

Rohversion Skriptum zur UV
Abhängigkeitsmodellierung

LV-Nr. 405.552

Assoz.-Prof. Dr. Wolfgang Trutschnig

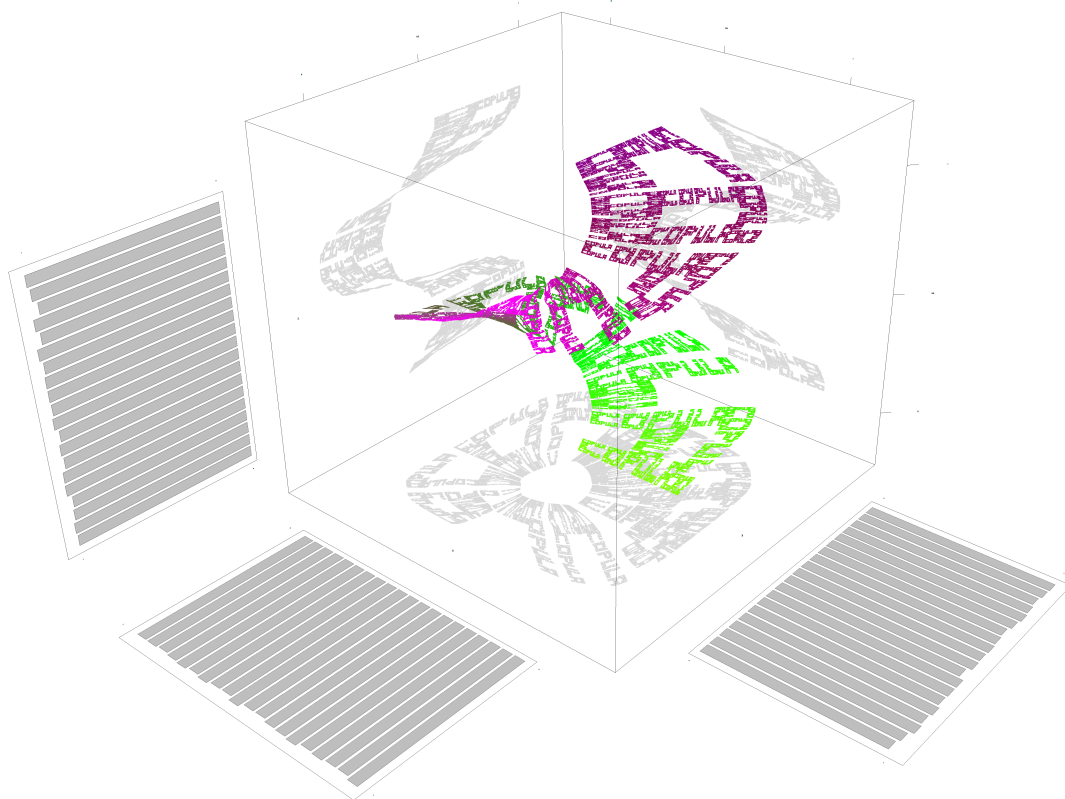
www.trutschnig.net

Fachbereich Mathematik

Paris Lodron Universität Salzburg

Hellbrunner Strasse 34

A-5020 Salzburg



Inhaltsverzeichnis

1	Ergänzungen zur Maß/Wahrscheinlichkeitstheorie	3
1.1	Lebesgue-Zerlegung und der Satz von Radon-Nikodym	3
1.2	Bedingte Erwartungen	6
1.3	Disintegration und Markov Kerne	11
2	Analytische Eigenschaften zweidimensionaler Copulas	18
2.1	Definition und alternative Darstellungen	18
2.2	Der Satz von Sklar	22
2.3	Metriken auf \mathcal{C}	25
3	Konkordanzmaße	29
4	Wichtige Copula Klassen	35
4.1	Empirische Copulas	35
4.2	Checkerboard Copulas	37
4.3	Extreme-Value Copulas	39
5	Übungsblätter	45
	Literatur	43

Kapitel 1

Ergänzungen zur Maß/Wahrscheinlichkeitstheorie

Um die größtenteils maßtheoretisch formulierten Resultate in [5] sauber verstehen und beweisen zu können, bedarf es insbesondere zweier Haupttheoreme der Maßtheorie/Wahrscheinlichkeitstheorie, die in den Grundvorlesungen leider oft zu kurz kommen.

1.1 Lebesgue-Zerlegung und der Satz von Radon-Nikodym

Eine sehr spezielle Version der Lebesgue-Zerlegung und des Satzes von Radon-Nikodym kennen Sie aus der Statistik Bachelor Vorlesung: Jede eindimensionale Verteilungsfunktion lässt sich als Konvexkombination einer absolut stetigen, einer diskreten und einer singulären Verteilungsfunktion darstellen. Im Folgenden betrachten wir diese Zerlegungseigenschaft auf einem allgemeinen Messraum (Ω, \mathcal{A}) ohne jegliche topologische Struktur. Mit $\mathcal{M}(\Omega, \mathcal{A})$ bezeichnen wir (im gesamten Skriptum) die Menge aller endlichen Maße auf \mathcal{A} . Die Resultate in diesem Kapitel gelten großteils auch für σ -endliche positive und für komplexe Maße - nachdem der Fokus der LV auf Copulas, und damit auf speziellen Wahrscheinlichkeitsmaßen liegt, reicht es, mit $\mathcal{M}(\Omega, \mathcal{A})$ zu arbeiten.

Definition 1.1 Gegeben seien zwei endliche Maße $\nu, \mu \in \mathcal{M}(\Omega, \mathcal{A})$. Dann heißt ν absolut stetig bezüglich μ (wir schreiben $\nu \ll \mu$) genau dann, wenn für jedes $E \in \mathcal{A}$ aus $\mu(E) = 0$ auch $\nu(E) = 0$ folgt.

Definition 1.2 Sei $\mu \in \mathcal{M}(\Omega, \mathcal{A})$. Falls eine Menge $E \in \mathcal{A}$ existiert, sodass für jedes $F \in \mathcal{A}$ die Gleichheit

$$\mu(F) = \mu(E \cap F)$$

gilt, dann sagen wir, dass μ auf E konzentriert ist (oft auch: μ 'lebt' auf E).

Bemerkung 1.3 Gelte $\mu(E) = 0$, dann ist μ offensichtlich auf E^c konzentriert. Zusätzlich gilt: μ ist auf E konzentriert genau dann wenn $\mu(E) = \mu(\Omega) =: \|\mu\|$.

Definition 1.4 Zwei Maße $\nu, \mu \in \mathcal{M}(\Omega, \mathcal{A})$ heißen singulär zueinander (wir schreiben $\nu \perp \mu$), falls zwei disjunkte Mengen $E_1, E_2 \in \mathcal{A}$ existieren, sodass ν auf E_1 und μ auf E_2 konzentriert ist.

Bemerkung 1.5 Offensichtlich gilt $\nu \perp \mu$ genau dann wenn eine Menge $E \in \mathcal{A}$ existiert, für die $\mu(E) = \|\mu\|$ (also $\mu(E^c) = 0$) und $\nu(E) = 0$ gilt.

Das folgende Lemma fasst einige grundlegende Eigenschaften der gerade eingeführten Begriffe zusammen:

Lemma 1.6 Für $\nu, \nu_1, \nu_2, \mu \in \mathcal{M}(\Omega, \mathcal{A})$ gelten folgende Aussagen:

1. Aus $\nu_1 \ll \mu$ und $\nu_2 \ll \mu$ folgt $\nu_1 + \nu_2 \ll \mu$.
2. Aus $\nu_1 \perp \mu$ und $\nu_2 \perp \mu$ folgt $\nu_1 + \nu_2 \perp \mu$.
3. Aus $\nu_1 \ll \mu$ und $\nu_2 \perp \mu$ folgt $\nu_1 \perp \nu_2$.
4. Aus $\nu \ll \mu$ und $\nu \perp \mu$ folgt $\nu = 0$ (also $\nu(E) = 0$ für jedes $E \in \mathcal{A}$).

Beweis: Einfache Übungsaufgabe

Satz 1.7 (Lebesgue Zerlegung und Satz von Radon-Nikodym für endliche Maße)
Seien $\nu, \mu \in \mathcal{M}(\Omega, \mathcal{A}) =: \mathcal{M}$, dann gelten folgende Aussagen:

1. Es existiert ein eindeutiges Paar $(\nu_a, \nu_s) \in \mathcal{M}^2$ mit

$$\nu = \nu_a + \nu_s, \quad \nu_a \ll \mu, \quad \nu_s \perp \mu \quad (1.1)$$

2. Es existiert ein eindeutiges $f \in L^1(\Omega, \mathcal{A}, \mu)$ mit

$$\nu_a(E) = \int_E f d\mu \quad (1.2)$$

für jedes $E \in \mathcal{A}$. Dieses f heißt Radon-Nikodym Ableitung/Dichte von ν_a bezüglich μ .

Beweis: Wir definieren ein neues Maß $\varphi \in \mathcal{M}$ durch

$$\varphi(E) := \nu(E) + \mu(E)$$

für jedes $E \in \mathcal{A}$. Für jedes $h \in L^2(\Omega, \mathcal{A}, \varphi)$ gilt offensichtlich

$$\left| \int_{\Omega} h d\nu \right| \leq \int_{\Omega} |h| d\nu \leq \int_{\Omega} |h| d\varphi \leq \underbrace{\left(\int_{\Omega} h^2 d\varphi \right)^{1/2}}_{=:\|h\|_{L^2(\Omega, \mathcal{A}, \varphi)}} (\varphi(\Omega))^{1/2} < \infty. \quad (1.3)$$

Die Abbildung $T : L^2(\Omega, \mathcal{A}, \varphi) \rightarrow \mathbb{R}$, definiert durch $T(h) = \int_{\Omega} h d\nu$, ist wohldefiniert und linear. Für $h_1, h_2 \in L^2(\Omega, \mathcal{A}, \varphi)$ gilt wegen (1.3)

$$|T(h_1) - T(h_2)| \leq \|h_1 - h_2\|_{L^2(\Omega, \mathcal{A}, \varphi)} (\varphi(\Omega))^{1/2},$$

T ist also (Lipschitz-)stetig, und damit insgesamt ein stetiges lineares Funktional auf dem Hilbertraum $L^2(\Omega, \mathcal{A}, \varphi)$. Anwendung des Darstellungssatzes von Riesz liefert die Existenz eines eindeutigen $g \in L^2(\Omega, \mathcal{A}, \varphi)$ mit

$$\int_{\Omega} h d\nu = T(h) = \langle h, g \rangle_{L^2(\Omega, \mathcal{A}, \varphi)} = \int_{\Omega} h g d\varphi \quad (1.4)$$

für jedes $h \in L^2(\Omega, \mathcal{A}, \varphi)$. Setzen wir $h = \mathbf{1}_E$ für ein $E \in \mathcal{A}$ mit $\varphi(E) > 0$, dann folgt via (1.4) sofort

$$\nu(E) = \int_{\Omega} \mathbf{1}_E d\nu = \int_{\Omega} \mathbf{1}_E g d\varphi = \int_E g d\varphi,$$

und damit

$$0 \leq \frac{1}{\varphi(E)} \int_E g d\varphi = \frac{\nu(E)}{\varphi(E)} \leq 1.$$

Nachdem $E \in \mathcal{A}$ mit $\varphi(E) > 0$ beliebig war, erhalten wir $g(x) \in [0, 1]$ für φ -fast jedes $x \in \Omega$. O.B.d.A. können wir annehmen (falls notwendig, ändern wir g auf einer φ -Nullmenge), dass $g(x) \in [0, 1]$ sogar für jedes $x \in \Omega$ gilt.

Unter Verwendung der Definition von φ kann Gleichung (1.4) umgeformt werden zu

$$\int_{\Omega} h(1 - g) d\nu = \int_{\Omega} hg d\mu. \tag{1.5}$$

Wir setzen wir $A = g^{-1}([0, 1]) \in \mathcal{A}$ (also $A^c = g^{-1}(\{1\})$) und definieren $\nu_a, \nu_s \in \mathcal{M}$ durch

$$\nu_a(E) = \nu(A \cap E), \quad \nu_s(E) = \nu(A^c \cap E). \tag{1.6}$$

Das Maß ν_a ist konzentriert auf A , ν_s auf A^c , und offensichtlich gilt $\nu = \nu_a + \nu_s$. Für $h = \mathbf{1}_{A^c}$ liefert (1.5)

$$0 = \int_{A^c} 1 - g d\nu = \int_{A^c} g d\mu = \mu(A^c),$$

es gilt also $\nu_s \perp \mu$.

Um $\nu_a \ll \mu$ und die Darstellung (1.2) zu beweisen, gehen wir wie folgt vor: Für beliebiges $E \in \mathcal{A}$, $n \in \mathbb{N}$ und $h = (1 + g + g^2 + \dots + g^n) \mathbf{1}_E$ liefert (1.5)

$$L_n := \int_E (1 - g^{n+1}) d\nu = \int_E \underbrace{g(1 + g + g^2 + \dots + g^n)}_{=: I_n} d\mu =: R_n. \tag{1.7}$$

Für $x \in E \cap A$ gilt $1 - g^{n+1}(x) \rightarrow 1$ für $n \rightarrow \infty$, für $x \in E \cap A^c$ gilt offensichtlich $1 - g^{n+1}(x) = 0$, insgesamt erhalten wir also (Satz von der majorisierten Konvergenz oder Satz von der monotonen Konvergenz) $\lim_{n \rightarrow \infty} L_n = \nu(E \cap A) = \nu_a(E)$. Der Integrand I_n der rechten Seite ist monoton nicht fallend in n und konvergiert (in $[0, \infty]$) punktweise gegen eine nicht-negative messbare Funktion $f : \Omega \rightarrow [0, \infty]$ (beachten Sie, dass wir f sogar explizit berechnen können - wie?). Anwendung des Satzes von der monotonen Konvergenz liefert $\lim_{n \rightarrow \infty} R_n = \int_E f d\mu$, und wir erhalten insgesamt

$$\infty > \nu_a(E) = \int_E f d\mu,$$

also genau die gewünschte Darstellung für ν_a ($E \in \mathcal{A}$ war beliebig). Für den Fall $E = \Omega$ folgt wegen $\int_{\Omega} |f| d\mu = \int_{\Omega} f d\mu = \nu_a(\Omega) < \infty$ die Integrierbarkeit von f . Damit ist der Beweis bis auf die in der ersten Aussage behauptete Eindeutigkeit des Paares $(\nu_a, \nu_s) \in \mathcal{M}^2$ (Übungsaufgabe) komplett. ■

Bemerkung 1.8 Für $\nu \ll \mu$ schreiben wir in der Folge auch oft $f = \frac{d\nu}{d\mu}$ für die Radon-Nikodym Dichte/Ableitung von ν bezüglich μ , und verwenden zusätzlich die in der Literatur übliche Schreibweise $d\nu = f d\mu$.

Satz 1.9 (Rechenregeln für Radon-Nikodym Ableitungen/Dichten)

Seien $\mu_1, \mu_2, \mu_3 \in \mathcal{M}(\Omega, \mathcal{A})$ und es gelte $\mu_1 \ll \mu_2 \ll \mu_3$. Dann gilt die sogenannte Kettenregel für Radon-Nikodym Ableitungen, i.e.

$$\frac{d\mu_1}{d\mu_3} = \frac{d\mu_1}{d\mu_2} \cdot \frac{d\mu_2}{d\mu_3} \quad [\mu_3] \quad (1.8)$$

Falls $\mu_1 \ll \mu_2$ und $\mu_2 \ll \mu_1$ gilt, dann folgt

$$\frac{d\mu_2}{d\mu_1} = \left(\frac{d\mu_1}{d\mu_2}\right)^{-1} = \frac{1}{\frac{d\mu_1}{d\mu_2}} \quad [\mu_1]. \quad (1.9)$$

Beweis: Übungsaufgabe

Der folgende abschließende Satz begründet die Verwendung des Wortes 'Stetig' im Zusammenhang mit $\nu \ll \mu$.

Satz 1.10 Für $\nu, \mu \in \mathcal{M}(\Omega, \mathcal{A})$ sind die folgenden beiden Aussagen äquivalent:

1. $\nu \ll \mu$
2. Für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ sodass $\nu(E) < \varepsilon$ für jedes $E \in \mathcal{A}$ mit $\mu(E) < \delta$.

Beweis: Übungsaufgabe

1.2 Bedingte Erwartungen

Wir betrachten ein motivierendes (nicht-diskretes) Beispiel, in der die Berechnung des bedingten Erwartungswerts zumindest intuitiv klar ist.

Beispiel 1.11 Der Zufallsvektor (X, Y) sei absolut stetig (bezüglich des Lebesgue-Maßes λ_2 auf $\mathcal{B}(\mathbb{R}^2)$) mit Dichte

$$f(x, y) = \frac{1}{\pi} \mathbf{1}_{B(0,1)}(x, y),$$

wobei $B(0, 1) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ die offene Einheitskugel bezeichnet. Wie würden Sie $\mathbb{P}(Y \in B | X = x)$ sowie $\mathbb{E}(Y | X = x)$ für $B \in \mathcal{B}(\mathbb{R})$ und $x \in \mathbb{R}$ definieren? Intuitiv würde man zweifelsohne

$$Y | (X = x) \sim \mathcal{U}(-\sqrt{1-x^2}, \sqrt{1-x^2})$$

für jedes $x \in [-1, 1]$ setzen und damit $\mathbb{E}(Y | X = x) = 0$ erhalten. Intuitiv ebenfalls naheliegender wäre, dass

$$\mathbb{P}((X, Y) \in E \times F) = \mathbb{P}^{(X, Y)}(E \times F) = \int_E \mathbb{P}(Y \in F | X = x) d\mathbb{P}^X(x)$$

für jedes $G \in \mathcal{B}(\mathbb{R}^2)$ sowie

$$\mathbb{E}(Y) = \int_{[-1,1]} \mathbb{E}(Y | X = x) d\mathbb{P}^X(x)$$

gilt. Tatsächlich lassen sich die letzten beiden Identitäten auch formal beweisen - beides sind Spezialfälle der sogenannten Disintegration, die im nächsten Kapitel behandelt wird.

Das nachfolgende Beispiel deutet an, in welche Richtung es gehen wird.

Beispiel 1.12 Wir betrachten den Wurf mit einem fairen Würfel, modelliert durch $(\Omega, \mathcal{A}, \mathbb{P}) = (\{1, 2, \dots, 6\}, \mathcal{P}(\{1, 2, \dots, 6\}), \mathbb{P})$ mit $\mathbb{P}(\{i\}) = \frac{1}{6}$ für jedes $i \in \Omega$ sowie die Ereignisse $A = \{1, 2, 3\}$, $B = \{1, 3, 5\}$. Die bedingte Wahrscheinlichkeit von A gegeben B berechnet sich bekanntlich zu

$$\mathbb{P}_B(A) := \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

Für jede andere Teilmenge $E \in \mathcal{A}$ kann $\mathbb{P}_B(E)$ analog berechnet werden. Sei nun X eine Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Wie kann der Erwartungswert $\mathbb{E}(X|B)$ von X unter B berechnet werden? Naheliegenderweise ist $\mathbb{E}(X|B)$ definiert als der Erwartungswert von X unter dem (neuen) Wahrscheinlichkeitsmaß \mathbb{P}_B . Wir erhalten also

$$\begin{aligned} \mathbb{E}(X|B) &= \int_{\Omega} X d\mathbb{P}_B = \sum_{\omega=1}^6 X(\omega) \mathbb{P}_B(\{\omega\}) = \sum_{\omega \in B} X(\omega) \frac{\mathbb{P}(\{\omega\})}{\mathbb{P}(B)} \\ &= \frac{1}{\mathbb{P}(B)} \sum_{\omega \in \Omega} X(\omega) \mathbf{1}_B(\omega) \mathbb{P}(\{\omega\}) \\ &= \frac{1}{\mathbb{P}(B)} \mathbb{E}(X \mathbf{1}_B) = \frac{(X(1) + X(3) + X(5)) \frac{1}{6}}{\frac{1}{2}} = \frac{X(1) + X(3) + X(5)}{3} \end{aligned}$$

Für B^c erhalten wir analog

$$\mathbb{E}(X|B^c) = \frac{1}{\mathbb{P}(B^c)} \mathbb{E}(X \mathbf{1}_{B^c}).$$

Setzen wir $\mathcal{B} := \mathcal{A}_{\sigma}(\{B, B^c\}) = \{\emptyset, \Omega, B, B^c\} \subseteq \mathcal{A}$ und definieren wir $\mathbb{E}(X|\mathcal{B}) : \Omega \rightarrow \mathbb{R}$ durch

$$\mathbf{E}(X|\mathcal{B})(\omega) = \begin{cases} \mathbb{E}(X|B) & \text{für } \omega \in B, \\ \mathbb{E}(X|B^c) & \text{für } \omega \in B^c, \end{cases} \quad (1.10)$$

dann hat $\mathbf{E}(X|\mathcal{B})$ die folgenden Eigenschaften:

1. $\mathbf{E}(X|\mathcal{B})$ ist \mathcal{B} -messbar (und ein Element von $L^1(\Omega, \mathcal{B}, \mathbb{P})$).
2. Für jedes $C \in \mathcal{B}$ gilt

$$\int_C \mathbf{E}(X|\mathcal{B}) d\mathbb{P} = \int_C X d\mathbb{P}. \quad (1.11)$$

Tatsächlich liefert die Wahl von $C = B$ sofort

$$\begin{aligned} \int_B \mathbf{E}(X|\mathcal{B}) d\mathbb{P} &= \int_B \mathbb{E}(X|B) d\mathbb{P} = \int_B \frac{\mathbb{E}(X \mathbf{1}_B)}{\mathbb{P}(B)} d\mathbb{P} = \frac{\mathbb{E}(X \mathbf{1}_B)}{\mathbb{P}(B)} \int_B \mathbf{1} d\mathbb{P} \\ &= \int_B X d\mathbb{P}, \end{aligned}$$

die Gleichheit für $C = B^c$ und $C = \Omega$ folgt analog, und für $C = \emptyset$ ist die Gleichheit trivial. $\mathbf{E}(X|\mathcal{B})$ kann also interpretiert werden als geglättete (aggregierte) Version von X .

Im obigen Beispiel hätten wir jeden beliebigen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ und eine Menge $B \in \mathcal{A}$ mit $\mathbb{P}(B) > 0$ nehmen können - die Konstruktion der *bedingten Erwartung* $\mathbb{E}(X|\mathcal{B})$ wäre gleich verlaufen.

