

# Modellierung stochastischer Abhängigkeiten von Risiken in der Sachversicherung

Prof. Dr. Dietmar Pfeifer  
Institut für Mathematik

Schwerpunkt Versicherungs- und Finanzmathematik

Kolloquium Versicherungs- und Finanzmathematik  
Leibniz Universität & Concordia Versicherungen  
Hannover, 13. Januar 2012

## Agenda

- Copulas
- Die empirische Copula
- Bernstein-Copulas
- Von der empirischen zur Bernstein-Copula
- Simulationen mit Bernstein-Copulas
- Anwendungsbeispiele
- Literatur

## Copulas

Motivation:

**Stochastische Abhängigkeiten** entstehen im Versicherungsgeschehen auf natürliche Weise z.B.

- zwischen ähnlichen Sparten (etwa VGV / VHV)
- aufgrund räumlicher Kohärenz (etwa Sturm)
- durch gemeinsame Trigger (etwa Sturm- / Hagel- / Hochwasserschäden).

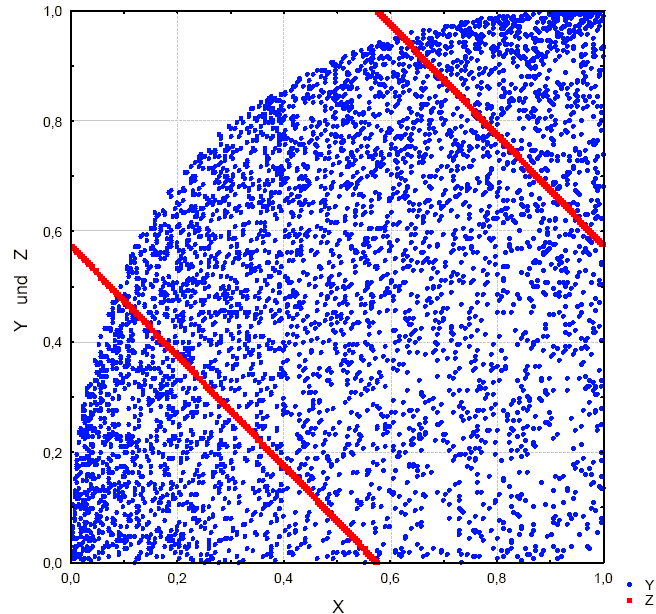
## Copulas

Beispiel:

Zufallsvariablen  $X$ ,  $Y$  und  $Z$   
mit *identischen* (uniformen)  
Randverteilungen und *gleichen*  
Korrelationen

$$\rho^L(X, Y) = \rho^L(X, Z) = 7/15,$$

aber *verschiedener* gemein-  
samer Verteilungsstruktur



## Copulas

Satz von Sklar (1958): Jede mehrdimensionale Verteilungsfunktion  $F$  kann zerlegt werden in eine Copula  $C$  und ihre eindimensionalen Randverteilungsfunktionen  $F_1, \dots, F_d$ :

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

Dabei bedeutet eine Copula anschaulich eine  $d$ -dimensionale Verteilungsfunktion mit stetigen uniformen Randverteilungsfunktionen.

## Copulas

Satz von Sklar (1958): Jede mehrdimensionale Verteilungsfunktion  $F$  kann zerlegt werden in eine Copula  $C$  und ihre eindimensionalen Randverteilungsfunktionen  $F_1, \dots, F_d$ :

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

Die Copula ist eindeutig bestimmt, wenn die Randverteilungsfunktionen *stetig* sind. In diesem Fall gilt auch die Umkehrung

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

mit den Quantilfunktionen  $F_1^{-1}, \dots, F_d^{-1}$  (Pseudo-Inverse).

## Copulas

Satz von Sklar (1958): Jede mehrdimensionale Verteilungsfunktion  $F$  kann zerlegt werden in eine Copula  $C$  und ihre eindimensionalen Randverteilungsfunktionen  $F_1, \dots, F_d$ :

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

Interpretation:

- Die **Randverteilungsfunktionen** beschreiben das stochastische Verhalten der *univariaten* Risiken.
- Die **Copula** beschreibt die stochastische *Abhängigkeit zwischen* den Risiken.

## Copulas

Vorteile des Copula-Ansatzes:

- Die Copula hängt nicht von den Randverteilungen der Einzelrisiken ab.
- Die Copula charakterisiert bei Stetigkeit der Randverteilungen die gemeinsame Abhängigkeitsstruktur der Einzelrisiken eindeutig.
- Die Copula ist invariant gegen alle gleichsinnig monotonen Transformationen der Einzelrisiken, insbesondere nichtlineare.
- Paarweise Korrelationen zwischen den Einzelrisiken können über die Copula ausgedrückt werden, aber nicht umgekehrt.



