

Copulae: Grundlagen und eine wichtige Anwendung

Teil der Vorlesung
Mathematische Statistik (LV-Nr. 405.160)

Assoz.Prof. Dr. Wolfgang Trutschnig
Fachbereich Mathematik y
Universität Salzburg x
www.trutschnig.net

Salzburg, 2019-11-11





- ▶ Eindimensionale Verteilungsfunktionen \longleftrightarrow Verteilungen (Wahrscheinlichkeitsmaße) auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
- ▶ NB: X ZV auf $(\Omega, \mathcal{A}, \mathbb{P})$. Für jedes $x \in \mathbb{R}$ gilt

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}^X((-\infty, x]). \quad (1)$$

- ▶ Eigenschaften der Quasiinversen F^- einer Verteilungsfunktion F .
- ▶ Erzeugung von Stichproben einer Zufallsvariable $X \sim F$ mittels $F^- \circ Z$ mit $Z \sim \mathcal{U}(0, 1)$.
- ▶ 2-dimensionale Verteilungsfunktionen \longleftrightarrow Verteilungen (Wahrscheinlichkeitsmaße) auf $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.
- ▶ NB: X_1, X_2 ZV auf $(\Omega, \mathcal{A}, \mathbb{P})$. Für jedes $(x, y) \in \mathbb{R}^2$ gilt

$$H_{(X_1, X_2)}(x_1, x_2) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbb{P}^{(X_1, X_2)}((-\infty, x_1] \times (-\infty, x_2]). \quad (2)$$

- ▶ $(X_1, X_2) \sim H$, dann gilt für die Randverteilungen:

$$F_{X_1}(x_1) = \lim_{x_2 \rightarrow \infty} H(x_1, x_2) =: H(x_1, \infty), \quad F_{X_2}(x_2) = H(\infty, x_2)$$



- ▶ Seien $F_1, F_2 \in \mathcal{F}$ beliebige Verteilungsfunktionen
- ▶ Leicht zu zeigen (Übungsaufgabe): H_1 und H_2 , definiert durch

$$H_1(x_1, x_2) = \min \{F_1(x_1), F_2(x_2)\}, \quad H_2(x_1, x_2) = F_1(x_1)F_2(x_2)$$

sind 2-dim Verteilungsfunktionen, die F_1 und F_2 als Randverteilungen haben.

- ▶ → Aus der Kenntnis der Randverteilungen kann nicht auf die gemeinsame Verteilung geschlossen werden.
- ▶ Grundlegende Frage (Fréchet, 1951):
Wie schaut die Menge aller 2-dimensionalen Verteilungsfunktionen H aus, die F_1 und F_2 als Randverteilungsfunktionen haben ?
- ▶ Copulae sind der Schlüssel zur Beantwortung obiger Frage.
- ▶ **Plan für ersten Teil heute:** (i) Wichtigste Eigenschaften von Copulae.
- ▶ (ii) Angewandte Problemstellung, die die Wichtigkeit von Copulae unterstreicht.



Definition 1 (Copula, Sklar, 1959)

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

1. $\forall x \in [0, 1] : A(x, 1) = A(1, x) = x, A(x, 0) = A(0, x) = 0.$
2. Für $0 \leq x_1 \leq x_2 \leq 1$ und $0 \leq y_1 \leq y_2 \leq 1$ gilt

$$\mu_A((x_1, x_2] \times (y_1, y_2]) = A(x_2, y_2) - A(x_1, y_2) - A(x_2, y_1) + A(x_1, y_1) \geq 0$$

Die Familie aller Copulae wird im Folgenden mit \mathcal{C} bezeichnet.
Für jedes $A \in \mathcal{C}$ heißt das entsprechende Wahrscheinlichkeitsmaß μ_A doppelt stochastisch.