Ist nun I eine abzählbare Menge, $(B_i)_{i \in \mathbb{N}}$ eine messbare Partition von Ω bestehend aus Mengen mit echt positiver Wahrscheinlichkeit und X eine integrierbare Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$, dann kann $\mathbb{E}(X|\mathcal{B})$ mit $\mathcal{B} = \mathcal{A}_\sigma(\{B_1, B_2, \dots\})$ naheliegenderweise definiert werden durch

$$\mathbb{E}(X|\mathcal{B})(\omega) = \sum_{i \in I} \mathbb{E}(X|B_i) \mathbf{1}_{B_i}(\omega) = \sum_{i \in I} \frac{\mathbb{E}(X \mathbf{1}_{B_i})}{\mathbb{P}(B_i)} \mathbf{1}_{B_i}(\omega). \quad (1.12)$$

Das nachfolgende Lemma bestätigt, dass die so erhaltene Zufallsvariable $\mathbb{E}(X|\mathcal{B}) : \Omega \rightarrow \mathbb{R}$ dieselben Eigenschaften hat wie $\mathbb{E}(X|\mathcal{B})$ im vorigen Beispiel:

Lemma 1.13 *Die Zufallsvariable $\mathbb{E}(X|\mathcal{B})$ gemäß (1.12) ist diskret und hat folgende Eigenschaften:*

1. $\mathbb{E}(X|\mathcal{B})$ ist \mathcal{B} -messbar (i.e. $\mathbb{E}(X|\mathcal{B})^{-1}(E) \in \mathcal{B}$ für jede Borel-Menge $E \in \mathcal{B}(\mathbb{R})$).
2. $\mathbb{E}(X|\mathcal{B})$ ist integrierbar und erfüllt

$$\int_C \mathbb{E}(X|\mathcal{B}) d\mathbb{P} = \int_C X d\mathbb{P} \quad (1.13)$$

für jedes $C \in \mathcal{B}$.

Beweis: Übungsaufgabe

Der finale Schritt von der abzählbar erzeugten σ -Algebra \mathcal{B} auf eine beliebige sub- σ -Algebra $\mathcal{C} \subseteq \mathcal{A}$ ist nun wenig überraschend:

Definition 1.14 *Sei X eine Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$ und \mathcal{C} eine sub- σ -Algebra von \mathcal{A} . Dann heißt eine Zufallsvariable X_0 (eine Version der) bedingte Erwartung von X unter \mathcal{C} (wir schreiben $X_0 = \mathbb{E}(X|\mathcal{C})$) genau dann, wenn sie die folgenden zwei Eigenschaften hat:*

(CE1) X_0 ist \mathcal{C} -messbar.

(CE2) Für jedes $C \in \mathcal{C}$ gilt die Gleichheit

$$\int_C X_0 d\mathbb{P} = \int_C X d\mathbb{P}.$$

Satz 1.15 *Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, \mathcal{C} eine sub- σ -Algebra von \mathcal{A} und X eine integrierbare Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Dann existiert $\mathbb{E}(X|\mathcal{C})$ und ist eindeutig \mathbb{P} -fast überall.*

Beweis: Zum Beweise der Existenz folgen wir, wie gewohnt, dem Integralaufbau und starten mit einer nicht-negativen Zufallsvariable X . Das Maß μ , definiert durch $\mu(C) = \int_C X d\mathbb{P}$ ist (wegen der vorausgesetzten Integrierbarkeit von X) endlich, und wird offensichtlich von \mathbb{P} dominiert. Nach dem Satz von Radon-Nikodym existiert daher eine \mathcal{C} -messbare, integrierbare (\mathbb{P} -fast überall eindeutig bestimmte) Zufallsvariable X_0 mit

$$\int_C X_0 d\mathbb{P} = \mu(C) = \int_C X d\mathbb{P}$$

für alle $C \in \mathcal{C}$. X_0 hat also die gewünschten Eigenschaften und der Satz ist für nicht-negatives X bewiesen.

Für allgemeines (integrierbares X) verwenden wir die Zerlegung $X = X^+ - X^-$, bestimmen die entsprechenden bedingten Erwartungen X_0^+, X_0^- und setzen einfach $X_0 = X_0^+ - X_0^-$. Zum Beweis der behaupteten Eindeutigkeit nehmen wir an, X_0 und X'_0 erfüllen (CE1) und (CE2), und setzen $E := \{\omega \in \Omega : X_0(\omega) > X'_0(\omega)\} \in \mathcal{C}$. Eigenschaft (CE2) liefert

$$0 = \int_E X_0 - X'_0 d\mathbb{P} = \int_{\Omega} \underbrace{\mathbf{1}_E (X_0 - X'_0)}_{\geq 0} d\mathbb{P},$$

woraus sofort $\mathbb{P}(E) = 0$, und damit $X_0 \leq X'_0$ $[\mathbb{P}]$ folgt. Nachdem sich $X'_0 \leq X_0$ $[\mathbb{P}]$ analog ergibt, folgt insgesamt $\mathbb{P}(\{\omega \in \Omega : X_0(\omega) = X'_0(\omega)\}) = 1$. ■

Einige der wichtigsten Eigenschaften bedingter Erwartungen sind im folgenden Satz zusammengefasst.

Satz 1.16 *Seien X, Y integrierbare Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$ und $\mathcal{D} \subseteq \mathcal{C}$ sub- σ -Algebren von \mathcal{A} . Dann gilt (alle Aussagen sind $[\mathbb{P}]$ zu interpretieren):*

1. $\mathbb{E}(aX + bY|\mathcal{C}) = a\mathbb{E}(X|\mathcal{C}) + b\mathbb{E}(Y|\mathcal{C})$ (Linearität)
2. $Y \leq X$ impliziert $\mathbb{E}(Y|\mathcal{C}) \leq \mathbb{E}(X|\mathcal{C})$ (Monotonie)
3. Ist X sogar \mathcal{C} -messbar, dann gilt $\mathbb{E}(X|\mathcal{C}) = X$
4. Falls $\mathbb{E}(|XY|) < \infty$ und Y sogar \mathcal{C} -messbar ist, dann folgt $\mathbb{E}(XY|\mathcal{C}) = Y\mathbb{E}(X|\mathcal{C})$.
5. $\mathbb{E}(\mathbb{E}(X|\mathcal{C})|\mathcal{D}) = \mathbb{E}(\mathbb{E}(X|\mathcal{D})|\mathcal{C}) = \mathbb{E}(X|\mathcal{D})$ (Turmeigenschaft)
6. $\mathbb{E}(|X||\mathcal{C}) \geq |\mathbb{E}(X|\mathcal{C})|$ (Dreiecksungleichung)
7. Falls $\mathcal{A}_\sigma(X)$ und \mathcal{C} unabhängig sind, gilt $\mathbb{E}(X|\mathcal{C}) = \mathbb{E}(X)$
8. Fall \mathcal{C} nur aus 0- oder 1-Mengen besteht folgt $\mathbb{E}(X|\mathcal{C}) = \mathbb{E}(X)$
9. Für eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsvariablen, die \mathbb{P} -fast sicher gegen X konvergiert und $|X_n| \leq Y$ $[\mathbb{P}]$ erfüllt, folgt

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n|\mathcal{C}) = \mathbb{E}(X|\mathcal{C}) \quad [\mathbb{P}]$$

und

$$\lim_{n \rightarrow \infty} \int_{\Omega} |\mathbb{E}(X_n|\mathcal{C}) - \mathbb{E}(X|\mathcal{C})| d\mathbb{P} = 0.$$

Beweis: Linearität folgt sofort aus der Tatsache, dass $a\mathbb{E}(X|\mathcal{C}) + b\mathbb{E}(Y|\mathcal{C})$ \mathcal{C} -messbar ist und dass für jedes $C \in \mathcal{C}$

$$\begin{aligned} \int_C a\mathbb{E}(X|\mathcal{C}) + b\mathbb{E}(Y|\mathcal{C}) d\mathbb{P} &= a \int_C \mathbb{E}(X|\mathcal{C}) d\mathbb{P} + b \int_C \mathbb{E}(Y|\mathcal{C}) d\mathbb{P} = a \int_C X d\mathbb{P} + b \int_C Y d\mathbb{P} \\ &= \int_C aX + bY d\mathbb{P} \end{aligned}$$

gilt.

Zum Beweis der Monotonie setzen wir $C := \{\omega \in \Omega : \mathbb{E}(Y|\mathcal{C}) > \mathbb{E}(X|\mathcal{C})\} \in \mathcal{C}$ und erhalten (unter Verwendung der Linearität und (CE2))

$$\int_{\Omega} \underbrace{\mathbf{1}_C \mathbb{E}(Y|\mathcal{C}) - \mathbb{E}(X|\mathcal{C})}_{\geq 0} d\mathbb{P} = \int_C \mathbb{E}(Y|\mathcal{C}) - \mathbb{E}(X|\mathcal{C}) d\mathbb{P} = \int_C Y - X d\mathbb{P} \leq 0,$$

und damit $\mathbb{P}(C) = 0$. Der Beweis der restlichen Aussagen ist eine Übungsaufgabe.

Satz 1.16 besagt insbesondere, dass für jede integrierbare Zufallsvariable X die Eigenschaft $\mathbb{E}(\mathbb{E}(X|\mathcal{C})|\mathcal{C}) = \mathbb{E}(X|\mathcal{C})$ gilt, die Abbildung $X \mapsto \mathbb{E}(X|\mathcal{C})$ also idempotent ist. Idempotenz kennen wir insbesondere von orthogonalen Projektionen (i.e. Bestapproximierende in Unterräumen). Zusätzlich wurde vorhin schon erwähnt wurde, dass $\mathbb{E}(X|\mathcal{C})$ als geglättete Version von X gesehen werden kann. Mit dieser Interpretation im Hinterkopf ist das folgende Resultat nicht mehr allzu überraschend:

Satz 1.17 *Seien X quadratisch integrierbar und \mathcal{C} eine sub- σ -Algebren von \mathcal{A} . Dann gilt für jede quadratisch integrierbare, \mathcal{C} -messbare Zufallsvariable Y die Ungleichung*

$$\mathbb{E}((X - Y)^2) \geq \mathbb{E}((X - \mathbb{E}(X|\mathcal{C}))^2) \tag{1.14}$$

mit Gleichheit, dann und nur dann, wenn $Y = \mathbb{E}(X|\mathcal{C})$ $[\mathbb{P}]$.

Mit anderen Worten: $\mathbb{E}(X|\mathcal{C})$ ist die orthogonale Projektion von $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ auf den (abgeschlossenen) Unterraum $L^2(\Omega, \mathcal{C}, \mathbb{P})$.

Beweis: Nachdem aus der quadratischen Integrierbarkeit von X die quadratische Integrierbarkeit von $\mathbb{E}(X|\mathcal{C})^2$ folgt (Übungsaufgabe), reicht es, Ungleichung (1.14) zu beweisen.

Sei nun Y \mathcal{C} -messbar und quadratisch integrierbar. Anwendung der Cauchy-Schwarz Ungleichung liefert sofort $\mathbb{E}(|XY|) < \infty$, weiters können lt. Punkt 4. in Satz 1.16 $\mathbb{E}(XY)$ und $\mathbb{E}(X\mathbb{E}(X|\mathcal{C}))$ umgeschrieben werden zu

$$\begin{aligned} \mathbb{E}(XY) &= \mathbb{E}\left(\mathbb{E}(XY|\mathcal{C})\right) \stackrel{4.}{=} \mathbb{E}(Y\mathbb{E}(X|\mathcal{C})) \\ \mathbb{E}(X\mathbb{E}(X|\mathcal{C})) &= \mathbb{E}\left(\mathbb{E}(X\mathbb{E}(X|\mathcal{C})|\mathcal{C})\right) \stackrel{4.}{=} \mathbb{E}(\mathbb{E}(X|\mathcal{C})^2). \end{aligned}$$

Unter Verwendung dieser Gleichheiten folgt die gewünschte Ungleichung direkt aus

$$\begin{aligned} \mathbb{E}(X - Y)^2 - \mathbb{E}((X - \mathbb{E}(X|\mathcal{C}))^2) &= \mathbb{E}\left(X^2 - 2XY + Y^2 - X^2 + 2X\mathbb{E}(X|\mathcal{C}) - \mathbb{E}(X|\mathcal{C})^2\right) \\ &= \mathbb{E}(Y^2) - 2\mathbb{E}(XY) + 2 \underbrace{\mathbb{E}(X\mathbb{E}(X|\mathcal{C}))}_{=\mathbb{E}(\mathbb{E}(X\mathbb{E}(X|\mathcal{C})|\mathcal{C}))} - \mathbb{E}(\mathbb{E}(X|\mathcal{C})^2) \\ &= \mathbb{E}(Y^2) - 2\mathbb{E}(Y\mathbb{E}(X|\mathcal{C})) + \mathbb{E}(\mathbb{E}(X|\mathcal{C})^2) \\ &= \mathbb{E}(Y - \mathbb{E}(X|\mathcal{C}))^2 \geq 0. \quad \blacksquare \end{aligned}$$

Die von der Zufallsvariable X erzeugte σ -Algebra $\mathcal{C} = \mathcal{A}_\sigma(X) = X^{-1}(\mathcal{B}(\mathbb{R}))$ ist (aufgrund der Messbarkeit von X) offensichtlich eine sub- σ -Algebra von \mathcal{A} . Für eine integrierbare Zufallsvariable Y auf $(\Omega, \mathcal{A}, \mathbb{P})$ und eine weitere (nicht notwendigerweise integrierbare) Zufallsvariable X nennen wir daher

$$\mathbb{E}(Y|X) := \mathbb{E}(Y|\mathcal{A}_\sigma(X)), \tag{1.15}$$

bedingte Erwartung von Y unter X .

Das folgende Hilfsresultat impliziert, dass $\mathbb{E}(Y|X)$ sich immer in der Form $\mathbb{E}(Y|X) = h \circ X$ mit einer (Borel-) messbaren Funktion $h : \mathbb{R} \rightarrow \mathbb{R}$ darstellen lässt.

Lemma 1.18 (Faktorisierungslemma) *Seien X und Y Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Dann sind die folgenden beiden Bedingungen äquivalent:*

1. Y ist $\mathcal{A}_\sigma(X)$ messbar.
2. Es existiert eine (Borel) messbare Transformation $h : \mathbb{R} \rightarrow \mathbb{R}$ mit $Y = h \circ X$.

Zusatz: Ist Y nicht-negativ, dann kann auch h nicht-negativ gewählt werden. Weiters ist h eindeutig bestimmt \mathbb{P}^X -fast überall.

Beweis: Dem Integralaufbau folgend eine einfach Übungsaufgabe.

1.3 Disintegration und Markov Kerne

Wir wissen aus dem vorigen Abschnitt, dass sich für (reelle) Zufallsvariable X, Y auf $(\Omega, \mathcal{A}, \mathbb{P})$ die bedingte Erwartung $\mathbb{E}(Y|X)$ von Y unter X in der Form $\mathbb{E}(Y|X) = h \circ X$ darstellen lässt, wobei $h : \mathbb{R} \rightarrow \mathbb{R}$ eine messbare (und i.A. nicht eindeutige) Funktion ist. Wie am Beginn von Abschnitt 1.2 skizziert, ist unser Hauptziel, allgemeine bedingte Verteilungen (i.e. die Verteilung von Y gegeben $X = x$ für $x \in \mathbb{R}$ und beliebige, nicht notwendigerweise diskrete Zufallsvariable X) zu konstruieren. Die folgenden Definition ist wenig überraschend:

Definition 1.19 *Seien X, Y Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$, zusätzlich sei Y integrierbar. Dann ist EINE Version der bedingte Erwartung $\mathbb{E}(Y|X = x)$ von Y gegeben $X = x$ gegeben durch*

$$\mathbb{E}(Y|X = x) = h(x), \quad (1.16)$$

wobei h EINE lt. Faktorisierungslemma existierende Funktion mit $\mathbb{E}(Y|X) = h \circ X$ ist.

Frage 1.20 Stimmt die Gleichheit $\mathbb{E}(Y|X = X(\omega)) = \mathbb{E}(Y|X)(\omega)$ für jedes ω , für \mathbb{P} -fast-jedes ω , oder i.A. nicht für \mathbb{P} -fast-jedes ω ?

Haben wir mit Definition 1.19 unser Ziel der Konstruktion bedingter Verteilungen schon erreicht? Naheliegenderweise würden wir nun für jedes $A \in \mathcal{A}$ die bedingte Wahrscheinlichkeit $\mathbb{P}(A|X = x)$ von A gegeben $X = x$ definieren durch (h_A erfülle $\mathbb{E}(\mathbf{1}_A|X) = h_A \circ X$)

$$\mathbb{P}(A|X = x) := \mathbb{E}(\mathbf{1}_A|X = x) = h_A(x). \quad (1.17)$$

Ist die dadurch definierte Abbildung $A \mapsto \mathbb{P}(A|X = x)$ für jedes x ein Wahrscheinlichkeitsmaß auf \mathcal{A} ? Die Frage ist nicht trivial zu beantworten, da die involvierten Funktionen h_A nur \mathbb{P}^X -überall eindeutig bestimmt sind (für absolut stetiges X können wir h_A also auf beliebigen Lebesgue-Nullmengen abändern), und die Nullmengen von A abhängen können. Beispielsweise gilt

$$\mathbb{P}(\emptyset|X = x) = 0$$

nur für \mathbb{P}^X -fast-alle $x \in \mathbb{R}$. Im Allgemeinen kann und wird \mathcal{A} überabzählbar viele Mengen enthalten, wir können also nicht einfach auf den Nullmengen modifizieren, die entsprechenden Nullmengen vereinigen und argumentieren, dass die resultierende Menge wieder eine

Nullmenge ist. Wenn die σ -Algebra sich 'gut' durch abzählbar viele 'brave' Mengen approximieren lässt, kann aber tatsächlich eine Version von $\mathbb{P}(\cdot|X = x)$ bestimmt werden, sodass $A \mapsto \mathbb{P}(A|X = x)$ ein Wahrscheinlichkeitsmaß für jedes feste $x \in \mathbb{R}$, und die Abbildung $x \mapsto \mathbb{P}(A|X = x)$ messbar in x für jedes feste $A \in \mathcal{A}$ ist. $(x, A) \mapsto \mathbb{P}(A|X = x)$ ist dann ein sogenannter Markov Kern.

Definition 1.21 Seien $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$ Messräume. Eine Abbildung $K : \Omega_1 \times \mathcal{A}_2 \rightarrow [0, 1]$ heißt Markov Kern von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$ (oft auch 'stochastischer Kern' oder 'Übergangswahrscheinlichkeit') genau dann, wenn die folgenden zwei Bedingungen erfüllt sind:

- Für jedes feste $A_2 \in \mathcal{A}_2$ gilt: Die Abbildung $\omega_1 \mapsto K(\omega_1, A_2)$ ist $\mathcal{A}_1 - \mathcal{B}(\mathbb{R})$ -messbar.
- Für jedes feste $\omega_1 \in \Omega_1$ ist $A_2 \mapsto K(\omega_1, A_2)$ ein Wahrscheinlichkeitsmaß auf \mathcal{A}_2 .

Beispiel 1.22 Wir betrachten $\Omega_1 = \Omega_2 = \Omega = \{1, 2, \dots, N\}$ mit $N \geq 2$ und $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A} = \mathfrak{p}(\Omega)$. Weiters sei $M = (m_{i,j}) \in [0, 1]^{N \times N}$ eine sogenannte stochastische Matrix, i.e. eine Matrix mit nicht-negativen Einträgen, die

$$\sum_{j=1}^N m_{i,j} = 1, \quad i \in \{1, 2, \dots, N\}$$

erfüllt. Definieren wir $K : \Omega \times \mathcal{A} \rightarrow [0, 1]$ durch

$$K(i, A) = \sum_{j \in A} m_{i,j},$$

dann ist offensichtlich K ein Markov Kern von Ω nach Ω (beachten Sie, dass die Messbarkeit wegen $\mathcal{A} = \mathfrak{p}(\Omega)$ trivial ist). Mit anderen Worten: Jede stochastische Matrix induziert einen Markov Kern. Umgekehrt (einfache Übungsaufgabe) induziert jeder Markov Kern eine stochastische Matrix.

Beispiel 1.23 Wir betrachten $\Omega_1 = (0, \infty), \mathcal{A}_1 = \mathcal{B}((0, \infty)), \Omega_2 = \mathbb{N}, \mathcal{A}_2 = \mathfrak{p}(\Omega_2)$ und setzen (für $B = \emptyset$ sei die Summe per definitionem gleich 0)

$$K(\theta, B) = \sum_{k \in B} \frac{\theta^k}{k!} e^{-\theta}$$

für jedes $\theta \in \Omega$ und jedes $B \in \mathcal{A}$. Dann ist K ein Markov Kern von $(\Omega_1, \mathcal{A}_1)$ nach $(\Omega_2, \mathcal{A}_2)$ (wie lässt sich die Messbarkeit zeigen?), und das Maß $K(\theta, \cdot)$ entspricht der Poisson-Verteilung mit Parameter θ .

Beispiel 1.24 (Fortsetzung von Beispiel 1.11) Sei (X, Y) stetig gleichverteilt auf dem Einheitskreis $B(0, 1)$, i.e. die Dichte von (X, Y) ist gegeben durch $f = \frac{1}{\pi} \mathbf{1}_{B(0,1)}$. Wir setzen $\Omega := \Omega_1 = \Omega_2 = (-1, 1), \mathcal{A} = \mathcal{A}_1 = \mathcal{A}_2 = \mathcal{B}((-1, 1))$, sowie $I_x := (-\sqrt{1-x^2}, \sqrt{1-x^2})$ für jedes $x \in (-1, 1)$. Definieren wir $K : (-1, 1) \times \mathcal{B}((-1, 1)) \rightarrow [0, 1]$ durch

$$K(x, F) = \frac{1}{2\sqrt{1-x^2}} \int_E \mathbf{1}_{I_x}(y) d\lambda(y) = \frac{1}{2\sqrt{1-x^2}} \lambda(F \cap I_x), \quad (1.18)$$

dann lässt sich unschwer zeigen, dass K ein Markov Kern von (Ω, \mathcal{A}) nach (Ω, \mathcal{A}) ist (Übungsaufgabe!). Woher kommt die Idee, den Kern wie in Gleichung (1.18) zu wählen?

Definition 1.25 Seien X, Y Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$ und \mathcal{C} eine sub- σ -Algebra von \mathcal{A} . Dann heißt ein Markov Kern K von $(\Omega, \mathcal{C})^\dagger$ nach $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ eine reguläre bedingte Verteilung von Y unter \mathcal{C} genau dann wenn

$$K(\omega, B) = \mathbb{E}(\mathbf{1}_B \circ Y | \mathcal{C})(\omega) \quad [\mathbb{P}]$$

für jedes $B \in \mathcal{B}(\mathbb{R})$ gilt.