# Copulas

Bekannte Typen von Copulas:

parametrische Copulas:

- Elliptische Copulas (→ Gauß-Copula,  $t$ -Copula)
- Archimedische Copulas

nichtparametrische Copulas:

- Bernstein-Copulas
- Zerlegung-der-Eins-Copulas (→ B-Spline-Copulas)

# Copulas

Bekannte Typen von Copulas:

genestete Copulas:

- Hierarchische Archimedische Copulas
- D-Vine Copulas

Mischung singularer Copulas

Generelles Problem: „Fluch der Dimension“:

Handhabbarkeit der Modellierung in hohen Dimensionen

## Copulas

Gauß-Copula  $C_{\Sigma}^G$ :

$$C_{\Sigma}^G(u_1, \dots, u_n) = \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_n)} \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2} \mathbf{v}^{\text{tr}} \Sigma^{-1} \mathbf{v}\right) dv_1 \cdots dv_n$$

$t$ -Copula  $C_{\Sigma}^{t_{\nu}}$  mit  $\nu \in \mathbb{N}$  Freiheitsgraden:

$$C_{\Sigma}^{t_{\nu}}(u_1, \dots, u_n) = \int_{-\infty}^{t_{\nu}^{-1}(u_1)} \cdots \int_{-\infty}^{t_{\nu}^{-1}(u_n)} \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{(\pi\nu)^n \det(\Sigma)}} \left(1 + \frac{1}{\nu} \mathbf{v}^{\text{tr}} \Sigma^{-1} \mathbf{v}\right)^{\left(-\frac{\nu+n}{2}\right)} dv_1 \cdots dv_n$$

## Copulas

*Archimedische Copulas* sind charakterisiert durch ihren so genannten *Erzeuger*  $\varphi$  vermöge

$$C_n(u_1, \dots, u_n) = \varphi^{-1}\left(\sum_{i=1}^n \varphi(u_i)\right) \quad \text{für } u_1, \dots, u_n \in [0, 1].$$

Wichtigster Spezialfall:

$$\varphi(x) = -\ln x$$

mit

$$C_n(u_1, \dots, u_n) = \varphi^{-1}\left(\sum_{i=1}^n \varphi(u_i)\right) = \exp\left(\sum_{i=1}^n \ln u_i\right) = \prod_{i=1}^n u_i$$

(Unabhängigkeitscopula  $\Pi$ )

## Copulas

Charakterisierung “geeigneter” Erzeuger:

Sei  $\varphi: (0,1] \rightarrow \mathbb{R}$  stetig und streng monoton fallend mit  $\varphi(1) = 0$  and  $\lim_{z \downarrow 0} \varphi(z) = \infty$ ;  $\varphi^{-1}$  bezeichne die zugehörige Inverse auf dem Intervall  $[0, \infty)$ . Dann ist die durch

$$C_n(u_1, \dots, u_n) = \varphi^{-1} \left( \sum_{i=1}^n \varphi(u_i) \right) \quad \text{für } u_1, \dots, u_n \in [0,1]$$

gegebene Abbildung  $C_n$  eine Copula für  $n \geq 2$ . Sie ist eine Copula für alle  $n \geq 2$  genau dann, wenn  $\varphi^{-1}$  *total monoton* ist, d.h. es gilt

$$(-1)^k \frac{d^k}{ds^k} \varphi^{-1}(s) \geq 0 \quad \text{für alle } k \in \mathbb{N} \text{ und } s > 0.$$

## Copulas

Nach einem klassischen Satz von *Bernstein* können die Inversen solcher Erzeuger als Laplace-Transformierte nicht-negativer Zufallsvariablen  $Z$  dargestellt werden vermöge

$$\varphi^{-1}(s) = E(e^{-sZ}), \quad s \geq 0,$$

denn:

$$\varphi^{-1}(0) = 1, \quad \lim_{s \rightarrow \infty} \varphi^{-1}(s) = 0, \quad \text{und} \quad (-1)^k \frac{d^k}{ds^k} \varphi^{-1}(s) = E(Z^k e^{-sZ}) \geq 0$$

für alle  $k \in \mathbb{N}$  und  $s > 0$ .

