$

Ziel: Analoge Ergebnisse für eine zufällige Anzahl an Summanden. Dies hat wichtige Anwendungen in der Sequentialstatistik.

Satz 2.48 (1. Wald'sche Gleichung)

Sei $(X_n)_{n \geq 1} \subset L_1(\Omega, \mathcal{F}, \mathbb{P})$ eine Folge von iid. Zufallsvariablen mit $X_n \geq 0$ und $\mathbb{E}[X_1] = \mu$.
Bezeichne $\forall n \in \mathbb{N} : S_n := \sum_{i=1}^n X_i$, $\mathcal{F}_n := \sigma(X_1, \dots, X_n)$, und sei τ eine Stoppzeit bezüglich $(\mathcal{F}_n)_{n \geq 1}$.

Dann gilt:

$$\mathbb{E}[S_\tau] = \mathbb{E}[\tau]\mu,$$

wobei beide Seiten gleich $+\infty$ sein können.

Beweis: Für $n \geq 1$ setze $S'_n := S_n - n\mu$. Dann ist nach Beispiel 2.4.(a) die Folge $(S'_n, \mathcal{F}_n)_{n \geq 1}$ ein Martingal. Darüber hinaus ist gemäß Satz 2.28 (Transformation durch Stoppzeiten) auch $(S'_{\tau \wedge n}, \mathcal{F}_n)_{n \geq 1}$ ein Martingal. Es gilt daher:

$$\mathbb{E}[S'_{\tau \wedge n}] = \mathbb{E}[S'_{\tau \wedge 1}] = \mathbb{E}[S'_1] = 0.$$

Nun ist aber nach Definition

$$\mathbb{E}[S'_{\tau \wedge n}] = \mathbb{E}[S_{\tau \wedge n}] - \mu \mathbb{E}[\tau \wedge n] \Rightarrow \forall n \geq 1 : \mathbb{E}[S_{\tau \wedge n}] = \mu \mathbb{E}[\tau \wedge n].$$

Lässt man nun $n \rightarrow \infty$ laufen ergibt sich das Ergebnis aus dem Satz von der monotonen Konvergenz. ■

Korollar 2.49

Sei $(X_n)_{n \geq 1}$ wie in Satz 2.48, jedoch nicht zwingend nicht-negativ. Sei ferner τ eine Stoppzeit bezüglich der natürlichen Filtration von $(X_n)_{n \geq 1}$ mit $\mathbb{E}[\tau] < \infty$. Dann ist $\mathbb{E}[|S_\tau|] < \infty$ und es gilt

$$\mathbb{E}[S_\tau] = \mathbb{E}[\tau]\mu.$$

Beweis: Dass $S_\tau \in L_1(\Omega, \mathcal{F}, \mathbb{P})$ ist verifiziert man durch

$$\mathbb{E}[|S_\tau|] \leq \mathbb{E} \left[\sum_{i=1}^{\tau} |X_i| \right] = \mathbb{E}[\tau] \mathbb{E}[|X_1|] < \infty$$

nach 1. Wald'scher Gleichung.

Der Rest ergibt sich durch die Zerlegung $\forall i \geq 1 : X_i = X_i^+ - X_i^-$ und separate Anwendung von Satz 2.48. ■

Satz 2.50 (2. Wald'sche Gleichung)

Sei $(X_n)_{n \geq 1}$ eine iid. - Folge von Zufallsvariablen aus $L_2(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}[X_1] = 0$ (o.B.d.A.) und $\mathbb{E}[X_1^2] = \sigma^2$. Seien ferner S_n und \mathcal{F}_n definiert wie in Satz 2.48, und sei τ eine Stoppzeit zur Filtration $(\mathcal{F}_n)_{n \geq 1}$ mit $\mathbb{E}[\tau] < \infty$. Dann ist

$$\mathbb{E}[S_\tau^2] = \mathbb{E}[\tau]\sigma^2.$$

Beweis: Folgt aus Korollar 2.49 durch Übergang von X_i zu X_i^2 , $i \geq 1$. Für Details siehe Bauer (2002), Satz 17.7. ■

Anwendung 2.51 (Elementares Erneuerungstheorem)

Bezeichne X_1 die (zufällige) Wartezeit bis zum ersten Ausfall eines Bauteils, welches dann sofort erneuert wird. Die Wartezeit nach den ersten Ausfall bis zum zweiten Ausfall werde durch die Zufallsvariable X_2 beschrieben, usw.

Annahme: $(X_n)_{n \in \mathbb{N}}$ bildet eine iid. - Folge von Zufallsvariablen aus $L_1(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}[X_1] = \mu$. Der Wert μ stellt somit die mittlere Wartezeit zwischen zwei Ausfällen des Bauteils dar. Sei ferner $S_n = \sum_{i=1}^n X_i$.

Uns interessiert die (zufällige) Anzahl der notwendigen Erneuerungen des Bauteils bis zum Zeitpunkt t .

Fassen wir diese Anzahl als Funktion von t auf, so ergibt sich ein zeitstetiger stochastischer Prozess

$$\begin{aligned} N : \Omega \times \mathbb{R}_+ &\rightarrow \mathbb{N}_0 \\ (\omega, t) &\mapsto N(\omega, t) = |\{n : S_n(\omega) \leq t\}|. \end{aligned}$$

Falls $0 < \mu < \infty$ ist, so gilt

$$\frac{\mathbb{E}[N(\cdot, t)]}{t} \rightarrow \frac{1}{\mu}, \quad t \rightarrow \infty.$$

Beweis: Wir definieren für gegebenes $t > 0$ die Stoppzeit $\tau_t := \inf\{n \geq 1 : S_n > t\}$ bezüglich der natürlichen Filtration. Offenbar ist $N(\cdot, t) = \tau_t - 1$.

Genügt zu zeigen:

$$\frac{\mathbb{E}[\tau_t]}{t} \rightarrow \frac{1}{\mu}, \quad t \rightarrow \infty.$$

Aus der Definition von τ_t und der 1. Wald'schen Gleichung folgt

$$\forall t > 0 : t \leq \mathbb{E}[S_{\tau_t}] = \mu \mathbb{E}[\tau_t] \Rightarrow \liminf_{t \rightarrow \infty} \frac{\mathbb{E}[\tau_t]}{t} \geq \frac{1}{\mu}.$$

Bleibt zu zeigen:

$$\limsup_{t \rightarrow \infty} \frac{\mathbb{E}[\tau_t]}{t} \leq \frac{1}{\mu}. \quad (2.19)$$

Zum Nachweis von (2.19) betrachte zunächst beschränkte Wartezeiten, d.h., $0 < X_n \leq c$, $c > 0$. Dann ist $S_{\tau_t} \leq t + c$ und die 1. Wald'sche Gleichung liefert

$$t + c \geq \mathbb{E}[S_{\tau_t}] = \mu \mathbb{E}[\tau_t].$$

Dies impliziert indes (2.19).

Im allgemeinen Fall gehen wir zunächst zu den trunkeierten Zufallsvariablen $X_n^c := X_n \wedge c$, $n \geq 1$, über.

Wegen $S_n^c := \sum_{i=1}^n X_i^c \leq S_n$ ist $\tau_t \leq \tau_t^c := \inf\{n \geq 1 : S_n^c > t\}$.

Da für $\{X_n^c\}$ indes (2.19) schon nachgewiesen ist, erhalten wir

$$\limsup_{t \rightarrow \infty} \frac{\mathbb{E}[\tau_t]}{t} \leq \limsup_{t \rightarrow \infty} \frac{\mathbb{E}[\tau_t^c]}{t} \stackrel{(2.19)}{\leq} \frac{1}{\mathbb{E}[X_1^c]}.$$

Lassen wir nun $c \rightarrow \infty$ laufen, so ergibt sich schließlich die Aussage im allgemeinen Fall. ■

2.6 Optimales Stoppen (bei endlichem Zeithorizont)

Problemstellung 2.52

Sei $(S_n, \mathcal{F}_n)_{n \geq 1}$ eine adaptierte Folge in $L_1(\Omega, \mathcal{F}, \mathbb{P})$. Bezeichne

$$C := \{\tau : \tau \text{ endliche Stoppzeit bzgl. } (\mathcal{F}_n)_{n \geq 1}, S_\tau \text{ ist integrierbar}\}.$$

Anmerkung: Falls τ nur endlich viele Werte annehmen kann, sagen wir $\tau(\omega) \in \{1, \dots, N\}$, so ist S_τ stets integrierbar, denn dann gilt

$$\mathbb{E}[|S_\tau|] \leq \max_{1 \leq n \leq N} \mathbb{E}[|S_n|] < \infty.$$

Betrachte nun

$$V := \sup_{\tau \in C} \mathbb{E}[S_\tau].$$

Das Ziel des optimalen Stoppen ist es, ein $\tau_0 \in C$ zu finden mit $V = \mathbb{E}[S_{\tau_0}]$, d.h., die im Mittel beste Strategie.

Typische Fragen dabei sind:

1) Wie lässt sich V berechnen?

2) Gibt es ein optimales τ_0 ?

3) Welche Struktur besitzt τ_0 ?

Beispiel 2.53

Sei $S_n(\omega) \equiv 1 - \frac{1}{n} \forall \omega \in \Omega$ und $n \geq 1$. Dann ist offensichtlich $V = 1$, aber $\forall \tau \in C : \mathbb{E}[S_\tau] < V = 1$. Also existiert keine optimale Stoppzeit in C für dieses Beispiel.

Wir beschränken uns im Folgenden auf endliche Zeithorizonte $1 \leq n \leq N$.

Korollar 2.54 (zu Korollar 2.23 und Bemerkung 2.24)

Sei $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ eine adaptierte Folge in $L_1(\Omega, \mathcal{F}, \mathbb{P})$ mit endlichem Zeithorizont. Dann gilt:

- (a) Ist $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ ein Supermartingal, so ist $V = \mathbb{E}[S_1]$, also $\tau_0 \equiv 1$.
- (b) Ist $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ ein Martingal, so ist $V = \mathbb{E}[S_\tau]$ unabhängig von τ , und daher sind alle Strategien gleichwertig hinsichtlich des Zielkriteriums V .
- (c) Ist $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ ein Submartingal, so ist $V = \mathbb{E}[S_N]$, also $\tau_0 \equiv N$.

Korollar 2.54 zeigt, dass das Problem des optimalen Stoppens bei endlichem Zeithorizont und Sub- / Super- / Martingalstruktur von $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ bereit gelöst ist.

Definition 2.55

Unter den Voraussetzungen von Problemstellung 2.52 sei

$$C_n^N := \{\tau : \tau \text{ Stoppzeit mit } n \leq \tau \leq N\}.$$

Da wir nur durch N beschränkte Stoppzeiten überhaupt betrachten, gilt unter dieser Maßgabe $C_1^N = C$.

Definiere ferner

$$V_n^N := \sup_{\tau \in C_n^N} \mathbb{E}[S_\tau] \Rightarrow V = V_1^N.$$

Satz 2.56 (Rückwärtsinduktion, Dynamisches Programm, Bellmann-Gleichung)

Für festes $N \in \mathbb{N}$ und unter den Voraussetzungen von Problemstellung 2.52 sei sukzessive die Folge $(\gamma_n^N)_{1 \leq n \leq N}$ definiert durch

$$\gamma_N^N := S_N, \quad \gamma_n^N := \max(S_n, \mathbb{E}[\gamma_{n+1}^N | \mathcal{F}_n]), \quad n = N - 1, \dots, 1. \quad (2.20)$$

Definiere anschließend für $1 \leq n \leq N$:

$$\tau_n^N := \inf\{i \geq n : S_i = \gamma_i^N\}.$$

Dann gilt für alle $1 \leq n \leq N$:

$$(i) \tau_n^N \in C_n^N$$

$$(ii) \forall \tau \in C_n^N : \mathbb{E}[S_{\tau_n^N} | \mathcal{F}_n] = \gamma_n^N \geq \mathbb{E}[S_\tau | \mathcal{F}_n]$$

$$(iii) V_n^N = \mathbb{E}[S_{\tau_n^N}].$$

Damit löst diese Konstruktion das Problem des optimalen Stoppens bei beschränktem Zeithorizont.

Beweis:

zu (i): Übungsaufgabe.

zu (ii): Wir führen den Beweis per Rückwärtsinduktion. Der Induktionsanfang $n = N$ ist trivial, da $C_N^N = \{\tau_N^N\}$ mit $\tau_N^N \equiv N$ ist.

Nehmen wir jetzt an, dass (ii) korrekt ist für $n, n+1, \dots, N$.

Seien dann $\tau \in C_{n-1}^N$ und $A \in \mathcal{F}_{n-1}$ beliebig ausgewählt. Dann ist $\tau' := \max(\tau, n) \in C_n^N$ und es gilt aufgrund der Induktionsannahme, dass

$$\begin{aligned} \int_A S_\tau d\mathbb{P} &= \int_{A \cap \{\tau=n-1\}} S_\tau d\mathbb{P} + \int_{A \cap \{\tau \geq n\}} S_{\tau'} d\mathbb{P} \\ &= \int_{A \cap \{\tau=n-1\}} S_{n-1} d\mathbb{P} + \int_{A \cap \{\tau \geq n\}} \mathbb{E}[S_{\tau'} | \mathcal{F}_{n-1}] d\mathbb{P}, \end{aligned} \quad (2.21)$$

da $A \cap \{\tau \geq n\} \in \mathcal{F}_{n-1}$.

Die Tower equation liefert nun, dass die rechte Seite von (2.21) gleich

$$\begin{aligned} &\int_{A \cap \{\tau=n-1\}} S_{n-1} d\mathbb{P} + \int_{A \cap \{\tau \geq n\}} \mathbb{E}[\mathbb{E}[S_{\tau'} | \mathcal{F}_n] | \mathcal{F}_{n-1}] d\mathbb{P} \\ &\stackrel{\text{(Induktionsvoraussetzung)}}{\leq} \int_{A \cap \{\tau=n-1\}} S_{n-1} d\mathbb{P} + \int_{A \cap \{\tau \geq n\}} \mathbb{E}[\gamma_n^N | \mathcal{F}_{n-1}] d\mathbb{P} \\ &\stackrel{\text{(Definition von } \gamma_{n-1}^N)}{\leq} \int_A \gamma_{n-1}^N d\mathbb{P} \text{ ist.} \end{aligned}$$

Also ist für jede Menge $A \in \mathcal{F}_{n-1}$

$$\int_A S_\tau d\mathbb{P} \leq \int_A \gamma_{n-1}^N d\mathbb{P}.$$

Daraus folgt aber, dass $\mathbb{E}[S_\tau | \mathcal{F}_{n-1}] \leq \gamma_{n-1}^N$ ist, denn γ_{n-1}^N ist \mathcal{F}_{n-1} -messbar.

Bleibt zu zeigen: Es gilt in obiger Rechnung überall Gleichheit für $\tau = \tau_{n-1}^N$.

Beachte dazu, dass

$$\mathbf{1}_{\{\tau_{n-1}^N \geq n\}} \left(\tau_{n-1}^N \right)' = \mathbf{1}_{\{\tau_{n-1}^N \geq n\}} \left(\tau_n^N \right) \text{ ist.}$$

Also kann das erste " \leq " für $\tau = \tau_{n-1}^N$ durch " $=$ " ersetzt werden.

Ferner ist

$$\gamma_{n-1}^N = \begin{cases} S_{n-1} & \text{auf } \{\tau_{n-1}^N = n-1\}, \\ \mathbb{E}[\gamma_n^N | \mathcal{F}_{n-1}] & \text{auf } \{\tau_{n-1}^N \geq n\}, \end{cases}$$

so dass für $\tau = \tau_{n-1}^N$ auch das zweite " \leq " durch " $=$ " ersetzt werden kann.

Dies komplettiert den Beweis von (ii).

zu (iii): Folgt aus (i) und (ii) durch Integration (Tower equation). ■

Bemerkung 2.57

Die Folge $(\gamma_n^N)_{1 \leq n \leq N}$ ist das kleinste Supermartingal, welches S_1, \dots, S_N majorisiert. Ferner ist τ_1^N der erste Zeitpunkt, zu dem sich beide Prozesse treffen.

Beweis: Übungsaufgabe ■

Anwendung 2.58 (Sekretärsproblem)

Gegeben seien N Objekte, die ihrer Qualität nach unterscheidbar sind. Wir können ihnen also aufgrund ihrer Qualität einen Rang zuordnen, wobei der Rang 1 der höchsten Qualität entsprechen soll. Die endgültige Zuordnung der Ränge ist aber erst möglich, nachdem alle N Objekte geprüft worden sind.

Um dieses Verfahren abzukürzen, soll direkt nach Prüfung des n -ten Objekts entschieden werden, dieses auszuwählen oder nicht. Dabei ist die Prüfungsreihenfolge der N Objekte rein zufällig.

Ziel: Optimale Strategie (d.h. Stoppzeit), die die Wahrscheinlichkeit dafür maximiert, das beste Objekt auszuwählen.

Mathematisches Modell (Laplace'scher Wahrscheinlichkeitsraum):

$\Omega = \mathcal{S}_N$ (Menge aller Permutationen von $\{1, \dots, N\}$)

$\mathcal{F} = 2^\Omega$ (Potenzmenge)

$\mathbb{P}(\{\omega\}) = \frac{1}{N!} \forall \omega \in \Omega.$

Für $\omega = (a_1, \dots, a_N)$ heißt $a_i = 1$, dass das beste Objekt an der i -ten Stelle erscheint, $1 \leq n \leq N$.

Bezeichne für ein solches ω die Zahl $y_n(\omega)$ den relativen (bzw. sequentiellen) Rang von a_n , $1 \leq n \leq N$, also $y_n(\omega) := \sum_{i=1}^n \mathbf{1}_{\{a_i \leq a_n\}}$.

Setze ferner $\mathcal{F}_n = \sigma(y_1, \dots, y_n) \forall 1 \leq n \leq N$.

Gesucht wird eine Stoppzeit τ_0 zu $(\mathcal{F}_n)_{1 \leq n \leq N}$ mit $\mathbb{P}(a_{\tau_0} = 1) = \sup_{\tau \in C} \mathbb{P}(a_\tau = 1)$, mit $C = C_1^N$ wie in Definition 2.55.

Beachte dazu, dass

$$\begin{aligned}\mathbb{P}(a_\tau = 1) &= \sum_{n=1}^N \int_{\{\tau=n\}} \mathbf{1}\{a_n = 1\} d\mathbb{P} \\ &= \sum_{n=1}^N \int_{\{\tau=n\}} \mathbb{E}[\mathbf{1}\{a_n = 1\} | \mathcal{F}_n] d\mathbb{P} = \mathbb{E}[S_\tau]\end{aligned}$$

mit

$$\forall 1 \leq n \leq N : S_n = \mathbb{P}(a_n = 1 | \mathcal{F}_n) = \mathbb{E}[\mathbf{1}\{a_n = 1\} | \mathcal{F}_n].$$

Damit ist das Sekretärsproblem auf die Voraussetzungen von Satz 2.56 zurückgeführt und es genügt, zur adaptierten Folge $(S_n, \mathcal{F}_n)_{1 \leq n \leq N}$ die Größen $(\gamma_n^N)_{1 \leq n \leq N}$ zu bestimmen.

Lemma 2.59

Unter dem Modell aus Anwendung 2.58 gilt:

- (i) $\forall 1 \leq n \leq N : \forall 1 \leq j \leq n : \mathbb{P}(y_n = j) = \frac{1}{n}$.
- (ii) Die sequentiellen Ränge y_1, \dots, y_N sind stochastisch unabhängig.
- (iii) $S_n = \mathbb{P}(a_n = 1 | \mathcal{F}_n) = \frac{n}{N} \mathbf{1}_{\{y_n=1\}}$.

Beweis: Übungsaufgabe ■

Satz 2.60 (Lösung des Sekretärsproblem)

Wir betrachten zunächst $(\gamma_n^N)_{1 \leq n \leq N}$. Da N fest vorgegeben ist schreiben wir abkürzend γ_n statt γ_n^N . Nach (2.20) in Satz 2.56 ist $\gamma_N = S_N$ und $\gamma_{N-1} = \max(S_{N-1}, \mathbb{E}[S_N | \mathcal{F}_{N-1}])$.

Nun folgt aber aus Lemma 2.59, dass $\mathbb{E}[S_N | \mathcal{F}_{N-1}] = \mathbb{E}[S_N]$, also $\gamma_{N-1} = \max(S_{N-1}, \mathbb{E}[\gamma_N])$.

Analog ermittelt man weiter $\forall 1 \leq n \leq N - 1: \gamma_n = \max(S_n, \mathbb{E}[\gamma_{n+1}])$. Damit ist die optimale Stoppzeit $\tau_0 = \tau_1^N$ gegeben durch

$$\tau_0 = \inf\{1 \leq n \leq N : S_n = \gamma_n\} = \inf\{1 \leq n \leq N : S_n \geq \mathbb{E}[\gamma_{n+1}]\},$$

mit $\mathbb{E}[\gamma_{N+1}] := 0$.