Für den Spezialfall $\mathcal{C} = \mathcal{A}_\sigma(X)$ heißt ein Markov Kern K von $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ nach $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ eine reguläre bedingte Verteilung von Y unter X genau dann, wenn für jedes $B \in \mathcal{B}(\mathbb{R})$ die Gleichheit

$$K(X(\omega), B) = \mathbb{E}(\mathbf{1}_B \circ Y | X)(\omega)$$

für \mathbb{P} -fast alle ω gilt (mit anderen Worten: wenn $(\omega, B) \mapsto K(X(\omega), B)$ eine Version der bedingten Erwartung von $\mathbf{1}_B \circ Y$ unter $\mathcal{A}_\sigma(X)$ ist.)

Bevor wir die Existenz regulärer bedingter Verteilungen beweisen, unterstreichen wir deren Wichtigkeit für die Zerlegung von Maßen:

Bemerkung 1.26 Falls K eine reguläre bedingte Verteilung von Y unter X ist, dann gilt also per definitionem für alle $E, B \in \mathcal{B}(\mathbb{R})$ die folgende Gleichheit

$$\int_{X^{-1}(E)} K(X(\omega), B) d\mathbb{P}(\omega) = \int_{X^{-1}(E)} \mathbf{1}_B \circ Y d\mathbb{P} = \mathbb{P}(X^{-1}(E) \cap Y^{-1}(B)) = \mathbb{P}^{(X,Y)}(E \times B).$$

Nachdem sich der erste Ausdruck mittels Change of Coordinates umschreiben lässt zu

$$\int_{X^{-1}(E)} K(X(\omega), B) d\mathbb{P}(\omega) = \int_{\Omega} \mathbf{1}_E \circ X(\omega) K(X(\omega), B) d\mathbb{P}(\omega) = \int_E K(x, B) d\mathbb{P}^X(x)$$

erhalten wir insgesamt

$$\mathbb{P}^{(X,Y)}(E \times B) = \int_E K(x, B) d\mathbb{P}^X(x), \quad (1.19)$$

das Wahrscheinlichkeitsmaß $\mathbb{P}^{(X,Y)}$ wird also zerlegt (disintegriert) in eine Familie eindimensionaler Maße $K(x, \cdot)$ und die Randverteilung \mathbb{P}^X von X .

Gleichung (1.19) charakterisiert sogar reguläre bedingte Verteilungen von Y unter X - es gilt das folgende Lemma:

Lemma 1.27 Ein Markov Kern K von $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ nach $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ist eine reguläre bedingte Verteilung von Y unter X genau dann, wenn Gleichung (1.19) für alle $E, B \in \mathcal{B}(\mathbb{R})$ gilt.

Beweis: Einfache Übungsaufgabe.

Beispiel 1.28 (Fortsetzung von Beispiel 1.24) Wir bestätigen die in Bemerkung 1.26 hergeleitete Gleichung (1.19) und berechnen zuerst die (Rand-)Verteilung \mathbb{P}^X von X : Für $x \in (-1, 1)$ gilt

$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}(X \leq x, Y \in (-1, 1)) = \int_{(-1,x] \times (-1,1)} f(t, s) d\lambda_2(t, s) \\ &= \int_{(-1,x]} \frac{1}{\pi} \int_{(-1,1)} \mathbf{1}_{I_x}(y) d\lambda(y) d\lambda(x) = \frac{1}{\pi} \int_{(-1,x]} 2\sqrt{1-x^2} d\lambda(x). \end{aligned}$$

[†]Beachten Sie, dass hier die Meßbarkeit bezüglich der sub- σ -Algebra \mathcal{C} gefordert wird

X ist also absolut stetig mit Dichte $\frac{2}{\pi}\sqrt{1-x^2}\mathbf{1}_{(-1,1)}(x)$. Für jedes ganz in $(-1,1)^2$ liegende, messbare Rechteck $R = E \times F$ erhalten wir

$$\begin{aligned} \int_E K(x, F) d\mathbb{P}^X(x) &= \int_E \frac{1}{2\sqrt{1-x^2}} \lambda(F \cap I_x) d\mathbb{P}^X(x) \\ &= \int_E \frac{1}{2\sqrt{1-x^2}} \lambda(F \cap I_x) \frac{2}{\pi} \sqrt{1-x^2} \mathbf{1}_{(-1,1)}(x) d\lambda(x) \\ &= \frac{1}{\pi} \lambda(F) \lambda(E) = \frac{1}{\pi} \lambda_2(E \times F) = \mathbb{P}^{(X,Y)}(E \times F). \end{aligned}$$

Satz 1.29 (Existenz regulärer bedingter Verteilungen) *Sei Y eine Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$ und \mathcal{C} eine sub- σ -Algebra von \mathcal{A} . Dann existiert eine reguläre bedingte Verteilung von Y unter \mathcal{C} .*

Beweis: Die Beweisidee besteht darin, Versionen der bedingten Verteilungsfunktionen geschickt zu wählen, und daraus einen Markov Kern zu basteln.

Für jedes $r \in \mathbb{Q}$ bezeichne $F(r, \cdot)$ eine Version von $\mathbb{E}(\mathbf{1}_{(-\infty, r]} \circ Y | \mathcal{C})$. Nachdem für $r < s, r, s \in \mathbb{Q}$ offensichtlich $\mathbf{1}_{(-\infty, r]} \circ Y(\omega) \leq \mathbf{1}_{(-\infty, s]} \circ Y(\omega)$ für jedes $\omega \in \Omega$ gilt, folgt aus Satz 1.16 Punkt 2. sofort die Existenz einer Nullmenge $A_{r,s} \in \mathcal{C}$ mit

$$F(r, \omega) \leq F(s, \omega) \quad \text{für jedes } \omega \in A_{r,s}^c.$$

Weiters existiert lt. Satz 1.16 Punkt 8. eine Nullmenge $B_r \in \mathcal{C}$ mit

$$\lim_{n \rightarrow \infty} F\left(r + \frac{1}{n}, \omega\right) = F(r, \omega) \quad \text{für jedes } \omega \in B_r^c$$

sowie eine weitere Nullmenge $D \in \mathcal{C}$ mit

$$\lim_{n \rightarrow \infty} F(-n, \omega) = 0, \quad \text{und} \quad \lim_{n \rightarrow \infty} F(n, \omega) = 1 \quad \text{für jedes } \omega \in D^c$$

Setzen wir

$$N := \bigcup_{r,s \in \mathbb{Q}, r < s} A_{r,s} \cup \bigcup_{r \in \mathbb{Q}} B_r \cup D$$

dann folgt sofort $N \in \mathcal{C}$ sowie $\mathbb{P}(N) = 0$. Für jedes $\omega \in N^c$ ist $r \mapsto F(r, \omega)$ die Einschränkung einer Verteilungsfunktion auf \mathbb{Q} . Diese Verteilungsfunktion $\bar{F}(\cdot, \omega)$ ist offensichtlich gegeben durch

$$\bar{F}(z, \omega) = \inf\{F(r, \omega) : r \in \mathbb{Q} \cap (z, \infty)\}.$$

Wir wählen eine beliebige eindimensionale Verteilungsfunktion F_0 und definieren \bar{F} auf $\mathbb{R} \times N^c$ durch $\bar{F}(z, \omega) = F_0(z)$. Ausgehend von \bar{F} definieren wir K für $\omega \in \Omega$ und $B = (-\infty, r]$ durch

$$K(\omega, B) = \bar{F}(r, \omega). \tag{1.20}$$

und zeigen, dass K zu einem Markov-Kern mit den gewünschten Eigenschaften fortgesetzt werden kann. Für $r \in \mathbb{Q}$ gilt nach Konstruktion

$$K(\omega, (-\infty, r]) = \mathbf{1}_{N^c}(\omega) F(r, \omega) + \mathbf{1}_N(\omega) F_0(r),$$

die Fortsetzung von K auf ganz $\Omega \times \mathcal{B}(\mathbb{R})$ ist trivial[†]. Nachdem alle vier auf der rechten Seite vorkommenden Funktionen \mathcal{C} -messbar sind, ist auch $\omega \mapsto K(\omega, (-\infty, r])$ \mathcal{C} -messbar für jedes $r \in \mathbb{Q}$. Wir wissen, dass $\{(-\infty, r] : r \in \mathbb{Q}\}$ ein durchschnittsstabiler Erzeuger der Borelschen σ -Algebra $\mathcal{B}(\mathbb{R})$ ist; zusätzlich kann leicht überprüft werden, dass

$$\mathcal{D} = \{B \in \mathcal{B}(\mathbb{R}) : \omega \mapsto K(\omega, B) \text{ messbar}\}$$

ein Dynkin-System ist. Insgesamt folgt also (π - λ -Theorem, Übungsaufgabe) $\mathcal{D} = \mathcal{B}(\mathbb{R})$. K ist also ein Markov-Kern und der Beweis ist komplett, sobald gezeigt ist, dass für jedes $B \in \mathcal{B}(\mathbb{R})$ die \mathcal{C} -messbare Abbildung $\omega \mapsto K(\omega, B)$ eine Version der bedingten Erwartung von Y unter \mathcal{C} ist. Für $B = (-\infty, r]$ und beliebiges $C \in \mathcal{C}$ folgt die gewünschte Gleichheit aus

$$\begin{aligned} L(B) &:= \int_C K(\omega, B) d\mathbb{P}(\omega) = \int_C \mathbf{1}_{N^c}(\omega) F(r, \omega) d\mathbb{P}(\omega) = \int_{C \cap N^c} \underbrace{F(r, \omega)}_{=\mathbb{E}(\mathbf{1}_{(-\infty, r]} \circ Y | \mathcal{C}) [\mathbb{P}]} d\mathbb{P}(\omega) \\ &= \int_{C \cap N^c} \mathbf{1}_{(-\infty, r]} \circ Y d\mathbb{P} = \mathbb{P}(C \cap Y^{-1}(B)) =: R(B) \end{aligned}$$

Nachdem $B \mapsto L(B)$ und $B \mapsto R(B)$ endliche Maße sind, die auf $\{(-\infty, r] : r \in \mathbb{Q}\}$ übereinstimmen, folgt $L = R$ auf ganz $\mathcal{B}(\mathbb{R})$. ■

Satz 1.29 gilt nicht nur für reellwertige Zufallsvariable Y , sondern ganz allgemein für Zufallsvariable Y mit Werten in einem sogenannten *Borel Raum*. Nachdem jeder separable, vollständige metrischen Raum (M, d) ein Borel Raum ist, existieren reguläre bedingte Verteilungen also insbesondere für Zufallsvariable Y mit Werten in $(\mathbb{R}^d, \|\cdot\|_2)$, in $(C[a, b], \|\cdot\|_\infty)$, etc.:

Korollar 1.30 (Existenz regulärer bedingter Verteilungen, allgemein) *Sei Y eine Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$ mit Werten in einem vollständigen, separablen metrischen Raum (M, d) und \mathcal{C} eine sub- σ -Algebra von \mathcal{A} . Dann existiert eine reguläre bedingte Verteilung von Y unter \mathcal{C} .*

Beweis: Mit guten Maßtheoriekenntnissen leicht machbar, siehe [8, 9].

Die Kenntnis einer reguläre bedingte Verteilung $K(\cdot, \cdot)$ von Y unter \mathcal{C} ermöglicht die Berechnung einer bedingten Erwartung $g \circ Y$ unter \mathcal{C} - es gilt folgendes Resultat:

Satz 1.31 *Sei Y eine Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$ mit Werten in einem vollständigen, separablen metrischen Raum (M, d) , \mathcal{C} eine sub- σ -Algebra von \mathcal{A} , und $g : M \rightarrow \mathbb{R}$ Borel-messbar mit $\mathbb{E}(|g \circ Y|) < \infty$. Dann gilt*

$$\mathbb{E}(g \circ Y | \mathcal{C})(\omega) = \int_M g(y) K(\omega, dy) \quad [\mathbb{P}]$$

Beweis: Übungsaufgabe

Für den Fall $\mathcal{C} = \mathcal{A}_\sigma(X)$ lässt sich noch mehr machen. Der Beweis des folgenden Resultats verläuft vollkommen analog zu dem von Satz 1.29 - anstatt $(\Omega, \mathcal{A}, \mathbb{P})$ arbeitet man gleich auf dem Wahrscheinlichkeitsraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}^X)$.

[†]1:1 Zusammenhang 1-dimensionaler Verteilungsfunktionen und der entsprechenden Wahrscheinlichkeitsmaße

Satz 1.32 (Existenz regulärer bedingter Verteilungen von Y unter X) Seien X, Y Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Dann existiert eine reguläre bedingte Verteilung $K(\cdot, \cdot)$ von Y unter X . Dieses $K(x, \cdot)$ ist für \mathbb{P}^X -fast alle $x \in \mathbb{R}$ eindeutig bestimmt. Falls X und Y unabhängig sind, ist $K(x, B) = \mathbb{P}^Y(B)$ eine reguläre bedingte Verteilung von Y unter X .

In [8] ist die folgende, wesentlich allgemeinere Version dieses Resultats bewiesen (selbe Beweisstruktur und Verwendung der Isomorphie von Borel Räumen und $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$):

Korollar 1.33 (Existenz reg. bedingter Verteilungen von Y unter X , allgemein) Sei (M, d) ein separabler, vollständiger metrischer Raum und (N, \mathcal{N}) ein Messraum. Weiters sei X eine N -wertige, und Y eine M -wertige Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Dann existiert eine reguläre bedingte Verteilung $K(\cdot, \cdot)$ von Y unter X . Dieses $K(x, \cdot)$ ist für \mathbb{P}^X -fast alle $x \in N$ eindeutig bestimmt. Falls X und Y unabhängig sind, ist $K(x, B) = \mathbb{P}^Y(B)$ eine reguläre bedingte Verteilung von Y unter X .

Nach mehr oder weniger aufwändigen Vorbereitungen können wir nun endlich das Disintegrationstheorem formulieren und beweisen - wir formulieren das Resultat für das einfache Setting in Satz 1.32, die Übertragung auf den allgemeinen Fall ist problemlos.

Satz 1.34 (Disintegrationstheorem) Seien X, Y Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$ und $K(\cdot, \cdot)$ eine reguläre bedingte Verteilung von Y unter X . Weiters sei $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ messbar mit $\mathbb{E}(|g(X, Y)|) < \infty$. Dann folgt:

$$\mathbb{E}(g(X, Y)|X)(\omega) = \int_{\mathbb{R}} g(X(\omega), y)K(X(\omega), dy) [\mathbb{P}] \quad (1.21)$$

$$\mathbb{E}(g(X, Y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} g(x, y)K(x, dy) d\mathbb{P}^X(x) \quad (1.22)$$

Beweisskizze: Offensichtlich folgt (1.22) sofort mittels Integration bezüglich \mathbb{P} und Change of Coordinates aus (1.21). Letztere Gleichung folgt, indem man, wie gewohnt, dem Integralaufbau folgt und mit $g = \mathbf{1}_{E \times F}$ startet (Übungsaufgabe).

Nachdem Satz 1.34 das Hauptziel dieses Abschnitts war, einige ergänzende Beobachtungen:

(B1) Der elementare Satz von der totalen Wahrscheinlichkeit ist ein Spezialfall von Gleichung (1.22): Sei X eine diskrete Zufallsvariable mit Wertebereich $\{\alpha_1, \alpha_2, \dots\}$ und $\mathbb{P}(X = \alpha_i) > 0$ für jedes i . Für $B \in \mathcal{B}(\mathbb{R})$ und $g(x, y) = \mathbf{1}_B(x)$ erhalten wir aus (1.22)

$$\begin{aligned} \mathbb{P}(Y \in B) &= \mathbb{E}(\mathbf{1}_B(Y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_B(y)K(x, dy) d\mathbb{P}^X(x) = \int_{\mathbb{R}} K(x, B) d\mathbb{P}^X(x) \\ &= \sum_{i=1}^{\infty} K(\alpha_i, B)\mathbb{P}(X = \alpha_i) = \sum_{i=1}^{\infty} \mathbb{P}(Y \in B|X = \alpha_i)\mathbb{P}(X = \alpha_i). \end{aligned}$$

(B2) Für den Fall, dass X und Y unabhängig sind, liefert (1.22) unter Verwendung von Change of Coordinates sofort

$$\begin{aligned} \int_{\mathbb{R}^2} g(x, y)d\mathbb{P}^{(X, Y)}(x, y) &= \mathbb{E}(g(X, Y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} g(x, y)K(x, dy) d\mathbb{P}^X(x) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} g(x, y)d\mathbb{P}^Y(y) d\mathbb{P}^X(x), \end{aligned}$$

also genau den Satz von Fubini (für das Produktmaß aus \mathbb{P}^X und \mathbb{P}^Y).

(B3) Sei $G \in \mathcal{B}(\mathbb{R}^2)$ beliebig. Definieren wir den x -Schnitt G_x von G durch $G_x := \{y \in \mathbb{R} : (x, y) \in G\} \in \mathcal{B}(\mathbb{R})$, dann liefert (1.22)

$$\begin{aligned} \mathbb{P}^{(X,Y)}(G) &= \mathbb{E}(\mathbf{1}_G(X, Y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_G(x, y) K(x, dy) d\mathbb{P}^X(x) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{G_x}(y) K(x, dy) d\mathbb{P}^X(x) = \int_{\mathbb{R}} K(x, G_x) d\mathbb{P}^X(x) \end{aligned}$$

Beispiel 1.35 (Fortsetzung von Beispiel 1.24) Wir bestätigen die in Bemerkung (B3) hergeleitete Gleichheit für den Fall, dass (X, Y) stetig gleichverteilt am Einheitskreis $B(0, 1)$ ist. Es wurde schon gezeigt, dass X absolut stetig mit Dichte $\frac{2}{\pi} \sqrt{1-x^2} \mathbf{1}_{(-1,1)}(x)$ ist. Wir berechnen abschließend die Wahrscheinlichkeit dafür, dass (X, Y) im Kreis $B(0, r)$ mit Radius $r \in [0, 1]$ liegt und setzen $I_x^r := (-\sqrt{r^2-x^2}, \sqrt{r^2-x^2})$ für jedes $x \in (0, 1)$:

$$\begin{aligned} \mathbb{P}^{(X,Y)}(B(0, r)) &= \int_{\mathbb{R}} K(x, B(0, r)_x) d\mathbb{P}^X(x) = \int_{(-r,r)} K(x, I_x^r) d\mathbb{P}^X(x) \\ &= \int_{(-r,r)} \frac{2\sqrt{r^2-x^2}}{2\sqrt{1-x^2}} d\mathbb{P}^X(x) = \int_{(-r,r)} \frac{2\sqrt{r^2-x^2}}{2\sqrt{1-x^2}} \frac{2}{\pi} \sqrt{1-x^2} d\lambda(x) \\ &= \frac{2}{\pi} \int_{(-r,r)} \sqrt{r^2-x^2} d\lambda(x) = \frac{r^2\pi}{\pi} = r^2 \end{aligned}$$

Kapitel 2

Analytische Eigenschaften zweidimensionaler Copulas

Wir betrachten ob des geringeren technischen Aufwands nur zweidimensionale Copulas, studieren deren analytische Eigenschaften, und beweisen zuerst den Satz von Sklar.

2.1 Definition und alternative Darstellungen

Die einfachstmögliche (implizite) Definition einer zweidimensionalen Copula ist wie folgt: Seien X und Y auf $[0, 1]$ stetig gleichverteilte Zufallsvariable. Dann heißt die Verteilungsfunktion von (X, Y) eingeschränkt auf $[0, 1]^2$ eine Copula. Setzen wir für $(x, y) \in [0, 1]^2$

$$A(x, y) := \mathbb{P}^{(X, Y)}([0, x] \times [0, y]) = \mathbb{P}(X \leq x, Y \leq y)$$

dann hat $A : [0, 1]^2 \rightarrow [0, 1]$ unter anderem folgende Eigenschaften:

- Für alle $x, y \in [0, 1]$ gilt $A(x, 1) = x$, $A(1, y) = y$, $A(x, 0) = 0 = A(0, y)$.
Beweis: Die erste Gleichheit folgt sofort aus

$$A(x, 1) = \mathbb{P}(X \leq x, Y \leq 1) = \mathbb{P}(X \leq x) = x,$$

die anderen Gleichheiten lassen sich analog herleiten.

- Für $0 \leq x_1 \leq x_2 \leq 1$ und $0 \leq y_1 \leq y_2 \leq 1$ gilt

$$\begin{aligned} A(x_2, y_2) - A(x_1, y_2) - A(x_2, y_1) + A(x_1, y_1) &= \mathbb{P}^{(X, Y)}((x_1, x_2] \times (y_1, y_2]) \\ &= \mathbb{P}^{(X, Y)}([x_1, x_2] \times [y_1, y_2]) \geq 0. \end{aligned}$$

A ist also 2-monoton auf $[0, 1]^2$ wie Sie es von allgemeinen zweidimensionalen Verteilungsfunktionen aus der VO Statistik kennen.

Die zwei soeben erwähnten Eigenschaften liefern die in der Literatur übliche Definition zweidimensionaler Copulas:

Definition 2.1 *Eine Copula ist eine Funktion $A : [0, 1]^2 \rightarrow [0, 1]$ mit den folgenden Eigenschaften:*

$$1. \forall x \in [0, 1] : A(x, 1) = A(1, x) = x, \quad A(x, 0) = A(0, x) = 0.$$

2. Für $0 \leq x_1 \leq x_2 \leq 1$ und $0 \leq y_1 \leq y_2 \leq 1$ gilt

$$A(x_2, y_2) - A(x_1, y_2) - A(x_2, y_1) + A(x_1, y_1) \geq 0. \quad (2.1)$$

\mathcal{C} bezeichnet im Folgenden die Familie aller zweidimensionalen Copulas.

Bemerkung 2.2 Obwohl Copulas nur auf $[0, 1]^2$ definiert sind, können Sie mit eindeutigen zweidimensionalen Verteilungsfunktionen H_A identifiziert, bzw. zu zweidimensionalen Verteilungsfunktionen fortgesetzt werden. Wie funktioniert diese Fortsetzung?

Beispiel 2.3 Die folgenden Funktionen sind Copulas (direktes Nachrechnen):

$$M(x, y) = \min\{x, y\}, \quad \Pi(x, y) = xy, \quad W(x, y) = \max\{x + y - 1, 0\}$$

Lemma 2.4 Sei $A \in \mathcal{C}$ beliebig, dann gilt:

1. A ist monoton nicht-fallend in jeder Koordinate.

2. Für $x_1, x_2, y_1, y_2 \in [0, 1]$ gilt

$$|A(x_2, y_2) - A(x_1, y_1)| \leq |x_2 - x_1| + |y_2 - y_1|. \quad (2.2)$$

Jede Copula ist also (Lipschitz-) stetig.

Beweis: Einfache Übungsaufgabe.