## Copulas

Beispiele:

Name	Copula $C_\theta$	Erzeuger $\varphi_\theta$	Mischver- teilung
Clayton	$\left[ \sum_{i=1}^n u_i^{-\theta} - n + 1 \right]^{-1/\theta}, \theta > 0$	$\frac{1}{\theta}(t^{-\theta} - 1)$	$\Gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$
Gumbel	$\exp\left(-\left\{\sum_{i=1}^n (-\ln(u_i))^\theta\right\}^{1/\theta}\right), \theta \geq 1$	$(-\ln t)^\theta$	positiv stabil
Frank	$-\frac{1}{\theta} \ln\left(1 - (1 - e^{-\theta}) \prod_{i=1}^n \left\{\frac{1 - e^{-\theta u_i}}{1 - e^{-\theta}}\right\}\right), \theta > 0$	$-\ln \frac{1 - e^{-\theta t}}{1 - e^{-\theta}}$	$\mathcal{LS}(e^{-\theta})$ auf $\mathbb{N}$

## Die empirische Copula

- Analogon zur empirischen Verteilungs-/Quantilfunktion

Gegeben:  $n$  unabhängige Beobachtungen  $z_{j1}, \dots, z_{jn}$  je Dimension  $j$ , mit  $j = 1, \dots, d$

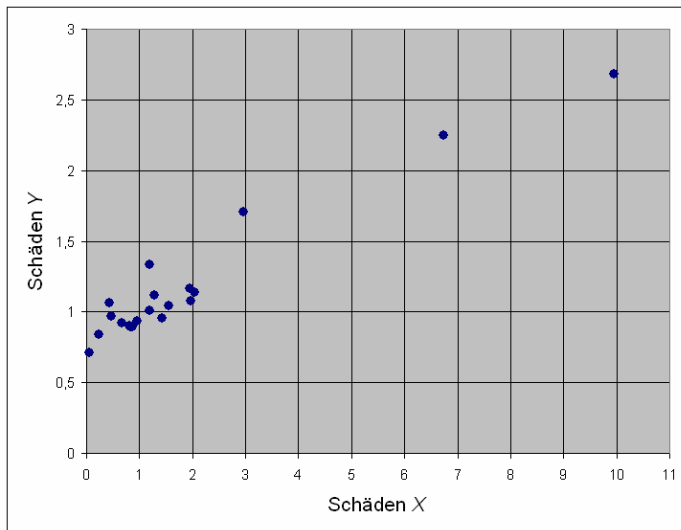
- Ansatz: Transformation der Daten  $z_{j1}, \dots, z_{jn}$  auf ihre relativen Ränge  $r_{j1}, \dots, r_{jn} \in \left\{ \frac{1}{n+1}, \dots, \frac{n}{n+1} \right\}$
- Verwendung der  $\mathbf{r}_i = (r_{i1}, \dots, r_{di})$  als *empirische Stichprobe* aus der zu Grunde liegenden Copula.



# Die empirische Copula

Beispieldatensatz ( $d = 2$ ,  $n = 20$ ):

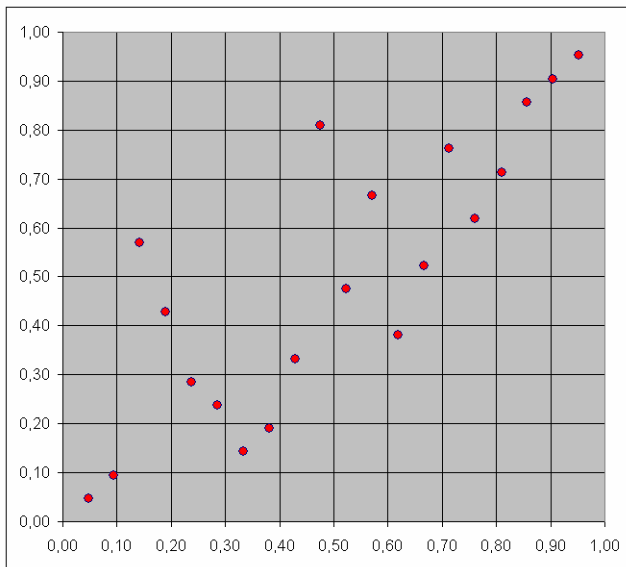
Nr.	Schäden X	Schäden Y
1	0,468	0,966
2	9,951	2,679
3	0,866	0,897
4	6,731	2,249
5	1,421	0,956
6	2,040	1,141
7	2,967	1,707
8	1,200	1,008
9	0,426	1,065
10	1,946	1,162
11	0,676	0,918
12	1,184	1,336
13	0,960	0,933
14	1,972	1,077
15	1,549	1,041
16	0,819	0,899
17	0,063	0,710
18	1,280	1,118
19	0,824	0,894
20	0,227	0,837