Es gilt:

$$\mathbb{E}[\gamma_n] = \frac{n-1}{N} \left(\sum_{k=n-1}^{N-1} \frac{1}{k} \right) \tag{2.22}$$

für alle n mit $\mathbb{E}[\gamma_{n+1}] \leq n/N$. Bezeichne n^* das kleinste n mit dieser Eigenschaft. Dann ist folglich unter Beachtung von Lemma 2.59.(iii)

$$\forall n < n^* : S_n \leq \frac{n}{N} < \mathbb{E}[\gamma_{n+1}]$$

und damit ist Stoppen vor dem Zeitpunkt n^* nicht möglich.

Wegen der Antitonicität der $\mathbb{E}[\gamma_{n+1}]$ mit $\mathbb{E}[\gamma_{n+1}] \leq n/N$ lässt sich τ_0 also wie folgt beschreiben:

1) Prüfe $n^* - 1$ Objekte, ohne Entscheidung für eines dieser Objekte.

2) Anschließend wähle das Objekt, das zum ersten Mal besser als alle vorangegangenen Objekte ist.

Beweis: Offenbar bleibt nur (2.22) zu zeigen. Dies geschieht mit Rückwärtsinduktion.

Induktionsanfang: $n = N$ mit $\mathbb{E}[\gamma_{N+1}] = 0 \leq 1 = \frac{N}{N}$:

$$\mathbb{E}[\gamma_N] = \mathbb{E}[S_N] \stackrel{\text{Lemma 2.59.(iii)}}{=} \mathbb{E}[\mathbf{1}_{\{y_N=1\}}] = \mathbb{P}(y_N = 1) \stackrel{\text{Lemma 2.59.(i)}}{=} \frac{1}{N} = \frac{N-1}{N} \frac{1}{N-1}.$$

Induktionsschritt:

$$\begin{aligned} \mathbb{E}[\gamma_n] &= \mathbb{E}[\max(S_n, \mathbb{E}[\gamma_{n+1}])] \\ &= \int_{\{y_n=1\}} S_n d\mathbb{P} + \int_{\{y_n \neq 1\}} \mathbb{E}[\gamma_{n+1}] d\mathbb{P} \\ &\stackrel{\text{(Induktionsvoraussetzung)}}{=} \frac{n}{N} \mathbb{P}(y_n = 1) + \frac{n}{N} \sum_{k=n}^{N-1} \frac{1}{k} \cdot \mathbb{P}(y_n \neq 1) \\ &\stackrel{\text{Lemma 2.59.(i)}}{=} \frac{n}{N} \cdot \frac{1}{n} + \frac{n}{N} \sum_{k=n}^{N-1} \frac{1}{k} \left(\frac{n-1}{n} \right) \\ &= \frac{1}{N} + \frac{n-1}{N} \sum_{k=n}^{N-1} \frac{1}{k} \\ &= \frac{n-1}{N} \left[\frac{1}{n-1} + \sum_{k=n}^{N-1} \frac{1}{k} \right] = \frac{n-1}{N} \sum_{k=n-1}^{N-1} \frac{1}{k} \end{aligned}$$

wie gewünscht. ■

Kapitel 3

Copula-Theorie

3.1 Das Prinzip der Quantilstransformation

Definition 3.1 (verallgemeinerte Inverse, Quantilsfunktion)

Sei F eine Verteilungsfunktion auf \mathbb{R} . Dann heißt

$$F^{-1} : (0, 1) \rightarrow \bar{\mathbb{R}}$$
$$t \mapsto F^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}$$

die (linksseitig stetige) verallgemeinerte Inverse bzw. die Quantilsfunktion zu F .

Falls F strikt isoton und stetig ist, so stimmt F^{-1} mit der Umkehrfunktion von F überein.

Satz 3.2 (Quantilstransformation)

Sei $U \sim \text{UNI}[0, 1]$ und F eine beliebige Verteilungsfunktion auf \mathbb{R} . Dann hat die Zufallsvariable $X := F^{-1}(U)$ die Verteilungsfunktion F , in Zeichen: $X = F^{-1}(U) \sim F$.

Es gelten sogar die stärkeren Aussagen

$$\{X \leq x\} = \{U \leq F(x)\}, \quad (3.1)$$

$$\mathbf{1}_{\{X \leq x\}} = \mathbf{1}_{\{U \leq F(x)\}}, \quad x \in \mathbb{R}. \quad (3.2)$$

Beweis: Nach Definition von F^{-1} gilt

$$U \leq F(x) \Rightarrow X = F^{-1}(U) = \inf\{z | F(z) \geq U\} \leq x.$$

Aus $X = F^{-1}(U) \leq x$ folgt aber andererseits $\forall \varepsilon > 0 : F(x + \varepsilon) \geq U$.

Wegen der Rechtsstetigkeit von F liefert der Grenzübergang $\varepsilon \rightarrow 0$ damit $F(x) \geq U$. Damit ist (3.1) gezeigt, was (3.2) sofort impliziert.

Die Verteilungsaussage über X folgt schließlich aus

$$\mathbb{P}(X \leq x) \stackrel{(3.1)}{=} \mathbb{P}(U \leq F(x)) \stackrel{U \sim \text{UNI}[0,1]}{=} F(x).$$

■

Bemerkung 3.3

Aus (3.1) folgt, dass die folgenden Aussagen für jede Verteilungsfunktion F auf $\mathbb{R} \ni x$ und jedes $t \in]0, 1[$ gelten.

(a) $F(x) \geq t \Leftrightarrow F^{-1}(t) \leq x$

(b) $F(x) < t \Leftrightarrow F^{-1}(t) > x$

(c) $F(x_1) < t \leq F(x_2) \Leftrightarrow x_1 < F^{-1}(t) \leq x_2$.

Satz 3.4

Für jede Verteilungsfunktion F auf \mathbb{R} gilt:

$$\forall 0 \leq t \leq 1 : (F \circ F^{-1})(t) \geq t$$

mit Gleichheit genau dann, wenn t im Bild von F auf $\bar{\mathbb{R}}$ liegt. Ist F stetig, so gilt $F \circ F^{-1} = id_{[0,1]}$.

Beweis: Wir betrachten in Bemerkung 3.3.(a) den Spezialfall $x = F^{-1}(t)$.

Dies führt zu $F(F^{-1}(t)) \geq t \Leftrightarrow F^{-1}(t) \leq F^{-1}(t)$, was stets wahr ist.

Sei nun t im Bild von F . Dann folgt

$$F^{-1}(t) = \inf\{x | F(x) \geq t\} = \inf\{x | F(x) = t\} \Rightarrow F(F^{-1}(t)) = t.$$

Liegt t indes nicht im Bild von F , so ist

$$F^{-1}(t) = \inf\{x | F(x) \geq t\} = \inf\{x | F(x) > t\} \Rightarrow F(F^{-1}(t)) > t.$$

■

Satz 3.5 (Quantilstransformation, Rückrichtung)

Sei X eine reellwertige Zufallsvariable mit Verteilungsfunktion F . Dann gilt

$$\mathbb{P}(F(X) \leq t) \leq t \text{ für alle } 0 \leq t \leq 1,$$

mit Gleichheit genau dann, wenn t im Abschluss des Bildes von F liegt. Falls also F stetig ist, so ist $U := F(X) \sim UNI[0, 1]$.

Beweis: Sei t im Abschluss des Bildes von F . Dann liefert Satz 3.4, dass

$$\mathbb{P}(F(X) \leq t) = \mathbb{P}(X \leq F^{-1}(t)) = F(F^{-1}(t)) = t.$$

Sei nun t nicht im Abschluss des Bildes von F und sei $\varepsilon > 0$ so gewählt, dass $t - \varepsilon$ im Abschluss des Bildes von F liegt, nicht jedoch $t - \varepsilon + \delta$, für jedes $\delta > 0$. Dann gilt

$$\begin{aligned}\mathbb{P}(F(X) \leq t) &= \mathbb{P}(F(X) \leq t - \varepsilon) \\ &= \mathbb{P}(X \leq F^{-1}(t - \varepsilon)) \\ &= F(F^{-1}(t - \varepsilon)) = t - \varepsilon < t.\end{aligned}$$

■

Satz 3.6

Sei F eine beliebige Verteilungsfunktion auf \mathbb{R} . Dann gilt:

$$\forall x \in \mathbb{R} : (F^{-1} \circ F)(x) \leq x,$$

wobei genau dann Gleichheit herrscht, wenn $\forall \varepsilon > 0 : F(x - \varepsilon) \neq F(x)$.

Es folgt, dass $\mathbb{P}((F^{-1} \circ F)(X) \neq X) = 0$, falls $X \sim F$.

Beweis: Übungsaufgabe.

■

Korollar 3.7

(i) F ist genau dann stetig, wenn F^{-1} strikt isoton ist.

(ii) F ist genau dann strikt isoton, wenn F^{-1} stetig ist.

Satz 3.8

Sei F eine stetige Verteilungsfunktion auf \mathbb{R} und gelte $U := F(X) \sim UNI[0, 1]$. Dann ist $X \sim F$.

Beweis: Es gilt stets $\{X \leq x\} \subseteq \{F(X) \leq F(x)\}$, da F isoton, aber nicht zwingenderweise strikt isoton ist. Also ist

$$\mathbb{P}(X \leq x) \leq \mathbb{P}(F(X) \leq F(x)) = \mathbb{P}(U \leq F(x)) = F(x)$$

nach Voraussetzung.

Andererseits ist

$$\begin{aligned}F(x) &= \mathbb{P}(U \leq F(x)) \\ &= \mathbb{P}(F^{-1}(U) \leq x) \\ &= \mathbb{P}((F^{-1} \circ F)(X) \leq x) = \mathbb{P}(X \leq x),\end{aligned}$$

es sei denn, $F(x - \varepsilon) = F(x)$ für ein $\varepsilon > 0$, nach Satz 3.6.

Da F stetig ist, folgt, dass $x \mapsto \mathbb{P}(X \leq x)$ mit F übereinstimmen muss.

■

3.2 Allgemeine Eigenschaften von Copulae

Definition 3.9 (Axiome einer d -dimensionalen Verteilungsfunktion)

Sei $F : \mathbb{R}^d \rightarrow [0, 1]$ eine Abbildung. Dann heißt F eine Verteilungsfunktion auf \mathbb{R}^d , falls die folgenden drei Bedingungen erfüllt sind.

- (i) Für alle $\mathbf{x} = (x_1, \dots, x_d)^\top \in \mathbb{R}^d$ mit $\exists 1 \leq i \leq d : x_i = -\infty$ gilt $F(\mathbf{x}) = 0$.
- (ii) Für jedes d -dimensionale Intervall $I = [\mathbf{a}, \mathbf{b}] = [a_1, b_1] \times \dots \times [a_d, b_d]$ ist das d -dimensionale Volumen $V_F(I) \geq 0$. Dabei ist

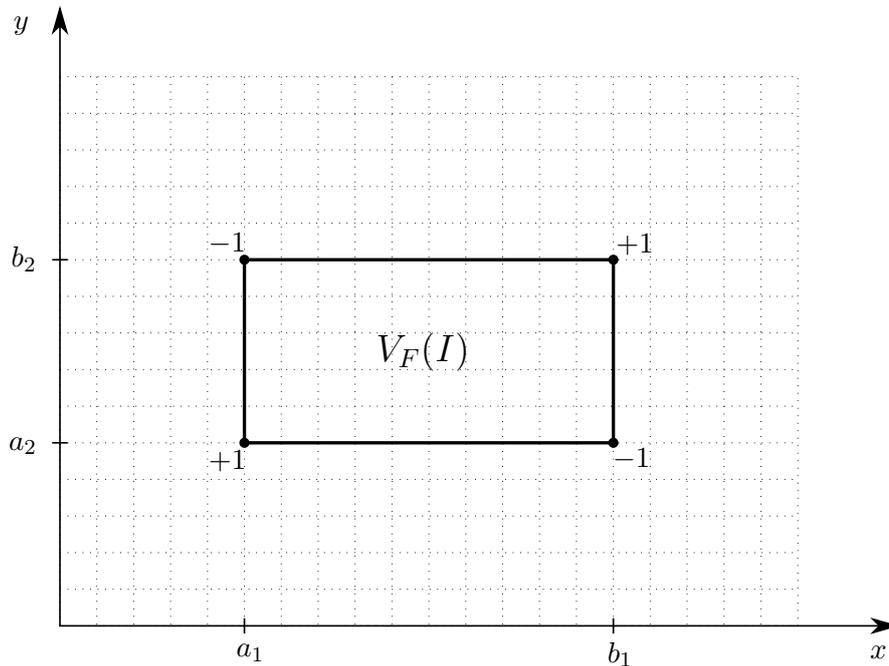
$$V_F(I) = \sum_{c: c \text{ Eckpunkt von } I} \text{sgn}(c)F(c)$$

mit

$$\text{sgn}(c) = \begin{cases} +1, & \text{falls } \#\{1 \leq k \leq d : c_k = a_k\} \text{ gerade} \\ -1, & \text{falls } \#\{1 \leq k \leq d : c_k = a_k\} \text{ ungerade} \end{cases}$$

- (iii) $F(\infty, \dots, \infty) = 1$.

Abbildung 3.10



Definition 3.11

Sei F eine d -dimensionale Verteilungsfunktion. Dann heißt $F_k : \mathbb{R} \rightarrow [0, 1]$ mit

$$F_k(x) = F(+\infty, \dots, +\infty, x, +\infty, \dots, +\infty)$$

die k -te (univariate) Randverteilungsfunktion von F , für $1 \leq k \leq d$.

Definition 3.12

Der Träger $\text{supp}(F)$ einer d -dimensionalen Verteilungsfunktion F ist das Komplement der Vereinigung aller offenen Teilmengen des \mathbb{R}^d mit F -Maß gleich Null. Dabei meint F -Maß das durch F induzierte Wahrscheinlichkeitsmaß \mathbb{P} auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, so dass

$$\mathbb{P}((-\infty, x_1] \times (-\infty, x_2] \times \dots \times (-\infty, x_d]) = F(x_1, x_2, \dots, x_d).$$

Bemerkung 3.13

Besitzt eine Verteilungsfunktion F den Träger $\text{supp}(F)$, so schreiben wir auch $F : \text{supp}(F) \rightarrow [0, 1]$.

Definition 3.14

Eine Copula $C : [0, 1]^d \rightarrow [0, 1]$ ist eine d -dimensionale Verteilungsfunktion, so dass alle zugehörigen Randverteilungen gleich $UNI[0, 1]$ sind, d.h. $\forall 1 \leq k \leq d : \forall u \in [0, 1] : C_k(u) = u$.

Satz 3.15 (Satz von Sklar)

Sei F eine d -dimensionale Verteilungsfunktion mit Randverteilungsfunktionen F_1, \dots, F_d . Dann existiert eine Copula C , so dass

$$\forall \mathbf{x} = (x_1, \dots, x_d)^\top \in \mathbb{R}^d : F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d)). \quad (3.3)$$

Falls F_k stetig ist für alle $1 \leq k \leq d$, so ist C eindeutig bestimmt, ansonsten nur auf dem kartesischen Produkt der Bildmengen von $(F_k)_{1 \leq k \leq d}$.

Ist umgekehrt $C : [0, 1]^d \rightarrow [0, 1]$ (irgend) eine Copula und sind $(F_k : 1 \leq k \leq d)$ Verteilungsfunktionen auf \mathbb{R} , dann ist die durch (3.3) gegebene Abbildung F eine d -dimensionale Verteilungsfunktion.

Anmerkung: Die Copula C koppelt die Randverteilungsfunktionen und beschreibt somit eine Abhängigkeitsstruktur. Manche Autor/innen bezeichnen C daher auch als „Abhängigkeitsfunktion“.

Beweis: Sei $\mathbf{X} = (X_1, \dots, X_d)^\top : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ein d -dimensionaler Zufallsvektor, der die Verteilungsfunktion F besitzt.

Setze in dem Fall, dass alle F_k stetig sind, $U_k := F_k(X_k)$, $1 \leq k \leq d$, und definiere C als die gemeinsame Verteilungsfunktion von $\mathbf{U} = (U_1, \dots, U_d)^\top$. Mit Hilfe des Prinzips der Quantilstransformation verifizieren wir nun für alle $\mathbf{x} = (x_1, \dots, x_d)^\top$, dass

$$\begin{aligned} F(\mathbf{x}) &= \mathbb{P}(\forall 1 \leq k \leq d : X_k \leq x_k) \\ &= \mathbb{P}(\forall 1 \leq k \leq d : F_k^{-1}(U_k) \leq x_k) \\ &= \mathbb{P}(\forall 1 \leq k \leq d : U_k \leq F_k(x_k)) \\ &= C(F_1(x_1), \dots, F_d(x_d)) \end{aligned}$$

wie gewünscht.

Im allgemeinen Fall gehen wir zu den (verallgemeinerten) Verteilungstransformierten $(\tilde{U}_k)_{1 \leq k \leq d}$ nach Rüschendorf (2009) über. Definiere dazu für eine reellwertige Zufallsvariable Y mit Verteilungsfunktion G die Funktion

$$\tilde{G}(\cdot, \cdot) : \mathbb{R} \times [0, 1] \rightarrow [0, 1]$$

vermittels

$$\tilde{G}(y, \lambda) := \mathbb{P}(Y < y) + \lambda \mathbb{P}(Y = y).$$

Dann ist die (verallgemeinerte) Verteilungstransformierte \tilde{U} von Y gegeben durch

$$\tilde{U} := \tilde{G}(Y, V) = G(Y-) + V(G(Y) - G(Y-)),$$

wobei $V \sim UNI[0, 1]$ stochastisch unabhängig von Y gewählt ist.

Nach Übungsaufgabe gelten

$$\tilde{U} = \tilde{G}(Y, V) \sim UNI[0, 1], \quad (3.4)$$

$$Y = G^{-1}(\tilde{U}) \text{ fast sicher.} \quad (3.5)$$

Damit kann die Copula C als gemeinsame Verteilungsfunktion der (verallgemeinerten) Verteilungstransformierten $(\tilde{U}_k)_{1 \leq k \leq d}$ der $(X_k : 1 \leq k \leq d)$ gewählt werden und die Verifikation erfolgt analog zum stetigen Fall. ■

Anmerkung: Die Nicht-Eindeutigkeit der Copula C im Falle nicht notwendigerweise stetiger F_k liegt daran, dass beim Bilden der \tilde{U}_k stets das gleiche V oder aber unterschiedliche $V_k \sim UNI[0, 1]$ benutzt werden können, $1 \leq k \leq d$.

Korollar 3.16

Sei F eine d -dimensionale Verteilungsfunktion mit stetigen Randverteilungsfunktionen F_1, \dots, F_d und Copula C gemäß (3.3) im Satz von Sklar. Dann gilt

$$\forall \mathbf{u} = (u_1, \dots, u_d)^\top \in [0, 1]^d : C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)).$$

Beispiele 3.17

(a) Produkt-Copula:

Gelte $C(u_1, \dots, u_d) = \prod_{k=1}^d u_k$, so ist nach dem Satz von Sklar

$$F(\mathbf{x}) = \prod_{k=1}^d F_k(x_k), \quad \mathbf{x} = (x_1, \dots, x_d)^\top \in \mathbb{R}^d.$$

Dies entspricht der stochastischen Unabhängigkeit.

(b) Gauß-Copula:

Sei Φ die Verteilungsfunktion der Standardnormalverteilung auf \mathbb{R} und Φ_R die (gemeinsame) Verteilungsfunktion der d -dimensionalen Normalverteilung mit Erwartungswertvektor 0 und Kovarianz- und Korrelationsmatrix gleich R (d.h., alle d Varianzen sind gleich 1).

Dann heißt die durch

$$C(u_1, \dots, u_d) = \Phi_R(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$$

gegebene Funktion C die d -dimensionale Gauß-Copula mit Korrelationsmatrix R .