Definition 2.5 Ein Wahrscheinlichkeitsmaß μ auf $([0, 1]^2, \mathcal{B}([0, 1]^2))$ heißt doppelt stochastisch, genau dann, wenn seine eindimensionalen Randverteilungen mit λ zusammenfallen, i.e. wenn für jedes $E \in \mathcal{B}([0, 1])$ die Gleichheit

$$\mu(E \times [0, 1]) = \lambda(E) = \mu([0, 1] \times E) \quad (2.3)$$

gilt. $\mathcal{P}_{\mathcal{C}}$ bezeichnet im Folgenden die Menge aller doppelt stochastischen Maße.

Offensichtlich entspricht jeder Copula A genau ein doppelt stochastisches Maß μ_A und umgekehrt - es gilt folgender Satz:

Lemma 2.6 Für jede Copula $A \in \mathcal{C}$ ist das durch $\mu_A([0, x] \times [0, y]) = A(x, y)$ auf $\mathcal{B}([0, 1]^2)$ definierte Wahrscheinlichkeitsmaß[†] doppelt stochastisch. Umgekehrt ist für jedes doppelt stochastische Maß $\mu \in \mathcal{P}_{\mathcal{C}}$ die durch $A(x, y) = \mu([0, x] \times [0, y])$ auf $[0, 1]^2$ definierte Funktion eine Copula.

Beweis: Sei $A \in \mathcal{C}$ beliebig, μ_A für $x, y \in [0, 1]$ definiert durch $\mu_A([0, x] \times [0, y]) = A(x, y)$ und in gewohnter Manier (eindeutig!) fortgesetzt auf ganz $\mathcal{B}([0, 1]^2)$. Das Wahrscheinlichkeitsmaß $\nu_1 : \mathcal{B}([0, 1]) \rightarrow [0, 1]$ (a.k.a. erste Randverteilung von μ_A), definiert durch

$$\nu_1(E) := \mu_A(E \times [0, 1])$$

erfüllt $\nu_1([0, x]) = \mu_A([0, x] \times [0, 1]) = A(x, 1) = x$ für jedes $x \in [0, 1]$, woraus sofort $\nu_1 = \lambda$ folgt. Die zweite Aussage des Satzes ist einfach nachzurechnen. ■

[†]Die Eindeutigkeit der Fortsetzung ist klar aus dem bijektiven Zusammenhang zweidimensionaler Verteilungsfunktionen und zweidimensionalen Wahrscheinlichkeitsmaßen auf $\mathcal{B}(\mathbb{R}^2)$

Bemerkung 2.7 In der Folge bezeichnet μ_A das zur Copula $A \in \mathcal{C}$ gehörige doppelt stochastische Maß und $A_\mu \in \mathcal{C}$ die zum doppelt stochastischen Maß $\mu \in \mathcal{P}_{\mathcal{C}}$ gehörige Copula.

Copulas lassen sich weiters identifizieren mit speziellen Markov Kernen von $([0, 1], \mathcal{B}([0, 1]))$ nach $([0, 1], \mathcal{B}([0, 1]))$ (siehe [16]):

Lemma 2.8 Sei A eine beliebige zweidimensionale Copula. Dann existiert ein Markov Kern $K(\cdot, \cdot)$ von $([0, 1], \mathcal{B}([0, 1]))$ nach $([0, 1], \mathcal{B}([0, 1]))$ mit

$$\mu_A(E \times F) = \int_E K(x, F) d\lambda(x) \tag{2.4}$$

für alle $E, F \in \mathcal{B}([0, 1])$. Dieser Kern ist für λ -fast jedes $x \in [0, 1]$ eindeutig bestimmt (wir nennen ein solches $K(\cdot, \cdot)$ im Folgenden kurz 'Markov Kern von A ').

Erfüllt umgekehrt ein Markov Kern $K(\cdot, \cdot)$ von $([0, 1], \mathcal{B}([0, 1]))$ nach $([0, 1], \mathcal{B}([0, 1]))$

$$\int_{[0,1]} K(x, F) d\lambda(x) = \lambda(F) \tag{2.5}$$

für jedes $F \in \mathcal{B}([0, 1])$, dann ist $K(\cdot, \cdot)$ Markov Kern der Copula $A \in \mathcal{C}$, definiert durch

$$A(x, y) = \int_{[0,x]} K(t, [0, y]) d\lambda(t) \tag{2.6}$$

für $x, y \in [0, 1]$.

Beweis: Seien X, Y ($\mathcal{U}(0, 1)$ -verteilte) Zufallsvariable mit $(X, Y) \sim A^\dagger$. Direkte Anwendung von Satz 1.32 und dem Disintegrationstheorem mit $g := \mathbf{1}_{E \times F}$ liefert die ersten beiden Aussagen. Die verbleibende dritte Aussage ist eine einfache Rechenaufgabe, es müssen nur die in Definition 2.1 angegebenen Eigenschaften nachgerechnet werden. ■

Ein vierter alternativer Zugang zu Copulas ist mittels sog. Markov Operatoren und geht zurück auf [2]: Ein linearer operator T auf $L^1([0, 1], \mathcal{B}([0, 1]), \lambda)$ heißt *Markov Operator* (im engeren Sinne), falls er die folgenden drei Eigenschaften erfüllt (alles im L^1 -Sinne zu interpretieren):

1. T ist positive, i.e. $T(g) \geq 0$ falls $g \geq 0$
2. $T(\mathbf{1}_{[0,1]}) = \mathbf{1}_{[0,1]}$
3. $\int_{[0,1]} (Tg)(x) d\lambda(x) = \int_{[0,1]} g(x) d\lambda(x)$

Die Menge aller Markov Operatoren auf $L^1([0, 1], \mathcal{B}([0, 1]), \lambda)$ wird im Folgenden mit \mathcal{M} bezeichnet. Markov Operatoren haben offensichtlich Operatornorm 1 (einfache Übungsaufgabe).

Lemma 2.9 Sei $A \in \mathcal{C}$ beliebig, dann ist T_A , definiert durch $(A_{,2}(x, t) := \frac{\partial A}{\partial y}(x, t))$

$$(T_A g)(x) := \frac{d}{dx} \int_{[0,1]} A_{,2}(x, t) g(t) d\lambda(t) \tag{2.7}$$

[†]i.e. (X, Y) hat Verteilungsfunktion A

für jedes $g \in L^1([0, 1])$ ein Markov Operator, i.e. $T_A \in \mathcal{M}$.
 Umgekehrt ist für jeden Markov Operator $T \in \mathcal{M}$ die Abbildung A_T , definiert durch

$$A_T(x, y) := \int_{[0, x]} (T\mathbf{1}_{[0, y]})(t) d\lambda(t) \tag{2.8}$$

für alle $x, y \in [0, 1]$ eine Copula.

Beweis: Übungsaufgabe

Gleichung (2.7) lässt sich mittels bedingter Erwartungen und Markov Kernen vereinfachen zu (siehe [16])

$$(T_A g)(x) = \mathbb{E}(g \circ Y | X = x) = \int_{[0, 1]} g(y) K_A(x, dy). \tag{2.9}$$

Der Markov Operator beschreibt also nichts anderes als den bedingten Erwartungswert von $g \circ Y$ gegeben $X = x$, wobei $(X, Y) \sim A$.

Beispiel 2.10 Die Wahrscheinlichkeitsdichte f sei gegeben durch

$$f(x, y) = (1 + \theta(1 - 2x)(1 - 2y)) \mathbf{1}_{[0, 1]^2}(x, y)$$

mit $\theta \in [-1, 1]$. Dann ist das zu f gehörige Wahrscheinlichkeitsmaß μ doppelt stochastisch: Offensichtlich gilt für jedes $E \in \mathcal{B}([0, 1])$

$$\mu([0, 1] \times E) = \int_{[0, 1] \times E} f(x, y) d\lambda_2(x, y) = \int_E \int_{[0, 1]} f(x, y) d\lambda(x) d\lambda(y) = \int_E 1 d\lambda(y) = \lambda(E),$$

und wegen der Symmetrie der Dichte daher auch $\lambda(E) = \mu([0, 1] \times E) = \mu(E \times [0, 1])$. $A \in \mathcal{C}$ bezeichne die entsprechende Copula. Beachten Sie: Für $\theta = 0$ gilt $A = \Pi$.

Wir berechnen den Markov Kern $K_A(\cdot, \cdot)$ und den Markov Operator T_A und erhalten

$$\begin{aligned} K_A(x, [0, y]) &= \int_{[0, y]} f(x, s) d\lambda(s) = y + \theta(1 - 2x) \int_{[0, y]} (1 - 2s) d\lambda(s) \\ &= y + \theta(1 - 2x)(y - y^2) \end{aligned}$$

und

$$(T_A g)(x) = \int_{[0, 1]} g(y) K_A(x, dy) = \int_{[0, 1]} g(y) f(x, y) d\lambda(y).$$

Beispiel 2.11 Eine messbare Abbildung $h : [0, 1] \rightarrow [0, 1]$ heißt λ -treu (oder λ -erhaltend) genau dann, wenn $\lambda^h = \lambda$ gilt. Für $X \sim \mathcal{U}(0, 1)$ setzen wir $Y = h \circ X$. Dann folgt $Y \sim \mathcal{U}(0, 1)$, die (auf $[0, 1]^2$ betrachtete) Verteilungsfunktion A_h von (X, Y) ist daher eine Copula (die wir später als 'vollständig abhängig' bezeichnen werden). Markov Kern K_{A_h} und Markov Operator T_{A_h} sind in diesem Fall besonders einfach - es gilt

$$K_{A_h}(x, F) = \mathbf{1}_F(h(x)) = \delta_{h(x)}(F)$$

und

$$(T_{A_h} g)(x) = \int_{[0, 1]} g(y) K_{A_h}(x, dy) = g \circ h(x).$$

Copulas der Form A_h mögen auf den ersten Blick viel zu speziell/zu exotisch und daher nutzlos erscheinen - wir im Laufe der LV sehen, dass dieser erste Eindruck ganz und gar falsch ist.

2.2 Der Satz von Sklar

Das wohl wichtigste und bekannteste Resultat im Zusammenhang mit Copulas ist der Satz von Sklar, der insbesondere besagt, dass Copulas der 'Link' zwischen zweidimensionalen Verteilungsfunktionen und deren Randverteilungen sind.

Satz 2.12 (Satz von Sklar) *Gelte $(X, Y) \sim H$, sowie $X \sim F$ und $Y \sim G$. Dann existiert eine Copula $A \in \mathcal{C}$ sodass für alle $x, y \in \mathbb{R}$ die folgende Gleichheit gilt:*

$$H(x, y) = A(F(x), G(y)). \quad (2.10)$$

Falls F und G stetig sind, ist die Copula A eindeutig bestimmt[†].

Umgekehrt ist für jede Copula $A \in \mathcal{C}$ und eindimensionale Verteilungsfunktionen F, G die Funktion $H(x, y) = A(F(x), G(y))$ eine zweidimensionale Verteilungsfunktion.

Es gibt mittlerweile zahlreiche verschiedene Beweise dieses Resultats (siehe [5]). Wir konzentrieren uns im Folgenden auf den Standardbeweis (für einen Spezialfall) und einen (etwas eleganteren) Beweis basierend auf der verallgemeinerten PIT.

Standardbeweis (Spezialfall): Die dritte Aussage des Satzes kann direkt nachgerechnet werden, die Hauptaussage des Satzes ist der erste Teil.

Wir beweisen die Aussage nur für den Fall, dass F und G stetig sind, und damit offensichtlich $(0, 1) \subseteq \text{Range}(F), \text{Range}(G)$ gilt. Im folgenden bezeichnet F^{-} die (schon aus der Statistik VO bekannte) Quasiinverse von F . Für $u, v \in (0, 1)$ definieren wir A durch

$$A(u, v) := H(F^{-}(u), G^{-}(v)), \quad (2.11)$$

und setzen A mittels $A(x, 1) = A(1, x) = x$, $A(x, 0) = A(0, x) = 0$ für jedes $x \in [0, 1]$ auf ganz $[0, 1]^2$ fort. Es bleibt zu zeigen, dass A eine Copula ist und dass Gleichung (2.10) erfüllt ist. Für $0 < x_1 \leq x_2 < 1$ und $0 < y_1 \leq y_2 < 1$ folgt aus der Monotonie der Quasiinversen sofort $F^{-}(x_1) \leq F^{-}(x_2)$ und $G^{-}(y_1) \leq G^{-}(y_2)$, und wir erhalten unter Verwendung der 2-Monotonie zweidimensionaler Verteilungsfunktionen

$$\begin{aligned} m &:= A(x_2, y_2) - A(x_1, y_2) - A(x_2, y_1) + A(x_1, y_1) \\ &= H(F^{-}(x_2), G^{-}(y_2)) - H(F^{-}(x_1), G^{-}(y_2)) - H(F^{-}(x_2), G^{-}(y_1)) + H(F^{-}(x_1), G^{-}(y_1)) \\ &\geq 0. \end{aligned}$$

Für den Fall, dass mindestens eine Seite des Rechtecks $[x_1, x_2] \times [y_1, y_2]$ am Rand von $[0, 1]^2$ liegt, kann analog vorgegangen werden (selbst durchdenken!).

Um Gleichung (2.10) zu zeigen betrachten wir zuerst Punkte x, y von der Form $x = F^{-}(u), y = F^{-}(v)$ für $u, v \in (0, 1)$. In diesem Fall gilt

$$\begin{aligned} A(F(x), G(y)) &= H(F^{-}(F(x)), G^{-}(G(y))) = H(F^{-} \circ F \circ F^{-}(u), G^{-} \circ G \circ G^{-}(v)) \\ &= H(F^{-}(u), G^{-}(v)) = H(x, y) \end{aligned}$$

Für den Fall $F(x) = 0$ ist die Aussage offensichtlich erfüllt da dann $H(x, y) = 0$ gilt. Für $F(x) \in (0, 1)$ und $x \notin F^{-}((0, 1))$ setzen wir $\underline{x} := F^{-}(F(x))$ und erhalten $F(x) = F(\underline{x})$ sowie $H(x, y) = H(\underline{x}, y)$. Analoge Argumentation für $y \notin G^{-}((0, 1))$ beendet den Beweis. ■

[†]Wir nennen A dann die (X, Y) (bzw. die H) zugrundeliegende Copula

Frage 2.13 Wo im Beweis geht die Stetigkeit von F und G überhaupt ein? Haben wir doch schon den allgemeinen Fall bewiesen?

Vorbereitend auf den angekündigten, eleganten Beweis betrachten wir die folgende Verallgemeinerung der PIT: Sei X eine Zufallsvariable mit Verteilungsfunktion F (nicht notwendigerweise stetig). Dann heißt die Funktion $\Phi : \mathbb{R} \times [0, 1] \rightarrow [0, 1]$, definiert durch

$$\Phi(x, r) = F(x-) + r(F(x) - F(x-)) = \mathbb{P}(X < x) + r \mathbb{P}(X = x)$$

die modifizierte Verteilungsfunktion von F . Offensichtlich gilt $\Phi(x, r) \in [F(x-), F(x)]$ für jedes $x \in \mathbb{R}$. Für stetiges F folgt $\Phi(x, r) = F(x)$ für jedes $r \in [0, 1]$ und $x \in \mathbb{R}$. Aus der VO Statistik ist bekannt, dass für stetiges F die Zufallsvariable $F \circ X$ (PIT) stetig gleichverteilt auf $[0, 1]$ ist. Letztere Aussage gilt auch für die verallgemeinerte PIT, die wie folgt definiert ist.

Definition 2.14 Für $X \sim F$, $R \sim \mathcal{U}(0, 1)$ und X, R unabhängig heißt die Zufallsvariable U , definiert durch

$$U := \Phi(X, R) = F(X-) + R(F(X) - F(X-)), \quad (2.12)$$

die verallgemeinerte PIT von X .

Lemma 2.15 Unter den Voraussetzungen von Definition 2.14 ist U stetig gleichverteilt auf $[0, 1]$ und es gilt $F^-(U) = X$ $[\mathbb{P}]$.

Beweis: Sei $\alpha \in (0, 1)$ beliebig, aber fix; q_α bezeichne das α -Quantil von F , i.e.

$$q_\alpha = F^-(\alpha) = \min\{x \in \mathbb{R} : F(x) \geq \alpha\} = \sup\{x \in \mathbb{R} : F(x) < \alpha\}$$

Wir unterscheiden zwei Fälle: (1) Angenommen, $\mathbb{P}(X = q_\alpha) > 0$. Dann gilt offensichtlich $\Phi_F(x, r) < \alpha$ dann und nur dann, wenn

$$(x, r) \in (-\infty, q_\alpha) \times [0, 1] \cup \{q_\alpha\} \times \left[0, \frac{\alpha - F(q_\alpha-)}{\mathbb{P}(X = q_\alpha)}\right].$$

Für U erhalten wir damit unter Verwendung der Unabhängigkeit von X und R sofort

$$\begin{aligned} \mathbb{P}(U < \alpha) &= \mathbb{P}(X < q_\alpha, R \in [0, 1]) + \mathbb{P}\left(X = q_\alpha, R \leq \frac{\alpha - F(q_\alpha-)}{\mathbb{P}(X = q_\alpha)}\right) \\ &= F(q_\alpha-) + \mathbb{P}(X = q_\alpha) \frac{\alpha - F(q_\alpha-)}{\mathbb{P}(X = q_\alpha)} = \alpha \end{aligned}$$

(2) Für den Fall, dass $\mathbb{P}(X = q_\alpha) = 0$ gilt (q_α also Stetigkeitsstelle von F ist), ist $\Phi_F(x, r) < \alpha$ offensichtlich äquivalent zu $x < q_\alpha$, und wir erhalten

$$\mathbb{P}(U < \alpha) = \mathbb{P}(X < q_\alpha) = \mathbb{P}(X \leq q_\alpha) = F(q_\alpha) = F(F^-(\alpha)) = \alpha.$$

In beiden Fällen haben wir also $\mathbb{P}(U < \alpha) = \alpha$ für jedes $\alpha \in (0, 1)$ gezeigt, $\mathbb{P}(U \in [0, 1])$ ist klar aus der Konstruktion. Nachdem die Stetigkeitsstellen der Verteilungsfunktion G von U dicht in $[0, 1]$ liegen, und für jede Stetigkeitsstelle α von G die Gleichheit $G(\alpha) = \mathbb{P}(U < \alpha) = \alpha$ gilt, folgt die gewünschte Eigenschaft $U \sim \mathcal{U}(0, 1)$ unmittelbar.

Der noch ausstehende Beweis der Gleichheit $F^-(U) = X [\mathbb{P}]$ ist eine Übungsaufgabe. ■

Beweis des Satzes von Sklar, allgemeiner Fall: Seien (X, Y) und $R \sim \mathcal{U}(0, 1)$ unabhängig. Setzen wir $U = \Phi(X, R)$ und $V = \Phi(Y, R)$, dann impliziert Lemma 2.15 $U, V \sim \mathcal{U}(0, 1)$ und $X = F^-(U), Y = G^-(V) [\mathbb{P}]$. Insbesondere ist die Verteilungsfunktion A von (U, V) (eingeschränkt auf $[0, 1]^2$) eine Copula. Für $x, y \in \mathbb{R}$ erhalten wir insgesamt

$$\begin{aligned} H(x, y) &= \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(F^- \circ U \leq x, G^- \circ V \leq y) \\ &= \mathbb{P}(U \leq F(x), V \leq G(y)) = A(F(x), G(y)), \end{aligned}$$

womit der Beweis komplett ist. ■

Bemerkung 2.16 Beachten Sie, dass für unstetige Verteilungsfunktionen F, G mehrere Copulas A gefunden werden können, für die $H = A \circ (F, G)$ gilt, konkrete Beispiele werden in den Übungen diskutiert.

Bemerkung 2.17 Hauptaussage des Satzes von Sklar ist, dass eine Faktorisierung in Copula und Randverteilungen möglich ist und dass selbige im Fall stetiger Randverteilungen eindeutig ist. Als Konsequenz daraus können zweidimensionale Verteilungen alternativ auch anhand von Copulas und zweier eindimensionaler Verteilungsfunktionen studiert werden.

Aufgrund der im vorigen Abschnitt erwähnten alternativen Charakterisierung von Copulas als spezielle Markov Kerne ist das folgende Resultat wenig überraschend:

Satz 2.18 (Satz von Sklar für Markov Kerne) (X, Y) habe Verteilungsfunktion H , stetige Randverteilungen F und G , und Copula $A \in \mathcal{C}$. Weiters sei $K_A(\cdot, \cdot)$ ein Markov Kern von A mit $K_A(x, \{0, 1\}) = 0$ für jedes $x \in [0, 1]^\dagger$. Dann ist $K_H(\cdot, \cdot)$, definiert durch

$$K_H(x, (-\infty]) = K_A(F(x), [0, G(y)])$$

eine reguläre bedingte Verteilung von (X, Y) .

Beweis: Übungsaufgabe

[†]Existiert ein solcher Kern immer?

2.3 Metriken auf \mathcal{C}

Üblicherweise wird auf \mathcal{C} (ausgehend von der Interpretation als Familie spezieller zweidimensionaler Verteilungsfunktionen) die schwache Konvergenz, also die punktweise Konvergenz in Stetigkeitsstellen der Zielfunktion betrachtet. Aufgrund der Lipschitzstetigkeit von Copulas entspricht dies der punktweisen Konvergenz auf ganz $[0, 1]^2$, und, wie in den Übungen gezeigt, der gleichmäßigen Konvergenz auf $[0, 1]^2$. Definieren wir die Metrik d_∞ auf \mathcal{C} durch

$$d_\infty(A, B) = \max\{|A(x, y) - B(x, y)| : x, y \in [0, 1]\},$$

dann gilt also folgendes Resultat:

Satz 2.19 *Für $A, A_1, A_2, \dots \in \mathcal{C}$ sind die folgenden zwei Aussagen äquivalent:*

1. $(A_n)_{n \in \mathbb{N}}$ konvergiert schwach gegen A .
2. $\lim_{n \rightarrow \infty} d_\infty(A_n, A) = 0$.