## Die empirische Copula

Beispieldatensatz ( $d = 2$ ,  $n = 20$ ):

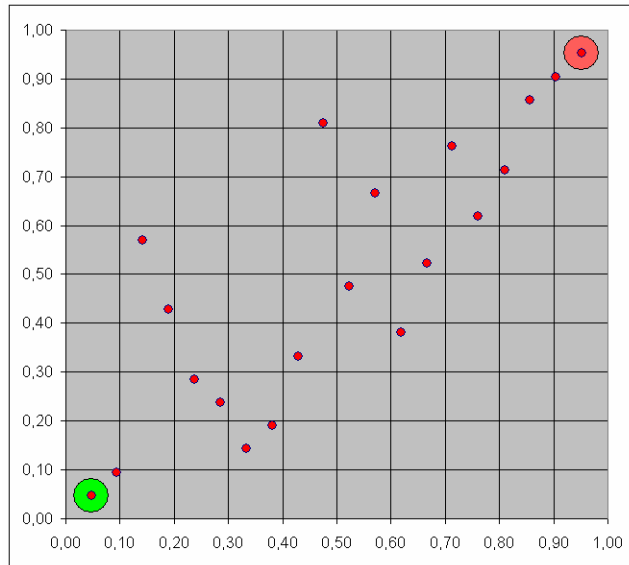
Nr.	Rang X	Rang Y
1	0,18967	0,42852
2	0,95238	0,95238
3	0,38057	0,18967
4	0,90443	0,90443
5	0,61852	0,38057
6	0,80943	0,71352
7	0,85648	0,85648
8	0,52352	0,47557
9	0,14262	0,57057
10	0,71352	0,76148
11	0,23762	0,28557
12	0,47557	0,80943
13	0,42852	0,33262
14	0,76148	0,61852
15	0,66648	0,52352
16	0,28557	0,23762
17	0,04762	0,04762
18	0,57057	0,66648
19	0,33262	0,14262
20	0,09467	0,09467



# Die empirische Copula

Beispieldatensatz ( $d = 2$ ,  $n = 20$ ):

Nr.	Rang X	Rang Y
1	0,18967	0,42852
2	0,95238	0,95238
3	0,38057	0,18967
4	0,90443	0,90443
5	0,61852	0,38057
6	0,80943	0,71352
7	0,85648	0,85648
8	0,52352	0,47557
9	0,14262	0,57057
10	0,71352	0,76148
11	0,23762	0,28557
12	0,47557	0,80943
13	0,42852	0,33262
14	0,76148	0,61852
15	0,66648	0,52352
16	0,28557	0,23762
17	0,04762	0,04762
18	0,57057	0,66648
19	0,33262	0,14262
20	0,09467	0,09467



## Die empirische Copula

Vorteile:

- Die empirische Copula ist oft leicht direkt aus einer gegebenen multivariaten Verteilung simulierbar (z.B. elliptische Verteilungen;  $t$ -Copula)
- Mit Hilfe der empirischen Copula (d.h. der damit assoziierten Ränge) können Ausgaben von simulierten Daten aus unterschiedlichen externen Quellen passgenau „zusammengesetzt“ werden (z.B. Outputs aus unterschiedlichen geophysikalischen Modellen)

## Bernstein-Copulas

Idee: Glättung der Daten durch geeignete Interpolation (ähnlich den höherdimensionalen Bernstein-Polynomen)

- Die „Steuerpunkte“ des mehrdimensionalen Ansatzes werden durch Gewichte aus Wahrscheinlichkeiten geeigneter diskreter multivariater Verteilungen mit uniformen Rändern gebildet.
- Die Gewichte werden aus den Daten z.B. mittels der empirischen Copula geschätzt (→ multivariate Kontingenztafel).
- Es ist eine explizite Darstellung der Copula-Dichte möglich.