(c) Archimedische Copulae:

Sei $\psi : [0, \infty) \rightarrow [0, 1]$ mit $\psi(0) = 1$, $\lim_{x \rightarrow \infty} \psi(x) = 0$ eine Funktion, deren erste d Ableitungen alternierendes Verzeichen haben, also

$$\forall 1 \leq k \leq d : \forall x \in [0, \infty) : (-1)^k \psi^{(k)}(x) \geq 0.$$

Dann ist die Archimedische Copula mit Generatorfunktion ψ gegeben durch

$$C_\psi(\mathbf{u}) = \psi \left(\sum_{k=1}^d \psi^{-1}(u_k) \right), \quad \mathbf{u} = (u_1, \dots, u_d)^\top \in [0, 1]^d.$$

Spezielle parametrische Generatoren sind gegeben durch

$$\psi(x) = (1 + \eta x)^{-\frac{1}{\eta}} \Leftrightarrow \psi^{-1}(t) = \eta^{-1}(t^{-\eta} - 1), \quad \eta > 0,$$

so dass

$$C_\eta(\mathbf{u}) = \left(1 + \eta \sum_{k=1}^d \frac{1}{\eta} (u_k^{-\eta} - 1) \right)^{-\frac{1}{\eta}} = \left(\sum_{k=1}^d u_k^{-\eta} - d + 1 \right)^{-\frac{1}{\eta}}$$

(Familie der Clayton-Copulae mit Parameter $\eta > 0$) sowie

$$\psi(x) = \exp(-x^{1/\eta}) \Leftrightarrow \psi^{-1}(t) = (-\log(t))^\eta, \quad \eta \geq 1,$$

so dass

$$C_\eta(\mathbf{u}) = \exp \left(- \left\{ \sum_{k=1}^d [-\log(u_k)]^\eta \right\}^{\frac{1}{\eta}} \right)$$

(Familie der Gumbel-Hougaard-Copulae mit Parameter $\eta \geq 1$).

Gilt sogar $\forall k \in \mathbb{N} : \forall x \in [0, \infty) : (-1)^k \psi^{(k)}(x) \geq 0$, so ist ψ die Laplace-Transformierte L_Z einer nicht-negativen Zufallsvariable Z , vgl. Bernstein (1929). Damit lässt sich das Archimedische Copula-Modell mit einem solchen Generator auch wie folgt motivieren.

Es gelte

$$F_{\mathbf{X}}(\mathbf{x}) = \int \prod_{k=1}^d [G_k(x_k)]^z dF_Z(z) \quad (3.6)$$

für univariate Verteilungsfunktionen ($G_k : 1 \leq k \leq d$).

Dann ist die k -te Randverteilungsfunktion gegeben durch

$$\begin{aligned} F_k(x_k) &= \int [G_k(x_k)]^z dF_Z(z) \\ &= \int \exp(z \log G_k(x_k)) dF_Z(z) \\ &= L_Z(-\log G_k(x_k)), \end{aligned}$$

denn

$$L_Z(s) = \int \exp(-sz) dF_Z(z);$$

vgl. Stochastik I - Vorlesung.

Einsetzen von $G_k(x_k) = \exp(-L_Z^{-1}(F_k(x_k)))$ in (3.6) liefert

$$F_{\mathbf{X}}(\mathbf{x}) = L_Z \left(\sum_{k=1}^d L_Z^{-1}(F_k(x_k)) \right).$$

Somit besitzt \mathbf{X} nach dem Satz von Sklar eine Archimedische Copula mit Generator $\psi = L_Z$.

(d) Extremwert-Copulae:

Sei $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,d})^\top$ für $1 \leq i \leq n$ ein d -dimensionaler Zufallsvektor, so dass $(\mathbf{X}_i : 1 \leq i \leq n)$ iid. mit Copula C_F sind. Sei $\mathbf{M}_n = (M_{n,1}, \dots, M_{n,d})^\top$ der d -dimensionale Vektor der komponentenweisen Maxima $M_{n,1}, \dots, M_{n,d}$ der $(\mathbf{X}_i : 1 \leq i \leq n)$.

Dann ist die Copula C_n von \mathbf{M}_n gegeben durch $C_n(\mathbf{u}) = C_F^n(u_1^{1/n}, \dots, u_d^{1/n})$, $\mathbf{u} = (u_1, \dots, u_d)^\top \in [0, 1]^d$; vgl. Übungsaufgabe.

Eine Copula C heißt nun Extremwert-Copula, falls

$$\forall n \in \mathbb{N} : \exists \text{ Copula } C_F : \forall \mathbf{u} = (u_1, \dots, u_d)^\top \in [0, 1]^d : C_F^n(u_1^{1/n}, \dots, u_d^{1/n}) = C(\mathbf{u}).$$

Es ist bekannt, dass C genau dann eine Extremwert-Copula ist, wenn C Maximums-stabil ist in dem Sinne, dass $C(\mathbf{u}) = C^m(u_1^{1/m}, \dots, u_d^{1/m}) \forall m \in \mathbb{N}$. Zum Beispiel ist die Gumbel-Hougaard-Copula C_η aus Beispiel 3.17.(c) eine Extremwert-Copula, vgl. Übungsaufgabe.

Wir betrachten im Folgenden drei spezielle Funktionen $M^{(d)}$, $\Pi^{(d)}$ und $W^{(d)}$ auf $[0, 1]^d$, die wie folgt gegeben sind.

$$\begin{aligned} M^{(d)}(\mathbf{u}) &= \min(u_1, \dots, u_d) \\ \Pi^{(d)}(\mathbf{u}) &= \prod_{i=1}^d u_i \\ W^{(d)}(\mathbf{u}) &= \max(0, \sum_{k=1}^d u_k - d + 1). \end{aligned}$$

Lemma 3.18

(a) Die Funktionen $M^{(d)}$ und $\Pi^{(d)}$ sind Copulae, für alle $d \geq 2$.

(b) Die Funktion $W^{(d)}$ ist für $d \geq 3$ keine Copula.

Beweis: Zu Teil (a) beachten wir, dass $\min(1, \dots, 1, u_k, 1, \dots, 1) = u_k$ sowie $1 \cdot \dots \cdot u_k \cdot 1 \cdot \dots \cdot 1 = u_k$ ist, $\forall 1 \leq k \leq d$.

Zu Teil (b) betrachten wir das d -dimensionale Intervall $[1/2, 1]^d$. Dessen $W^{(d)}$ -Volumen berechnet sich zu

$$\begin{aligned} V_{W^{(d)}}([1/2, 1]^d) &= \max(1 + \dots + 1 - d + 1, 0) \\ &\quad - d \cdot \max\left(\frac{1}{2} + 1 + \dots + 1 - d + 1, 0\right) \\ &\quad + \binom{d}{2} \max\left(\frac{1}{2} + \frac{1}{2} + 1 + \dots + 1 - d + 1, 0\right) \\ &\quad + \dots + \max\left(\frac{1}{2} + \dots + \frac{1}{2} - d + 1, 0\right) \\ &= 1 - \frac{d}{2} + 0 + \dots + 0. \end{aligned}$$

■

Satz 3.19 (Fréchet-Hoeffding-Schranken)

Für jede d -dimensionale Copula C gilt:

$$\forall \mathbf{u} \in [0, 1]^d : W^{(d)}(\mathbf{u}) \leq C(\mathbf{u}) \leq M^{(d)}(\mathbf{u}). \quad (3.7)$$

Ferner existiert für jedes $d \geq 3$ und $\mathbf{u} \in [0, 1]^d$ eine Copula C (die von \mathbf{u} abhängt) mit $C(\mathbf{u}) = W^{(d)}(\mathbf{u})$, so dass die untere Schranke aus (3.7) nicht gleichmäßig verbessert werden kann.

Beweis: Für die obere Schranke beachten wir, dass

$$\forall \mathbf{u} = (u_1, \dots, u_d)^\top \in [0, 1]^d : \forall 1 \leq k \leq d : C(\mathbf{u}) \leq C(1, \dots, 1, u_k, 1, \dots, 1) = u_k.$$

Also ist $C(\mathbf{u}) \leq \min_{1 \leq k \leq d} u_k = M^{(d)}(\mathbf{u})$.

Zum Nachweis der unteren Schranke benutzen wir, dass nach Übungsaufgabe

$$\begin{aligned} C(1, \dots, 1) - C(\mathbf{u}) &\leq \sum_{k=1}^d C_k(1) - C_k(u_k) \\ \Leftrightarrow 1 - C(\mathbf{u}) &\leq \sum_{k=1}^d (1 - u_k) \\ \Leftrightarrow C(\mathbf{u}) &\geq \sum_{k=1}^d u_k - d + 1. \end{aligned}$$

Da zudem $C(\mathbf{u}) \geq 0$ ist, folgt $C(\mathbf{u}) \geq W^{(d)}(\mathbf{u})$. Die Konstruktion einer Copula, die die untere Schranke in einem gegebenen Punkte erreicht, ist eine einfache Interpolation, siehe Theorem 2.10.13 in Nelsen (2006). ■

Bemerkung 3.20

(a) Die Copula $M^{(d)}$ entspricht „totaler positiver Abhängigkeit“ in dem folgenden Sinne.

Sei $U_k = V \sim \text{UNI}[0, 1] \mathbb{P}$ - fast sicher für alle $1 \leq k \leq d$, so ergibt sich

$$\begin{aligned} C(\mathbf{u}) &= \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) \\ &= \mathbb{P}(V \leq u_1, \dots, V \leq u_d) \\ &= \mathbb{P}(V \leq \min_{1 \leq k \leq d} u_k) = \min_{1 \leq k \leq d} u_k \\ &= M^{(d)}(\mathbf{u}). \end{aligned}$$

(b) In Dimension $d=2$ entspricht die Copula $W^{(d)}$ „totaler negativer Abhängigkeit“ in dem folgenden Sinne.

Sei $U_2 = 1 - U_1 \sim \text{UNI}[0, 1] \mathbb{P}$ - fast sicher, so gilt

$$\begin{aligned} C(\mathbf{u}) &= \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2) \\ &= \mathbb{P}(U_1 \leq u_1, 1 - U_1 \leq u_2) \\ &= \mathbb{P}(U_1 \leq u_1, U_1 \geq 1 - u_2) \\ &= \mathbb{P}(1 - u_2 \leq U_1 \leq u_1) \\ &= u_1 - (1 - u_2) = u_1 + u_2 - 1 \\ &= W^{(2)}(\mathbf{u}). \end{aligned}$$

Definition 3.21

Falls $\mathbf{X} = (X_1, \dots, X_d)^\top$ die Verteilungsfunktion F besitzt, so nennen wir die durch $\bar{F}(x_1, \dots, x_d) = \mathbb{P}(X_1 > x_1, \dots, X_d > x_d)$ gegebene Funktion $\bar{F} : \mathbb{R}^d \rightarrow [0, 1]$ die (gemeinsame) Survivalfunktion von \mathbf{X} .

Ferner nennen wir $\check{C} : [0, 1]^d \rightarrow [0, 1]$ mit $\bar{F}(x_1, \dots, x_d) = \check{C}(\bar{F}_1(x_1), \dots, \bar{F}_d(x_d))$ die Survival-Copula von \mathbf{X} . Besitzt $\mathbf{U} = (U_1, \dots, U_d)^\top$ mit $U_k \sim \text{UNI}[0, 1] \forall 1 \leq k \leq d$ die gemeinsame Verteilungsfunktion (und Copula) C , so ist $\bar{C}(u_1, \dots, u_d) = \check{C}(1 - u_1, \dots, 1 - u_d)$.

Satz 3.22

Sei $\mathbf{X} = (X_1, \dots, X_d)^\top : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ein d -dimensionaler Vektor stetig (d. h., absolut stetig bezüglich des Lebesguemaßes) verteilter Zufallsvariablen mit Copula C .

Für $1 \leq k \leq d$ sei $T_k : \mathbb{R} \rightarrow \mathbb{R}$ eine streng monoton wachsende Funktion. Dann besitzt auch $\mathbf{T} = (T_1(X_1), \dots, T_d(X_d))^\top$ die Copula C .

Beweis: Bezeichne $C_{\mathbf{T}}$ die Copula von \mathbf{T} . Wir berechnen zunächst für jedes $1 \leq k \leq d$ und $x \in \mathbb{R}$:

$$F_{T_k}(x) = \mathbb{P}(T_k(X_k) \leq x) = \mathbb{P}(X_k \leq T_k^{-1}(x)) = F_{X_k}(T_k^{-1}(x)).$$

Daraus folgt

$$\begin{aligned} C_{\mathbf{T}}(F_{T_1}(x_1), \dots, F_{T_d}(x_d)) &= \mathbb{P}(T_1(X_1) \leq x_1, \dots, T_d(X_d) \leq x_d) \\ &= \mathbb{P}(X_1 \leq T_1^{-1}(x_1), \dots, X_d \leq T_d^{-1}(x_d)) \\ &= C(F_{X_1}(T_1^{-1}(x_1)), \dots, F_{X_d}(T_d^{-1}(x_d))) \\ &= C(F_{T_1}(x_1), \dots, F_{T_d}(x_d)). \end{aligned}$$

Da die $(X_k : 1 \leq k \leq d)$ stetig verteilt sind, stimmen somit $C_{\mathbf{T}}$ und C überein. ■

Satz 3.23

Sei $\mathbf{X} = (X_1, \dots, X_d)^\top : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ein d -dimensionaler Vektor stetig verteilter Zufallsvariablen mit Copula C_{X_1, \dots, X_d} . Für $2 \leq k \leq d$ sei $T_k : \mathbb{R} \rightarrow \mathbb{R}$ strikt isoton, und sei $T_1 : \mathbb{R} \rightarrow \mathbb{R}$ strikt antiton.

Bezeichne C_{T_1, \dots, T_d} die Copula von $(T_1(X_1), \dots, T_d(X_d))^\top$. Dann gilt für $\mathbf{u} = (u_1, \dots, u_d)^\top \in [0, 1]^d$:

$$C_{T_1, \dots, T_d}(\mathbf{u}) = C_{T_2, \dots, T_d}(u_2, \dots, u_d) - C_{X_1, T_2, \dots, T_d}(1 - u_1, u_2, \dots, u_d).$$

Ist T_j antiton für $j \neq 1$, so gilt die analoge Formel durch Koordinatentausch.

Beweis:

$$\begin{aligned} &C_{T_1, \dots, T_d}(F_{T_1}(x_1), \dots, F_{T_d}(x_d)) \\ &= \mathbb{P}(T_1(X_1) \leq x_1, \dots, T_d(X_d) \leq x_d) \\ &= \mathbb{P}(T_2(X_2) \leq x_2, \dots, T_d(X_d) \leq x_d) - \mathbb{P}(X_1 \leq T_1^{-1}(x_1), T_2(X_2) \leq x_2, \dots, T_d(X_d) \leq x_d) \\ &= C_{T_2, \dots, T_d}(F_{T_2}(x_2), \dots, F_{T_d}(x_d)) - C_{X_1, T_2, \dots, T_d}(F_{X_1}(T_1^{-1}(x_1)), F_{T_2}(x_2), \dots, F_{T_d}(x_d)) \\ &= C_{T_2, \dots, T_d}(F_{T_2}(x_2), \dots, F_{T_d}(x_d)) - C_{X_1, T_2, \dots, T_d}(1 - F_{T_1}(x_1), F_{T_2}(x_2), \dots, F_{T_d}(x_d)). \end{aligned}$$

Da X_1, \dots, X_d stetig verteilt sind impliziert dies die Aussage. ■

Beispiel 3.24 (in Dimension $d = 2$)

(a) Sei T_1 strikt antiton und T_2 strikt isoton. Dann ist

$$\begin{aligned} C_{T_1, T_2}(u_1, u_2) &= u_2 - C_{X_1, T_2}(1 - u_1, u_2) \\ &= u_2 - C_{X_1, X_2}(1 - u_1, u_2) \end{aligned}$$

wegen Satz 3.22.

(b) Seien T_1 und T_2 beide strikt antiton. Dann ist

$$\begin{aligned} C_{T_1, T_2}(u_1, u_2) &= u_2 - C_{X_1, T_2}(1 - u_1, u_2) \\ &= u_2 - (1 - u_1 - C_{X_1, X_2}(1 - u_1, 1 - u_2)) \\ &= u_1 + u_2 - 1 + C_{X_1, X_2}(1 - u_1, 1 - u_2). \end{aligned}$$

Da hier beide Transformationen T_1 und T_2 jeweils antiton sind, ist $C_{T_1, T_2} = \check{C}_{X_1, X_2}$, also die Survival-Copula von $(X_1, X_2)^\top$.

Bezeichne $\mathbf{U} = (U_1, U_2)^\top$ den bivariaten Vektor der Verteilungstransformierten von $(X_1, X_2)^\top$, so verifizieren wir mit Hilfe von Definition 3.21:

$$\begin{aligned} \check{C}_{X_1, X_2}(u_1, u_2) &= \check{C}_{U_1, U_2}(u_1, u_2) \\ &= \bar{C}_{U_1, U_2}(1 - u_1, 1 - u_2) \\ &= \mathbb{P}(U_1 > 1 - u_1, U_2 > 1 - u_2) \\ &= 1 - \mathbb{P}(\{U_1 \leq 1 - u_1\} \cup \{U_2 \leq 1 - u_2\}) \\ &= 1 - [\mathbb{P}(U_1 \leq 1 - u_1) + \mathbb{P}(U_2 \leq 1 - u_2) - C_{X_1, X_2}(1 - u_1, 1 - u_2)] \\ &= 1 - [1 - u_1 + 1 - u_2 - C_{X_1, X_2}(1 - u_1, 1 - u_2)] \\ &= u_1 + u_2 - 1 + C_{X_1, X_2}(1 - u_1, 1 - u_2). \end{aligned}$$

Beispiel 3.24 zeigt, dass die Copula von $T_1(X_1), \dots, T_d(X_d)$ für strikt monotone Transformationen T_1, \dots, T_d durch die Copula von X_1, \dots, X_d bereits festgelegt ist, denn in höheren Dimensionen kann rekursiv in analoger Weise gerechnet werden.

3.3 Abhängigkeitsmaße

3.3.1 Der (lineare) Korrelationskoeffizient

Definition 3.25

(a) Es sei $(X_1, X_2)^\top: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ ein bivariater Zufallsvektor mit $\mathbb{E}[X_i^2] < \infty$, $i = 1, 2$. Dann heißt

$$\begin{aligned} \text{Cov}(X_1, X_2) &:= \mathbb{E}[(X_1 - \mathbb{E}[X_1])(X_2 - \mathbb{E}[X_2])] \\ &= \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2] \end{aligned}$$

die Kovarianz von X_1 und X_2 . Sind zusätzlich $\text{Var}(X_1) > 0$ und $\text{Var}(X_2) > 0$, so heißt

$$\rho(X_1, X_2) := \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} \in [-1, 1]$$

der Korrelationskoeffizient von X_1 und X_2 .

(b) Ist $\mathbf{X} = (X_1, \dots, X_d)^\top$ ein d -dimensionaler Zufallsvektor mit nicht-trivialen, endlichen Varianzen, so heißt

$$\text{Cov}(\mathbf{X}) := (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq d} \in \mathbb{R}^{d \times d}$$

die Kovarianzmatrix von \mathbf{X} und

$$R \equiv R(\mathbf{X}) := (\rho(X_i, X_j))_{1 \leq i, j \leq d} \in [-1, 1]^{d \times d}$$

die Korrelationsmatrix von \mathbf{X} .

Satz 3.26 (Eigenschaften der Kovarianz)

Unter den Voraussetzungen von Definition 3.25 gelten die folgenden Aussagen.

(a) Falls $X_2 = aX_1 + b$ \mathbb{P} -fast sicher mit $a \neq 0$ und $b \in \mathbb{R}$, so ist $\rho(X_1, X_2) = \text{sgn}(a)$. Anderenfalls ist $-1 < \rho(X_1, X_2) < +1$. Der Korrelationskoeffizient misst also den Grad der linearen Abhängigkeit von X_1 und X_2 .

(b) Die Kovarianz ist eine symmetrische Bilinearform, d.h. $\text{Cov}(X_1, X_2) = \text{Cov}(X_2, X_1)$ und $\text{Cov}(aX_1 + b, cX_2 + d) = ac \cdot \text{Cov}(X_1, X_2)$. Analog gilt $\rho(\alpha X_1 + \beta, \gamma X_2 + \delta) = \text{sgn}(\alpha\gamma)\rho(X_1, X_2)$.

(c) Sind A und B zwei $(m \times d)$ -Matrizen, $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$, und \mathbf{X} und \mathbf{Y} zwei d -dimensionale Zufallsvektoren, so ist

$$\text{Cov}(\mathbf{X}, \mathbf{Y}) = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{Y} - \mathbb{E}[\mathbf{Y}])^\top]$$

sowie

$$\text{Cov}(A\mathbf{X} + \mathbf{a}, B\mathbf{Y} + \mathbf{b}) = A \cdot \text{Cov}(\mathbf{X}, \mathbf{Y})B^\top.$$

Ferner ist $\text{Cov}(\mathbf{X}) = \text{Cov}(\mathbf{X}, \mathbf{X})$ die Varianz- / Kovarianzmatrix von \mathbf{X} .