Den folgenden Begriff haben wir schon in Beispiel 2.11 kennengelernt:

Definition 2.20 *Eine Copula A heißt vollständig abhängig genau dann, wenn eine λ -treue Transformation $h : [0, 1] \rightarrow [0, 1]$ existiert, sodass $\mu_A(\Gamma(h)) = 1$ gilt. Die Menge aller vollständig abhängigen Copulas bezeichnen wir im Folgenden mit \mathcal{C}_d .*

Lemma 2.21 *$A \in \mathcal{C}$ ist genau dann vollständig abhängig, wenn eine λ -treue Transformation $h : [0, 1] \rightarrow [0, 1]$ existiert, sodass $K(x, E) = \mathbf{1}_E(h(x))$ ein Markov Kern von A ist.*

Beweis: Übungsaufgabe

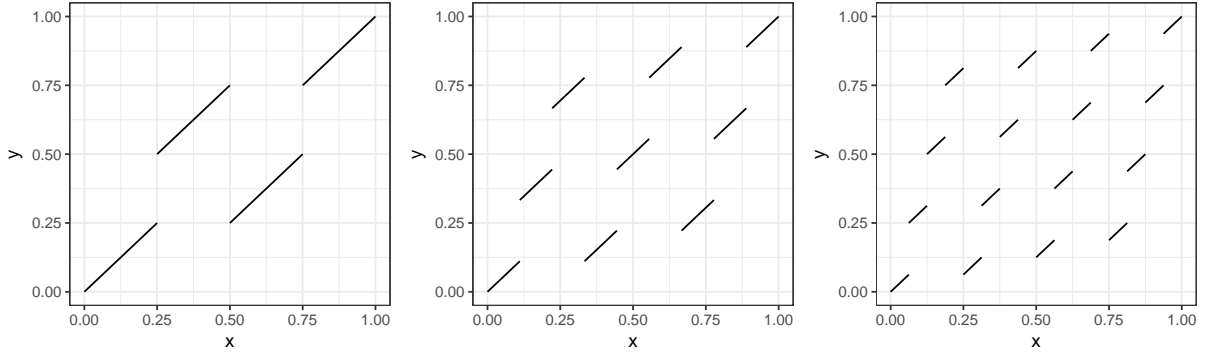
Wir werden in der Folge beweisen, dass die Menge aller vollständig abhängigen Copulas dicht in (\mathcal{C}, d_∞) ist. Direkte Konsequenz daraus ist, dass sich die Metrik d_∞ nicht dafür eignet, Abhängigkeit zu quantifizieren - die gerade erwähnte Dichtheit impliziert nämlich insbesondere, dass es eine Folge $(A_n)_{n \in \mathbb{N}}$ vollständig abhängiger Copulas gibt, die bezüglich d_∞ gegen die Produktcopula Π (Unabhängigkeit!) konvergiert. Für den Beweis verwenden wir λ -treue Transformationen von sehr einfacher Struktur:

Definition 2.22 (Gerader Shuffle) *Eine λ -treue Transformation $h : [0, 1] \rightarrow [0, 1]$ heißt ein (klassischer) gerader Shuffle genau dann, wenn $n \in \mathbb{N}$ und $0 = s_0 < s_1 < \dots < s_{n-1} < s_n = 1$ existieren, sodass h auf jedem Intervall (s_{i-1}, s_i) linear mit Steigung 1 und auf $\bigcup_{i=1}^n (s_{i-1}, s_i)$ injektiv ist. Die zu einem geraden Shuffle h gehörige Copula A_h heißt (gerader) Shuffle von M . Die Menge aller geraden Shuffles wird in der Folge mit \mathcal{S}^+ bezeichnet.*

Lemma 2.23 *h ist ein (klassischer) gerader Shuffle genau dann wenn $n \in \mathbb{N}$ und $0 = s_0 < s_1 < \dots < s_{n-1} < s_n = 1$, $0 = t_0 < t_1 < \dots < t_{n-1} < t_n = 1$ und eine Permutation $\sigma \in \Sigma_n$ existieren, sodass für jedes $i \in \{1, \dots, n\}$ die folgenden beiden Eigenschaften erfüllt sind:*

- $t_{\sigma(i)} - t_{\sigma(i)-1} = s_i - s_{i-1}$
- Für jedes $x \in (s_{i-1}, s_i)$ gilt $h(x) = t_{\sigma(i)-1} + (x - s_{i-1})$

Beweis: Übungsaufgabe

Abbildung 2.1: Beispiele gerader Shuffles, $n = 4, 9, 16$.

Beispiel 2.24 Abbildung 2.3 zeigt drei Beispiele gerader Shuffles.

Satz 2.25 ([11], 1992) \mathcal{S}^+ ist dicht in (\mathcal{C}, d_∞) .

Beweis: Sei $C \in \mathcal{C}$ fest und $\varepsilon > 0$. Wir beweisen die Existenz eines geraden Shuffles A_h von M mit $d_\infty(C, A_h) < \varepsilon$ und wählen dazu $N \in \mathbb{N}$ so, dass $\frac{4}{N} \leq \varepsilon$ gilt. Lemma 2.4 impliziert, dass für jede(!) Copula $A \in \mathcal{C}$ und $x_0, x_1, y_0, y_1 \in [0, 1]$

$$|A(x_0, y_0) - A(x_1, y_1)| < \frac{2}{N} \leq \frac{\varepsilon}{2}$$

folgt, falls

$$|x_0 - x_1| < \frac{1}{N} \quad \text{und} \quad |y_0 - y_1| < \frac{1}{N}$$

gilt.

Für jedes $i \in \{1, \dots, N\}$ definieren wir die (kompakten) vertikalen und horizontalen Streifen V_i bzw. H_i durch

$$V_i = \left[\frac{i-1}{N}, \frac{i}{N} \right] \times [0, 1], \quad H_i = [0, 1] \times \left[\frac{i-1}{N}, \frac{i}{N} \right]$$

und setzen $R_{i,j} = V_i \cap H_j = \left[\frac{i-1}{N}, \frac{i}{N} \right] \times \left[\frac{j-1}{N}, \frac{j}{N} \right]$ für alle $i, j \in \{1, \dots, N\}$. Nachdem jede Copula A die Gleichheit $\mu_A(\partial R_{i,j}) = 0$ erfüllt (warum?), folgt sofort, dass die Matrix $M = (m_{i,j})_{i,j \in \{1, \dots, N\}}$, definiert durch

$$m_{i,j} = \mu_C(R_{i,j})$$

folgende Eigenschaft hat: $NM = (Nm_{i,j})_{i,j \in \{1, \dots, N\}}$ ist doppelt stochastisch.

Die Grundidee für den Beweis besteht nun darin, alle vertikalen und horizontalen Streifen gemäß $(m_{i,j})_{i,j \in \{1, \dots, N\}}$ geschickt weiter zu unterteilen: Wegen $\sum_{k=1}^N m_{i,k} = \frac{1}{N} = \lambda_2(V_i) = \lambda_2(H_i)$ für jedes $i \in \{1, \dots, N\}$ kann der Streifen V_i unterteilt werden in N Streifen V_i^k der Breite $m_{i,k}$, und der Streifen H_j in N Streifen H_j^k der Höhe $m_{k,j}$ (kein Indexfehler, essentiell für den Beweis, eine Skizze hilft). Es folgt direkt aus der Konstruktion, dass $V_i^j \cap H_j^i$ ein (kompaktes) Quadrat der Seitenlänge $m_{i,j}$ ist. Fall V_i^j nichtleeres Inneres hat, definieren

wir die Transformation h auf dem Inneren von V_i^j so, dass der Graph von h genau mit der Diagonale von $V_i^j \cap H_j^i$ übereinstimmt. Setzen wir weiters $h(x) = 0$ für jeden Eckpunkt x eines der Intervalle V_i^j , dann ist h offensichtlich ein gerader Shuffle. Bezeichnen wir mit A_h den von h induzierten Shuffle von M , dann erfüllt A_h daher

$$\mu_{A_h}(R_{i,j}) = \mu_{A_h}(V_i^j \cap H_j^i) = \lambda(V_i^j) = \lambda(H_j^i) = m_{i,j}$$

für alle $i, j \in \{1, \dots, N\}$. Als direkte Konsequenz daraus stimmen C und A_h auf dem Gitter

$$\left\{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}, 1\right\}^2 \subseteq [0, 1]^2$$

überein. Da wir für jeden Punkt $(x, y) \in [0, 1]^2$ Indices $i, j \in \{1, \dots, N\}$ finden können, sodass $|x - \frac{i}{N}| < \frac{1}{N}$ und $|y - \frac{j}{N}| < \frac{1}{N}$ gilt, folgt insgesamt

$$\begin{aligned} |C(x, y) - A_h(x, y)| &\leq |C(x, y) - C(\frac{i}{N}, \frac{j}{N})| + |C(\frac{i}{N}, \frac{j}{N}) - A_h(\frac{i}{N}, \frac{j}{N})| \\ &\quad + |A_h(\frac{i}{N}, \frac{j}{N}) - A_h(x, y)| \\ &< \frac{2}{N} + 0 + \frac{2}{N} = \frac{4}{N} \leq \varepsilon, \end{aligned}$$

und der Beweis ist komplett. ■

Um Metriken zu konstruieren, die in der Lage sind, vollständige Abhängigkeit und Unabhängigkeit klar zu trennen, kann wie folgt vorgegangen werden (alternative Konstruktionen sind in [5, 16] skizziert): Für jedes $p \in [1, \infty)$ definieren wir D_p durch

$$D_p^p(A, B) := \int_{[0,1]} \int_{[0,1]} |K_A(x, [0, y]) - K_B(x, [0, y])|^p d\lambda(x) d\lambda(y), \quad (2.13)$$

für $p = \infty$ setzen wir

$$D_\infty(A, B) := \sup_{y \in [0,1]} \int_{[0,1]} |K_A(x, [0, y]) - K_B(x, [0, y])| d\lambda(x). \quad (2.14)$$

Das folgende Resultat fasst die wichtigsten Eigenschaften der Metriken D_p zusammen:

Satz 2.26 *Für jedes $p \in [1, \infty]$ ist D_p eine Metrik auf \mathcal{C} . Die resultierenden metrischen Räume (\mathcal{C}, D_p) sind vollständig und separabel. Weiters gilt*

$$D_p^p(A, B) \leq D_1(A, B) \leq D_p(A, B) \quad (2.15)$$

$$d_\infty(A, B) \leq D_\infty(A, B) \leq 2\sqrt{D_1(A, B)} \quad (2.16)$$

für alle $A, B \in \mathcal{C}$.

Beweis: Die behaupteten Eigenschaften sind für den Fall $p = 1$ in [16] bewiesen. Ungleichung 2.15 ist eine direkte Konsequenz der Hölderschen Ungleichung, Ungleichung 2.15 ist ebenfalls in [16] enthalten. Mit Hilfe der Ungleichungen folgen die behaupteten Eigenschaften des metrischen Raums (\mathcal{C}, D_p) direkt aus den entsprechenden Eigenschaften von (\mathcal{C}, D_1) . Der handliche Spezialfall $p = 2$ wird in den Übungen ausführlich behandelt. ■

Satz 2.27 Für jedes $p \in [1, \infty)$ gilt folgende Aussage: $D_p(A, \Pi)$ ist maximal genau dann wenn $A \in \mathcal{C}$.

Beweis: Mittels leichter Modifikation des Beweises für den Fall $p = 1$ in [16]. Wir beweisen den elegant beweisbaren Spezialfall $p = 2$:

Sei $A \in \mathcal{C}$ und $y \in [0, 1]$. Wir definieren die Zufallsvariable Z_y auf dem Wahrscheinlichkeitsraum $([0, 1], \mathcal{B}([0, 1]), \lambda) \rightarrow [0, 1]$ durch

$$Z_y(x) := K_A(x, [0, y])$$

und erhalten

$$\mathbb{E}(Z_y) = \int_{[0,1]} K_A(x, [0, y]) d\lambda(x) = \mu_A([0, 1] \times [0, y]) = y.$$

Wegen

$$\begin{aligned} \int_{[0,1]} (K_A(x, [0, y]) - y)^2 d\lambda(x) &= \mathbb{V}(Z_y) = \mathbb{E}(Z_y^2) - (\mathbb{E}(Z_y))^2 = \mathbb{E}(Z_y^2) - y^2 \\ &\leq \mathbb{E}(Z_y) - y^2 = y - y^2 \end{aligned} \quad (2.17)$$

ergibt sich $D_2^2(A, \Pi) \leq \int_{[0,1]} (y - y^2) d\lambda(y) = \frac{1}{6}$.

Nachdem Ungleichung (2.17) genau dann zur Gleichung wird, wenn $\mathbb{E}(Z_y^2) = \mathbb{E}(Z_y)$ gilt, folgt sofort, dass $D_2^2(A, \Pi) = \frac{1}{6}$ genau dann gilt, wenn $Z_y^2(x) = Z_y(x)$ für λ -fast alle $x \in [0, 1]$. Letzteres ist offensichtlich äquivalent zur Existenz einer Menge $\Lambda_y \in \mathcal{B}([0, 1])$ mit $\lambda(\Lambda_y) = 1$, sodass $Z_y(x) = K_A(x, [0, y]) \in \{0, 1\}$ für jedes $x \in \Lambda_y$.

Gelte nun $D_2^2(A, \Pi) = 1/6$. Wiederholung des vorhergehenden Arguments liefert eine Menge $\Lambda \in \mathcal{B}([0, 1])$ mit $\lambda(\Lambda) = 1$ sodass $F_x^A(y) = K_A(x, [0, y]) \in \{0, 1\}$ für jedes $x \in \Lambda$ und jedes $y \in \mathbb{Q} \cap [0, 1]$. Unter Verwendung der Rechtsstetigkeit von Verteilungsfunktionen erhalten wir sofort, dass λ -fast jede bedingte Verteilungsfunktion F_x^A nur $\{0, 1\}$ -wertig, und damit die Copula A vollständig abhängig ist. Nachdem andererseits die Gleichheit $D_2^2(A_h, \Pi) = \frac{1}{6}$ für jedes $h \in \mathcal{T}$ leicht nachgerechnet werden kann (siehe Übungen), ist der Beweis komplett. ■

Kapitel 3

Konkordanzmaße

Hauptziel dieses kurzen Abschnitts ist, die zwei wichtigsten Konkordanzmaße - Kendall's τ und Spearman's ρ - einzuführen und deren wichtigste Eigenschaften zu studieren. Wir starten mit Konkordanz von Punkten und erweitern dann auf Zufallsvariable.

Sei (X, Y) ein Zufallsvektor mit stetiger Verteilungsfunktion H und $(x_1, y_1), (x_2, y_2)$ eine Stichprobe von (X, Y) . Wir nennen $(x_1, y_1), (x_2, y_2)$ konkordant genau dann, wenn entweder $x_1 < x_2$ und $y_1 < y_2$ oder $x_1 > x_2$ und $y_1 > y_2$ gilt. Die Punkte $(x_1, y_1), (x_2, y_2)$ heißen diskordant genau dann wenn entweder $x_1 < x_2$ und $y_1 > y_2$ oder $x_1 > x_2$ und $y_1 < y_2$ gilt. Mit anderen Worten: $(x_1, y_1), (x_2, y_2)$ ist konkordant genau dann, wenn $(x_2 - x_1)(y_2 - y_1) > 0$ und diskordant wenn $(x_2 - x_1)(y_2 - y_1) < 0$.

Bemerkung 3.1 Beachten Sie, dass wegen der vorausgesetzten Stetigkeit von H das Paar $(x_1, y_1), (x_2, y_2)$ mit Wahrscheinlichkeit 1 entweder konkordant oder diskordant ist.

Intuitiv beschreibt Konkordanz also gemeinsames Wachsen, nicht notwendigerweise auf lineare Art und Weise (wie es die klassische Pearson Korrelation tut). Die Übertragung des Konzepts von Paaren von Stichproben auf die Verteilung von (X, Y) ist wenig überraschend:

Definition 3.2 (Kendall τ) Sei $(X_1, Y_1), (X_2, Y_2)$ eine Zufallsstichprobe des Zufallsvektors $(X, Y) \sim H$ mit H stetig. Dann ist Kendall's τ von (X, Y) definiert als die Wahrscheinlichkeit von Konkordanz minus die Wahrscheinlichkeit von Diskordanz, i.e.

$$\tau_{X,Y} = \mathbb{P}((X_2 - X_1)(Y_2 - Y_1) > 0) - \mathbb{P}((X_2 - X_1)(Y_2 - Y_1) < 0). \quad (3.1)$$

Kendall's τ hängt nur von der zugrundeliegenden Copula ab (wenig überraschend in Anbetracht der Tatsache, dass Konkordanz per definitionem skaleninvariant ist) - wir folgen [12] und zeigen ein etwas allgemeineres Resultat, das in der Folge eine wichtige Rolle spielen wird und zudem einige nützliche Nebenprodukte liefert.

Lemma 3.3 Gelte $(X_1, Y_1) \sim H_1, (X_2, Y_2) \sim H_2, X_1, X_2 \sim F$ sowie $Y_1, Y_2 \sim G$ mit stetigen eindimensionalen Verteilungsfunktionen F, G . Weiters seien (X_1, Y_1) und (X_2, Y_2) unabhängig. Setzen wir

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

dann folgt

$$Q = 4 \int_{[0,1]^2} A_2 d\mu_{A_1} - 1 = 4 \int_{[0,1]^2} A_1 d\mu_{A_2} - 1 \quad (3.2)$$

Beweis: Aus der Stetigkeit von F und G folgt sofort $\mathbb{P}((X_2 - X_1)(Y_2 - Y_1) = 0) = 0$, also

$$\mathbb{P}((X_2 - X_1)(Y_2 - Y_1) > 0) = 1 - \mathbb{P}((X_2 - X_1)(Y_2 - Y_1) < 0)$$

und damit

$$Q = 2\mathbb{P}((X_2 - X_1)(Y_2 - Y_1) > 0) - 1.$$

(i) Wir vereinfachen zuerst den Ausdruck $\mathbb{P}(X_1 > X_2, Y_1 > Y_2)$ und gehen wie folgt vor: Die allgemeine Version des Disintegrationstheorems (Satz 1.34) in Kombination mit dem Satz von Sklar liefert aufgrund der vorausgesetzten Unabhängigkeit von (X_1, Y_2) und (X_2, Y_1) sofort (warum genau?)

$$\begin{aligned} \mathbb{P}(X_1 > X_2, Y_1 > Y_2) &= \mathbb{P}(X_2 < X_1, Y_2 < Y_1) = \mathbb{P}(X_2 \leq X_1, Y_2 \leq Y_1) \\ &= \int_{\mathbb{R}^2} \mathbb{P}(X_2 \leq x_1, Y_2 \leq y_1) d\mathbb{P}^{(X_1, Y_1)}(x_1, y_1) \\ &= \int_{\mathbb{R}^2} A_2(F(x_1), G(y_1)) d\mathbb{P}^{(X_1, Y_1)}(x_1, y_1). \end{aligned}$$

Definieren wir $\Psi : \mathbb{R}^2 \rightarrow [0, 1]^2$ durch $\Psi(x, y) = (F(x), G(y))$, dann folgt für das Bildmaß von $\mathbb{P}^{(X_1, Y_1)}$ unter Ψ sofort

$$\left(\mathbb{P}^{(X_1, Y_1)}\right)^\Psi = \mathbb{P}^{\Psi \circ (X_1, Y_1)} = \mathbb{P}^{(F \circ X_1, G \circ Y_1)} = \mu_{A_1}.$$

Change of Coordinates liefert insgesamt

$$\begin{aligned} \mathbb{P}(X_1 > X_2, Y_1 > Y_2) &= \int_{\mathbb{R}^2} A_2(\Psi(x_1, y_1)) d\mathbb{P}^{(X_1, Y_1)}(x_1, y_1) \\ &= \int_{[0, 1]^2} A_2(u, v) d\left(\mathbb{P}^{(X_1, Y_1)}\right)^\Psi(u, v) = \int_{[0, 1]^2} A_2 d\mu_{A_1}. \end{aligned}$$

(ii) Der Ausdruck $\mathbb{P}(X_1 < X_2, Y_1 < Y_2)$ kann analog (abermals unter Verwendung der Funktion Ψ) vereinfacht werden:

$$\begin{aligned} \mathbb{P}(X_1 < X_2, Y_1 < Y_2) &= \mathbb{P}(X_2 > X_1, Y_2 > Y_1) = \int_{\mathbb{R}^2} \mathbb{P}(X_2 > x_1, Y_2 > y_1) d\mathbb{P}^{(X_1, Y_1)}(x_1, y_1) \\ &= \int_{\mathbb{R}^2} 1 - F(x_1) - G(y_1) + A_2(F(x_1), G(y_1)) d\mathbb{P}^{(X_1, Y_1)}(x_1, y_1) \\ &= \int_{[0, 1]^2} 1 - u - v + A_2(u, v) d\mu_{A_1}(u, v) \\ &= 1 - \frac{1}{2} - \frac{1}{2} + \int_{[0, 1]^2} A_2 d\mu_{A_1} = \int_{[0, 1]^2} A_2 d\mu_{A_1} \end{aligned}$$

Insgesamt erhalten wir daher wie gewünscht

$$Q = 2\mathbb{P}((X_2 - X_1)(Y_2 - Y_1) > 0) - 1 = 4 \int_{[0, 1]^2} A_2 d\mu_{A_1} - 1.$$

Hätten wir anstatt auf (X_1, Y_1) auf (X_2, Y_2) bedingt, wäre $4 \int_{[0, 1]^2} A_1 d\mu_{A_2} - 1$ das Ergebnis gewesen. Damit ist der Beweis komplett. ■

Bevor wir die Konsequenz von Lemma 3.3 für Kendall's τ zusammenfassen, formulieren wir das folgende wichtige Nebenprodukt des vorangehenden Beweises:

Folgerung 3.4 Für beliebige Copulas $A, B \in \mathcal{C}$ gilt die Gleichheit.

$$\int_{[0,1]^2} Ad\mu_B = \int_{[0,1]^2} Bd\mu_A$$

Satz 3.5 Sei (X, Y) ein Zufallsvektor mit stetiger Verteilungsfunktion H und Copula A . Dann kann Kendall's τ (unter alleiniger Verwendung von A) berechnet werden durch

$$\tau_{X,Y} = 4 \int_{[0,1]^2} Ad\mu_A - 1 \quad (3.3)$$

Statt $\tau_{X,Y}$ schreiben wir in der Folge oft auch einfach $\tau(A)$ (wenn A die (X, Y) zugrunde liegende Copula ist).

Definition 3.6 Eine Abbildung κ , die jedem Paar (X, Y) stetiger Zufallsvariable[†] eine reelle Zahl $\kappa_{X,Y}$ zuweist heißt Konkordanzmaß genau dann, wenn sie die folgenden sechs Bedingungen erfüllt:

1. $\kappa_{X,Y} \in [-1, 1]$, $\kappa_{X,X} = 1$, $\kappa_{X,-X} = -1$.
2. $\kappa_{X,Y} = \kappa_{Y,X}$.
3. Falls X und Y unabhängig sind folgt $\kappa_{X,Y} = 0$.
4. $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$.
5. Falls die Copulas A_1 und A_2 von (X_1, Y_1) und (X_2, Y_2) die Ungleichung $A_1 \leq A_2$ (punktweise) erfüllen, dann folgt auch $\kappa_{X_1, Y_1} \leq \kappa_{X_2, Y_2}$.
6. Seien $(X, Y), (X_1, Y_1), (X_2, Y_2), \dots$ Paare stetiger Zufallsvariable mit Copulas A, A_1, A_2, \dots . Dann folgt aus $\lim_{n \rightarrow \infty} d_\infty(A_n, A) = 0$ auch $\lim_{n \rightarrow \infty} \kappa_{X_n, Y_n} = \kappa_{X, Y}$.