## Bernstein-Copulas

$\mathbf{U} = (U_1, \dots, U_d)$  sei ein Zufallsvektor mit Komponenten  $U_i$ , die einer **diskreten Gleichverteilung** über  $T_i := \{0, 1, \dots, m_i - 1\}$  mit  $m_i \in \mathbb{N}$ ,  $i = 1, \dots, d$  genügen. Ferner bezeichne

$$p(k_1, \dots, k_d) := P\left(\bigcap_{i=1}^d \{U_i = k_i\}\right) \text{ für alle } (k_1, \dots, k_d) \in \prod_{i=1}^d T_i.$$

Dann wird über  $B(m, k, z) = \binom{m}{k} z^k (1-z)^{m-k}$  durch

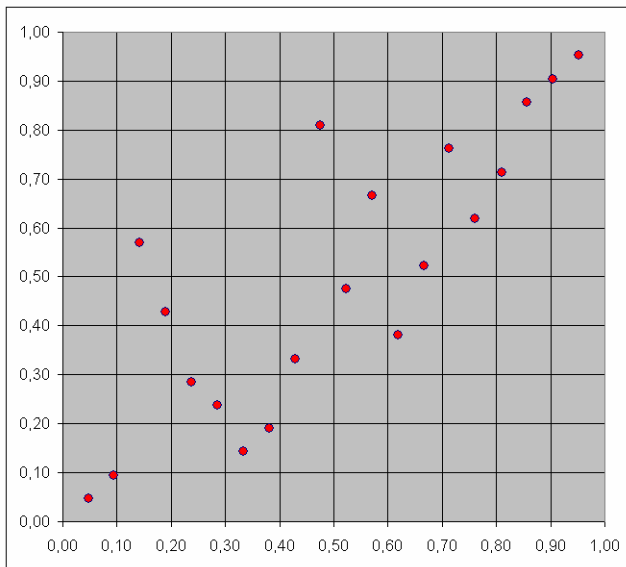
$$c(u_1, \dots, u_d) := \sum_{k_1=0}^{m_1-1} \cdots \sum_{k_d=0}^{m_d-1} p(k_1, \dots, k_d) \prod_{i=1}^d m_i B(m_i - 1, k_i, u_i), \mathbf{u} \in [0, 1]^d$$

die Dichte einer  $d$ -dimensionalen Bernstein-Copula definiert.

# Von der empirischen zur Bernstein-Copula

Beispieldatensatz ( $d = 2$ ,  $n = 20$ ):

Nr.	Rang X	Rang Y
1	0,18967	0,42852
2	0,95238	0,95238
3	0,38057	0,18967
4	0,90443	0,90443
5	0,61852	0,38057
6	0,80943	0,71352
7	0,85648	0,85648
8	0,52352	0,47557
9	0,14262	0,57057
10	0,71352	0,76148
11	0,23762	0,28557
12	0,47557	0,80943
13	0,42852	0,33262
14	0,76148	0,61852
15	0,66648	0,52352
16	0,28557	0,23762
17	0,04762	0,04762
18	0,57057	0,66648
19	0,33262	0,14262
20	0,09467	0,09467

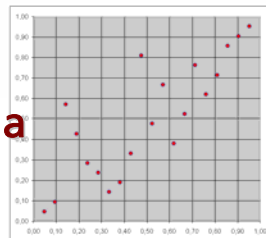




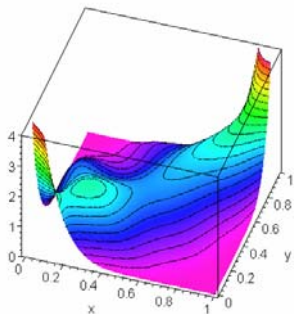


# Von der empirischen zur Bernstein-Copula

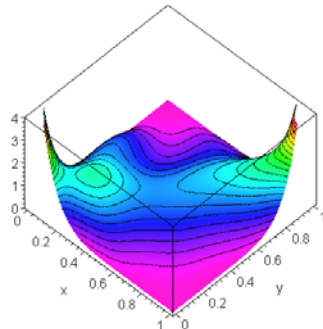
## Bernstein-Copula-Dichte:



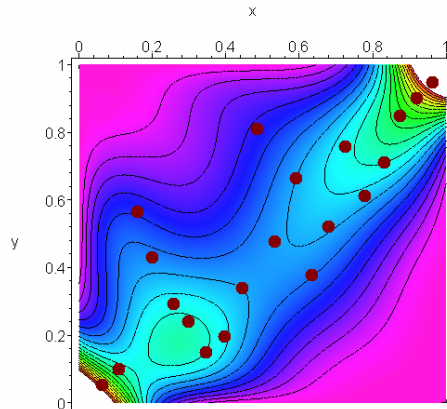
$$\begin{aligned}
 c(x, y) = & 10(1-x)^9(1-y)^9 + 7560x^3(1-x)^6y(1-y)^8 + 12960x^2(1-x)^7y^2(1-y)^7 \\
 & + 52920x^4(1-x)^5y^3(1-y)^6 + 35280x^6(1-x)^3y^3(1-y)^6 + 5670x(1-x)^8y^4(1-y)^5 \\
 & + 79380x^5(1-x)^4y^4(1-y)^5 + 5670x(1-x)^8y^5(1-y)^4 + 52920x^6(1-x)^3y^5(1-y)^4 \\
 & + 52920y^6(1-y)^3x^5(1-x)^4 + 15120y^6(1-y)^3x^7(1-x)^2 + 6480y^7(1-y)^2x^7(1-x)^2 \\
 & + 1620x^8(1-x)y^7(1-y)^2 + 5670x^4(1-x)^5y^8(1-y) + 405x^8(1-x)y^8(1-y) + 10x^9y^9
 \end{aligned}$$



$c(x, y)$



## Von der empirischen zur Bernstein-Copula



Höhenlinien der Bernstein-Copula-Dichte  
überlagert mit Punkten der empirischen Copula

## Von der empirischen zur Bernstein-Copula

**Problem:** Randsummen eventuell nicht „passend“

- Optimierungsproblem zur Approximation durch eine “zulässige” Kontingenztafel  $[x_{ij}]$  (mit  $m \times m$ -Gitterung):

$$\min! \sum_{i=1}^m \sum_{j=1}^m (x_{ij} - a_{ij})^2 \text{ unter den Nebenbedingungen}$$

$$\sum_{i=1}^m x_{ik} = \sum_{j=1}^m x_{\ell j} = \frac{1}{m} \text{ und } x_{\ell,k} \geq 0 \text{ für } k, \ell = 1, \dots, m.$$

- explizite Lösung rechnerisch aufwändig ( $\rightarrow$  Karush-Kuhn-Tucker-Theorem), **hoher Speicheraufwand bei großem  $d$**

## Von der empirischen zur Bernstein-Copula

Suboptimale Lösung: äquivalentes Lagrange-Problem ohne die Nichtnegativitätsbedingung:

$$x_{ij} = a_{ij} - \frac{a_{\cdot j}}{m} - \frac{a_{i \cdot}}{m} + \frac{2}{m^2} \quad \text{für } i, j = 1, \dots, m,$$

Positivitäts-Korrektur:

$$[y_{ij}] = \left[ \frac{x_{ij} - a}{1 - m^2 \cdot a} \right] \quad \text{mit } a := \min\{0, x_{ij} \mid 1 \leq i, j \leq m\}$$

- Explizite suboptimale Lösung in jeder Dimension bekannt!

**! Problem tritt nicht auf, wenn  $m$  Teiler der Datenzahl ist !**

## Simulationen mit Bernstein-Copulas

Mehrdimensionales Problem: analoges Lagrange-Problem:

$$\min! \sum_{(i_1 \dots i_d) \in I^d} (x_{i_1 \dots i_d} - a_{i_1 \dots i_d})^2 \quad \text{unter } x_{\bullet[k]}(i_k) := \sum_{(i_1 \dots i_d) \in I_k^d(i_k)} x_{i_1 \dots i_d} = \frac{1}{m}$$

mit  $I_k^d(j) := \{1, \dots, m\}^{k-1} \times \{j\} \times \{1, \dots, m\}^{d-k}$

## Simulationen mit Bernstein-Copulas

Allgemeine suboptimale Lösung (vor Positivitätskorrektur):

$$x_{i_1 \dots i_d} = a_{i_1 \dots i_d} - \frac{1}{m^{d-1}} \sum_{k=1}^d a_{\bullet [k]}(i_k) + \frac{d}{m^d} \text{ für } (i_1, \dots, i_d) \in \{1, \dots, m\}^d$$

$$[y_{i_1 \dots i_d}] = \left[ \frac{x_{i_1 \dots i_d} - a}{1 - m^d a} \right] \text{ mit } a := \min \{0, x_{i_1 \dots i_d} \mid 1 \leq i_1, \dots, i_d \leq m\}$$