(d) $\text{Var}(X_1) = \text{Cov}(X_1, X_1)$. Für $\mathbf{X} : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ und $\boldsymbol{\alpha} \in \mathbb{R}^d$ fest vorgegeben ist

$$\text{Var}(\boldsymbol{\alpha}^\top \mathbf{X}) = \boldsymbol{\alpha}^\top \text{Cov}(\mathbf{X})\boldsymbol{\alpha}.$$

Damit ist $\text{Cov}(\mathbf{X})$ stets symmetrisch und positiv (semi-) definit.

Satz 3.27 (Lemma von Hoeffding)

Sei $\mathbf{X} = (X_1, X_2)^\top$ ein bivariater Zufallsvektor mit existierender Kovarianz, bivariater Copula C und Randverteilungsfunktionen F_1 und F_2 . Dann gilt:

$$\text{Cov}(X_1, X_2) = \int_{\mathbb{R}^2} [C(F_1(x_1), F_2(x_2)) - F_1(x_1)F_2(x_2)] dx_1 dx_2.$$

Der Beweis beruht auf dem folgenden Lemma.

Lemma

Seien $y_1, y_2, z_1, z_2 \in \mathbb{R}$ fest vorgegeben. Dann gilt

$$(*) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\mathbf{1}_{\{u < y_1\}} - \mathbf{1}_{\{u < z_1\}}] [\mathbf{1}_{\{v < y_2\}} - \mathbf{1}_{\{v < z_2\}}] dudv = (y_1 - z_1)(y_2 - z_2).$$

Beweis: Nach dem Satz von Fubini ist

$$(*) = \int_{-\infty}^{\infty} [\mathbf{1}_{\{v < y_2\}} - \mathbf{1}_{\{v < z_2\}}] \left[\int_{-\infty}^{\infty} [\mathbf{1}_{\{u < y_1\}} - \mathbf{1}_{\{u < z_1\}}] du \right] dv.$$

Nun ist aber

$$\mathbf{1}_{\{u < y_1\}} - \mathbf{1}_{\{u < z_1\}} = \begin{cases} 0, & y_1 = z_1 \\ 0, & u < \min(y_1, z_1) \\ 0, & u \geq \max(y_1, z_1) \\ +1, & z_1 \leq u < y_1, z_1 < y_1 \\ -1, & y_1 \leq u < z_1, y_1 < z_1 \end{cases}$$

Also kann o.B.d.A. $z_1 < y_1$ angenommen werden und es folgt

$$\int_{-\infty}^{\infty} [\mathbf{1}_{\{u < y_1\}} - \mathbf{1}_{\{u < z_1\}}] du = \int_{z_1}^{y_1} +1 du = (y_1 - z_1),$$

was die Aussage impliziert, da die Integration bezüglich v analog gehandhabt werden kann. ■

Beweis: (von Satz 3.27)

Seien $\mathbf{Y} = (Y_1, Y_2)^\top$ und $\mathbf{Z} = (Z_1, Z_2)^\top$ stochastisch unabhängig und identisch verteilt, jeweils mit der gleichen Verteilung wie $\mathbf{X} = (X_1, X_2)^\top$.

Dann ist

$$2 \cdot \text{Cov}(X_1, X_2) = 2 \{ \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2] \} = \mathbb{E}[(Y_1 - Z_1)(Y_2 - Z_2)]. \quad (3.8)$$

Nach dem vorangegangenen Lemma ist (3.8) aber gleich

$$\mathbb{E} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\mathbf{1}_{\{u < Y_1\}} - \mathbf{1}_{\{u < Z_1\}}] [\mathbf{1}_{\{v < Y_2\}} - \mathbf{1}_{\{v < Z_2\}}] dudv \right]. \quad (3.9)$$

Da $\mathbf{X} = (X_1, X_2)^\top$ endliche zweite Momente besitzt, ist vertauschen von $\mathbb{E}[\cdot]$ mit dem Doppelin-

tegral zulässig. Damit ist (3.9) gleich

$$\begin{aligned}
& \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{E} \left[[\mathbf{1}_{\{u < Y_1\}} - \mathbf{1}_{\{u < Z_1\}}] [\mathbf{1}_{\{v < Y_2\}} - \mathbf{1}_{\{v < Z_2\}}] \right] dudv \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbb{E} \left[\mathbf{1}_{\{u < Y_1, v < Y_2\}} - \mathbf{1}_{\{u < Y_1, v < Z_2\}} - \mathbf{1}_{\{u < Z_1, v < Y_2\}} + \mathbf{1}_{\{u < Z_1, v < Z_2\}} \right] dudv \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [2\mathbb{P}(X_1 > u, X_2 > v) - 2\mathbb{P}(X_1 > u)\mathbb{P}(X_2 > v)] dudv \\
&= 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\mathbb{P}(X_1 > u, X_2 > v) - \mathbb{P}(X_1 > u)\mathbb{P}(X_2 > v)] dudv.
\end{aligned}$$

Beachte schließlich, dass

$$\begin{aligned}
& \mathbb{P}(X_1 > u, X_2 > v) - \mathbb{P}(X_1 > u)\mathbb{P}(X_2 > v) \\
&= 1 - \mathbb{P}(\{X_1 \leq u\} \cup \{X_2 \leq v\}) - [1 - \mathbb{P}(X_1 \leq u)][1 - \mathbb{P}(X_2 \leq v)] \\
&= 1 - [\mathbb{P}(X_1 \leq u) + \mathbb{P}(X_2 \leq v) - \mathbb{P}(X_1 \leq u, X_2 \leq v)] \\
&\quad - [1 - \mathbb{P}(X_2 \leq v) - \mathbb{P}(X_1 \leq u) + \mathbb{P}(X_1 \leq u)\mathbb{P}(X_2 \leq v)] \\
&= \mathbb{P}(X_1 \leq u, X_2 \leq v) - \mathbb{P}(X_1 \leq u)\mathbb{P}(X_2 \leq v) \\
&= C(F_1(u), F_2(v)) - F_1(u)F_2(v) \text{ ist.}
\end{aligned}$$

■

3.3.2 Konkordanzmaße

Satz 3.28

Seien $(X, Y)^\top$ und $(\tilde{X}, \tilde{Y})^\top$ zwei stochastisch unabhängige, bivariate Vektoren stetig verteilter Zufallsvariablen mit identischen Randverteilungsfunktionen F (von X bzw. \tilde{X}) und G (von Y bzw. \tilde{Y}), die auf dem gleichem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ definiert sind.

Bezeichne C die Copula von $(X, Y)^\top$ und \tilde{C} die Copula von $(\tilde{X}, \tilde{Y})^\top$. Definiere

$$Q := \mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) > 0) - \mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) < 0).$$

Dann gilt:

$$Q = Q(C, \tilde{C}) = 4 \int \int_{[0,1]^2} \tilde{C}(u, v) dC(u, v) - 1.$$

Beweis: Wegen der angenommenen Stetigkeit ist

$$\mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) > 0) = 1 - \mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) < 0),$$

also

$$Q = 2\mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) > 0) - 1.$$

Darüber hinaus ist

$$\mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) > 0) = \mathbb{P}(X > \tilde{X}, Y > \tilde{Y}) + \mathbb{P}(X < \tilde{X}, Y < \tilde{Y}).$$

Wir berechnen nun beide Summanden separat. Zunächst erhalten wir für den ersten Summanden, dass

$$\begin{aligned} \mathbb{P}(X > \tilde{X}, Y > \tilde{Y}) &= \int \int_{\mathbb{R}^2} \mathbb{P}(\tilde{X} < x, \tilde{Y} < y) dC(F(x), G(y)) \\ &= \int \int_{\mathbb{R}^2} \tilde{C}(F(x), G(y)) dC(F(x), G(y)). \end{aligned}$$

Gehen wir zu den verteilungstransformierten Werten $u = F(x)$ und $v = G(y)$ über, so erhalten wir

$$\mathbb{P}(X > \tilde{X}, Y > \tilde{Y}) = \int \int_{[0,1]^2} \tilde{C}(u, v) dC(u, v).$$

Analog rechnen wir

$$\begin{aligned} \mathbb{P}(X < \tilde{X}, Y < \tilde{Y}) &= \int \int_{\mathbb{R}^2} \mathbb{P}(\tilde{X} > x, \tilde{Y} > y) dC(F(x), G(y)) \\ &= \int \int_{\mathbb{R}^2} [1 - F(x) - G(y) + \tilde{C}(F(x), G(y))] dC(F(x), G(y)) \\ &= \int \int_{[0,1]^2} [1 - u - v + \tilde{C}(u, v)] dC(u, v). \end{aligned} \tag{3.10}$$

Wir beachten, dass C die gemeinsame Verteilungsfunktion eines bivariaten Zufallsvektor $(U, V)^\top$ mit $\mathbb{E}[U] = \mathbb{E}[V] = \frac{1}{2}$ ist und folgern, dass die rechte Seite von (3.10) gleich

$$1 - \frac{1}{2} - \frac{1}{2} + \int \int_{[0,1]^2} \tilde{C}(u, v) dC(u, v) = \int \int_{[0,1]^2} \tilde{C}(u, v) dC(u, v).$$

ist. Also ist insgesamt

$$\mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) > 0) = 2 \int \int_{[0,1]^2} \tilde{C}(u, v) dC(u, v)$$

und damit

$$Q = 4 \int \int_{[0,1]^2} \tilde{C}(u, v) dC(u, v) - 1$$

wie gewünscht. ■

Korollar 3.29

Seien C, \tilde{C} und Q wie in Satz 3.28. Dann gelten die folgenden drei Aussagen.

- (a) Q ist symmetrisch, $Q(C, \tilde{C}) = Q(\tilde{C}, C)$.
- (b) Q ist nicht-fallend in jedem Argument.
- (c) Die Copulae C und \tilde{C} können durch ihre Survival-Copulae ersetzt werden. Das heißt,
 $Q(C, \tilde{C}) = Q(\check{C}, \check{\check{C}})$.

Definition 3.30

Ein reellwertiges Funktional κ zweier stetig verteilter Zufallsvariablen X und Y (bzw. ihrer gemeinsamen Verteilung $\mathcal{L}(X, Y)$ mit Copula C) heißt ein Konkordanzmaß, falls gilt:

- (i) κ ist wohldefiniert für alle stetig verteilten bivariaten Zufallsvektoren $(X, Y)^\top$.
- (ii) $-1 \leq \kappa_{X,Y} \leq 1$, $\kappa_{X,X} = 1$, $\kappa_{X,-X} = -1$.
- (iii) $\kappa_{X,Y} = \kappa_{Y,X}$.
- (iv) Sind X und Y stochastisch unabhängig, so ist $\kappa_{X,Y} = \kappa_{\Pi^{(2)}} = 0$.
- (v) $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$.
- (vi) κ ist nicht-fallend in C .
- (vii) Ist $\{(X_n, Y_n)^\top\}_{n \geq 1}$ eine Folge stetig verteilter bivariater Zufallsvektoren mit Copulae $\{C_n\}_{n \geq 1}$ und konvergiert C_n punktweise gegen C für $n \rightarrow \infty$, so ist $\lim_{n \rightarrow \infty} \kappa_{C_n} = \kappa_C$.

Bemerkung 3.31

Falls unter den Voraussetzungen von Definition 3.30 Y fast sicher eine wachsende (fallende) Funktion von X ist, so ist $\kappa_{X,Y} = \kappa_{M^{(2)}} = 1$ ($\kappa_{X,Y} = \kappa_{W^{(2)}} = -1$).
 Ferner ist $\kappa_{X,Y}$ invariant gegenüber komponentenweise strikt isotonen Transformationen.

Definition 3.32

Unter den Voraussetzungen von Definition 3.30 ist das Kendall'sche Konkordanzmaß $\tau(X, Y)$ gegeben durch

$$\tau(X, Y) = \mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) > 0) - \mathbb{P}((X - \tilde{X})(Y - \tilde{Y}) < 0),$$

wobei $(\tilde{X}, \tilde{Y})^\top$ eine unabhängige Kopie von $(X, Y)^\top$ ist, d.h. $(\tilde{X}, \tilde{Y})^\top \perp\!\!\!\perp (X, Y)^\top$ und $\mathcal{L}((\tilde{X}, \tilde{Y})^\top) = \mathcal{L}((X, Y)^\top)$.

Korollar 3.33

Unter den Voraussetzungen von Definition 3.32 ist

$$\tau(X, Y) = Q(C, C) = 4 \int \int_{[0,1]^2} C(u, v) dC(u, v) - 1 = 4\mathbb{E}[C(U, V)] - 1$$

mit $(U, V)^\top \sim C$.

Definition 3.34

Unter den Voraussetzungen von Definition 3.30 ist der Spearman'sche Rangkorrelationskoeffizient ρ_S gegeben durch

$$\rho_S(X, Y) = 3[\mathbb{P}((X - \tilde{X})(Y - Y') > 0) - \mathbb{P}((X - \tilde{X})(Y - Y') < 0)],$$

wobei $(\tilde{X}, \tilde{Y})^\top$ und $(X', Y')^\top$ unabhängige Kopien von $(X, Y)^\top$ sind.

Korollar 3.35

Unter den Voraussetzungen von Definition 3.34 gilt:

$$\begin{aligned} \rho_S(X, Y) &= 3Q(C, \Pi) & (3.11) \\ &= 12 \int \int_{[0,1]^2} uv dC(u, v) - 3 \\ &= 12 \int \int_{[0,1]^2} C(u, v) dudv - 3, \end{aligned}$$

wobei C die Copula von $(X, Y)^\top$ bezeichnet. Falls X die Verteilungsfunktion F und Y die Verteilungsfunktion G besitzt, so definieren wir $U := F(X)$ und $V := G(Y)$ und erhalten

$$\begin{aligned} \rho_S(X, Y) &= 12\mathbb{E}[UV] - 3 & (3.12) \\ &= \frac{\mathbb{E}[UV] - 1/4}{1/12} \\ &= \frac{\text{Cov}(U, V)}{\sqrt{\text{Var}(U)\text{Var}(V)}} = \rho(F(X), G(Y)) \end{aligned}$$

mit $\rho(\cdot, \cdot)$ wie in Definition 3.25.

Beweis: \tilde{X} und Y' sind stochastisch unabhängig und besitzen daher die Produkt-Copula Π . Damit folgt (3.11) direkt aus Satz 3.28. Die Gleichheit (3.12) ergibt sich aus der Tatsache, dass $(U, V)^\top$ die gemeinsame Verteilungsfunktion C besitzen. ■

Satz 3.36

Der lineare Korrelationskoeffizient $\rho(\cdot, \cdot)$ ist nicht invariant gegenüber komponentenweise strikt isotonen Transformationen, und daher kein Konkordanzmaß im Sinne von Definition 3.30.

Beweis: (durch Gegenbeispiel)

Sei die Copula von $(X, Y)^\top$ gegeben durch $C(u, v) = uv + \eta uv(1-u)(1-v)$ für einen Copula-Parameter $\eta \in [-1, 1]$.

Ferner sei die gemeinsame Verteilungsfunktion F von $(X, Y)^\top$ gegeben durch

$$F(x, y) = C(1 - e^{-x}, 1 - e^{-y}), \quad (x, y)^\top \in \mathbb{R}_{\geq 0}^2.$$

Dann ist $X \sim \text{Exp}(1)$ und $Y \sim \text{Exp}(1)$ mit $\mathbb{E}[X] = \mathbb{E}[Y] = 1 = \text{Var}(X) = \text{Var}(Y)$ und damit $\rho(X, X) = \mathbb{E}[XY] - 1$.

Wir rechnen:

$$\begin{aligned}\mathbb{E}[XY] &= \int_0^\infty \int_0^\infty xy dF(x, y) \\ &= \int_0^\infty \int_0^\infty xy \frac{\partial^2 F(x, y)}{\partial x \partial y} dx dy = 1 + \frac{\eta}{4}\end{aligned}$$

(siehe Live-Präsentation mit Maple-Software). Damit ist $\rho(X, Y) = \eta/4$.

Nun ist aber

$$\begin{aligned}\rho(1 - e^{-X}, 1 - e^{-Y}) \\ &= \rho_S(X, Y) \\ &= 12 \int \int_{[0,1]^2} C(u, v) dudv - 3 \\ &= 12 \left(\frac{1}{4} + \frac{\eta}{36} \right) - 3 = \frac{\eta}{3}.\end{aligned}$$

■

3.4 Tail-Abhängigkeit

Definition 3.37

Sei $(X, Y)^\top$ ein Vektor stetig verteilter reellwertiger Zufallsvariablen mit Randverteilungsfunktionen F und G auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann heißt

$$\lambda_U := \lim_{u \uparrow 1} \mathbb{P}(Y > G^{-1}(u) | X > F^{-1}(u))$$

der obere Tailabhängigkeitskoeffizient von $(X, Y)^\top$, sofern dieser Grenzwert in $[0, 1]$ existiert.

Falls $\lambda_U = 0$ ist, so heißen X und Y *asymptotisch unabhängig im oberen Tail*, anderenfalls heißen sie *asymptotisch abhängig im oberen Tail*.

Lemma 3.38

Unter den Voraussetzungen von Definition 3.37 existiere λ_U in $[0, 1]$. Dann gilt

$$\begin{aligned}\lambda_U &= \lim_{u \uparrow 1} \frac{1 - \mathbb{P}(X \leq F^{-1}(u)) - \mathbb{P}(Y \leq G^{-1}(u)) + \mathbb{P}(X \leq F^{-1}(u), Y \leq G^{-1}(u))}{1 - \mathbb{P}(X \leq F^{-1}(u))} \\ &= \lim_{u \uparrow 1} \frac{1 - 2u + C(u, u)}{1 - u} = \lim_{u \uparrow 1} \frac{\bar{C}(u, u)}{1 - u},\end{aligned}$$

wobei C die Copula von $(X, Y)^\top$ bezeichnet.

Beweis: Sei $A := \{Y > G^{-1}(u)\}$ und $B := \{X > F^{-1}(u)\}$. Dann ist

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1 - \mathbb{P}(A^c \cup B^c)}{1 - \mathbb{P}(B^c)} \\ &= \frac{1 - \{\mathbb{P}(A^c) + \mathbb{P}(B^c) - \mathbb{P}(A^c \cap B^c)\}}{1 - \mathbb{P}(B^c)},\end{aligned}$$

was die Aussagen vermittels des Prinzips der Quantilstransformation impliziert. ■

Anmerkung:

Unter den Voraussetzungen von Lemma 3.38 sagt man auch, dass die Copula C obere Tailabhängigkeit besitzt, falls $\lambda_U = \lambda_U(C) > 0$ ist.

Beispiele 3.39

(a) Gumbel-Hougaard-Copulae

Sei

$$C_\eta(u, v) = \exp(-[(-\ln(u))^\eta + (-\ln(v))^\eta]^{\frac{1}{\eta}})$$

für $\eta \geq 1$, vgl. Beispiel 3.17.(c).

Dann ist

$$\frac{1 - 2u + C_\eta(u, u)}{1 - u} = \frac{1 - 2u + \exp(2^{\frac{1}{\eta}} \ln(u))}{1 - u}$$

und damit

$$\begin{aligned}\lambda_U(C_\eta) &= \lim_{u \uparrow 1} \frac{1 - 2u + C_\eta(u, u)}{1 - u} \\ &= \lim_{u \uparrow 1} \frac{1 - 2u + u^{2^{\frac{1}{\eta}}}}{1 - u} \\ &\stackrel{L'Hospital}{=} 2 - \lim_{u \uparrow 1} 2^{\frac{1}{\eta}} u^{2^{\frac{1}{\eta}} - 1} = 2 - 2^{\frac{1}{\eta}}.\end{aligned}$$

Also besitzt C_η genau dann obere Tailabhängigkeit, wenn $\eta > 1$ ist. Damit kann der Parameter η auch als Tailabhängigkeitsmaß angesehen werden.

(b) Gauß-Copulae:

Sei

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

mit linearem Korrelationskoeffizient $-1 < R_{12} < +1$ von

$$(X, Y)^\top \sim \mathcal{N}_2(\mathbf{0}, R) \text{ mit } R = \begin{pmatrix} 1 & R_{12} \\ R_{12} & 1 \end{pmatrix}.$$

Bezeichne $(U, V)^\top$ den Vektor der Verteilungstransformierten von $(X, Y)^\top$ und beachte, dass

$$\begin{aligned}\mathbb{P}(V \leq v|U = u) &= \frac{\partial C_R(u, v)}{\partial u}, \\ \mathbb{P}(V > v|U = u) &= 1 - \frac{\partial C_R(u, v)}{\partial u},\end{aligned}$$

und analog bei Bedingen auf die Realisierung von v . Daraus folgt

$$\begin{aligned}\lambda_U(C_R) &= \lim_{u \uparrow 1} \frac{\bar{C}_R(u, u)}{1 - u} \\ &= -\lim_{u \uparrow 1} \frac{d\bar{C}_R(u, u)}{du} \quad (L'Hospital) \\ &= -\lim_{u \uparrow 1} \left(-2 + \frac{\partial}{\partial s} C_R(s, t) \Big|_{s=t=u} + \frac{\partial}{\partial t} C_R(s, t) \Big|_{s=t=u} \right) \\ &= \lim_{u \uparrow 1} [\mathbb{P}(V > u|U = u) + \mathbb{P}(U > u|V = u)] \\ &= 2 \lim_{u \uparrow 1} \mathbb{P}(V > u|U = u) \\ &= 2 \lim_{x \rightarrow \infty} \mathbb{P}(Y > x|X = x).\end{aligned}$$

Es ist bekannt, dass $\mathcal{L}(Y|X = x) = \mathcal{N}_1(R_{12}x, 1 - R_{12}^2)$.