- ▶ Eine Copula ist per definitionem eine Funktion auf $[0, 1]^2$.
- ▶ Jedes $A \in \mathcal{C}$ kann aber zu einer Verteilungsfunktion (Definitionsbereich \mathbb{R}^2) \tilde{A} fortgesetzt werden:



- ▶ Definiere $\tilde{A} : \mathbb{R}^2 \rightarrow [0, 1]$ durch (Skizze !)

$$\tilde{A}(x_1, x_2) = \begin{cases} 0 & \text{für } x_1 \leq 0 \text{ oder } x_2 \leq 0 \\ A(x_1, x_2) & \text{für } x_1, x_2 \in [0, 1] \\ x_1 & \text{für } (x_1, x_2) \in [0, 1] \times (1, \infty) \\ x_2 & \text{für } (x_1, x_2) \in (1, \infty) \times [0, 1] \\ 1 & \text{für } x_1, x_2 \geq 1. \end{cases} \quad (3)$$

- ▶ Leicht nachzurechnen: $\tilde{A} \in \mathcal{F}_2$.
- ▶ Für die Randverteilungen F_1, F_2 erhalten wir sofort

$$F_1(x_1) = \begin{cases} 0 & \text{für } x_1 \leq 0 \\ x_1 & \text{für } x_1 \in [0, 1] \\ 1 & \text{für } x_1 \geq 1. \end{cases}$$

und $F_1 = F_2$.

- ▶ Mit anderen Worten: Eine Copula ist eine 2-dimensionale Verteilungsfunktion mit $\mathcal{U}(0, 1)$ -Randverteilungen eingeschränkt auf $[0, 1]^2$.



Beispiel 2 (Π, M)

$\Pi(x_1, x_2) = x_1 x_2$ und $M(x_1, x_2) := \min\{x_1, x_2\}$ sind Copulae, genannt *Produktcopula* und *Minimum Copula*.

- ▶ $(X_1, X_2) \sim \Pi$ entspricht dem Fall, dass X_1 und X_2 unabhängig sind,
- ▶ $(X_1, X_2) \sim M$ dem Fall $\mathbb{P}(X_1 = X_2) = 1$ (siehe Übungen).

Beispiel 3 (Marshall-Olkin Familie)

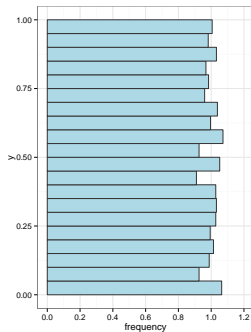
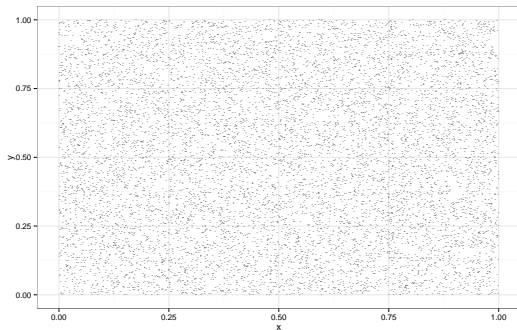
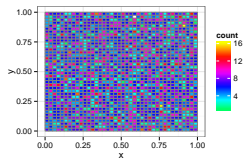
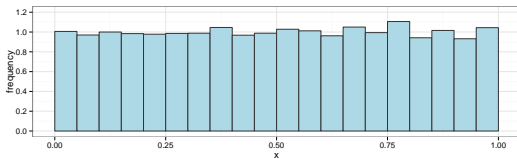
Die MO-Familie $(M_{\alpha, \beta})_{(\alpha, \beta) \in [0, 1]^2}$ von Copulae ist gegeben durch

$$M_{\alpha, \beta}(x_1, x_2) = \begin{cases} x_1^{1-\alpha} x_2 & \text{für } x_1^\alpha \geq x_2^\beta \\ x_1 x_2^{1-\beta} & \text{für } x_1^\alpha \leq x_2^\beta. \end{cases} \quad (4)$$