**! Problem tritt nicht auf, wenn  $m$  Teiler der Datenzahl ist !**

## Simulationen mit Bernstein-Copulas (I)

- Bernstein-Copula-Dichte  $c$  ist nach oben beschränkt (etwa durch  $M = m^d \max\{y_{i_1 \dots i_d} \mid 1 \leq i_1, \dots, i_d \leq m\}$ )
- Verwerfungsmethode:
  1. Erzeuge  $d + 1$  unabhängige Zufallszahlen  $u_1, \dots, u_{d+1}$ .
  2. Falls  $c(u_1, \dots, u_d) > M u_{d+1}$ , gehe zu 3, anderenfalls zu 1.
  3.  $(u_1, \dots, u_d)$  ist "Stichprobe" der Bernstein-Copula.
- durchschnittliche Stichprobenrate:  $1/M$

**„Fluch der Dimension“: geringe Effizienz mit wachsendem  $d$**

## Simulationen mit Bernstein-Copulas (II)

- Beobachtung: Bernstein-Copula-Dichte  $c$  ist Mischung von Dichten **unabhängiger, Beta-verteilter Zufallsvariablen**:

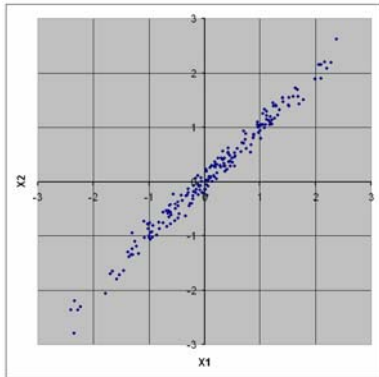
$$c(u_1, \dots, u_d) := \sum_{k_1=0}^{m_1-1} \cdots \sum_{k_d=0}^{m_d-1} p(k_1, \dots, k_d) \prod_{i=1}^d \frac{u_i^{k_i} (1-u_i)^{m_i-1-k_i}}{\text{Beta}(k_i+1, m_i-k_i)}, \mathbf{u} \in [0, 1]^d$$

- Vorteil: Direkte Inversionsmethode anwendbar!
- Gewichte  $p(k_1, \dots, k_d)$  aus empirischer Copula:  
→ Resampling der Rangzahlen
- **Aufwand extrem klein, wenn  $m$  Teiler der Datenzahl ist!**

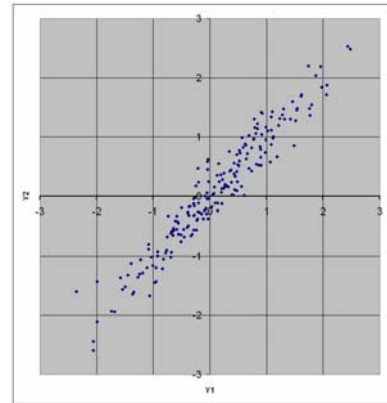


## Anwendungsbeispiele (I)

Implementierung in EXCEL, Fall  $d = 2$ :



200 Simulationen mit  
Normalverteilung,  $\rho = 0,99$



200 Simulationen mit  
Bernstein-Copula,  $m = 100$

## Anwendungsbeispiele (II)

Hochdimensionale Anwendungen:

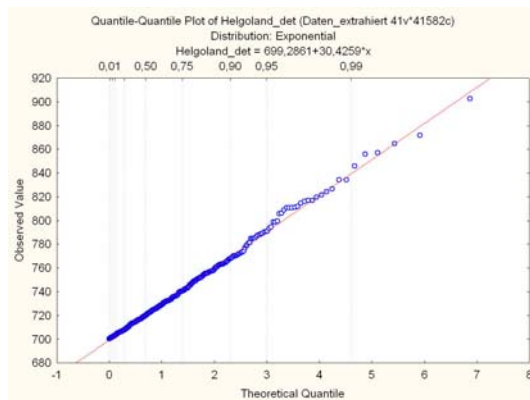
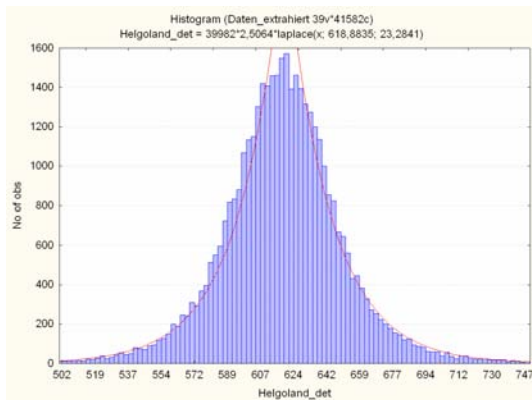
- Naturgefahren in der Schweiz: 19 Kantone mit öffentlich-rechtlicher Gebäudeversicherung, 6 Gefährdungstypen (Sturm, Hagel, Überschwemmung, Erdbeben, Lawinen, Schneedruck):