Also ist

$$\begin{aligned}\lambda_U(C_R) &= 2 \lim_{x \rightarrow \infty} \bar{\Phi} \left(\frac{x - R_{12}x}{\sqrt{1 - R_{12}^2}} \right) \\ &= 2 \lim_{x \rightarrow \infty} \bar{\Phi} \left(x \cdot \frac{\sqrt{1 - R_{12}^2}}{\sqrt{1 + R_{12}^2}} \right) = 0,\end{aligned}$$

da $-1 < R_{12} < +1$ vorausgesetzt war.

Definition 3.40

Eine Copula C besitzt den unteren Tailabhängigkeitskoeffizient $\lambda_L(C)$, falls

$$\lambda_L(C) = \lim_{u \downarrow 0} \frac{C(u, u)}{u}$$

in $[0, 1]$ existiert. Ist $\lambda_L(C) \in (0, 1]$, so sagen wir, dass C untere Tailabhängigkeit besitzt.

Lemma 3.41

(a) Falls $\lambda_L(C)$ in $[0, 1]$ existiert, so gilt

$$\lambda_L(C) = \lim_{u \downarrow 0} [\mathbb{P}(V < u|U = u) + \mathbb{P}(U < u|V = u)],$$

wobei $(U, V)^\top \sim C$.

(b) Der obere Tailabhängigkeitskoeffizient von C stimmt mit dem unteren Tailabhängigkeitskoeffizient der Survival-Copula zu C überein, d.h.,

$$\lambda_U(C) = \lambda_L(\check{C}).$$

Beweis: Für Teil (a) rechnen wir in Analogie zu Beispiel 3.39.(b):

$$\begin{aligned} \lambda_L(C) &= \lim_{u \downarrow 0} \frac{C(u, u)}{u} = \lim_{u \downarrow 0} \frac{dC(u, u)}{du} \\ &= \lim_{u \downarrow 0} \left\{ \left. \frac{\partial}{\partial s} C(s, t) \right|_{s=t=u} + \left. \frac{\partial}{\partial t} C(s, t) \right|_{s=t=u} \right\} \\ &= \lim_{u \downarrow 0} [\mathbb{P}(V < u | U = u) + \mathbb{P}(U < u | V = u)]. \end{aligned}$$

Für Teil (b) beachten wir, dass

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

gilt, siehe Definition 3.21 und Beispiel 3.24.(b).

Damit ist

$$\begin{aligned} \lambda_U(C) &= \lim_{u \uparrow 1} \frac{\bar{C}(u, u)}{1 - u} = \lim_{u \uparrow 1} \frac{\check{C}(1 - u, 1 - u)}{1 - u} \\ &= \lim_{w \downarrow 0} \frac{\check{C}(w, w)}{w} = \lambda_L(\check{C}). \end{aligned}$$

■

3.5 Elliptische Copulae

Definition 3.42

Sei \mathbf{X} ein d -dimensionaler Zufallsvektor. Falls $\boldsymbol{\mu} \in \mathbb{R}^d$ und $\Sigma \in \mathbb{R}^{d \times d}$ existieren, so dass

$$\varphi_{\mathbf{X} - \boldsymbol{\mu}}(\mathbf{t}) = h(\mathbf{t}^\top \Sigma \mathbf{t})$$

für alle $\mathbf{t} \in \mathbb{R}^d$ ist, so sagt man, dass \mathbf{X} eine elliptische Verteilung mit Parametern $\boldsymbol{\mu}$, Σ und h besitzt und schreibt $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, h)$. Dann heißt h die Generatorfunktion von $\mathcal{L}(\mathbf{X})$.

Wir starten unsere Untersuchungen mit dem Spezialfall $\Sigma = I_d$ (Identitätsmatrix im \mathbb{R}^d , sphärischer Fall).

Lemma 3.43 (Schoenberg (1938))

Sei ψ eine Funktion von $[0, \infty)$ nach \mathbb{R} . Dann ist $\mathbf{t} \rightarrow \psi(\mathbf{t}^\top \mathbf{t})$ für $\mathbf{t} \in \mathbb{R}^k$ genau dann die charakteristische Funktion eines k -dimensionalen Zufallsvektor \mathbf{X} , falls

$$\forall u \geq 0 : \psi(u) = \int_0^\infty \Omega_k(r^2 u) dF_R(r), \quad (3.13)$$

wobei F_R eine Verteilungsfunktion auf $[0, \infty)$ und $\mathbf{t} \rightarrow \Omega_k(\mathbf{t}^\top \mathbf{t})$, $\mathbf{t} \in \mathbb{R}^k$, die charakteristische Funktion eines Zufallsvektor \mathbf{U} ist, der gleichverteilt auf der Einheitskugel $S_k = \{\mathbf{z} \in \mathbb{R}^k : \mathbf{z}^\top \mathbf{z} = 1\}$ ist. Ferner ist F_R die Verteilungsfunktion von $R = (\mathbf{X}^\top \mathbf{X})^{\frac{1}{2}}$.

Beweis: Nehmen wir zunächst an, ψ besitzt die in (3.13) angegebene Darstellung. Sei $R \sim F_R$ und sei \mathbf{U} stochastisch unabhängig von R und auf S_k gleichverteilt. Setze $\mathbf{X} := R\mathbf{U}$. Dann gilt für die charakteristische Funktion von \mathbf{X} :

$$\begin{aligned}
 \varphi_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E}[\exp(i\mathbf{t}^\top \mathbf{X})] \\
 &= \mathbb{E}[\exp(i\mathbf{t}^\top R\mathbf{U})] \\
 &= \mathbb{E}[\mathbb{E}[\exp(i\mathbf{t}^\top R\mathbf{U})|R]] \\
 &= \int_0^\infty \mathbb{E}[\exp(i\mathbf{t}^\top R\mathbf{U})|R=r] dF_R(r) \\
 &= \int_0^\infty \varphi_{\mathbf{U}}(r\mathbf{t}) dF_R(r) \\
 &= \int_0^\infty \Omega_k(r^2 \mathbf{t}^\top \mathbf{t}) dF_R(r).
 \end{aligned}$$

Also induziert ψ tatsächlich $\varphi_{\mathbf{X}}$ in der behaupteten Art und Weise. Außerdem ist

$$\begin{aligned}
 F_R(r) &= \mathbb{P}(R \leq r) = \mathbb{P}(\sqrt{R^2} \leq r) \\
 &= \mathbb{P}(\sqrt{(R\mathbf{U})^\top (R\mathbf{U})} \leq r) \\
 &= \mathbb{P}((\mathbf{X}^\top \mathbf{X})^{\frac{1}{2}} \leq r).
 \end{aligned}$$

Dies komplettiert den Beweis der „Rückrichtung“ der behaupteten Äquivalenz.

Nehmen wir nun andererseits an, dass $\mathbf{t} \rightarrow \psi(\mathbf{t}^\top \mathbf{t})$ die charakteristische Funktion eines Zufallsvektor \mathbf{X} mit Verteilungsfunktion G ist. Bezeichne $A_k = 2\pi^{k/2}/\Gamma(k/2)$ die Oberfläche von S_k und $d\omega_k$ die Integration auf S_k .

Wir beachten, das $\psi(u) = \psi(u\mathbf{t}^\top \mathbf{t})$ ist, falls $\mathbf{t}^\top \mathbf{t} = 1$ gilt. Demnach gilt:

$$\begin{aligned}
\psi(u) &= A_k^{-1} \int_{S_k} \psi(u\mathbf{t}^\top \mathbf{t}) d\omega_k(\mathbf{t}) \\
&= A_k^{-1} \int_{S_k} \varphi_{\mathbf{X}}(\sqrt{u}\mathbf{t}) d\omega_k(\mathbf{t}) \\
&= A_k^{-1} \int_{S_k} \int_{\mathbb{R}^k} \exp(i\sqrt{u}\mathbf{t}^\top \mathbf{x}) dG(\mathbf{x}) d\omega_k(\mathbf{t}) \\
&= \int_{\mathbb{R}^k} \left[A_k^{-1} \int_{S_k} \exp(i\sqrt{u}\mathbf{x}^\top \mathbf{t}) d\omega_k(\mathbf{t}) \right] dG(\mathbf{x}) \\
&= \int_{\mathbb{R}^k} \Omega_k((\sqrt{u}\mathbf{x})^\top (\sqrt{u}\mathbf{x})) dG(\mathbf{x}) \\
&= \int_{\mathbb{R}^k} \Omega_k(u\mathbf{x}^\top \mathbf{x}) dG(\mathbf{x}) = \int_0^\infty \Omega_k(ur^2) dF_R(r)
\end{aligned}$$

mit $F_R(r) = \mathbb{P}((\mathbf{X}^\top \mathbf{X})^{\frac{1}{2}} \leq r)$. ■

Korollar 3.44

Es ist $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, h)$ mit $\text{rang}(\Sigma) = k$ genau dann, wenn es eine Zufallsvariable $R \geq 0$, einen auf S_k gleichverteilten Zufallsvektor \mathbf{U} mit $R \perp \mathbf{U}$, sowie eine $(d \times k)$ -Matrix A mit $AA^\top = \Sigma$ gibt, so dass

$$X \stackrel{\mathcal{D}}{=} \boldsymbol{\mu} + RA\mathbf{U}.$$

Beispiel 3.45 (Normalverteilungen)

Sei $\mathbf{X} = (X_1, \dots, X_d)^\top \sim \mathcal{N}_d(\mathbf{0}, I_d)$. Dann sind $(X_i)_{1 \leq i \leq d}$ i.i.d. mit $X_1 \sim \mathcal{N}(0, 1)$. Bekanntermaßen ist damit $\varphi_{X_1} = \varphi_{\mathcal{N}(0,1)}$ mit $\varphi_{X_1}(u) = \exp(-u^2/2)$.

Wegen der stochastischen Unabhängigkeit der $(X_i)_{1 \leq i \leq d}$ ist ferner

$$\begin{aligned}
\varphi_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E}[\exp(i\mathbf{t}^\top \mathbf{X})] \\
&= \mathbb{E} \left[\prod_{j=1}^d \exp(it_j X_j) \right] \\
&= \prod_{j=1}^d \mathbb{E}[\exp(it_j X_j)] \\
&= \prod_{j=1}^d \exp(-t_j^2/2) \\
&= \exp\left(-\frac{1}{2} \sum_{j=1}^d t_j^2\right) = \exp\left(-\frac{1}{2} \mathbf{t}^\top \mathbf{t}\right).
\end{aligned}$$

Also ist $\mathbf{X} \sim E_d(\mathbf{0}, I_d, h)$ mit $h(u) = \exp(-u/2)$.

Analog (siehe Beispiel 1.19.d)) verifiziert man, dass $\varphi_{\mathbf{X}-\boldsymbol{\mu}}(\mathbf{t}) = \exp(-\frac{1}{2}\mathbf{t}^\top \Sigma \mathbf{t})$ ist, falls $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$. Somit ist dann $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, h)$ mit dem selben Generator h wie zuvor.

Definition 3.46 (Sphärische Verteilung)

Gilt für einen \mathbb{R}^k -wertigen Zufallsvektor \mathbf{Y} , dass $\mathbf{Y} \sim E_k(\mathbf{0}, I_k, h)$ ist, so heißt $\mathcal{L}(\mathbf{Y})$ sphärische Verteilung mit Generatorfunktion h und wir schreiben abkürzend $\mathbf{Y} \sim S_k(h)$.

Unter den Voraussetzungen von Korollar 3.44 gilt mit $\mathbf{Y} \sim S_k(h)$, dass $\mathbf{X} \stackrel{\mathcal{D}}{=} \boldsymbol{\mu} + A\mathbf{Y}$.

Anmerkung:

Für $\mathbf{Y} \sim S_k(h)$ gilt $\varphi_{\mathbf{Y}}(\mathbf{t}) = h(\mathbf{t}^\top \mathbf{t})$, $\mathbf{t} \in \mathbb{R}^k$.

Satz 3.47

- (a) Sei $\mathbf{Y} \sim S_k(h)$. Dann besitzt $\mathbf{Y} \stackrel{\mathcal{D}}{=} R\mathbf{U}$ genau dann eine Lebesgue-dichte $\mathbf{y} \mapsto g(\mathbf{y}^\top \mathbf{y})$, wenn R eine Lebesgue-dichte f_R besitzt. Ferner gilt dann

$$f_R(r) = \frac{2\pi^{k/2}}{\Gamma(k/2)} r^{k-1} g(r^2).$$

- (b) Eine nicht-negative Funktion $g : [0, \infty) \rightarrow [0, \infty)$ definiert die Dichte $\mathbf{y} \mapsto c \cdot g(\mathbf{y}^\top \mathbf{y})$ einer sphärischen Verteilung genau dann, wenn

$$\int_0^\infty z^{k/2-1} g(z) dz < \infty.$$

Der Beweis von Satz 3.47 basiert auf dem folgenden Lemma, das wir ohne Beweis angeben.

Lemma 3.48 (siehe Abschnitt 1.4 in Fang et al. (1990))

Es sei f eine nicht-negative Borel-messbare Funktion. Dann gilt für alle $\mathbf{y} = (y_1, \dots, y_k)^\top \in \mathbb{R}^k$, dass

$$\int_{\mathbb{R}^k} f\left(\sum_{i=1}^k y_i^2\right) dy_1 \dots dy_k = \int_{\mathbb{R}^k} f(\mathbf{y}^\top \mathbf{y}) d\mathbf{y} = \frac{\pi^{k/2}}{\Gamma(k/2)} \int_0^\infty z^{k/2-1} f(z) dz.$$

Beweis: (Beweis von Satz 3.47)

zu Teil (a):

Angenommen, $f_{\mathbf{Y}}(\mathbf{y}) = g(\mathbf{y}^\top \mathbf{y})$. Sei $u(\cdot)$ eine beliebig ausgewählte, nicht-negative, Borel-messbare Funktion. Dann ist nach Lemma 3.48

$$\mathbb{E}[u(R)] = \int_{\mathbb{R}^k} u\left\{(\mathbf{y}^\top \mathbf{y})^{1/2}\right\} g(\mathbf{y}^\top \mathbf{y}) d\mathbf{y} = \frac{\pi^{k/2}}{\Gamma(k/2)} \int_0^\infty z^{k/2-1} u(z^{1/2}) g(z) dz. \quad (3.14)$$

Mit der Substitution

$$r := z^{1/2} \iff r^2 = z, \quad \frac{dr}{dz} = \frac{1}{2}z^{-1/2} \Rightarrow dz = 2z^{1/2} dr$$

ergibt sich die rechte Seite von (3.14) zu

$$\frac{2\pi^{k/2}}{\Gamma(k/2)} \int_0^\infty u(r)g(r^2)r^{k-1}dr = \mathbb{E}[u(R)].$$

Also hat R die Lebesguedichte $r \mapsto f_R(r) = \frac{2\pi^{k/2}}{\Gamma(k/2)}g(r^2)r^{k-1}$ wie gewünscht. Die Rückrichtung ist offensichtlich.

zu Teil (b):

Gelte $f_Y(\mathbf{y}) = g(\mathbf{y}^\top \mathbf{y})$. Dann ist nach Lemma 3.48

$$\int_{\mathbb{R}^k} g(\mathbf{y}^\top \mathbf{y})d\mathbf{y} = \frac{\pi^{k/2}}{\Gamma(k/2)} \int_0^\infty z^{k/2-1}g(z)dz,$$

was die Aussage impliziert. ■

Korollar 3.49

Es sei $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, h)$. Dann besitzt \mathbf{X} eine Lebesguedichte, falls $\text{rang}(\Sigma) = d$ ist und \mathbf{Y} in der stochastischen Darstellung $\mathbf{X} \stackrel{D}{=} \boldsymbol{\mu} + A\mathbf{Y}$ eine Lebesguedichte $\mathbf{y} \mapsto g(\mathbf{y}^\top \mathbf{y})$ besitzt, wobei g die Bedingung aus Satz 3.47.(b) erfüllt. Die Lebesguedichte von \mathbf{X} ist dann von der Form

$$\mathbf{x} \mapsto |\Sigma|^{-1/2}g((\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})).$$

Beispiel 3.50 (Skalen-Mischung von Normalverteilungen)

Sei \mathbf{X} ein \mathbb{R}^d -wertiger Zufallsvektor mit stochastischer Darstellung $\mathbf{X} \stackrel{D}{=} \boldsymbol{\mu} + \sqrt{W}A\mathbf{Z}$, wobei

- (i) $\mathbf{Z} \sim \mathcal{N}_k(\mathbf{0}, I_k)$,
- (ii) W eine nicht-negative, skalare Zufallsvariable, stochastisch unabhängig von \mathbf{Z} ,
- (iii) $A \in \mathbb{R}^{d \times k}$ und $\boldsymbol{\mu} \in \mathbb{R}^d$ konstant sind.

Wir beachten, dass

$$\mathcal{L}(\mathbf{X}|W = w) = \mathcal{N}_d(\boldsymbol{\mu}, w\Sigma)$$

mit $\Sigma = AA^\top$ ist, wobei wir annehmen, dass Σ positiv definit ist.

Die Verteilung von \mathbf{X} heißt Skalen-Mischung von Normalverteilungen, und wir schreiben $\mathbf{X} \sim \mathcal{M}_d(\boldsymbol{\mu}, \Sigma, W)$.

Falls $W > 0$ fast sicher gilt, so ist die Lebesguedichte von \mathbf{X} gegeben durch

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \int_0^{\infty} f_{\mathbf{X}|W}(\mathbf{x}|w) dF_W(w) \\ &= |\Sigma|^{-\frac{1}{2}} \int_0^{\infty} \frac{w^{-\frac{d}{2}}}{(2\pi)^{\frac{d}{2}}} \exp\left(-\frac{(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}{2w}\right) dF_W(w). \end{aligned}$$

Also besitzt \mathbf{X} eine elliptische Verteilung.

Falls W eine inverse Gammaverteilung mit Parametern $\nu/2$ und $\nu/2$ besitzt (bzw. äquivalenterweise $\nu/W \sim \chi_\nu^2$ gilt), so besitzt \mathbf{X} eine multivariate t -Verteilung mit Lebesguedichte $f_{\mathbf{X}}$, gegeben durch

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma((\nu + d)/2)}{\Gamma(\nu/2)(\pi\nu)^{d/2}|\Sigma|^{1/2}} \left(1 + \frac{(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}{\nu}\right)^{-\frac{\nu+d}{2}}.$$

Damit können Pseudo-Zufallsvektoren, die sich wie Realisierungen von \mathbf{X} verhalten, nach dem folgenden Algorithmus generiert werden:

- 1) Ermittle A aus der Cholesky-Zerlegung von Σ
- 2) Generiere $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, I_d)$
- 3) Generiere $S \sim \chi_\nu^2$ und setze $W = \nu/S$
- 4) Setze $\mathbf{X} = \boldsymbol{\mu} + \sqrt{W}AZ$.

Satz 3.51

Sei \mathbf{U} gleichverteilt auf der Einheitssphäre $S_k := \{\mathbf{z} \in \mathbb{R}^k : \mathbf{z}^\top \mathbf{z} = 1\}$. Dann gilt $\mathbb{E}[\mathbf{U}] = \mathbf{0}$, $Cov(\mathbf{U}) = k^{-1}I_k$.

Beweis: Sei $\mathbf{X} \sim \mathcal{N}_k(0, I_k)$. Dann ist $\mathbf{X} \stackrel{\mathcal{D}}{=} \|\mathbf{X}\|_2 \mathbf{U}$ mit $\mathbb{E}[\mathbf{X}] = \mathbf{0}$, $Cov(\mathbf{X}) = I_k$ und $\|\mathbf{X}\|_2^2 \perp \mathbf{U}$. Ferner ist $\|\mathbf{X}\|_2^2 \sim \chi_k^2$ mit $\mathbb{E}[\|\mathbf{X}\|_2^2] = k$ und es gilt $\mathbb{E}[\|\mathbf{X}\|_2] > 0$.