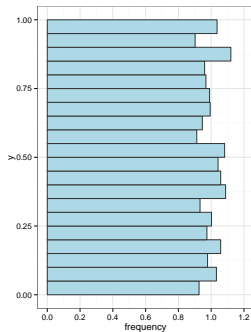
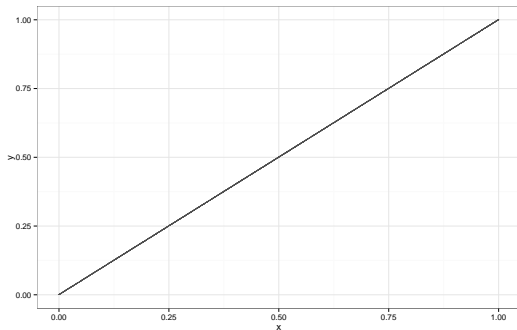
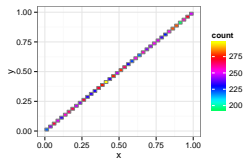
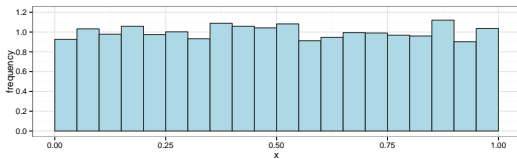
- ▶ Die folgenden Grafiken zeigen samples von $\Pi, M, M_{0.5, 1}, M_{1, 0.5}$ (in dieser Reihenfolge)



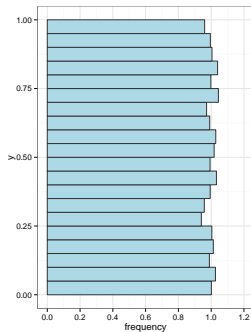
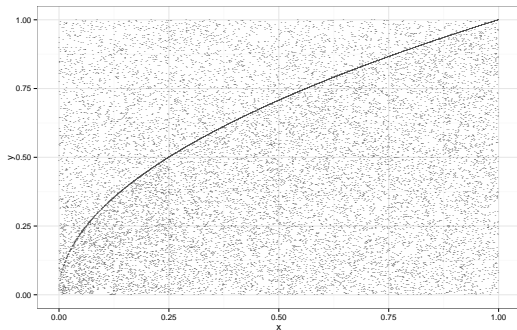
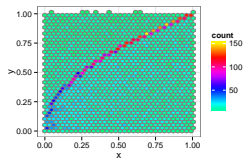
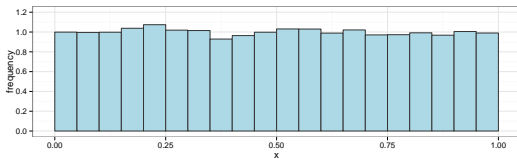
Definition und Beispiele



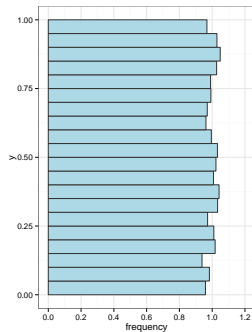
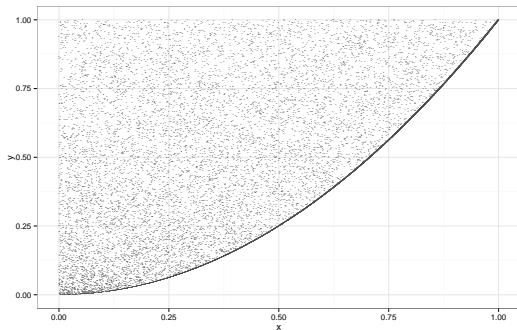
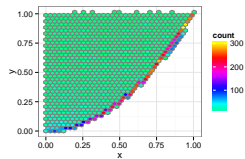
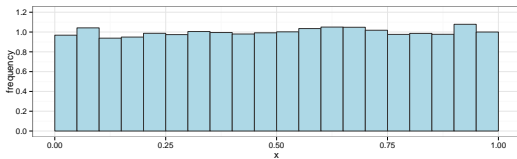
Definition und Beispiele



Definition und Beispiele



Definition und Beispiele



- ▶ Das berühmteste und wichtigste Resultat betreffend Copulae ist der folgende

Satz 4 (Satz von Sklar, 1959)

Seien X_1, X_2 Zufallsvariable mit Verteilungsfunktion F_1 bzw. F_2 und gemeinsamer Verteilungsfunktion H . Dann existiert eine Copula $A \in \mathcal{C}$ mit

$$H(x_1, x_2) = A(F_1(x_1), F_2(x_2)) \quad (5)$$

für alle $x_1, x_2 \in \mathbb{R}$. A ist nur eindeutig falls F_1 und F_2 stetig sind.