Satz 3.7 Kendall's τ ist ein Konkordanzmaß.

Beweis: Übungsaufgabe

Sei $A \in \mathcal{C}$ beliebig. Für jedes $t \in [0, 1]$ definieren wir den (unteren) t -Schnitt $[A]^t$ von A durch

$$[A]^t = \{(x, y) \in [0, 1]^2 : A(x, y) \leq t\}. \quad (3.4)$$

Dann ist $[A]^t$ für jedes $t \in [0, 1]$ eine nichtleere, kompakte Teilmenge von $[0, 1]^2$.

Definition 3.8 (Kendall distribution function) Für $A \in \mathcal{C}$ heißt die durch

$$F_K^A(t) := \mu_A([A]^t)$$

auf $[0, 1]$ definierte, $[0, 1]$ -wertige Funktion die Kendall Verteilungsfunktion der Copula A .

[†]Wir nennen eine Zufallsvariable (der Literatur folgend) stetig, wenn ihre Verteilungsfunktion stetig ist.

Satz 3.9 Für jede Copula A gilt der folgende Zusammenhang zwischen Kendall's τ und der Kendall Verteilungsfunktion F_K^A :

$$\tau(A) = 3 - 4 \int_{[0,1]} F_K^A(t) d\lambda(t)$$

Ein alternative Quantifizierung von Konkordanz ist Spearman's ρ (a.k.a. Spearman rank correlation):

Definition 3.10 (Spearman's ρ) Sei $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ eine Zufallsstichprobe des Zufallsvektors $(X, Y) \sim H$ mit stetigem H . Dann ist Spearman's ρ von (X, Y) definiert durch

$$\rho_{X,Y} = 3(\mathbb{P}((X_2 - X_1)(Y_3 - Y_1) > 0) - \mathbb{P}((X_2 - X_1)(Y_3 - Y_1) < 0)). \quad (3.5)$$

Wegen der Unabhängigkeit von X_2 und Y_3 liefert Lemma 3.3 sofort das folgende Resultat:

Satz 3.11 Sei (X, Y) ein Zufallsvektor mit stetiger Verteilungsfunktion H und Copula A . Dann kann Spearman's ρ (unter alleiniger Verwendung von A) berechnet werden durch

$$\rho_{X,Y} = 12 \int_{[0,1]^2} \Pi d\mu_A - 3 = 12 \int_{[0,1]^2} A d\lambda_2 - 3 \quad (3.6)$$

Statt $\rho_{X,Y}$ schreiben wir in der Folge oft auch einfach $\rho(A)$ (wenn A die (X, Y) zugrunde liegende Copula ist).

Beweis: Nachdem (X_1, Y_1) Copula A und (X_2, Y_3) Copula Π hat, liefert Lemma 3.3

$$\rho_{X,Y} = 3 \left(4 \int_{[0,1]^2} \Pi d\mu_A - 1 \right) = 3 \left(4 \int_{[0,1]^2} A d\lambda_2 - 1 \right). \blacksquare$$

Wie vorhin erwähnt, wird Spearman's ρ in der Literatur auch oft als Spearman rank correlation bezeichnet. Der Grund dahinter ist wie folgt: Betrachten wir die klassische Pearson Korrelation der auf $[0, 1]$ stetig gleichverteilten Zufallsvariable $F \circ X$ und $G \circ Y$, und verwenden, dass für $Z \sim \mathcal{U}(0, 1)$

$$\mathbb{E}(Z) = \frac{1}{2}, \quad \mathbb{V}(Z) = \frac{1}{12}$$

gilt, dann folgt mittels Change of Coordinates sofort

$$\begin{aligned} \text{Cor}(F \circ X, G \circ Y) &= \frac{\mathbb{E}(F \circ X \cdot G \circ Y) - \frac{1}{4}}{\frac{1}{12}} = 12 \int_{\Omega} F \circ X \cdot G \circ Y d\mathbb{P} - 3 \\ &= 12 \int_{[0,1]^1} uv d\mathbb{P}^{(F \circ X, G \circ Y)}(u, v) - 3 = 12 \int_{[0,1]^1} uv d\mu_A(u, v) - 3 \\ &= \rho_{X,Y} \end{aligned}$$

Wir haben also folgendes Resultat bewiesen:

Satz 3.12 Sei (X, Y) ein Zufallsvektor mit stetiger Verteilungsfunktion H , Copula A und Randverteilungen F, G . Dann ist Spearman's ρ die klassische Pearson Korrelation der $\mathcal{U}(0, 1)$ -verteilten Zufallsvariable $F \circ X, G \circ Y$, i.e.

$$\rho_{X,Y} = \text{Cor}(F \circ X, G \circ Y). \quad (3.7)$$

Satz 3.13 *Spearman's ρ ist ein Konkordanzmaß.*

Beweis: Übungsaufgabe

Wir wissen schon, dass Kendall's τ und Spearman's ρ Konkordanzmaße sind und daher für die Copulas Π, W, M die selben Werte liefern. Nachdem sie im Allgemeinen aber natürlich nicht übereinstimmen, ist die naheliegende Frage, wie stark sich die beiden Konkordanzmaße tatsächlich unterscheiden können, i.e. wie groß $|\rho(A) - \tau(A)|$ werden kann. Anfang der 1950er Jahren wurden die folgenden beiden Ungleichungen bewiesen - die erste geht zurück auf Daniels [1], die zweite auf Durbin und Stuart [4]:

$$|3\tau - 2\rho| \leq 1 \tag{3.8}$$

$$\frac{(1 + \tau)^2}{2} - 1 \leq \rho \leq 1 - \frac{(1 - \tau)^2}{2} \tag{3.9}$$

Daniel's Ungleichung ist scharf, von der zweiten Ungleichung war jedoch bis 2016 unklar, ob sie verbessert werden kann oder nicht - was bekannt war, ist die Existenz abzählbar viele Shuffles, für die die Ungleichung scharf ist. Abbildung 3.1 zeigt einige dieser Shuffles (in blau). Diese Shuffles (in [15] als Prototypen bezeichnet) bestehen aus n gleich langen, absteigenden

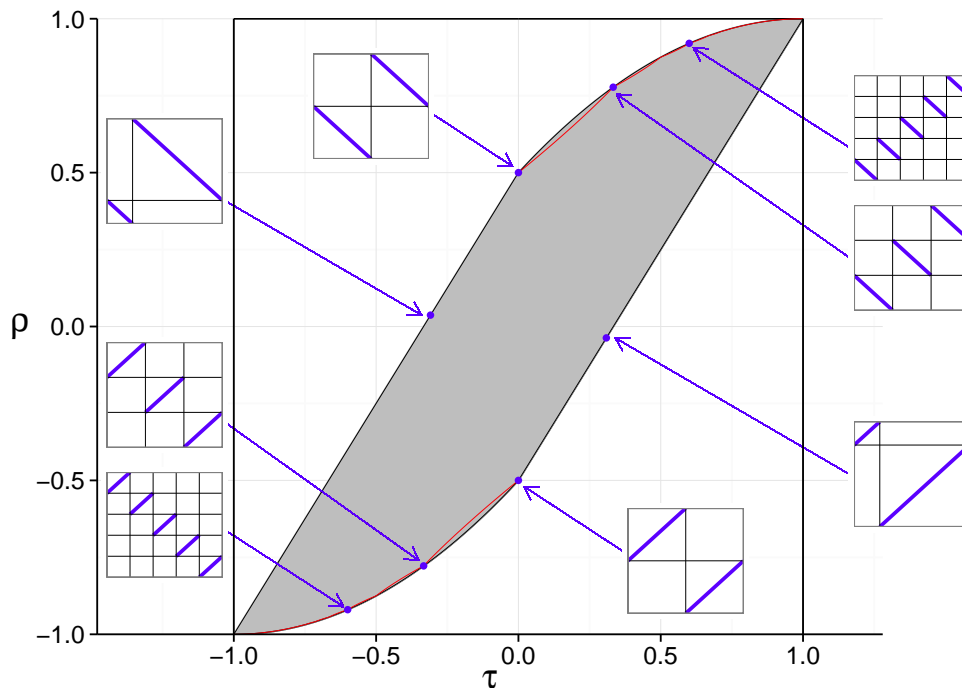


Abbildung 3.1: Der grau scharffierte Bereich ist die durch die beiden klassischen Ungleichungen von Daniels und Durbin & Stuart bestimmte (kompakte, konvexe) Menge. Die blauen Punkte am Rand der Menge sind (einige) jener Punkte, in denen die parabolische Ungleichung scharf ist.

(aufsteigenden) Stücken. Die Vermutung liegt daher nahe[†], dass alle Shuffles bestehend aus $n \geq 2$ gleich langen und einem kürzeren Stück die untere Grenze der τ - ρ -Region Ω , definiert durch

$$\Omega = \{(\tau(A), \rho(A)) : A \in \mathcal{C}\}, \quad (3.10)$$

bilden. Definieren wir für $n \in \mathbb{N}$ die Funktion $\Phi_n : [-1 + \frac{2}{n}, 1] \rightarrow [-1, 1]$ durch

$$\Phi_n(x) = -1 - \frac{4}{n^2} + \frac{3}{n} + \frac{3x}{n} - \frac{n-2}{\sqrt{2}n^2\sqrt{n-1}}(n-2+nx)^{3/2} \quad (3.11)$$

und setzen

$$\Phi(x) = \begin{cases} -1 & \text{if } x = -1, \\ \Phi_n(x) & \text{if } x \in \left[\frac{2-n}{n}, \frac{2-(n-1)}{n-1}\right] \text{ for some } n \geq 2. \end{cases} \quad (3.12)$$

dann gilt folgendes Resultat (die untere rote Linie in Abbildung 3.1 ist der Graph der Funktion Φ , eine animierte Darstellung finden Sie unter www.trutschnig.net/tau-rho-boundary.pdf):

Satz 3.14 ([15]) *Die τ - ρ -Region Ω ist vollständig durch die Funktion Φ bestimmt - es gilt*

$$\Omega = \{(x, y) \in [-1, 1]^2 : \Phi(x) \leq y \leq -\Phi(-x)\}, \quad (3.13)$$

Beweis: Nichttrivial und technisch aufwändig; im Wesentlichen Beweis der Ungleichungen für Shuffles via Induktion, gefolgt von einem Homotopieargument, siehe [15].

Folgerung 3.15 ([15]) *Die Ungleichung von Durbin und Stuart ist nur in den schon seit den 1950er Jahren bekannten Punkten scharf. Die Menge Ω ist kompakt, aber nicht konvex.*

Folgerung 3.16 ([15]) *Für jeden Punkt $(x, y) \in \Omega$ existiert ein Shuffle $h : [0, 1] \rightarrow [0, 1]$ mit $(x, y) = (\tau(A_h), \rho(A_h))$.*

Bemerkung 3.17 Satz 3.14 liefert eine handliche Charakterisierung von Ω . Für wichtige Teilfamilien von Copulas (wie Archimendische oder Extreme-Value-Copulas) ist eine vollständige Charakterisierung der entsprechenden τ - ρ -Region noch unbekannt (und Inhalt aktueller Forschung).

[†]Was ist der natürliche Übergang von n auf $n + 1$ gleich lange Stücke?

Kapitel 4

Wichtige Copula Klassen

Bisher haben wir größtenteils theoretische Eigenschaften von Copulas studiert und relativ wenige konkrete Copulas kennengelernt. Inhalt des aktuellen Abschnitts sind daher empirische Copulas und einige wichtige Copula Familien, die immer wieder in der Literatur auftauchen.

4.1 Empirische Copulas

Sei (X, Y) ein Zufallsvektor mit stetiger Verteilungsfunktion H , Randverteilungen F, G und Copula $A \in \mathcal{C}$. Weiters sei $(x_1, y_1), \dots, (x_n, y_n)$ eine Stichprobe von (X, Y) . O.B.d.A[†] nehmen wir an, dass die x -Werte und die y -Werte der Stichprobe paarweise verschieden sind. Bezeichnen wir mit H_n die zweidimensionale, empirische Verteilungsfunktion und mit F_n, G_n die entsprechenden eindimensionalen empirischen Verteilungsfunktionen, dann liefert der Satz von Sklar die Existenz einer (nicht eindeutig bestimmten) Copula $A_n \in \mathcal{C}$ mit

$$H_n(x, y) = A_n(F_n(x), G_n(x)) \quad (4.1)$$

für alle $x, y \in \mathbb{R}$. Die Copula A_n ist nur auf $\text{Range}(F_n) \times \text{Range}(G_n) = \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}^2$ eindeutig bestimmt - jede Copula, die A_n vom Raster $R_n := \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}^2$ auf ganz $[0, 1]^2$ fortsetzt erfüllt offensichtlich Gleichung (4.1) und kann daher als *empirische Copula* bezeichnet werden. In der Literatur üblich sind 2 verschiedene Fortsetzungen basierend auf Π und M , die sogenannte bilineare Fortsetzung (a.k.a. empirische Π Copula) und die empirische M Copula. Vor der Diskussion der Fortsetzungen macht es Sinn, Gleichung (4.1) näher zu betrachten. Auf der linken Seite steht mit der empirischen Verteilungsfunktionen eine Funktion, die genau die Werte $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$ annimmt, die Abbildung

$$A_n : R_n \longrightarrow \left\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\right\}$$

ist also surjektiv. Weiters ist A_n wegen der 2-Monotonie von H_n auch 2-monoton auf dem Raster $R_n = \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}^2$. Unter Verwendung von

$$A_n\left(\frac{i}{n}, 1\right) - A_n\left(\frac{i-1}{n}, 1\right) = A_n\left(1, \frac{i}{n}\right) - A_n\left(1, \frac{i-1}{n}\right) = \frac{1}{n} \quad \text{für jedes } i \in \{1, \dots, n\}$$

[†]Die Aussage gilt wegen der vorausgesetzten Stetigkeit von H mit Wahrscheinlichkeit 1

und der zuvor erwähnten Surjektivität von A_n auf dem Raster R_n folgt sofort, dass A_n (aufgefasst als Wahrscheinlichkeitsmaß ϑ_{A_n} auf der Potenzmenge von R_n) folgende Eigenschaften hat: Für jedes $i \in \{1, \dots, n\}$ gilt

$$\vartheta_{A_n}(\{\frac{i}{n}\} \times \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}) = \frac{1}{n} = \vartheta_{A_n}(\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\} \times \{\frac{i}{n}\})$$

und für jedes $(i, j) \in R_n$ gilt $\vartheta_{i,j} := \vartheta_{A_n}(\{(\frac{i}{n}, \frac{j}{n})\}) \in \{0, \frac{1}{n}\}$. Als Konsequenz daraus erhalten wir, dass die Matrix $(n \cdot \vartheta_{i,j})_{i,j=1}^n$ eine Permutationsmatrix ist. Ergänzend überlegt man sich leicht (siehe Übungen) dass folgender Zusammenhang gilt:

$$\vartheta_{A_n} = \frac{1}{n} \sum_{i=1}^n \delta_{(F_n(x_i), G_n(y_i))} \tag{4.2}$$

Mit anderen Worten: Die Massenverteilung von A_n auf R_n entspricht genau der diskreten Gleichverteilung auf der Menge $\{(F_n(x_i), G_n(y_i)) : i \in \{1, \dots, n\}\}$. In der Literatur werden die Punkte $(u_i, v_i) := (F_n(x_i), G_n(y_i))$ meist als *Pseudoobservationen* (oder als normalisierte Ränge) bezeichnet. A_n ist also nichts anderes als die empirische Verteilungsfunktion der Pseudoobservationen $(u_1, v_1), \dots, (u_n, v_n)$ eingeschränkt auf R_n . Abbildung 4.1 zeigt eine Stichprobe der Größe $n = 8$ (links) sowie die entsprechenden Pseudoobservationen (rechts). Der rechte Teil von Grafik 4.1 deutet auch an, wie A_n auf ganz $[0, 1]^2$ fortgesetzt werden

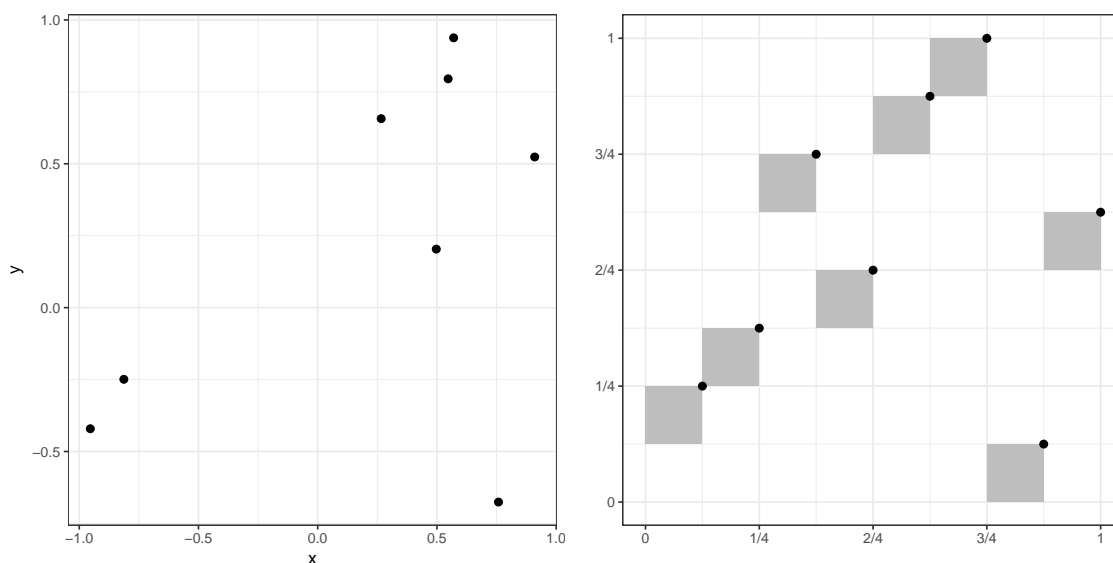


Abbildung 4.1: Stichprobe der Größe $n = 8$ (linker Panel), Pseudoobservationen und Massenverteilung der empirischen Π Copula (rechter Panel).

kann: Man wählt eine beliebige Copula $B \in \mathcal{C}$, definiert für jede Pseudoobservation (u_i, v_i) die Transformation $w_i : [0, 1]^2 \rightarrow (u_i - \frac{1}{n}, u_i] \times (v_i - \frac{1}{n}, v_i]$ durch

$$w_i(x, y) = (u_i - \frac{1}{n} + \frac{1}{n}x, v_i - \frac{1}{n} + \frac{1}{n}y),$$

und setzt schließlich

$$\mu_{A_n^B} := \frac{1}{n} \sum_{i=1}^n \mu_B^{w_i} \tag{4.3}$$

Mit anderen Worten: Die Massenverteilung von $A_n^B \in \mathcal{C}$ besteht aus n (um den Faktor $\frac{1}{n}$) geschrumpften Kopien von B . Wir nennen A_n^B die empirische B Copula. Wählen wir $B = \Pi$, dann heißt die entsprechende Copula A_n^Π konsequenterweise die empirische Π Copula (a.k.a. multilineare empirische Copula), wählen wir M , dann erhalten wir die anfangs erwähnte empirische M Copula. Beachten Sie, dass letztere ein (gerader) Shuffle ist.

Für $n \rightarrow \infty$ konvergiert die empirische Copula A_n^B gegen die zugrundeliegende Copula - es gilt folgendes Resultat:

Satz 4.1 *Sei $(X_1, Y_1), (X_2, Y_2), \dots$ eine Zufallsstichprobe von (X, Y) , wobei (X, Y) stetige Verteilungsfunktion H und Copula A habe. Weiters sei $B \in \mathcal{C}$ beliebig und A_n^B bezeichne die empirische B Copula der ersten n Elemente der Zufallsstichprobe. Dann gilt die folgende Aussage mit Wahrscheinlichkeit 1:*

$$d_\infty(A_n^B, A) = O\left(\sqrt{\frac{\log(\log(n))}{n}}\right) \quad (4.4)$$

Eine analoge Konsistenzaussage für die empirische Copula A_n^B kann für die Metrik D_1 nicht gelten. Beispielsweise lässt sich für den Fall $A = \Pi$ unschwer die folgende Aussage beweisen:

Satz 4.2 *Angenommen, es gelten die selben Voraussetzungen wie von Satz 4.1 und sowie $A = \Pi$. Dann gilt die folgende Aussage mit Wahrscheinlichkeit 1:*

$$\lim_{n \rightarrow \infty} D_1(A_n^B, \Pi) = \frac{1}{3}$$

Satz 4.2 impliziert insbesondere, dass empirische Copulas nicht ohne Weiteres für die Schätzung des Abhängigkeitsmaßes $D_1(A, \Pi)$ mittels $D_1(A_n^B, \Pi)$ verwendet werden können. Ein möglicher und gangbarer Ausweg ist, die empirischen Copulas zuerst zu glätten bzw. zu aggregieren, was (u.a.) zu sog. Bernstein bzw. Checkerboard Copulas führt. Wir werden in der Folge nur mit Checkerboard Copulas arbeiten, da selbige sowohl aus analytischer Sicht als auch implementationstechnisch einfach zu handhaben sind.

4.2 Checkerboard Copulas

Sei $N \in \mathbb{N}$ und die Quadrate $R_{i,j}^N \subseteq [0, 1]^2$ für $i, j \in \{1, \dots, N\}$ definiert durch

$$R_{i,j}^N = \left[\frac{i-1}{N}, \frac{i}{N}\right] \times \left[\frac{j-1}{N}, \frac{j}{N}\right].$$

Weiters bezeichne $w_{i,j}$ die affine Schrumpfung von $[0, 1]^2$ in $R_{i,j}^N$, i.e.

$$w_{i,j}^N(x, y) = \left(\frac{i-1}{N} + \frac{x}{N}, \frac{j-1}{N} + \frac{y}{N}\right)$$

Definition 4.3 *Eine Copula A heißt N -Checkerboard Copula, wenn sie absolut stetig ist, und die Dichte k_A von A konstant auf dem Inneren $\text{int}(R_{i,j}^N)$ jedes Quadrats $R_{i,j}^N$ ist. Das kleinstmögliche N mit dieser Eigenschaft heißt die Bandbreite von A . Die Menge aller N -Checkerboard Copulas (kurz N -Checkerboards) bezeichnen wir im Folgenden mit \mathcal{CB}_N und setzen $\mathcal{CB} = \bigcup_{N=1}^\infty \mathcal{CB}_N$.*

Definition 4.4 Für $A \in \mathcal{C}$ und $N \in \mathbb{N}$ heißt die (absolut stetige) Copula $\mathfrak{CB}_N(A) \in \mathcal{CB}_N$ mit Dichte

$$k_N(x, y) = N^2 \sum_{i,j=1}^N \mu_A(R_{i,j}^N) \mathbf{1}_{\text{int}(R_{i,j}^N)}(x, y)$$

die N -Checkerboard Π Approximation von A (kurz N -Checkerboard von A).