Dann folgt sofort, dass

$$0 = \mathbb{E}[\mathbf{X}] = \mathbb{E}[\|\mathbf{X}\|_2] \mathbb{E}[\mathbf{U}] \Rightarrow \mathbb{E}[\mathbf{U}] = \mathbf{0}.$$

Analog rechnen wir

$$\begin{aligned} I_k = Cov(\mathbf{X}) &= Cov(\|\mathbf{X}\|_2 \mathbf{U}) \\ &= \mathbb{E}[\|\mathbf{X}\|_2 \mathbf{U} (\|\mathbf{X}\|_2 \mathbf{U})^\top] \\ &= \mathbb{E}[\|\mathbf{X}\|_2^2 \mathbf{U} \mathbf{U}^\top] \\ &= k \mathbb{E}[\mathbf{U} \mathbf{U}^\top] = k \cdot Cov(\mathbf{U}). \end{aligned}$$

■

Korollar 3.52

(a) Ist $\mathbf{Y} \stackrel{D}{=} RU \sim S_k(h)$ mit $\mathbb{E}[R^2] < \infty$, so ist $\mathbb{E}[\mathbf{Y}] = 0$, $Cov(\mathbf{Y}) = \mathbb{E}[R^2]I_k/k$.

(b) Ist $\mathbf{X} \stackrel{D}{=} \boldsymbol{\mu} + A\mathbf{Y}$ mit $AA^\top = \Sigma \in \mathbb{R}^{d \times d}$, so ist

$$\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu} \text{ und } Cov(\mathbf{X}) = \Sigma \cdot \mathbb{E}[R^2]/k = \frac{\mathbb{E}[R^2]}{\text{rang}(\Sigma)} \cdot \Sigma.$$

Satz 3.53

Sei $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, h)$, $B \in \mathbb{R}^{q \times d}$ und $\mathbf{b} \in \mathbb{R}^q$. Dann ist $B\mathbf{X} + \mathbf{b} \sim E_q(\mathbf{b} + B\boldsymbol{\mu}, B\Sigma B^\top, h)$.

Beweis: Nach Korollar 3.44 gilt

$$\mathbf{X} \stackrel{D}{=} \boldsymbol{\mu} + RAU.$$

Also ist

$$\begin{aligned} B\mathbf{X} + \mathbf{b} &\stackrel{D}{=} B\boldsymbol{\mu} + BRAU + \mathbf{b} \\ &= \mathbf{b} + B\boldsymbol{\mu} + RBAU. \end{aligned}$$

Beachte schließlich, dass $BA(BA)^\top = B\Sigma B^\top$ ist. ■

Korollar 3.54

Sei $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, h)$. Wir partitionieren

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

mit $\mathbf{X}^{(1)} \in \mathbb{R}^q$, $\boldsymbol{\mu}^{(1)} \in \mathbb{R}^q$ und $\Sigma_{11} \in \mathbb{R}^{q \times q}$. Dann ist

$$\begin{aligned} \mathbf{X}^{(1)} &\sim E_q(\boldsymbol{\mu}^{(1)}, \Sigma_{11}, h) \text{ und} \\ \mathbf{X}^{(2)} &\sim E_{d-q}(\boldsymbol{\mu}^{(2)}, \Sigma_{22}, h). \end{aligned}$$

Korollar 3.55 (Elliptische Copula)

Sei $F(\cdot|\Sigma, h)$ die gemeinsame Verteilungsfunktion von $\mathbf{X} \sim E_d(\mathbf{0}, \Sigma, h)$ und $(F_j(\cdot|\Sigma, h) : 1 \leq j \leq d)$ die zugehörigen Randverteilungsfunktionen, vgl. Korollar 3.54.

Dann ist gemäß Korollar 3.16 die elliptische Copula von \mathbf{X} gegeben durch

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1|\Sigma, h), \dots, F_d^{-1}(u_d|\Sigma, h)|\Sigma, h).$$

Beispielweise ist die d -variate t -Copula mit ν Freiheitsgraden und Korrelationsmatrix R gegeben durch

$$\mathbf{u} \mapsto F_{t_d(\nu, R)}(F_{t_\nu}^{-1}(u_1), \dots, F_{t_\nu}^{-1}(u_d)),$$

wobei $F_{t_d(\nu, R)}$ die Verteilungsfunktion von $\sqrt{\nu}\mathbf{Y}/\sqrt{S}$ mit $S \sim \chi_\nu^2$ stochastisch unabhängig von $\mathbf{Y} \sim \mathcal{N}_d(\mathbf{0}, R)$, und F_{t_ν} die Verteilungsfunktion der univariaten Studentischen t -Verteilung bezeichnet.

Satz 3.56

Sei $\mathbf{Y} \sim S_d(h)$ mit $\mathbb{P}(\mathbf{Y} = \mathbf{0}) = 0$. Sei ferner T eine Abbildung mit der Eigenschaft $T(\alpha\mathbf{Y}) \stackrel{\mathcal{D}}{=} T(\mathbf{Y}) \forall \alpha > 0$. Dann ist $T(\mathbf{Y}) \stackrel{\mathcal{D}}{=} T(\mathbf{Z})$, $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, I_d)$.

Beweis: Nach Voraussetzung ist $\mathbf{Y} \stackrel{\mathcal{D}}{=} R\mathbf{U}$ mit $R \sim F_R$ stochastisch unabhängig von \mathbf{U} . Damit gilt

$$\begin{aligned} T(\mathbf{Y}) &\stackrel{\mathcal{D}}{=} T(R\mathbf{U}) = \int_0^\infty T(R\mathbf{U}|R=r) dF_R(r) \\ &= \int_0^\infty T(r\mathbf{U}) dF_R(r) \stackrel{\mathcal{D}}{=} \int_0^\infty T(\mathbf{U}) dF_R(r) \\ &= T(\mathbf{U}), \end{aligned}$$

unabhängig von F_R bzw h .

Da $\mathbf{Z} \sim S_d(h)$ mit $h(u) = \exp(-u/2)$ und $\mathbb{P}(\mathbf{Z} = \mathbf{0}) = 0$ ist, folgt die Aussage. ■

Lemma 3.57

Sei $\mathbf{X} = (X_1, X_2)^\top \sim \mathcal{N}_2(\boldsymbol{\mu}, \Sigma)$ und $\tilde{\mathbf{X}} = (\tilde{X}_1, \tilde{X}_2)^\top$ eine unabhängige Kopie von \mathbf{X} . Definiere $\mathbf{Y} = (Y_1, Y_2)^\top = \mathbf{X} - \tilde{\mathbf{X}}$.

Dann gelten:

(i) $\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, 2\Sigma)$

(ii) Schreiben wir $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$, so ist $\mathbf{Y} \stackrel{\mathcal{D}}{=} \sqrt{2}(\sigma_1 V \cos(\varphi) + \sigma_1 W \sin(\varphi), \sigma_2 W)$
mit $\varphi = \arcsin(\rho) \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ und $(V, W)^\top \sim \mathcal{N}_2(\mathbf{0}, I_2)$.

Beweis: Normalverteilung und Zentriertheit von \mathbf{Y} sind offensichtlich. Wir rechnen

$$\text{Var}(Y_1) = \text{Var}(X_1 - \tilde{X}_1) = \text{Var}(X_1) + \text{Var}(\tilde{X}_1) = 2\sigma_1^2,$$

$$\text{Var}(Y_2) = 2\sigma_2^2,$$

$$\text{Cov}(Y_1, Y_2) = \text{Cov}(X_1 - \tilde{X}_1, X_2 - \tilde{X}_2) = \text{Cov}(X_1, X_2) + \text{Cov}(\tilde{X}_1, \tilde{X}_2) = 2\sigma_1\sigma_2\rho,$$

was Aussage (i) impliziert.

Zum Nachweis von (ii) definiere

$$\begin{aligned} Z_1 &:= \sigma_1 V \cos(\varphi) + \sigma_1 W \sin(\varphi) \\ &= \sigma_1 \sqrt{1 - \rho^2} V + \sigma_1 \rho W \text{ und} \\ Z_2 &:= \sigma_2 W. \end{aligned}$$

Normalverteilung und Zentriertheit von $\mathbf{Z} = (Z_1, Z_2)^\top$ sind wiederum offensichtlich, und wir rechnen

$$\text{Var}(Z_1) = \sigma_1^2(1 - \rho^2) + \sigma_1^2\rho^2 = \sigma_1^2,$$

$$\text{Var}(Z_2) = \sigma_2^2,$$

$$\text{Cov}(Z_1, Z_2) = \text{Cov}(\sigma_1\sqrt{1 - \rho^2}V + \sigma_1\rho W, \sigma_2 W) = \sigma_1\sigma_2\rho \cdot \text{Var}(W) = \sigma_1\sigma_2\rho,$$

was (ii) impliziert. ■

Satz 3.58

Unter den Voraussetzungen von Lemma 3.57 gilt $\tau(X_1, X_2) = \frac{2}{\pi} \arcsin(\rho)$.

Beweis: Wir verwenden die Notation des Beweises von Lemma 3.57 und rechnen:

$$\begin{aligned} \tau(X_1, X_2) &= 2\mathbb{P}((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0) - 1 \\ &= 2\mathbb{P}(Y_1 Y_2 > 0) - 1 \\ &= 4\mathbb{P}(Y_1 > 0, Y_2 > 0) - 1 \\ &= 4\mathbb{P}(V \cos(\varphi) + W \sin(\varphi) > 0, W > 0) - 1. \end{aligned}$$

Wähle nun $\Phi \sim \text{UNI}[-\pi, \pi)$ stochastisch unabhängig von $R := \sqrt{V^2 + W^2}$. Dann ist $(V, W) \stackrel{D}{=} (R \cos(\Phi), R \sin(\Phi))$ nach Box and Muller (1958).

Es folgt

$$\begin{aligned} \tau(X_1, X_2) &= 4\mathbb{P}(\cos(\Phi) \cos(\varphi) + \sin(\Phi) \sin(\varphi) > 0, \sin(\Phi) > 0) - 1 \\ &= 4\mathbb{P}(\Phi \in (\varphi - \pi/2, \varphi + \pi/2) \cap (0, \pi)) - 1 \\ &= 4\mathbb{P}(\Phi \in (0, \varphi + \pi/2)) - 1 \\ &= 4 \frac{\varphi + \pi/2}{2\pi} - 1 = \frac{2\varphi}{\pi} = \frac{2}{\pi} \arcsin(\rho) \end{aligned}$$

wie gewünscht. ■

Korollar 3.59

Mit einer Argumentation analog zu der im Beweis von Satz 3.56 erhält man das folgende Resultat.

$\mathbf{X} = (X_1, X_2)^\top \sim E_2(\boldsymbol{\mu}, \Sigma, h)$ besitze endliche zweite Momente sowie eine Lebesgue-dichte.

Dann ist $\tau(X_1, X_2) = \frac{2}{\pi} \arcsin(\rho(X_1, X_2))$.

3.6 Archimedische Copulae

Erinnerung 3.60 (an Beispiel 3.17. (c))

Sei $\psi : [0, \infty) \rightarrow [0, 1]$ mit $\psi(0) = 1$, $\lim_{x \rightarrow \infty} \psi(x) = 0$ eine Funktion mit

$$\forall 1 \leq k \leq d : \forall x \in [0, \infty) : (-1)^k \psi^{(k)}(x) \geq 0.$$

dann heißt $C_\psi : [0, 1]^d \rightarrow [0, 1]$ mit $C_\psi(\mathbf{u}) = \psi\left(\sum_{k=1}^d \psi^{-1}(u_k)\right)$ Archimedische Copula mit Generatorfunktion ψ .

3.6.1 Eigenschaften bivariater Archimedischer Copulae

Erinnerung 3.61 (Archimedisches Axiom)

Für $a, b \in \mathbb{R}_+$ $\exists n \in \mathbb{N} : na > b$.

Definition 3.62

Sei C eine bivariate Copula und $u \in [0, 1]$. Dann ist die n -te C -Potenz u_C^n von u definiert durch

$$u_C^1 := u, \quad u_C^{n+1} := C(u, u_C^n), \quad n \geq 1.$$

Satz 3.63

Sei C_ψ eine bivariate Archimedische Copula mit Generatorfunktion ψ . Dann existiert für alle $(u, v) \in (0, 1)^2$ ein $n \in \mathbb{N}$ mit $u_C^n < v$.

Beweis: Sei $(u, v) \in (0, 1)^2$ beliebig ausgewählt. Offensichtlich ist $u_C^n = \psi(n\psi^{-1}(u))$. Da $\psi^{-1}(u)$ und $\psi^{-1}(v)$ positive reelle Zahlen sind, kann Erinnerung 3.61 angewendet werden. Also $\exists n \in \mathbb{N}$ mit $n\psi^{-1}(u) > \psi^{-1}(v)$. Da $v > 0$ ist, gilt $\psi^{-1}(v) < \psi^{-1}(0)$ und damit $v = \psi(\psi^{-1}(v)) > \psi(n\psi^{-1}(u)) = u_C^n$. ■

Beispiele 3.64

$$\begin{aligned} (a) \text{ Sei } \psi^{-1}(t) &= -\ln\left(\frac{\exp(-\eta t)-1}{\exp(-\eta)-1}\right), \quad \eta \in \mathbb{R} \setminus \{0\} \\ &\Rightarrow \psi(x) = -\frac{1}{\eta} \ln\{1 + \exp(-x)[\exp(-\eta) - 1]\} \\ &\Rightarrow C_\psi(u, v) = \psi(\psi^{-1}(u) + \psi^{-1}(v)) \\ &= -\frac{1}{\eta} \ln\left(1 + \frac{(e^{-\eta u} - 1)(e^{-\eta v} - 1)}{e^{-\eta} - 1}\right). \end{aligned}$$

Diese Copula-Familie heißt Familie der Frank-Copulae mit Parameter $\eta \in \mathbb{R} \setminus \{0\}$. Sie besitzt die Eigenschaft der Radialsymmetrie, d.h.,

$$C_\psi(u, v) = \check{C}_\psi(u, v).$$

In der Klasse der bivariaten Archimedischen Copulae ist dies eine ausgezeichnete Eigenschaft der Frank-Familie.

(b) Sei $\psi^{-1}(t) = 1 - t$, $t \in [0, 1]$. Dann ist

$$\psi(x) = \begin{cases} 1 - x, & x \in [0, 1] \\ 0, & x > 1 \end{cases} = \max(1 - x, 0)$$

$$\begin{aligned} \Rightarrow C_\psi(u, v) &= \max(1 - (1 - u + 1 - v), 0) \\ &= \max(u + v - 1, 0) \\ &= W^{(2)}(u, v). \end{aligned}$$

Also bilden die unteren Fréchet-Hoeffding-Schranken ebenfalls eine Archimedische Copula, $d = 2$.

Satz 3.65

Sei C_ψ eine Archimedische Copula mit Generator ψ . Dann gilt

- (i) C_ψ ist symmetrisch, d.h. $C_\psi(u, v) = C_\psi(v, u)$ für alle $(u, v) \in [0, 1]^2$.
- (ii) C_ψ ist assoziativ, d.h. $C_\psi(C_\psi(u, v), w) = C_\psi(u, C_\psi(v, w))$ für alle $u, v, w \in [0, 1]$.

Beweis: Teil (i) ist offensichtlich.

Für Teil (ii) rechnen wir

$$\begin{aligned} C_\psi(C_\psi(u, v), w) &= \psi(\psi^{-1}(C_\psi(u, v)) + \psi^{-1}(w)) \\ &= \psi(\psi^{-1}\{\psi(\psi^{-1}(u) + \psi^{-1}(v))\} + \psi^{-1}(w)) \\ &= \psi(\psi^{-1}(u) + \psi^{-1}(v) + \psi^{-1}(w)) \\ &= \psi(\psi^{-1}(u) + \psi^{-1}\{\psi(\psi^{-1}(v) + \psi^{-1}(w))\}) \\ &= \psi(\psi^{-1}(u) + \psi^{-1}(C_\psi(v, w))) \\ &= C_\psi(u, C_\psi(v, w)). \end{aligned}$$

■

Satz 3.66

Sei C_ψ eine bivariate Archimedische Copula mit Generator ψ .

Bezeichne für $\gamma \in (0, 1)$ die Menge $\{(u, v) \in [0, 1]^2 : C_\psi(u, v) = \gamma\}$ die Konturlinie von C_ψ zum Konturniveau γ .

Dann gelten:

- (a) Die Konturlinie von C_ψ zum Konturniveau γ ist gegeben durch die Gleichung

$$v = L_\gamma(u) = \psi(\psi^{-1}(\gamma) - \psi^{-1}(u)).$$

Also gilt $L_\gamma(1) = \gamma$ und $L_\gamma(\gamma) = 1$.

(b) Die Konturlinien von C_ψ sind konvex.

Beweis: zu (a):

$$\begin{aligned}
C_\psi(u, v) &= \gamma \\
\Leftrightarrow \psi(\psi^{-1}(u) + \psi^{-1}(v)) &= \gamma \\
\Leftrightarrow \psi^{-1}(u) + \psi^{-1}(v) &= \psi^{-1}(\gamma) \\
\Leftrightarrow \psi^{-1}(v) &= \psi^{-1}(\gamma) - \psi^{-1}(u) \\
\Leftrightarrow v &= \psi(\psi^{-1}(\gamma) - \psi^{-1}(u)).
\end{aligned}$$

zu (b): Wegen der Stetigkeit von L_γ reicht es zu zeigen, dass für beliebige $u_1, u_2 \in [0, 1]$ gilt:

$$L_\gamma\left(\frac{u_1 + u_2}{2}\right) \leq \frac{L_\gamma(u_1) + L_\gamma(u_2)}{2}, \quad \gamma \in (0, 1).$$

Da ψ fallend und konvex ist, ist auch ψ^{-1} konvex. Also ist

$$\begin{aligned}
\psi^{-1}(\gamma) - \psi^{-1}\left(\frac{u_1 + u_2}{2}\right) &\geq \psi^{-1}(\gamma) - \frac{\psi^{-1}(u_1) + \psi^{-1}(u_2)}{2} \\
&= \frac{[\psi^{-1}(\gamma) - \psi^{-1}(u_1)] + [\psi^{-1}(\gamma) - \psi^{-1}(u_2)]}{2} \quad (3.15)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow L_\gamma\left(\frac{u_1 + u_2}{2}\right) &= \psi\left\{\psi^{-1}(\gamma) - \psi^{-1}\left(\frac{u_1 + u_2}{2}\right)\right\} \\
&\stackrel{(3.15)}{\leq} \psi\left\{\frac{[\psi^{-1}(\gamma) - \psi^{-1}(u_1)] + [\psi^{-1}(\gamma) - \psi^{-1}(u_2)]}{2}\right\} \\
&\leq \frac{1}{2}\{\psi(\psi^{-1}(\gamma) - \psi^{-1}(u_1)) + \psi(\psi^{-1}(\gamma) - \psi^{-1}(u_2))\} \quad (3.16) \\
&= \frac{L_\gamma(u_1) + L_\gamma(u_2)}{2}.
\end{aligned}$$

Dabei gilt (3.16) wegen der Konvexität von ψ . ■

Satz 3.67

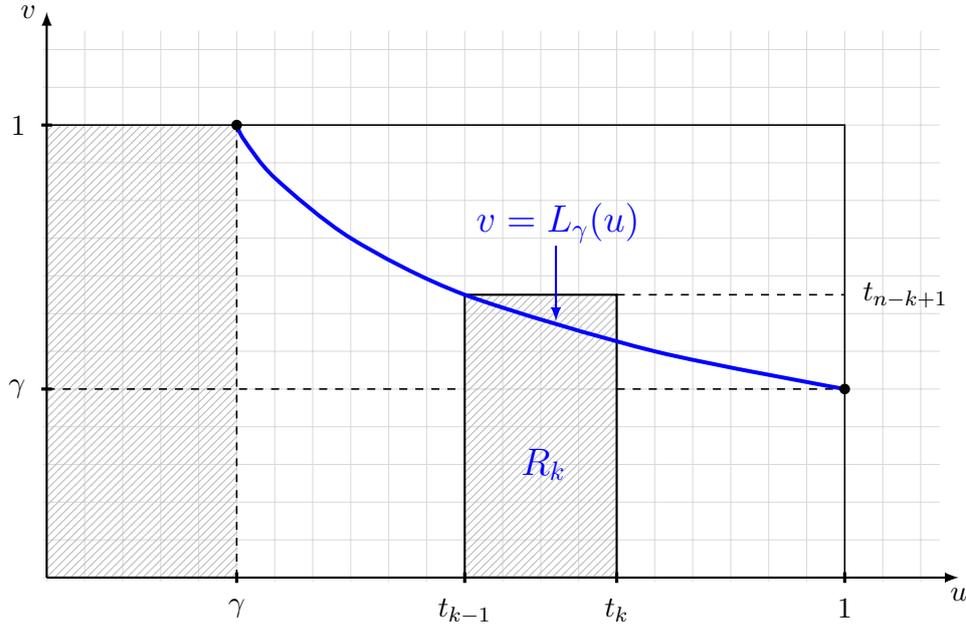
Unter den Voraussetzungen von Satz 3.66 bezeichne

$$K_{C_\psi}(\gamma) = V_{C_\psi}(\{(u, v) \in [0, 1]^2 : C_\psi(u, v) \leq \gamma\}), \quad \gamma \in (0, 1).$$

Dann gilt

$$K_{C_\psi}(\gamma) = \gamma - \frac{\psi^{-1}(\gamma)}{\left.\frac{d}{du}\psi^{-1}(u)\right|_{u=\gamma}}.$$

Beweis: Wir betrachten die Menge $\{(u, v) \in [0, 1]^2 : C_\psi(u, v) \leq \gamma\}$ grafisch:



Wir setzen $w := \psi^{-1}(\gamma)$ und betrachten für $n \in \mathbb{N}$ die Unterteilung $\{t_0 = \gamma, t_1, \dots, t_n = 1\}$ von $[\gamma, 1]$, wobei

$$t_{n-k} := \psi\left(\frac{k w}{n}\right), \quad 0 \leq k \leq n \Leftrightarrow t_\ell = \psi\left(\frac{(n-\ell)w}{n}\right), \quad 0 \leq \ell \leq n$$

und damit $\psi^{-1}(t_\ell) = (n-\ell)w/n, 0 \leq \ell \leq n$. Wir folgern, dass

$$\begin{aligned} & \psi^{-1}(t_{k-1}) + \psi^{-1}(t_{n-k+1}) \\ &= \frac{(n-k+1)w}{n} + \frac{(k-1)w}{n} = \frac{n w}{n} = w = \psi^{-1}(\gamma) \\ &\Rightarrow t_{n-k+1} = L_\gamma(t_{k-1}), \quad 1 \leq k \leq n, \end{aligned}$$

denn $\psi^{-1}(t_{n-k+1}) = \psi^{-1}(\gamma) - \psi^{-1}(t_{k-1})$.