Zusatz: Umgekehrt ist für alle Verteilungsfunktionen $F_1, F_2 \in \mathcal{F}$ und jede Copula $A \in \mathcal{C}$ die Funktion $H : \mathbb{R}^2 \rightarrow [0, 1]$, definiert durch $H(x_1, x_2) := A(F_1(x_1), F_2(x_2))$ eine 2-dimensionale Verteilungsfunktion.

Beweis: Elementar mit Hilfe der Anleitung am Übungsblatt.

Grundidee: (i) Verwende (5) als Definition für A auf $Rg(F_1) \times Rg(F_2)$,
(ii) Eigenschaften nachrechnen und (iii) A auf $[0, 1]^2$ fortsetzen.



- ▶ **Situation 1:** X_1, X_2 bezeichnen Lebensdauern von Bauteilen.
- ▶ **Situation 2:** X_1, X_2 bezeichne default times von Finanzinstrumenten (Finanzderivate, Collateralized Debt Obligations, Credit Default Swaps, etc.) oder Firmen (Kreditausfall, etc.).
- ▶ **Problemstellung:**
 1. Die gemeinsame Verteilungsfunktion H von (X_1, X_2) ist unbekannt, man kennt nur die Randverteilungsfunktionen F_1 und F_2 .
 2. Wie groß kann die Wahrscheinlichkeit für einen gemeinsamen Ausfall $\mathbb{P}(X_1 = X_2)$ maximal werden ?
- ▶ **Lösung ?**
 1. Schon gesehene Lösung: Schlimmstmöglicher Fall muss $A = M$ sein, da dies genau der Situation $X_1 = X_2$ entspricht.



- ▶ Wir überprüfen mit Hilfe von Simulationen die Qualität der Lösung $A = M$ für einen einfachen Fall von F_1, F_2 .

Beispiel 5 (Joint default probability)

X_1, X_2 Lebensdauern; $X_1 \sim F_1, X_2 \sim F_2$, wobei F_1 und F_2 gegeben sind durch:

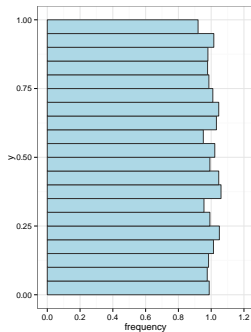
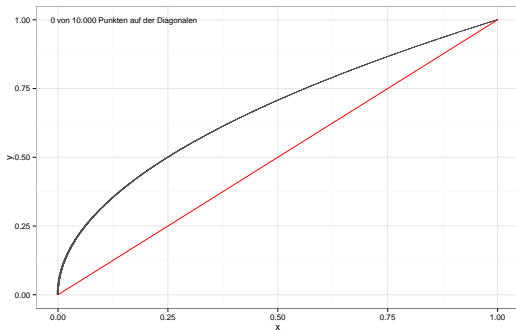
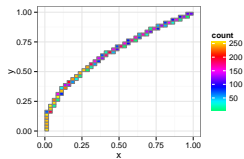
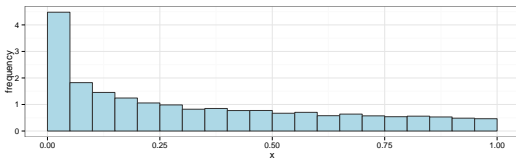
$$F_1(x_1) = \begin{cases} 0 & \text{für } x_1 < 0, \\ \sqrt{x_1} & \text{für } x_1 \in [0, 1] \\ 1 & \text{für } x_1 > 1. \end{cases} \quad F_2(x_2) = \begin{cases} 0 & \text{für } x_2 < 0, \\ x_2 & \text{für } x_2 \in [0, 1] \\ 1 & \text{für } x_2 > 1. \end{cases}$$

$A \in \mathcal{C}$ bezeichne die Copula von (X_1, X_2) .