Wir wissen schon aus den Übungen, dass \mathcal{CB} dicht in (\mathcal{C}, D_1) (und damit auch dicht in (\mathcal{C}, D_p)) ist.

Checkerboard Copulas lassen sich alternativ als Konvexkombinationen von Push-Forwards von Π darstellen - es gilt das folgende einfache Resultat:

Lemma 4.5 Sei $A \in \mathcal{C}$ und $N \in \mathbb{N}$. Die Matrix $(t_{i,j}^N)_{i,j=1}^N \in [0, 1]^{N \times N}$ sei definiert durch $t_{i,j}^N = \mu_A(R_{i,j}^N)$. Dann gilt

$$\mu_{\mathfrak{CB}_N(A)} = \sum_{i,j=1}^N t_{i,j}^N \mu_{\Pi}^{w_{i,j}^N}. \quad (4.5)$$

Beweis: Einfach Übungsaufgabe

Bemerkung 4.6 In der Folge werden wir in der Notation die 'Bandbreite' N weglassen, wenn keine Missinterpretation möglich ist. Gleichung (4.5) vereinfacht sich dann zu

$$\mu_{\mathfrak{CB}_N(A)} = \sum_{i,j=1}^N t_{i,j} \mu_{\Pi}^{w_{i,j}}.$$

Es lässt sich unschwer zeigen (Übungsaufgabe), dass für jede Copula $A \in \mathcal{C}$ die Checkerboard Approximationen gleichmäßig gegen A konvergieren, i.e.

$$\lim_{N \rightarrow \infty} d_{\infty}(\mathfrak{CB}_N(A), A) = 0.$$

Es gilt aber eine wesentlich stärkere Aussage:

Satz 4.7 ([10]) Für jede Copula $A \in \mathcal{C}$ gilt

$$\lim_{N \rightarrow \infty} D_1(\mathfrak{CB}_N(A), A) = 0 = \lim_{N \rightarrow \infty} D_1(\mathfrak{CB}_N(A)^t, A^t). \quad (4.6)$$

Beweis: Der Beweis eines noch allgemeineren Ergebnisses (für sogenannte Partitionen der Eins) ist technisch relativ aufwändig und findet sich [10]. Der Spezialfall $N = 2^n$ erlaubt aber einen recht einfachen Beweis mittels Markov Kernen, der wie folgt funktioniert: Sei $A \in \mathcal{C}$ fest und $K_A(\cdot, \cdot)$ ein Markov Kern von A . Für jedes y der Form $y = \frac{j}{2^m}$ mit $m \in \mathbb{N}$ und $j \in \{0, \dots, 2^m\}$ existiert eine Menge $\Lambda_y \in \mathcal{B}([0, 1])$ mit folgenden drei Eigenschaften:

1. $\Lambda_y \subseteq \left(\bigcup_{l=1}^{\infty} \{0, \frac{1}{2^l}, \frac{2}{2^l}, \dots, \frac{2^l-1}{2^l}, 1\} \right)^c$.
2. In jedem Punkt $x \in \Lambda_y$ gilt: Die Funktion $t \mapsto A(t, y)$ ist differenzierbar in x und erfüllt $\frac{\partial A}{\partial x}(x, y) = K_A(x, [0, y])$.
3. $\lambda(\Lambda_y) = 1$.

Wenn wir die Menge aller Punkte y der Form $y = \frac{j}{2^m}$ mit \mathcal{Q} bezeichnen, dann ist \mathcal{Q} abzählbar und wir erhalten $\lambda(\Lambda) = 1$ für $\Lambda = \bigcap_{y \in \mathcal{Q}} \Lambda_y$.

Sei nun $y = \frac{j}{2^m} \in \mathcal{Q}$ beliebig, aber fest. Für jedes $x \in \Lambda$ und jedes $n \geq m$ existiert genau ein Index $i_n(x) \in \{0, \dots, 2^n\}$ mit

$$x \in \left(\frac{i_n(x) - 1}{2^n}, \frac{i_n(x)}{2^n} \right).$$

Nachdem der Markov Kern $K_{\mathfrak{CB}_{2^n}(A)}(t, \cdot)$ auf dem Intervall $\left(\frac{i_n(x) - 1}{2^n}, \frac{i_n(x)}{2^n} \right)$ konstant ist[†] folgt mittels Disintegration sofort

$$\begin{aligned} A\left(\frac{i_n(x)}{2^n}, y\right) - A\left(\frac{i_n(x) - 1}{2^n}, y\right) &= \int_{\left(\frac{i_n(x) - 1}{2^n}, \frac{i_n(x)}{2^n}\right]} K_{\mathfrak{CB}_{2^n}(A)}(t, [0, y]) d\lambda(t) \\ &= \frac{1}{2^n} K_{\mathfrak{CB}_{2^n}(A)}(x, [0, y]), \end{aligned}$$

und wir erhalten

$$K_{\mathfrak{CB}_{2^n}(A)}(x, [0, y]) = \frac{A\left(\frac{i_n(x)}{2^n}, y\right) - A\left(\frac{i_n(x) - 1}{2^n}, y\right)}{\frac{1}{2^n}} \xrightarrow{n \rightarrow \infty} \frac{\partial A}{\partial x}(x, y) = K_A(x, [0, y]).$$

Insgesamt folgt daher (Satz von der majorisierten Konvergenz)

$$\lim_{n \rightarrow \infty} \underbrace{\int_{[0,1]} |K_{\mathfrak{CB}_{2^n}(A)}(x, [0, y]) - K_A(x, [0, y])| d\lambda(x)}_{\Phi_{\mathfrak{CB}_{2^n}(A), A}(y)} = 0.$$

Nachdem die vorige Aussage für beliebiges $y \in \mathcal{Q}$ gilt, und für beliebige Copulas $B, C \in \mathcal{C}$ die Funktion $\Phi_{A,B}(y) = \int_{[0,1]} |K_A(x, [0, y]) - K_B(x, [0, y])| d\lambda(x)$ gemäß [16] Lipschitz stetig ist, folgt daraus sofort die behauptete Konvergenzaussage

$$\lim_{n \rightarrow \infty} D_1(\mathfrak{CB}_{2^n}(A), A) = 0. \blacksquare$$

4.3 Extreme-Value Copulas

In der Literatur sehr beliebt sind Familien von Copulas, die über eindimensionale Funktionen beschreibbar sind. Dazu zählen insbesondere Archimedische Copulas, Extreme-Value Copulas, und Diagonalcopulas. Wir betrachten zum Abschluss die zweite Klasse, da selbige direkt aus folgender naheliegenden Fragestellung entsteht: Sei (X, Y) ein Zufallsvektor mit stetiger Verteilungsfunktion H , Randverteilungen F, G und Copula $A \in \mathcal{C}$. Weiters sei $(X_1, Y_1), \dots, (X_n, Y_n)$ eine Zufallsstichprobe von (X, Y) . Wie sieht die Copula $A_{(n)}$ von $(X_{(n)}, Y_{(n)}) = (\max_{i=1}^n X_i, \max_{i=1}^n Y_i)$ aus?

Unter Verwendung von $F_{(n)}(x) = F(x)^n, G_{(n)}(y) = G(y)^n$ und

$$H_{(n)}(x, y) = \mathbb{P}(X_{(n)} \leq x, Y_{(n)} \leq y) = H(x, y)^n$$

[†]Genauer: für jedes $t \in \left(\frac{i_n(x) - 1}{2^n}, \frac{i_n(x)}{2^n} \right)$ ist das Wahrscheinlichkeitsmaß $K_{\mathfrak{CB}_{2^n}(A)}(t, \cdot)$ gleich. Beachten Sie, dass ein solcher Kern direkt über die Integration der Dichte nach y entsteht.

erhalten wir aus dem Satz von Sklar sofort

$$H_{(n)}(x, y) = H(x, y)^n = (A(F(x), G(y)))^n = A^n((F_{(n)}(x))^{1/n}, (G_{(n)}(y))^{1/n}).$$

Die Copula $A_{(n)}$ ist daher gegeben durch

$$A_{(n)}(x, y) = A^n(\sqrt[n]{x}, \sqrt[n]{y}). \quad (4.7)$$

Definition 4.8 Eine Copula $A \in \mathcal{C}$ heißt Extreme-Value Copula (EVC) genau dann, wenn eine Copula $C \in \mathcal{C}$ existiert, sodass für jedes $(x, y) \in [0, 1]^2$

$$A(x, y) = \lim_{n \rightarrow \infty} C^n(\sqrt[n]{x}, \sqrt[n]{y})$$

gilt. Falls der eben erwähnte Zusammenhang gilt, sagen wir, dass C zum Anzugsgebiet von A gehört.

Offensichtlich ist jede Copula A , die sogar $A(x, y) = A^n(\sqrt[n]{x}, \sqrt[n]{y})$ für jedes $(x, y) \in [0, 1]^2$ erfüllt, eine EVC. Eine scheinbar noch etwas restriktivere Eigenschaft ist die folgende:

Definition 4.9 Eine Copula $A \in \mathcal{C}$ heißt maximumsstabil genau dann, wenn sie für jedes $r \in (0, \infty)$ und jedes $(x, y) \in [0, 1]^2$ die folgende Gleichheit erfüllt:

$$A(x, y) = A^r(\sqrt[r]{x}, \sqrt[r]{y})$$

Satz 4.10 Eine Copula A ist eine genau dann eine EVC wenn sie maximumsstabil ist.

Beweis: Falls A maximumsstabil ist, dann ist A offensichtlich eine EVC. Ist umgekehrt A eine EVC, dann existiert per definitionem eine Copula C im Anzugsgebiet von A . Für $r = \frac{p}{q} \in (0, \infty)$ setzen wir $n_k = qk$ für jedes $k \in \mathbb{N}$ und erhalten für jedes $(x, y) \in [0, 1]^2$ sofort

$$A^r(\sqrt[r]{x}, \sqrt[r]{y}) = \lim_{k \rightarrow \infty} C^{n_k r}(\sqrt[n_k]{x}, \sqrt[n_k]{y}) = \lim_{k \rightarrow \infty} C^{pk}(\sqrt[pk]{x}, \sqrt[pk]{y}) = A(x, y).$$

Für $r \in (0, \infty) \setminus \mathbb{Q}$ folgt die gewünschte Gleichheit nun einfach mittels Stetigkeitsargumenten.

■

Wesentlich überraschender als Satz 4.10 ist das folgende nichttriviale Resultat ([3, 13]) - dabei wir nennen eine Funktion $a : [0, 1] \rightarrow [0, 1]$ eine *Pickands Dependence Function*, wenn sie die folgenden zwei Eigenschaften erfüllt:

1. a ist konvex.
2. Für alle $x \in [0, 1]$ gilt $\max\{x, 1 - x\} \leq a(x) \leq 1$.

Satz 4.11 Eine Copula A ist eine genau dann eine EVC, wenn eine Pickands Dependence Function a existiert, sodass

$$A(x, y) = (xy)^{a\left(\frac{\ln(x)}{\ln(xy)}\right)} \quad (4.8)$$

für alle $(x, y) \in [0, 1]^2$ gilt.

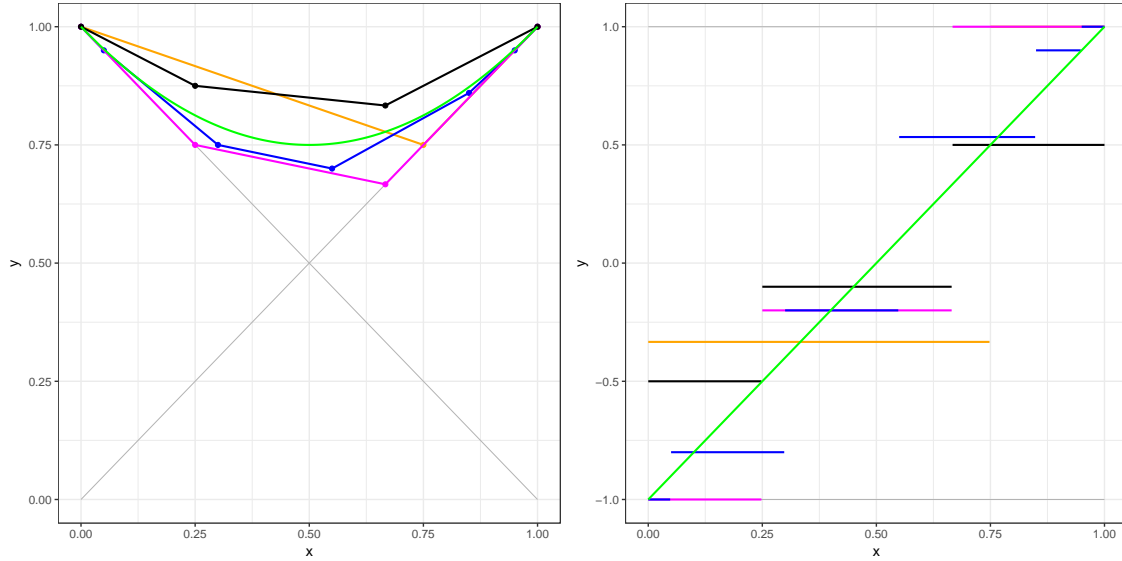


Abbildung 4.2: 5 Elemente von \mathcal{A} (linker Panel) und deren rechtsseitige Ableitungen (rechter Panel)

Im Folgenden bezeichnen wir mit \mathcal{A} die Menge aller Pickands Dependence Functions. Es lässt sich unschwer zeigen, dass $(\mathcal{A}, \|\cdot\|_\infty)$ ein kompakter metrischer Raum ist (direkte Folgerung des Satzes von Arzela-Ascoli, jedes a ist Lipschitz stetig mit Lipschitz-Konstante 1) und dass die rechtsseitige Ableitung D^+a von a rechtsseitig stetig und monoton wachsend auf $[0, 1)$ ist. D^+ kann unschwer zu einer maßerzeugenden Funktion auf ganz $[0, 1]$ fortgesetzt werden. Abbildung 4.2 zeigt 5 Beispiele von Pickands Dependence Functions und deren rechtsseitigen Ableitungen.

Die Menge \mathcal{A} hat sowohl ein maximales als auch ein minimales Element: Die Funktionen $a_M, a_\Pi \in \mathcal{A}$, definiert durch

$$a_M(x) = \max\{x, 1 - x\}, \quad a_\Pi(x) = 1$$

erfüllen offensichtlich $a_M \leq a \leq a_\Pi$. Als Konsequenz daraus erfüllt jede EVC die Ungleichung $\Pi \leq A \leq M$, woraus sich - unter Verwendung der Monotonie von Konkordanzmaßen - sofort folgendes Resultat ergibt:

Satz 4.12 Für jede EVC gilt $\tau(A) \geq 0$ und $\rho(A) \geq 0$.

Lt. [7] gelten die sogenannten Hutchinson-Lai Ungleichung, i.e.

$$-1 + \sqrt{1 + 3\tau(A)} \leq \rho(A) \leq \min \left\{ \frac{3\tau(A)}{2}, 2\tau(A) - \tau(A)^2 \right\} \quad (4.9)$$

für EVCs - der Originalbeweis ist allerdings nicht lückenlos (ein verwendetes Kompaktheitsargument ist falsch). Ausgehend von umfangreichen Simulationen liegt die Vermutung nahe, dass Ungleichung (4.9) nur für M und Π scharf ist, und dass für EVCs die folgenden, bisher unbewiesenen schärferen Ungleichungen gelten (siehe Abbildung 4.3):

$$\frac{3\tau(A)}{2 + \tau(A)} \leq \rho(A) \leq \frac{3\tau(A)}{2 + \tau(A)^3} - \frac{1}{3}(1 - \tau(A))^2\tau(A)^4 \quad (4.10)$$

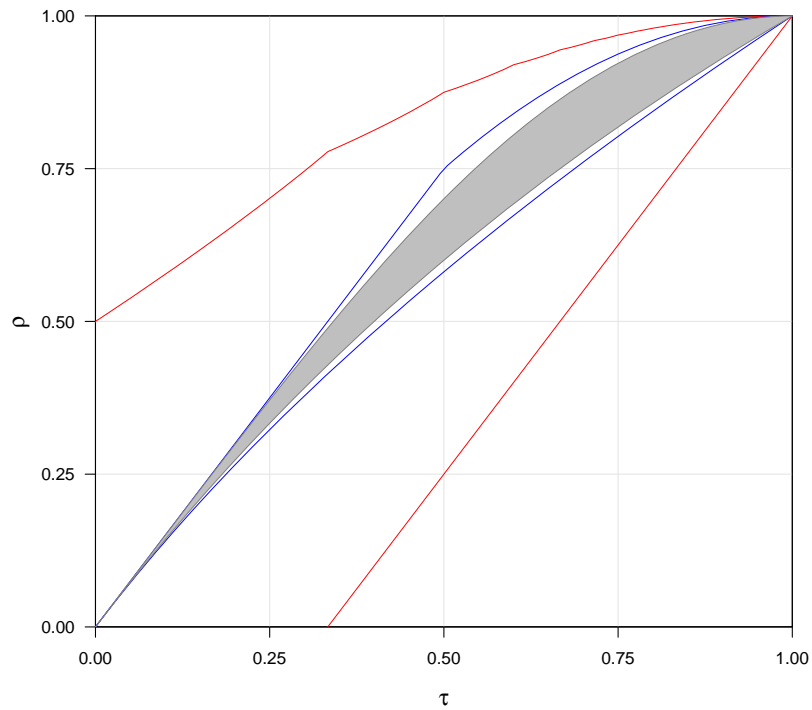


Abbildung 4.3: Die vermuteten Ungleichungen (4.10) begrenzen den grau schraffierten Bereich, die blauen Linien begrenzen die Hutchinson-Lai region, die roten Linien die im vorigen Kapitel beschriebene volle τ - ρ Region Ω gemäß [15].

Literatur

- [1] H. Daniels: Rank correlation and population models, *Journal of the Royal Statistical Society. Series B (Methodological)*, **12**, 171–191 (1950)
- [2] W.F. Darsow, B. Nguyen, E.T. Olsen: Copulas and Markov processes, *Illinois Journal of Mathematics* **36**, 600-642 (1992)
- [3] L. de Haan, S.I. Resnick: Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **40**, 317-337 (1977)
- [4] J. Durbin, A. Stuart: Inversions and rank correlation coefficients. *Journal of the Royal Statistical Society. Series B (Methodological)*, 303–309 (1951)
- [5] F. Durante, C. Sempi: *Principles of copula theory*, CRC Press Boca Raton (2016)
- [6] J. Elstrodt: *Maß- und Integrationstheorie*, Springer Verlag, Berlin Heidelberg New York, 1999
- [7] R. Hürlimann: Hutchinson-Lai’s conjecture for bivariate extreme value copulas, *Statistics and Probability Letters* **61**, 191-198 (2003)
- [8] O. Kallenberg: *Foundations of modern probability*, Springer New York (2002)
- [9] A. Klenke: *Probability Theory - a comprehensive course*, Springer, 2008
- [10] X. Li, P. Mikusinski, M.D. Taylor: Strong approximation of copulas, *Journal of Mathematical Analysis and Applications* **255**, 608-623 (1998)
- [11] P. Mikusinski, H. Sherwood, M.D. Taylor: Shuffles of Min, *Stochastica* **13**(1), 61-74 (1992)
- [12] R. B. Nelsen: *An introduction to copulas*, Springer New York (2006)
- [13] J. Pickands: Multivariate extreme value distributions, in: Proceedings 43rd Session International Statistical Institute **2**, 859-878 (1981)
- [14] W. Rudin: *Real and complex analysis*, McGraw-Hill New York (1987)
- [15] M. Schreyer, R. Paulin, W. Trutschnig: On the exact region determined by Kendall’s tau and Spearman’s rho, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **79** (2), 613-633 (2017)

-
- [16] W. Trutschnig: On a strong metric on the space of copulas and its induced dependence measure, *Journal of Mathematical Analysis and Applications* 384, 690-705 (2011)

Kapitel 5

Übungsblätter

01. Übung am 19. März 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 1 Beweisen Sie Lemma 1.6 (Aufwärmübung!)

Übungsaufgabe 2 Beweisen Sie die in Satz 1.7 behauptete Eindeutigkeit des Paares (ν_a, ν_s) (eine zweite Aufwärmübung!).

Übungsaufgabe 3 Beweisen Sie Satz 1.9.

Übungsaufgabe 4 Wir betrachten $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. X sei exponentialverteilt mit Parameter $\theta = 1$, Y sei stetig gleichverteilt auf $[0, 1]$. $\mu = \mathbb{P}^X$ und $\nu = \mathbb{P}^Y$ bezeichnen die entsprechenden Verteilungen. Bestimmen Sie die Lebesgue Zerlegung $\nu = \nu_a + \nu_s$ von ν bezüglich μ sowie die Lebesgue Zerlegung $\mu = \mu_a + \mu_s$ von μ bezüglich ν .

Übungsaufgabe 5 Die Definition der Absolut-Stetigkeit eines Maß ν bezüglich eines Maßes μ funktioniert auch für unendliche Maße. Wir betrachten die folgenden zwei Maße ν, μ auf $(\mathbb{R}, \mathcal{B}([0, 1]))$: μ sei das Zählmaß und ν sei das Lebesgue-Maß, i.e.

$$\mu(E) = \#E, \quad \nu(E) = \lambda(E).$$

für jedes $E \in \mathcal{B}([0, 1])$. Gilt dann $\nu \ll \mu$ (oder umgekehrt)? Existiert eine Radon-Nikodym Ableitung von ν bezüglich μ (oder umgekehrt)?

Übungsaufgabe 6 (just for fun*) † Die Folge $(a_i)_{i \in \mathbb{N}}$ konvergiere gegen $a \in \mathbb{R}$, $k \in \mathbb{N}$ sei beliebig. Berechnen Sie

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left(\frac{i}{n}\right)^k a_i$$

†eine meiner ehemaligen Prüfungsaufgaben in Maßtheorie, hat nichts mit Radon-Nikodym oder Lebesgue Zerlegung zu tun, ist aber eine gute Übung.

02. Übung am 09. April 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 7 Beweisen Sie Satz 1.10.

Übungsaufgabe 8 Berechnen Sie für $(\Omega, \mathcal{A}, \mathbb{P}) = ([0, 1], \mathcal{B}([0, 1]), \lambda)$, $X(\omega) = \omega^2$ und $\mathcal{C} = \mathcal{A}_\sigma(\{[0, 1/4], (1/4, 3/4], (3/4, 1]\})$ die bedingte Erwartung $\mathbb{E}(X|\mathcal{C})$ von X unter \mathcal{C} .