Definiere Rechtecke $R_k := [t_{k-1}, t_k] \times [0, t_{n-k+1}]$ für $1 \leq k \leq n$. Dann ist

$$\begin{aligned} V_{C_\psi}(R_k) &= C_\psi(t_k, t_{n-k+1}) - C_\psi(t_{k-1}, t_{n-k+1}) \\ &= C_\psi(t_k, t_{n-k+1}) - \gamma \\ &= \psi(\psi^{-1}(t_k) + \psi^{-1}(t_{n-k+1})) - \gamma \\ &= \psi\left(\frac{(n-k)w}{n} + \frac{(k-1)w}{n}\right) - \gamma \\ &= \psi\left(\frac{(n-1)w}{n}\right) - \gamma = \psi\left(w - \frac{w}{n}\right) - \psi(w) \end{aligned}$$

unabhängig von $1 \leq k \leq n$.

$$\begin{aligned}
\Rightarrow K_{C_\psi}(\gamma) &= \gamma + \lim_{n \rightarrow \infty} V_{C_\psi} \left(\bigcup_{k=1}^n R_k \right) \\
&= \gamma + \lim_{n \rightarrow \infty} \left\{ n \left[\psi \left(w - \frac{w}{n} \right) - \psi(w) \right] \right\} \\
&= \gamma + \lim_{n \rightarrow \infty} \left\{ -w \frac{\psi(w) - \psi \left(w - \frac{w}{n} \right)}{w/n} \right\} \\
&= \gamma - w \psi'(w) \\
&= \gamma - \psi^{-1}(\gamma) \psi'(\psi^{-1}(\gamma)) \\
&= \gamma - \frac{\psi^{-1}(\gamma)}{\left. \frac{d}{du} \psi^{-1}(u) \right|_{u=\gamma}}
\end{aligned}$$

wie gewünscht. ■

Korollar 3.68

Sei C_ψ eine bivariate Archimedische Copula mit Generatorfunktion ψ . Falls $(U, V)^\top \sim C_\psi$, so ist die Verteilungsfunktion der Zufallsvariable $C_\psi(U, V)$ gerade K_{C_ψ} aus Satz 3.67.

Lemma 3.69

Eine absolut stetige, bivariate Archimedische Copula C_ψ besitzt die Copula-Dichte

$$(u, v)^\top \mapsto c_\psi(u, v) = - \frac{(\psi^{-1})''(C_\psi(u, v)) (\psi^{-1})'(u) (\psi^{-1})'(v)}{[(\psi^{-1})'(C_\psi(u, v))]^3}.$$

Beweis: Wir müssen C_ψ zweimal partiell differenzieren.

$$\begin{aligned}
C_\psi(u, v) &= \psi(\psi^{-1}(u) + \psi^{-1}(v)) \\
\Rightarrow \frac{\partial}{\partial u} C_\psi(u, v) &= \psi'(\psi^{-1}(u) + \psi^{-1}(v)) \left[(\psi^{-1})'(u) \right] \\
&= \psi'(\psi^{-1}\{\psi(\psi^{-1}(u) + \psi^{-1}(v))\}) \left[(\psi^{-1})'(u) \right] \\
&= \psi'(\psi^{-1}\{C_\psi(u, v)\}) \left[(\psi^{-1})'(u) \right] \\
&= \frac{(\psi^{-1})'(u)}{(\psi^{-1})'(C_\psi(u, v))}.
\end{aligned}$$

Aus Symmetriegründen ist ebenso

$$\frac{\partial}{\partial v} C_\psi(u, v) = \frac{(\psi^{-1})'(v)}{(\psi^{-1})'(C_\psi(u, v))}.$$

Anwendung der Quotientenregel der Differentiation liefert schließlich

$$\begin{aligned}
c_\psi(u, v) &= \frac{\partial^2 C_\psi(u, v)}{\partial u \partial v} \\
&= \frac{-(\psi^{-1})'(u) \frac{\partial}{\partial v} (\psi^{-1})'(C_\psi(u, v))}{[(\psi^{-1})'(C_\psi(u, v))]^2} \\
&= -\frac{(\psi^{-1})'(u) (\psi^{-1})''(C_\psi(u, v)) \frac{\partial}{\partial v} C_\psi(u, v)}{[(\psi^{-1})'(C_\psi(u, v))]^2} \\
&= -\frac{(\psi^{-1})''(C_\psi(u, v)) (\psi^{-1})'(u) (\psi^{-1})'(v)}{[(\psi^{-1})'(C_\psi(u, v))]^3}.
\end{aligned}$$

■

Satz 3.70

Unter der Voraussetzungen von Lemma 3.69 sei $(U, V)^\top \sim C_\psi$. Definiere $S := \frac{\psi^{-1}(U)}{\psi^{-1}(U) + \psi^{-1}(V)}$ und $T := C_\psi(U, V)$. Dann gilt

$$F_{(S,T)}(s, t) = sK_{C_\psi}(t) \quad \forall (s, t)^\top \in [0, 1]^2.$$

Also sind S und T stochastisch unabhängig und S ist auf $[0, 1]$ gleichverteilt.

Beweis: Wir betrachten die Transformation

$$\begin{aligned}
\psi^{-1}(u) &= s\psi^{-1}(t) \Leftrightarrow u = \psi(s\psi^{-1}(t)), \\
\psi^{-1}(v) &= (1-s)\psi^{-1}(t) \Leftrightarrow v = \psi((1-s)\psi^{-1}(t)).
\end{aligned}$$

Deren Jacobi-Matrix ist gegeben durch

$$\begin{aligned}
\frac{du}{ds} &= \frac{d\psi^{-1}(u)}{ds} \cdot \frac{du}{d\psi^{-1}(u)} = \frac{\psi^{-1}(t)}{(\psi^{-1})'(u)}, \\
\frac{du}{dt} &= \frac{d\psi^{-1}(u)}{dt} \cdot \frac{du}{d\psi^{-1}(u)} = \frac{s(\psi^{-1})'(t)}{(\psi^{-1})'(u)}, \\
\frac{dv}{ds} &= -\frac{\psi^{-1}(t)}{(\psi^{-1})'(v)}, \quad \frac{dv}{dt} = \frac{(1-s)(\psi^{-1})'(t)}{(\psi^{-1})'(v)},
\end{aligned}$$

wobei wir die Rechenregel $(\psi^{-1})'(v) = [\psi'(\psi^{-1}(v))]^{-1}$ benutzt haben.

Die zugehörige Jacobi-Determinante ist gegeben durch

$$\begin{aligned}
&\frac{\psi^{-1}(t)[(1-s)(\psi^{-1})'(t)] + s(\psi^{-1})'(t)\psi^{-1}(t)}{(\psi^{-1})'(u)(\psi^{-1})'(v)} \\
&= \frac{\psi^{-1}(t)(\psi^{-1})'(t)}{(\psi^{-1})'(u)(\psi^{-1})'(v)} =: D(t, u, v).
\end{aligned}$$

Nach Transformationsformel ist also die Lebesgue-dichte $f_{(S,T)}$ von $(S, T)^\top$ gegeben durch

$$\begin{aligned} f_{(S,T)}(s, t) &= -\frac{(\psi^{-1})''(t)(\psi^{-1})'(u)(\psi^{-1})'(v)}{[(\psi^{-1})'(t)]^3} \times \{-D(t, u, v)\} \\ &= \frac{(\psi^{-1})''(t)\psi^{-1}(t)}{[(\psi^{-1})'(t)]^2} \end{aligned}$$

und damit ergibt sich

$$F_{(S,T)}(s, t) = \int_0^s \int_0^t \frac{(\psi^{-1})''(y)\psi^{-1}(y)}{[(\psi^{-1})'(y)]^2} dy dx = s \left[y - \frac{\psi^{-1}(y)}{(\psi^{-1})'(y)} \right]_0^t = sK_{C_\psi}(t). \quad (3.17)$$

Die Stammfunktion in (3.17) lässt sich durch Probe (mit Quotientenregel) verifizieren. ■

Korollar 3.71

Ein Zufallsvektor $(U, V)^\top \sim C_\psi$ kann wie folgt generiert werden:

- 1) Generiere $S \sim \text{UNI}[0, 1]$ und $Q \sim \text{UNI}[0, 1]$ stochastisch unabhängig.
- 2) Setze $T := K_{C_\psi}^{-1}(Q)$
- 3) Setze $U := \psi(S \cdot \psi^{-1}(T))$ sowie $V := \psi((1 - S) \cdot \psi^{-1}(T))$.

Anmerkung: Für höhere Dimensionen $d > 2$ existieren analoge Darstellungen und Algorithmen.

3.6.2 Kendall's τ für Archimedische Copulae

Satz 3.72

Sei $(X, Y)^\top$ ein bivariater Zufallsvektor mit Archimedischer Copula C_ψ . Dann gilt

$$\tau(X, Y) = 1 + 4 \int_0^1 \frac{\psi^{-1}(t)}{(\psi^{-1})'(t)} dt.$$

Beweis: Seien U und V die Verteilungstransformierten von X und Y . Dann ist $(U, V)^\top \sim C_\psi$.

Ferner sei K_{C_ψ} wie in Satz 3.67 und Korollar 3.68. Wir rechnen mit Hilfe von Korollar 3.33:

$$\begin{aligned}\tau(X, Y) &= 4\mathbb{E}[C_\psi(U, V)] - 1 = 4 \int_0^1 t dK_{C_\psi}(t) - 1 \\ &= 4 \left\{ [tK_{C_\psi}(t)]_0^1 - \int_0^1 K_{C_\psi}(t) dt \right\} - 1\end{aligned}\quad (3.18)$$

$$= 3 - 4 \int_0^1 K_{C_\psi}(t) dt = 3 - 4 \int_0^1 \left(t - \frac{\psi^{-1}(t)}{(\psi^{-1})'(t)} \right) dt \quad (3.19)$$

$$= 4 \int_0^1 \frac{\psi^{-1}(t)}{(\psi^{-1})'(t)} dt + 3 - 4 \int_0^1 t dt$$

$$= 4 \int_0^1 \frac{\psi^{-1}(t)}{(\psi^{-1})'(t)} dt + 3 - 2,$$

wobei wir (3.18) partiell integriert und in (3.19) beachtet haben, dass $[tK_{C_\psi}(t)]_0^1 = 1$ ist. ■

Beispiele 3.73

(a) Die in Beispiel 3.17.(c) vorgestellte Familie der Gumbel-Hougaard-Copulae ist gegeben durch

$$\psi^{-1}(t) = (-\log(t))^\eta, \quad \eta \geq 1.$$

Damit ist

$$(\psi^{-1})'(t) = \eta(-\log(t))^{\eta-1} \cdot \left(-\frac{1}{t}\right)$$

und

$$\frac{\psi^{-1}(t)}{(\psi^{-1})'(t)} = \frac{t \log(t)}{\eta}.$$

Es folgt, dass

$$\begin{aligned}\tau_\eta &= 1 + 4 \int_0^1 \frac{t \log(t)}{\eta} dt \\ &= 1 + \frac{4}{\eta} \left\{ \left[\frac{t^2}{2} \log(t) \right]_0^1 - \int_0^1 \frac{t}{2} dt \right\} \\ &= 1 + \frac{4}{\eta} \left\{ 0 - \frac{1}{4} \right\} \\ &= 1 - \frac{1}{\eta} = \frac{\eta - 1}{\eta}.\end{aligned}$$

(b) Die in Beispiel 3.17.(c) vorgestellte Familie der Clayton-Copulae ist gegeben durch $\psi^{-1}(t) = \eta^{-1}(t^{-\eta} - 1)$ für $\eta > 0$. Damit ist

$$(\psi^{-1})'(t) = -t^{-\eta-1}$$

und

$$\frac{\psi^{-1}(t)}{(\psi^{-1})'(t)} = \frac{t^{\eta+1} - t}{\eta}.$$

Es folgt, dass

$$\begin{aligned} \tau_\eta &= 1 + \frac{4}{\eta} \int_0^1 (t^{\eta+1} - t) dt \\ &= 1 + \frac{4}{\eta} \left\{ \left[\frac{1}{\eta+2} t^{\eta+2} \right]_0^1 - \frac{1}{2} \right\} \\ &= 1 + \frac{4}{\eta} \left[\frac{1}{\eta+2} - \frac{1}{2} \right] \\ &= \frac{\eta(\eta+2) + 4 - 2(\eta+2)}{\eta(\eta+2)} = \frac{\eta}{\eta+2}. \end{aligned}$$

3.6.3 Tailabhängigkeit bivariater Archimedischer Copulae

Satz 3.74

Sei C_ψ eine bivariate Archimedische Copula mit einem Generator ψ , der die Laplace-Transformierte einer strikt positiven Zufallsvariable Z ist, vgl. Beispiel 3.17.(c).

Falls $\psi'(0)$ endlich ist, so ist $\lambda_U(C_\psi) = 0$. Falls $\psi'(0) = -\infty$ ist, so ist

$$\lambda_U(C_\psi) = 2 - 2 \lim_{s \downarrow 0} \frac{\psi'(2s)}{\psi'(s)}.$$

Beweis: Nach Lemma 3.38 ist

$$\begin{aligned} \lambda_U(C_\psi) &= \lim_{u \uparrow 1} \frac{1 - 2u + C_\psi(u, u)}{1 - u} \\ &= \lim_{u \uparrow 1} \frac{1 - 2u + \psi(2\psi^{-1}(u))}{1 - u} \\ &= \lim_{u \uparrow 1} \frac{-2 + \psi'(2\psi^{-1}(u))2(\psi^{-1})'(u)}{-1} \quad (\text{nach L'Hospital}) \\ &= 2 - 2 \lim_{u \uparrow 1} \psi'(2\psi^{-1}(u))(\psi^{-1})'(u) \\ &= 2 - 2 \lim_{u \uparrow 1} \frac{\psi'(2\psi^{-1}(u))}{\psi'(\psi^{-1}(u))} \\ &= 2 - 2 \lim_{s \downarrow 0} \frac{\psi'(2s)}{\psi'(s)}. \end{aligned} \tag{3.20}$$

Falls nun $\psi'(0) \in (-\infty, 0)$ ist, so ist der Grenzwert in (3.20) gleich Eins und somit $\lambda_U(C_\psi) = 2 - 2 = 0$. Dies impliziert die Aussagen. ■

Beispiel 3.75

Nach Beispiel 3.17.(c) besitzt die Gumbel-Hougaard-Copula C_η mit Parameter $\eta \geq 1$ den Generator ψ mit $\psi(x) = \exp(-x^{\frac{1}{\eta}})$.

$$\begin{aligned} \Rightarrow \psi'(x) &= -\exp(-x^{\frac{1}{\eta}}) \frac{1}{\eta} x^{\frac{1}{\eta}-1} \\ \Rightarrow \lambda_U(C_\eta) &= 2 - 2 \lim_{s \downarrow 0} \frac{\psi'(2s)}{\psi'(s)} \\ &= 2 - 2 \lim_{s \downarrow 0} \frac{\exp(-(2s)^{\frac{1}{\eta}}) (2s)^{\frac{1}{\eta}-1}}{\exp(-s^{\frac{1}{\eta}}) s^{\frac{1}{\eta}-1}} \\ &= 2 - 2(2^{\frac{1}{\eta}-1}) = 2 - 2^{\frac{1}{\eta}}, \end{aligned}$$

vgl. auch Beispiel 3.39.(a).

Satz 3.76

Unter den Voraussetzungen von Satz 3.74 ist $\lambda_L(C_\psi) = 2 \lim_{s \rightarrow \infty} \frac{\psi'(2s)}{\psi'(s)}$.

Beweis: Analog zum Beweis von Satz 3.74. ■

Beispiel 3.77

(a) Nach Beispiel 3.17.(c) besitzt die Clayton-Copula C_η mit Parameter $\eta > 0$ den Generator ψ mit

$$\begin{aligned} \psi(x) &= (1 + \eta x)^{-\frac{1}{\eta}} \\ \Rightarrow \psi'(x) &= -\frac{1}{\eta} (1 + \eta x)^{-\frac{1}{\eta}-1} \cdot \eta \\ &= -(1 + \eta x)^{-\frac{1}{\eta}-1}. \\ \Rightarrow \lambda_L(C_\eta) &= 2 \lim_{s \rightarrow \infty} \frac{(1 + 2\eta s)^{-\frac{1}{\eta}-1}}{(1 + \eta s)^{-\frac{1}{\eta}-1}} \\ &= 2 \cdot 2^{-\frac{1}{\eta}-1} = 2^{-\frac{1}{\eta}}. \end{aligned}$$

(b) Wir betrachten noch einmal die Familie der bivariaten Frank-Copulae $\{C_\eta\}_\eta$, siehe Beispiel

3.64.(a). Für einen Parameter $\eta \in \mathbb{R} \setminus \{0\}$ besitzt sie den Generator ψ , gegeben durch

$$\begin{aligned}\psi(x) &= -\frac{1}{\eta} \ln\{1 + \exp(-x)[\exp(-\eta) - 1]\}. \\ \Rightarrow \psi'(x) &= -\frac{1}{\eta} \frac{\exp(-x)[1 - \exp(-\eta)]}{1 + \exp(-x)[\exp(-\eta) - 1]} \\ \Rightarrow \psi'(0) &= -\frac{1}{\eta} \frac{1 - \exp(-\eta)}{\exp(-\eta)} \\ &= -\frac{\exp(\eta) - 1}{\eta} > -\infty.\end{aligned}$$

Da $\psi'(0)$ endlich ist, ist $\lambda_U(C_\eta) = 0 \forall \eta \in \mathbb{R} \setminus \{0\}$ gemäß Satz 3.74. Da C_η zudem radialsymmetrisch ist (vgl. Beispiel 3.64.(a)), ist auch $\lambda_L(C_\eta) = 0 \forall \eta \in \mathbb{R} \setminus \{0\}$ nach Lemma 3.41.(b).

3.7 Marshall-Olkin Copulae

Modell 3.78

Seien $Z_1 \sim \text{Exp}(\lambda_1)$, $Z_2 \sim \text{Exp}(\lambda_2)$ und $Z_{12} \sim \text{Exp}(\lambda_{12})$ drei stochastisch unabhängige (zufällige) Ereigniszeiten.

Setze $X := \min(Z_1, Z_{12})$ und $Y := \min(Z_2, Z_{12})$.

Interpretation: $(X, Y)^\top$ modelliert die Überlebenszeiten eines Zweikomponentensystems, wobei die i -te Komponente durch den Eintritt des zu Z_i gehörigen Ereignisses außer Betrieb gesetzt wird, und ferner beide Komponenten gleichzeitig ausfallen, wenn das zu Z_{12} gehörige Ereignis eintritt.