Für den Fall $A = M$ liefert die Simulation das folgenden Resultate:



Wahrscheinlichkeit für joint default



Beispiel 5 (Joint default probability, cont)

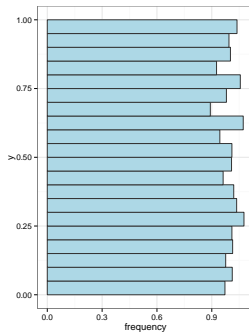
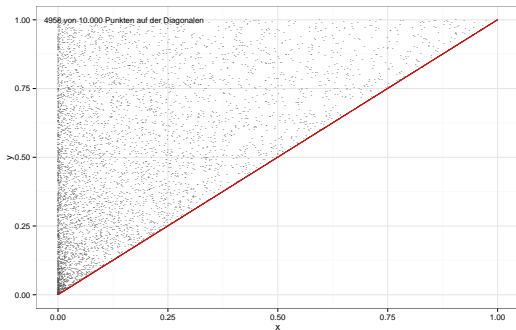
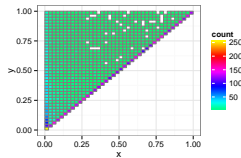
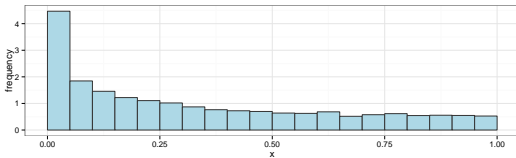
- ▶ $\mathbb{P}(X_1 = X_2) = 0$ unabhängig von der Copula A von (X_1, X_2) ?
- ▶ Für diese Wahl von F_1 und F_2 joint default unmöglich ?
- ▶ Saubere Lösung: Setze $T := F_2 \circ F_1^{-1} : [0, 1] \rightarrow [0, 1]$, wir erhalten $T(x) = x^2$. Falls (X_1, X_2) Copula $A \in \mathcal{C}$ hat dann folgt weiters (nichttrivial !)

$$\mathbb{P}(X_1 = X_2) = \mu_A(\Gamma(T)) = \mu_A\left(\{(x, x^2) : x \in [0, 1]\}\right)$$

- ▶ Wir haben heute schon eine Copula mit $\mu_A(\Gamma(T)) > 0$ gesehen.
- ▶ Marshall-Olkin Copula $M_{1,0.5}$!
- ▶ Für $A = M_{1,0.5}$ ergibt sich (analytisch oder mit Simulationen) $\mathbb{P}(X_1 = X_2) = 0.5 > 0$



Wahrscheinlichkeit für joint default



- ▶ Die vorhin vorgeschlagene Lösung $A = M$ ist falsch.
- ▶ Für 'böse' Verteilungsfunktionen H (mit Randverteilungen F_1 und F_2) gilt $\mathbb{P}(X_1 = X_2) \geq 0.5 > 0$.
- ▶ Ist 0.5 schon die obere Schranke ?
- ▶ Man kann zeigen, dass $\mathbb{P}(X_1 = X_2)$ maximal 0.75 sein kann (interessantes Thema für Bachelor-/Masterarbeit).
- ▶ Im schlimmsten Fall gibt es also eine 75%-tige Chance für einen joint default.
- ▶ Vorhin vorgeschlagene 'Lösung' führt zu einer katastrophalen Unterschätzung.
- ▶ **Kenntnis der Abhängigkeitsstruktur (i.e. der Copula) ist extrem wichtig für viele Anwendungen !**



Forschungsausblick:

Satz 6 (via 'optimal transport' oder zu Fuß)

$X_1 \sim F, X_2 \sim F_2$ mit F_1, F_2 stetig. Setze $T = F_2 \circ F_1^{-1}$. Dann ist die maximale joint-default Wahrscheinlichkeit gegeben durch

$$\bar{m}_{F_1, F_2} = \int_{[0,1]} \min\{T'(x), 1\} d\lambda(x)$$

- ▶ Angenommen x_1, \dots, x_n ist ein sample von $X_1 \sim F_1$ und y_1, \dots, y_m ist ein sample von $X_2 \sim F_2$
- ▶ Wie kann \bar{m}_{F_1, F_2} aus x_1, \dots, x_n und y_1, \dots, y_m bestmöglich geschätzt werden? Offene Problemstellung