Übungsaufgabe 9 Beweisen Sie mindestens zwei der Aussagen 4-9 in Satz 1.16.

Übungsaufgabe 10 Wir betrachten den Wahrscheinlichkeitsraum $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathbb{P})$, wobei $\mathbb{P} = \mu \otimes \mu^\dagger$ und μ ein beliebiges Wahrscheinlichkeitsmaß auf $\mathcal{B}(\mathbb{R})$ bezeichne. Die Familie $\mathcal{B}_s(\mathbb{R}^2)$ aller symmetrischen Borelmengen ist definiert durch[†]

$$\mathcal{B}_s(\mathbb{R}^2) = \{B \in \mathcal{B}(\mathbb{R}^2) : B^t = B\}.$$

Zeigen Sie, dass $\mathcal{B}_s(\mathbb{R}^2)$ eine σ -Algebra ist, und berechnen Sie für eine beliebige (integrierbare) Zufallsvariable $X : \mathbb{R}^2 \rightarrow \mathbb{R}$ die bedingte Erwartung $\mathbb{E}(X|\mathcal{B}_s(\mathbb{R}^2))$ von X unter $\mathcal{B}_s(\mathbb{R}^2)$.

Übungsaufgabe 11 Seien X_1, X_2, \dots, X_n i.i.d. und integrierbar und S_n definiert durch $S_n = X_1 + X_2 + \dots + X_n$. Berechnen Sie $\mathbb{E}(X_1|S_n) := \mathbb{E}(X_1|\mathcal{A}_\sigma(S_n))$.[†]

Übungsaufgabe 12 (elementar aber wichtig) Seien X und Y Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Dann sind die folgenden beiden Bedingungen äquivalent:

1. Y ist $\mathcal{A}_\sigma(X)$ messbar.
2. Es existiert eine (Borel) messbare Transformation $h : \mathbb{R} \rightarrow \mathbb{R}$ mit $Y = h \circ X$.

Zusatz: Ist Y nicht-negativ, dann kann auch h nicht-negativ gewählt werden. Warum ist dieses einfache Resultat beispielsweise nützlich für Aufgabe 11?

[†] \mathbb{P} ist also das Produktmaß von μ mit sich selbst

[†] $B^t = \{(x, y) \in \mathbb{R}^2 : (y, x) \in B\}$

[†]Wie gewohnt, bezeichnet $\mathcal{A}_\sigma(Z) = Z^{-1}(\mathcal{B}(\mathbb{R}))$ die von Z erzeugte σ -Algebra.

03. Übung am 16. April 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutchnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 13 Zeigen Sie den in Beispiel 1.22 behaupteten bijektiven Zusammenhang zwischen Markov Kernen und stochastischen Matrizen (Aufwärmübung).

Übungsaufgabe 14 Beweisen Sie Lemma 1.27 (Aufwärmübung).

Übungsaufgabe 15 Sei (X, Y) absolut stetig mit Dichte

$$f(x, y) = (1 + \theta(1 - 2x)(1 - 2y))\mathbf{1}_{[0,1]^2}(x, y),$$

wobei $\theta \in [-1, 1]$. Berechnen Sie eine reguläre bedingte Verteilung $K_{Y,X}$ von Y unter X .

Übungsaufgabe 16 Sei $X \sim \mathcal{U}(0, 1)$ und $Y = T \circ X$ mit

$$T(x) = 2x\mathbf{1}_{[0,1/2]}(x) + 2(1-x)\mathbf{1}_{(1/2,1]}(x).$$

Berechnen Sie eine reguläre bedingte Verteilung $K_{Y,X}$ von Y unter X .

Übungsaufgabe 17 Seien X, Y integrierbare Zufallsvariable auf $(\Omega, \mathcal{A}, \mathbb{P})$. Eine (Borel-) messbare Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ heißt Version der bedingten Erwartung von Y gegeben X , genau dann wenn $\mathbb{E}(Y|X)(\omega) = g \circ X(\omega)$ für \mathbb{P} -fast-jedes $\omega \in \Omega$ gilt.

Beweisen Sie, dass g genau dann eine Version der bedingten Erwartung von Y gegeben X ist, wenn die folgende Gleichheit für jedes $B \in \mathcal{B}(\mathbb{R})$ gilt:

$$\int_B g(x) d\mathbb{P}^X(x) = \int_{X^{-1}(B)} Y d\mathbb{P}$$

Übungsaufgabe 18 Seien die Voraussetzungen des zweiten Teils von Definition 1.25 erfüllt und K eine reguläre bedingte Erwartung von Y unter $\mathcal{C} = \mathcal{A}_\sigma(X)$. Wie kann aus K eine reguläre bedingte Erwartung $K_{Y,X}$ von Y unter X konstruiert werden?

04. Übung am 23. April 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 19 (Fortsetzung von Aufgabe 16) Sei $X \sim \mathcal{U}(0, 1)$ und $Y = T \circ X$ mit

$$T(x) = 2x\mathbf{1}_{[0, 1/2]}(x) + 2(1-x)\mathbf{1}_{(1/2, 1]}(x).$$

Wir setzen $\mu = \mathbb{P}^{(X, Y)}$ sowie $\Gamma(T) := \{(x, y) \in [0, 1]^2 : y = T(x)\}$. Welche Verteilung hat Y ? Berechnen Sie weiters (mittels Disintegration oder direkt) $\mu(\Gamma(T))$ und $\mu([0, x] \times [0, y])$.

Übungsaufgabe 20 Sei (X, Y) stetig gleichverteilt am Einheitskreis $B(0, 1)$. Die Zufallsvariablen R und Φ seien definiert durch

$$R = \sqrt{X^2 + Y^2}, \quad \Phi = \arctan\left(\frac{Y}{X}\right).$$

Berechnen Sie eine reguläre bedingte Verteilung von Φ unter R sowie eine reguläre bedingte Verteilung von R unter Φ .

Übungsaufgabe 21 Beweisen Sie Satz 1.31.

Übungsaufgabe 22 Sei (X, Y) absolut stetig mit Dichte f . Konstruieren Sie eine reguläre bedingte Verteilung $K(\cdot, \cdot)$ von Y unter X , die ausschließlich von f und der Randdichte f_1 von X abhängt.

Übungsaufgabe 23 X und Y seien stetig gleichverteilt auf $[0, 1]$, A bezeichne die Verteilungsfunktion von (X, Y) , $K(\cdot, \cdot)$ bezeichne eine reguläre bedingte Verteilung von Y unter X . Überlegen Sie sich, wie $K(\cdot, \cdot)$ aus A bestimmt werden kann.

Hinweis: Nehmen Sie vereinfachend an, dass die Abbildung $(x, y) \mapsto \frac{\partial A}{\partial x}(x, y)$ stetig ist.

05. Übung am 30. April 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 24 Bestimmen Sie den Markov Kern und den Markov Operator für jede der in Beispiel 2.3 erwähnten Copulas (Aufwärmübung 1).

Übungsaufgabe 25 Beweisen Sie Lemma 2.4 (Aufwärmübung 2).

Übungsaufgabe 26 Sei $h : [0, 1] \rightarrow [0, 1]$ λ -treu. Zeigen Sie, dass es genau eine Copula $A \in \mathcal{C}$ mit $\mu_A(\Gamma(h)) = 1$ gibt, wobei $\Gamma(h)$, wie gewohnt, den Graphen der Funktion h bezeichnet (Aufwärmübung 3).

Übungsaufgabe 27 Sei $(h_n)_{n \in \mathbb{N}}$ eine Folge λ -treuer Transformationen auf $[0, 1]$, die in $L^1([0, 1], \mathcal{B}([0, 1]), \lambda)$ gegen h konvergiert, $(A_{h_n})_{n \in \mathbb{N}}$ bezeichne die entsprechende Folge von Copulas. Zeigen Sie, dass h ebenfalls λ -treu ist und dass für alle $x, y \in [0, 1]$

$$\lim_{n \rightarrow \infty} A_{h_n}(x, y) = A_h(x, y)$$

gilt. L^1 -Konvergenz der Transformationen impliziert also Konvergenz der Copulas.

Übungsaufgabe 28 Zeigen Sie, dass eine Folge $(A_n)_{n \in \mathbb{N}}$ von Copulas genau dann punktweise gegen eine Copula A konvergiert, wenn sie gleichmäßig gegen A konvergiert.

Übungsaufgabe 29 Wir betrachten die absolut stetige Copula aus Beispiel 2.10 und den Fall $\theta = 1$. Überlegen Sie sich, wie mit Hilfe der bedingten Verteilungsfunktionen $y \mapsto K_A(x, [0, y])$ Stichproben von $(X, Y) \sim A$ erzeugt werden können. Implementieren Sie Ihre Methode in R.

06. Übung am 07. Mai 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 30 Gelte $(X, Y) \sim H$, $X \sim F, Y \sim G$ mit F, G stetig. A bezeichne die eindeutig bestimmte Copula von (X, Y) , T und S seien strikt wachsende Transformationen auf \mathbb{R} . Zeigen Sie, dass A auch die Copula von $(T \circ X, S \circ Y)$ ist.

Übungsaufgabe 31 Beweisen Sie die in Lemma 2.15 behauptete Eigenschaft $X = F^{-1}(U)$ [\mathbb{P}].

Übungsaufgabe 32 Gelte $(X, Y) \sim H$ mit

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

Berechnen Sie die (eindeutig bestimmte) Copula von (X, Y) .

Übungsaufgabe 33 Sei X_1, X_2, \dots, X_n i.i.d. mit stetiger Verteilungsfunktion F . Berechnen Sie die Copula A_n von $X_{(1)}, X_{(n)}$ mit $X_{(1)} = \min\{X_1, \dots, X_n\}$ und $X_{(n)} = \max\{X_1, \dots, X_n\}$.

Übungsaufgabe 34 Wir betrachten als Spezialfälle von Beispiel 2.11 die λ -treuen Transformationen h_n in Abbildung 5 (h_n besteht auch genau n^2 Segmenten gleicher Länge). A_{h_n} bezeichne die entsprechende Copula. Gegen welche Copula konvergiert die Folge $(A_n)_{n \in \mathbb{N}}$ (punktweise)?

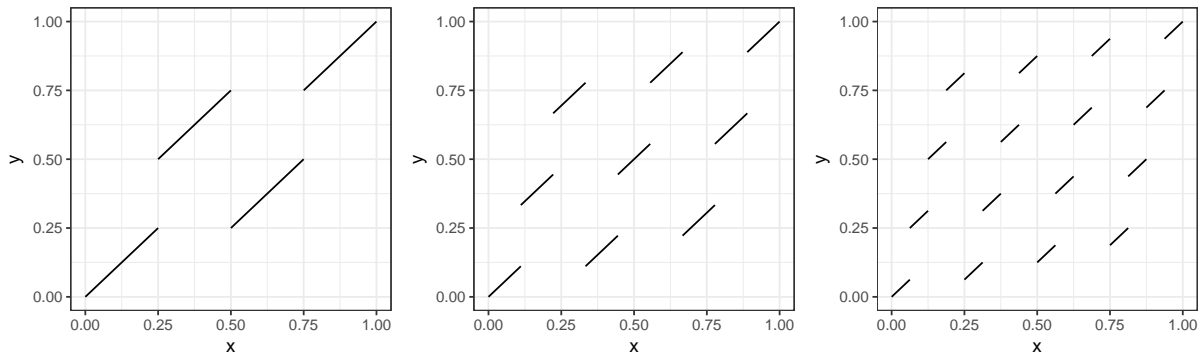


Abbildung 5.1: Die Fälle $n = 2, 3, 4$

Übungsaufgabe 35 (Fortsetzung von Aufgabe 33) Gegen welche Copula A konvergiert die Folge $(A_n)_{n \in \mathbb{N}}$ aus Beispiel 33?

07. Übung am 14. Mai 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 36 Beweisen Sie, dass D_p für jedes $p \in [1, \infty)$ eine Metrik ist (Aufwärmübung).

Übungsaufgabe 37 Beweisen Sie Lemma 2.23 (Aufwärmübung).

Übungsaufgabe 38 Seien $h_1, h_2 : [0, 1] \rightarrow [0, 1]$ λ -treu, $A_{h_1}, A_{h_2} \in \mathcal{C}_d$ bezeichne die entsprechenden vollständig abhängigen Copulas. Berechnen Sie $D_p(A_{h_1}, A_{h_2})$ für jedes $p \in [1, \infty)$.

Übungsaufgabe 39 Beweisen Sie Ungleichung 2.16 und folgern Sie daraus die zweite Aussage von Übungsaufgabe 27.

Übungsaufgabe 40 Berechnen Sie $D_p(A_h, \Pi)$ für jedes λ -treue $h : [0, 1] \rightarrow [0, 1]$ und jedes $p \in [1, \infty)$. Was fällt auf?

Übungsaufgabe 41 Verwenden Sie die vorhergehende Aufgabe um zu zeigen, dass die Metriken $(D_p)_{p \in [1, \infty)}$ nicht äquivalent sind[†]. Es reicht die Existenz von $p_1 \neq p_2$ zu zeigen, für die D_{p_1} und D_{p_2} nicht äquivalent sind.

[†]Zwei Metriken d_1, d_2 auf Ω heißen bekanntlich äquivalent genau dann, wenn Konstanten $a, b > 0$ existieren, sodass $ad_1(x, y) \leq d_2(x, y) \leq bd_1(x, y)$ für alle $x, y \in \Omega$ gilt

08. Übung am 28. Mai 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 42 Beweisen Sie, dass für jede Copula $A \in \mathcal{C}$ die folgende Ungleichung auf ganz $[0, 1]^2$ gilt: $W(x, y) \leq A(x, y) \leq M(x, y)$ (Aufwärmübung).

Übungsaufgabe 43 Sei A eine Copula. Angenommen, das entsprechende doppelt stochastische Maß μ_A ist absolut stetig (bezüglich λ_2) mit Dichte f . Welche Eigenschaften (zusätzlich zu $f \geq 0$ und $\int_{\mathbb{R}^2} f d\lambda_2 = 1$) erfüllt f ? (Aufwärmübung)

Übungsaufgabe 44 Seien $f, f_1, f_2, \dots [0, 1]^2 \rightarrow [0, \infty)$ Dichten der absolut stetigen Copulas A, A_1, A_2, \dots [†] und es gelte $\lim_{n \rightarrow \infty} f_n = f [\lambda_2]$. Folgt dann auch $\lim_{n \rightarrow \infty} D_1(A_n, A) = 0$ bzw. $\lim_{n \rightarrow \infty} d_\infty(A_n, A) = 0$?

Freiwilliger Zusatz*: Bleiben die Resultate erhalten, wenn die Dichten stattdessen im L^1 konvergieren?

Übungsaufgabe 45 Sei $N \in \mathbb{N}$ und die Quadrate $R_{i,j}^N \subseteq [0, 1]^2$ für $i, j \in \{1, \dots, N\}$ definiert durch

$$R_{i,j}^N = \left[\frac{i-1}{N}, \frac{i}{N} \right] \times \left[\frac{j-1}{N}, \frac{j}{N} \right].$$

Eine Copula A heißt *Checkerboard Copula*, wenn sie absolut stetig ist und ein $N \in \mathbb{N}$ existiert, sodass die Dichte k_A von A konstant auf dem Inneren $\text{int}(R_{i,j}^N)$ jedes Quadrats $R_{i,j}^N$ ist.

Beweisen Sie, dass die Menge aller Checkerboard Copulas dicht in (\mathcal{C}, D_1) ist.

Hinweis: Definieren Sie für $A \in \mathcal{C}$ die Dichte k_{B_n} der Checkerboard Approximation B_n von A durch

$$k_n(x, y) = N^2 \sum_{i,j=1}^N \mu_A(R_{i,j}^N) \mathbf{1}_{\text{int}(R_{i,j}^N)}(x, y)$$

Übungsaufgabe 46 Arbeiten Sie mit Checkerboards um zu zeigen, dass (\mathcal{C}, D_1) separabel ist.

Übungsaufgabe 47 Ein Punkt x einer konvexen, kompakten Menge F heißt *Extremalpunkt*, genau dann, wenn er nicht Mittelpunkt zweier von x verschiedener Punkte $y, z \in F$ ist.

Beweisen Sie, dass jede vollständig abhängige Copula A_h Extremalpunkt in \mathcal{C} ist und schließen Sie daraus, dass in (\mathcal{C}, d_∞) die Extremalpunkte dicht liegen.

[†]sauberer formuliert: die Dichten der entsprechenden doppelt stochastischen Maße $\mu_A, \mu_{A_1}, \mu_{A_2}, \dots$

09. Übung am 04. Juni 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 48 Beweisen Sie, dass für jede Copula $A \in \mathcal{C}$ und jedes $n \in \mathbb{N}$ die (im Beweis von Lemma 3.3 für den Spezialfall $n = 1$ verwendete) Gleichheit

$$\int_{[0,1]^2} u^n d\mu_A(u, v) = \int_{[0,1]^2} v^n d\mu_A(u, v) = \frac{1}{n+1}$$

gilt.

Übungsaufgabe 49 Beweisen Sie, dass Kendall's τ die ersten drei Eigenschaften eines Konkordanzmaßes gemäß Definition 3.6 erfüllt.

Hinweis: Arbeiten Sie - je nachdem, was für die konkrete Eigenschaft handlicher ist - mit (X, Y) oder gleich direkt mit der zugrunde liegenden Copula A und Gleichung 3.6.

Übungsaufgabe 50 Beweisen Sie, dass Kendall's τ die letzten drei Eigenschaften eines Konkordanzmaßes gemäß Definition 3.6 erfüllt.

Hinweis: Arbeiten Sie - je nachdem, was für die konkrete Eigenschaft handlicher ist - mit (X, Y) oder gleich direkt mit der zugrunde liegenden Copula A und Gleichung 3.6.

Zusatz*: Gilt auch die folgende, stärkere Aussage? Sei $((X_n, Y_n))_{n \in \mathbb{N}}$ eine Folge von Paaren stetiger Zufallsvariable, die schwach gegen das ebenfalls stetige Paar (X, Y) konvergiert. Dann folgt $\lim_{n \rightarrow \infty} \kappa_{X_n, Y_n} = \kappa_{X, Y}$.

Übungsaufgabe 51 Für jedes $z \in [0, 1]$ bezeichne R_z die λ -treue Rotation um z , i.e. $R_z(x) = x + z \pmod{1}$ für jedes $x \in [0, 1]$. $A_z \in \mathcal{C}_d$ bezeichne die entsprechende vollständig abhängige Copula. Berechnen Sie

$$\min_{y, z \in [0, 1]} \int_{[0, 1]^2} A_z d\mu_{A_y}.$$

Übungsaufgabe 52 Konstruieren Sie zwei Copulas A, B mit den folgenden Eigenschaften und erklären Sie, warum dieses Beispiel in Bezug auf die Konkordanz wichtig ist:

1. $A \leq B$ punktweise auf ganz $[0, 1]^2$ und $A(x, y) < B(x, y)$ für ein $(x, y) \in [0, 1]^2$.
2. $\tau(A) = \tau(B)$.

Hinweis: Es reicht, mit (geraden und ungeraden) Shuffles bestehend aus 2 Segmenten zu arbeiten.

Übungsaufgabe 53 (*) Beweisen Sie Satz 3.9.

10. Übung am 11. Juni 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 54 Berechnen Sie die Kendall Verteilungsfunktion der Copulas M, W, Π , sowie für einen Shuffle $A_h \notin \{M, W\}$.

Übungsaufgabe 55 Für $a, b \in [0, 1]$ und $a + b \leq 1$ sei die Copula $C_{a,b}$ definiert durch

$$C_{a,b} = aM + bW + (1 - a - b)\Pi.$$

Berechnen Sie $\tau(A)$ und $\rho(A)$.

Übungsaufgabe 56 Berechnen Sie Kendall's τ für sogenannte Prototypen, i.e. gerade Shuffles A_n bestehend aus n gleich langen, absteigenden Segmenten (A_2, A_3, A_4 sind in Abbildung 3.1 dargestellt).

Übungsaufgabe 57 Berechnen Sie Spearman's ρ für sogenannte Prototypen, i.e. gerade Shuffles A_n bestehend aus n gleich langen, absteigenden Segmenten (A_2, A_3, A_4 sind in Abbildung 3.1 dargestellt).

Übungsaufgabe 58 Rechnen Sie nach, dass die Ungleichung von Durbin und Stuart in den Punkte $(\tau(A_n), \rho(A_n))$ (mit A_n wie in den letzten beiden Aufgaben) scharf ist.

Übungsaufgabe 59 Beweisen Sie die folgende Gleichheit für beliebige Copulas $A, B \in \mathcal{C}$:

$$\int_{[0,1]^2} B d\mu_A = \frac{1}{2} - \int_{[0,1]^2} K_B(x, [0, y]) K_{A^t}(y, [0, x]) d\lambda_2(x, y)$$

11. Übung am 25. Juni 2018

[LVA 405.552 UV Abhängigkeitsmodellierung, Ankreuzliste siehe www.trutschnig.net/courses mit * versehene Aufgaben sind freiwillig]

Übungsaufgabe 60 Beweisen Sie Lemma 4.5.

Übungsaufgabe 61 Beweisen Sie Satz 4.2.

Übungsaufgabe 62 Beweisen Sie Gleichung 4.2.

Übungsaufgabe 63 Sei $N \in \mathbb{N}$ und $A \in \mathcal{CB}_N$. Zeigen Sie, dass für jedes $x \in [0, 1]$ die Abbildung $y \mapsto A(x, y)$ stückweise linear ist. Gilt die analoge Aussage auch für $x \mapsto A(x, y)$ und festes $y \in [0, 1]$?

Übungsaufgabe 64 (R) Verwenden Sie `marshall_olkin_sim.R` (auf der homepage) um Stichproben $(x_1, y_1), \dots, (x_n, y_n)$ der Marshall-Olkin Copula O mit Parameter $a = 1, b = 0.5$ zu generieren. Generieren Sie eine Stichprobe der Größe $n = 10.000$, berechnen Sie zuerst die empirische Π Copula A_n (am entsprechenden Gitter R_n), dann $\mathcal{CB}_N(A_n)$ mit $N = 100 = \sqrt{n}$ und schließlich (approximativ) $D_1(\mathcal{CB}_N(A_n), O)$. Was ist zu beobachten?

Übungsaufgabe 65 Verwenden Sie Satz 4.1 um zu beweisen, dass die Menge aller Shuffles mit gleich langen Streifen dicht in (\mathcal{C}, d_∞) ist.[†]

[†]Bisher haben wir nur bewiesen, dass die Menge aller Shuffles dicht liegt, die im Beweis konstruierten Shuffles haben aber i.A. nicht gleich lange Segmente.