Die gemeinsame Survivalfunktion von $(X, Y)^\top$ ist demnach gegeben durch

$$\begin{aligned}\bar{F}_{(X,Y)}(x, y) &= \mathbb{P}(X > x, Y > y) \\ &= \mathbb{P}(Z_1 > x, Z_2 > y, Z_{12} > \max(x, y)) \\ &= \exp(-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x, y)), \quad x, y > 0.\end{aligned}$$

Die zugehörigen Rand-Survivalfunktionen sind gegeben durch

$$\begin{aligned}\bar{F}_X(x) &= \mathbb{P}(X > x) = \mathbb{P}(Z_1 > x, Z_{12} > x) = \exp(-(\lambda_1 + \lambda_{12})x), \\ \bar{F}_Y(y) &= \exp(-(\lambda_2 + \lambda_{12})y). \\ \Rightarrow X &\sim \text{Exp}(\lambda_1 + \lambda_{12}), \quad Y \sim \text{Exp}(\lambda_2 + \lambda_{12}).\end{aligned}$$

Lemma 3.79

Sei $\alpha_1 = \frac{\lambda_{12}}{\lambda_1 + \lambda_{12}}$ und $\alpha_2 = \frac{\lambda_{12}}{\lambda_2 + \lambda_{12}}$. Dann ist, unter den Voraussetzungen von Modell 3.78, die

Survival-Copula \check{C} von $(X, Y)^\top$ gegeben durch

$$\begin{aligned}\check{C}(u, v) &= uv \min(u^{-\alpha_1}, v^{-\alpha_2}) \\ &= \min(u^{1-\alpha_1}v, uv^{1-\alpha_2}), \quad (u, v)^\top \in [0, 1]^2.\end{aligned}$$

Beweis: Nach Definition 3.21 ist

$$\bar{F}_{(X,Y)}(x, y) = \check{C}(\bar{F}_X(x), \bar{F}_Y(y)).$$

Wir beachten dass $\max(x, y) = x + y - \min(x, y)$ ist.

$$\begin{aligned}\Rightarrow \bar{F}_{(X,Y)}(x, y) &= \exp(-\lambda_1 x - \lambda_2 y - \lambda_{12}(x + y) + \lambda_{12} \min(x, y)) \\ &= \exp(-(\lambda_1 + \lambda_{12})x - (\lambda_2 + \lambda_{12})y + \lambda_{12} \min(x, y)) \\ &= \bar{F}_X(x) \bar{F}_Y(y) \min\{\exp(\lambda_{12}x), \exp(\lambda_{12}y)\}.\end{aligned}$$

Wir substituieren $u := \bar{F}_X(x)$, $v := \bar{F}_Y(y)$ und beachten, dass $\exp(\lambda_{12}x) = u^{-\alpha_1}$ sowie $\exp(\lambda_{12}y) = v^{-\alpha_2}$ ist, denn

$$\begin{aligned}u^{-\alpha_1} &= [\bar{F}_X(x)]^{-\frac{\lambda_{12}}{\lambda_1 + \lambda_{12}}} \\ &= [\exp(-(\lambda_1 + \lambda_{12})x)]^{-\frac{\lambda_{12}}{\lambda_1 + \lambda_{12}}} \\ &= \exp(\lambda_{12}x), \quad \text{und analog für } v^{-\alpha_2}.\end{aligned}$$

$$\begin{aligned}\Rightarrow \bar{F}_{(X,Y)}(x, y) &= \bar{F}_X(x) \bar{F}_Y(y) \min\left\{[\bar{F}_X(x)]^{-\alpha_1}, [\bar{F}_Y(y)]^{-\alpha_2}\right\} \\ \Rightarrow \check{C}(u, v) &= uv \min(u^{-\alpha_1}, v^{-\alpha_2}) = \min(u^{1-\alpha_1}v, uv^{1-\alpha_2})\end{aligned}$$

wie gewünscht. ■

Definition 3.80

Für Parameter $\alpha_1 \in [0, 1]$ und $\alpha_2 \in [0, 1]$ ist die Familie der Marshall-Olkin Copulae gegeben durch

$$C_{\alpha_1, \alpha_2}(u, v) = \min\{u^{1-\alpha_1}v, uv^{1-\alpha_2}\} = \begin{cases} u^{1-\alpha_1}v, & u^{\alpha_1} \geq v^{\alpha_2} \\ uv^{1-\alpha_2}, & u^{\alpha_1} \leq v^{\alpha_2} \end{cases}.$$

Satz 3.81

Der eindimensionale Unterraum $\{(u, v)^\top \in [0, 1]^2 : u^{\alpha_1} = v^{\alpha_2}\}$ von $[0, 1]^2$ hat das C_{α_1, α_2} -Volumen

$$\mathbb{P}(U^{\alpha_1} = V^{\alpha_2}) = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2 - \alpha_1 \alpha_2} > 0,$$

wobei $(U, V)^\top \sim C_{\alpha_1, \alpha_2}$.

Beweis: Wir berechnen äquivalenterweise $\mathbb{P}(X = Y) = \mathbb{P}(Z_{12} < Z_1, Z_{12} < Z_2)$ mit X, Y, Z_1, Z_2, Z_{12} wie in Modell 3.78. Es ist

$$\begin{aligned} \mathbb{P}(Z_{12} < Z_1, Z_{12} < Z_2) &= \int_0^{\infty} \mathbb{P}(Z_1 > z, Z_2 > z) d\mathbb{P}^{Z_{12}}(z) \\ &= \int_0^{\infty} \exp(-\lambda_1 z) \exp(-\lambda_2 z) \lambda_{12} \exp(-\lambda_{12} z) dz \\ &= \lambda_{12} \int_0^{\infty} \exp(-(\lambda_1 + \lambda_2 + \lambda_{12})z) dz \\ &= \frac{\lambda_{12}}{\lambda_1 + \lambda_2 + \lambda_{12}} \end{aligned}$$

wegen der Normierungsbedingung für die $\text{Exp}(\lambda_1 + \lambda_2 + \lambda_{12})$ -Verteilung. Einfache Rechnung ergibt schließlich, dass

$$\frac{\lambda_{12}}{\lambda_1 + \lambda_2 + \lambda_{12}} = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2 - \alpha_1 \alpha_2} \text{ ist.}$$

■

Lemma 3.82

Für zwei Copulae C_1 und C_2 gilt

$$\int \int_{[0,1]^2} C_1(u, v) dC_2(u, v) = \frac{1}{2} - \int \int_{[0,1]^2} \frac{\partial}{\partial u} C_1(u, v) \frac{\partial}{\partial v} C_2(u, v) du dv.$$

Beweis: Angenommen, C_1 und C_2 sind absolut stetig. Dann ist

$$\int \int_{[0,1]^2} C_1(u, v) dC_2(u, v) = \int_0^1 \int_0^1 C_1(u, v) \frac{\partial^2 C_2(u, v)}{\partial u \partial v} du dv.$$

Wir wenden den Satz von Fubini an und berechnen zunächst das innere Integral.

$$\begin{aligned} \int_0^1 C_1(u, v) \frac{\partial^2 C_2(u, v)}{\partial u \partial v} du &= \left[C_1(u, v) \frac{\partial C_2(u, v)}{\partial v} \right]_{u=0}^{u=1} - \int_0^1 \frac{\partial C_1(u, v)}{\partial u} \frac{\partial C_2(u, v)}{\partial v} du \\ &= v - \int_0^1 \frac{\partial C_1(u, v)}{\partial u} \frac{\partial C_2(u, v)}{\partial v} du. \end{aligned}$$

Anschließende äußere Integration bezüglich $v \in [0, 1]$ liefert das Ergebnis.

Der allgemeine Fall folgt durch Approximation von C_1 und C_2 durch Folgen absolut stetiger Copulae, siehe Li et al. (2002). ■

Satz 3.83

Für eine Marshall-Olkin Copula C_{α_1, α_2} sind Spearman's ρ und Kendall's τ gegeben durch

$$\rho_S(C_{\alpha_1, \alpha_2}) = \frac{3\alpha_1\alpha_2}{2\alpha_1 + 2\alpha_2 - \alpha_1\alpha_2},$$

$$\tau(C_{\alpha_1, \alpha_2}) = \frac{\alpha_1\alpha_2}{\alpha_1 + \alpha_2 - \alpha_1\alpha_2}.$$

Beweis:

$$\begin{aligned} \rho_S(C_{\alpha_1, \alpha_2}) &= 12 \int \int_{[0,1]^2} C_{\alpha_1, \alpha_2}(u, v) dudv - 3 \\ &= 12 \int_0^1 \left[\int_0^{u^{\frac{\alpha_1}{\alpha_2}}} u^{1-\alpha_1} v dv + \int_{u^{\frac{\alpha_1}{\alpha_2}}}^1 uv^{1-\alpha_2} dv \right] du - 3 \\ &= \frac{3\alpha_1\alpha_2}{2\alpha_1 + 2\alpha_2 - \alpha_1\alpha_2} \quad (\text{siehe Präsentation mit Maple}), \\ \tau(C_{\alpha_1, \alpha_2}) &= 4 \int \int_{[0,1]^2} C_{\alpha_1, \alpha_2}(u, v) dC_{\alpha_1, \alpha_2}(u, v) - 1. \end{aligned}$$

Wir wenden Lemma 3.82 an und erhalten

$$\tau(C_{\alpha_1, \alpha_2}) = 4 \left\{ \frac{1}{2} - \int \int_{[0,1]^2} \frac{\partial}{\partial u} C_{\alpha_1, \alpha_2}(u, v) \frac{\partial}{\partial v} C_{\alpha_1, \alpha_2}(u, v) dudv \right\} - 1.$$

Wir machen die Fallunterscheidung

$$\frac{\partial C_{\alpha_1, \alpha_2}(u, v)}{\partial u} = \begin{cases} (1 - \alpha_1)u^{-\alpha_1}v, & v \leq u^{\frac{\alpha_1}{\alpha_2}} \\ v^{1-\alpha_2}, & v > u^{\frac{\alpha_1}{\alpha_2}} \end{cases}$$

$$\frac{\partial C_{\alpha_1, \alpha_2}(u, v)}{\partial v} = \begin{cases} u^{1-\alpha_1}, & v \leq u^{\frac{\alpha_1}{\alpha_2}} \\ u(1 - \alpha_2)v^{-\alpha_2}, & v > u^{\frac{\alpha_1}{\alpha_2}} \end{cases}$$

Also erhalten wir

$$\frac{\partial}{\partial u} C_{\alpha_1, \alpha_2}(u, v) \frac{\partial}{\partial v} C_{\alpha_1, \alpha_2}(u, v) = \begin{cases} (1 - \alpha_1)u^{1-2\alpha_1}v =: g_1(u, v), & v \leq u^{\frac{\alpha_1}{\alpha_2}} \\ u(1 - \alpha_2)v^{1-2\alpha_2} =: g_2(u, v), & v > u^{\frac{\alpha_1}{\alpha_2}} \end{cases}$$

und damit ist

$$\begin{aligned} \tau(C_{\alpha_1, \alpha_2}) &= 4 \left\{ \frac{1}{2} - \int_0^1 \int_0^{u^{\frac{\alpha_1}{\alpha_2}}} g_1(u, v) dv du - \int_0^1 \int_{u^{\frac{\alpha_1}{\alpha_2}}}^1 g_2(u, v) dv du \right\} - 1 \\ &= \frac{\alpha_1\alpha_2}{\alpha_1 + \alpha_2 - \alpha_1\alpha_2} \quad (\text{siehe Präsentation mit Maple}). \end{aligned}$$

■

Satz 3.84

Sei C_{α_1, α_2} eine bivariate Marshall-Olkin Copula. Dann ist ihr oberer Tailabhängigkeitskoeffizient gegeben durch

$$\lambda_U(C_{\alpha_1, \alpha_2}) = \min(\alpha_1, \alpha_2).$$

Beweis: Sei (aus Symmetriegründen o.B.d.A.) $\alpha_1 > \alpha_2$. Wir rechnen

$$\begin{aligned} \lambda_U(C_{\alpha_1, \alpha_2}) &= \lim_{u \uparrow 1} \frac{\bar{C}_{\alpha_1, \alpha_2}(u, u)}{1 - u} \\ &= \lim_{u \uparrow 1} \frac{1 - 2u + u^2 \min(u^{-\alpha_1}, u^{-\alpha_2})}{1 - u} \\ &= \lim_{u \uparrow 1} \frac{1 - 2u + u^2 u^{-\alpha_2}}{1 - u} \\ &= \lim_{u \uparrow 1} \left[\frac{-2 + (2 - \alpha_2)u^{1-\alpha_2}}{-1} \right] \quad (\text{L'Hospital}) \\ &= \lim_{u \uparrow 1} [2 - 2u^{1-\alpha_2} + \alpha_2 u^{1-\alpha_2}] = \alpha_2. \end{aligned}$$

■

Bemerkung 3.85 (Höher-dimensionale Marshall-Olkin Copulae)

Betrachten wir im Allgemeinen ein d -Komponenten-System, $d \geq 2$, so kann eine zu Modell 3.78 analoge Modellbildung erfolgen, wobei insgesamt $2^d - 1$ zufällige Überlebenszeiten $(Z_S)_{S \in \mathcal{S}}$ betrachtet werden, wobei \mathcal{S} die Menge aller nicht-leeren Teilmengen S von $\{1, \dots, d\}$ bezeichnet, also $\mathcal{S} = 2^{\{1, \dots, d\}} \setminus \{\emptyset\}$.

Dabei sind alle $(Z_S)_{S \in \mathcal{S}}$ stochastisch unabhängig mit $Z_S \sim \text{Exp}(\lambda_S)$, $S \in \mathcal{S}$, und

$$X_j = \min_{S: j \in S} Z_S, \quad 1 \leq j \leq d.$$

Beispielweise ergibt sich für $d = 4$:

$$X_1 = \min(Z_1, Z_{12}, Z_{13}, Z_{14}, Z_{123}, Z_{124}, Z_{134}, Z_{1234}),$$

$$X_2 = \min(Z_2, Z_{12}, Z_{23}, Z_{24}, Z_{123}, Z_{124}, Z_{234}, Z_{1234}),$$

$$X_3 = \min(Z_3, Z_{13}, Z_{23}, Z_{34}, Z_{123}, Z_{134}, Z_{234}, Z_{1234}),$$

$$X_4 = \min(Z_4, Z_{14}, Z_{24}, Z_{34}, Z_{124}, Z_{134}, Z_{234}, Z_{1234}).$$

Es ergibt sich das folgende Schema mit $2^4 - 1 = 15$.

$$\begin{array}{cccccc} & & & & & & Z_{1234} \\ & & & & & & Z_{123} & Z_{124} & Z_{134} & Z_{234} \\ & & & & & & Z_{12} & Z_{13} & Z_{14} & Z_{23} & Z_{24} & Z_{34} \\ & & & & & & Z_1 & Z_2 & Z_3 & Z_4 \end{array}$$

Möchte man gewisse Einflüsse als nicht existent modellieren, so setzt man $\lambda_S = 0 \Rightarrow Z_S = +\infty$ fast sicher.

Algorithmus 3.86 (Generierung von Pseudo-Zufallsvektoren)

Pseudo-Zufallsvektoren, die sich wie unabhängige Realisierungen von d -dimensionalen Marshall-Olkin Copulae verhalten, lassen sich wie folgt am Computer generieren.

0. Ordne die $\ell := 2^d - 1$ nicht-leeren Teilmengen von $\{1, \dots, d\}$ in einer beliebigen Art und Weise an und erhalte S_1, \dots, S_ℓ mit zugehörigen Intensitätsparametern $\lambda_1, \dots, \lambda_\ell$.
1. Generiere ℓ i.i.d. Zufallszahlen v_1, \dots, v_ℓ mit $v_1 \sim \text{UNI}[0, 1]$.
2. Setze $x_i = \min_{\substack{1 \leq k \leq \ell: \\ i \in S_k, \\ \lambda_k \neq 0}} (-\ln(v_k)/\lambda_k)$, $1 \leq i \leq d$.
3. Setze $\Lambda_i = \sum_{k=1}^{\ell} \mathbf{1}_{\{i \in S_k\}} \lambda_k$, $1 \leq i \leq d$.
4. Setze $u_i = \exp(-\Lambda_i x_i)$, $1 \leq i \leq d$.

Dann folgt $\mathbf{x} = (x_1, \dots, x_d)^\top$ einer d -variaten Marshall-Olkin-Überlebenszeitverteilung und $\mathbf{u} = (u_1, \dots, u_d)^\top$ der zugehörigen Marshall-Olkin Copula.

Abbildungsverzeichnis

1.1	Skizze zur Veranschaulichung der bedingten Erwartung	6
1.2	Zusammenhang von Konvergenzarten	16

Literaturverzeichnis

- Ash, R. B. (1972). *Measure, integration, and functional analysis*. New York-London: Academic Press, Inc.
- Bauer, H. (1991). *Probability theory. (Wahrscheinlichkeitstheorie.) 4., völlig überarb. u. neugestaltete Aufl. des Werkes: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*. Berlin etc.: Walter de Gruyter.
- Bauer, H. (2002). *Wahrscheinlichkeitstheorie. 5., durchges. u. verbess. Aufl.* Berlin: de Gruyter.
- Bernstein, S. (1929). Sur les fonctions absolument monotones. *Acta Mathematica* 52(1), 1–66.
- Bickel, P. J. and K. A. Doksum (1977). *Mathematical statistics. Basic ideas and selected topics*. Holden-Day Series in Probability and Statistics. San Francisco etc.: Holden-Day, Inc.
- Boto von Querenburg (2001). *Mengentheoretische Topologie. 3., neu bearbeitete und erweiterte Aufl.* Berlin: Springer.
- Box, G. and M. E. Muller (1958). A note on the generation of random normal deviates. *Ann. Math. Stat.* 29, 610–611.
- Breiman, L. (1992). *Probability*. Philadelphia, PA: SIAM.
- Chung, K. L. (2000). *A course in probability theory. 3rd ed.* Orlando, FL: Academic Press.
- Cramér, H. (1946). *Mathematical methods of statistics*. Princeton Mathematical series. Princeton N. J.: Princeton University Press.
- Embrechts, P., F. Lindskog, and A. McNeil (2003). Modelling dependence with copulas and applications to risk management. In S. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*, pp. 329–384. Elsevier Science B.V.
- Etemadi, N. (1981). An elementary proof of the strong law of large numbers. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* 55, 119–122.
- Fang, K.-T., S. Kotz, and K.-W. Ng (1990). *Symmetric multivariate and related distributions*. London etc.: Chapman and Hall.

- Feller, W. (1971). *An introduction to probability theory and its applications. Vol II. 2nd ed.* Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley and Sons, Inc.
- Gaenssler, P. and W. Stute (1977). *Wahrscheinlichkeitstheorie.* Hochschultext. Berlin-Heidelberg-New York: Springer-Verlag.
- Gil-Pelaez, J. (1951). Note on the inversion theorem. *Biometrika* 38, 481–482.
- Hall, P. and C. Heyde (1980). *Martingale limit theory and its application.* Probability and Mathematical Statistics. New York etc.: Academic Press, A Subsidiary of Harcourt Brace Jovanovich, Publishers.
- Jacod, J. and P. Protter (2000). *Probability essentials.* Berlin: Springer.
- Klenke, A. (2008). *Probability theory. (Wahrscheinlichkeitstheorie.) 2nd revised ed.* Berlin: Springer.
- Lehmann, E. L. (1999). *Elements of large-sample theory.* New York, NY: Springer.
- Li, X., P. Mikusiński, and M. Taylor (2002). Some integration-by-parts formulas involving 2-copulas. In *Distributions with given marginals and statistical modelling. Papers presented at the meeting, Barcelona, Spain, July 17–20, 2000.*, pp. 153–159. Dordrecht: Kluwer Academic Publishers.
- Luschgy, H. (2013). *Martingale in diskreter Zeit. Theorie und Anwendungen.* Berlin: Springer Spektrum.
- Nelsen, R. B. (2006). *An introduction to copulas. 2nd ed.* Springer Series in Statistics. New York, NY: Springer.
- Rüschendorf, L. (2009). On the distributional transform, Sklar’s theorem, and the empirical copula process. *J. Stat. Plann. Inference* 139(11), 3921–3927.
- Schoenberg, I. (1938). Metric spaces and completely monotone functions. *Ann. Math. (2)* 39, 811–841.
- Serfling, R. J. (1980). *Approximation theorems of mathematical statistics.* Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley & Sons.
- Shorack, G. R. and J. A. Wellner (1986). *Empirical processes with applications to statistics.* Wiley Series in Probability and Mathematical Statistics. New York, NY: Wiley.