**19 x 6 = 114 Dimensionen!**

- Tägliche Pegelstände an Nord- und Ostsee (→ Sturmflutprojekt beim GDV, Berlin):

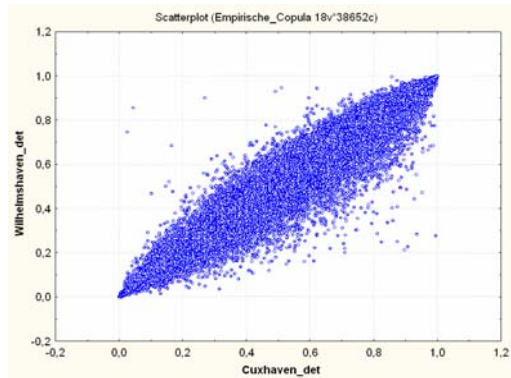
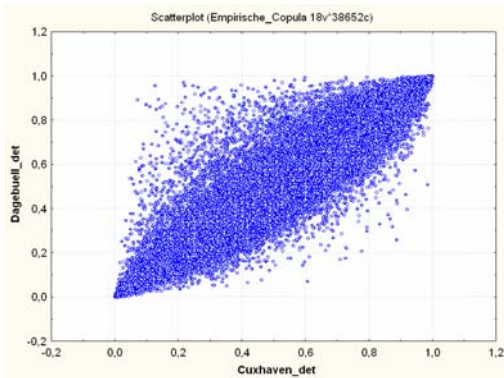
**30 Messstationen mit jeweils bis zu 40.000 Daten**

## Anwendungsbeispiele (III)



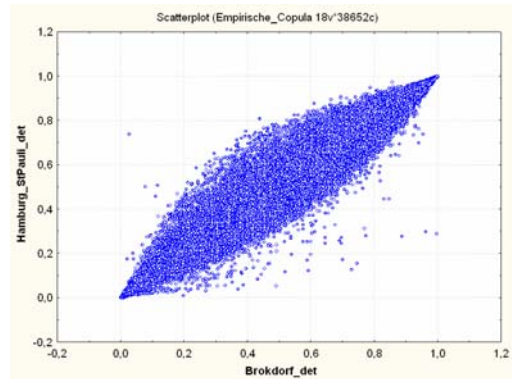
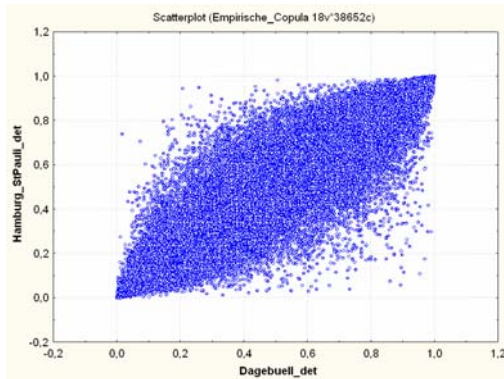
Histogramm von Pegelständen und Verteilungsanpassung mit Q-Q-Plot

## Anwendungsbeispiele (III)



Paarweise empirische Copulas

## Anwendungsbeispiele (III)



Paarweise empirische Copulas

## Literatur

- [1] AAS, K., CZADO, C., FRIGESSI A. AND BAKKEN, H. (2009): *Pair-copula constructions of multiple dependence*. Insurance, Mathematics and Economics 44, 182-198.
- [2] T. BOUEZMARNI, J. V.K. ROMBOUTS, A. TAAMOUTI (2008): *Asymptotic properties of the Bernstein density copula for dependent data*. CORE discussion paper 2008/45, Leuven University, Belgium.
- [3] V. DURRLEMAN, A. NIKEGHBALI, T. RONCALLI (2000): *Copulas approximation and new families*. Groupe de Recherche Opérationnelle, Crédit Lyonnais, France, Working Paper.
- [4] H. JOE (1997): *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.
- [5] G. KAUERMANN, C. SCHELLHASE, C. AND D. RUPPERT (2011): *Flexible Copula Density Estimation with Penalized Hierarchical B-Splines*. Preprint.
- [6] T. KULPA (1999): *On approximation of copulas*. Internat. J. Math. & Math. Sci. 22, 259 – 269.
- [7] A.J. MCNEIL , R. FREY,P. EMBRECHTS (2005): *Quantitative Risk Management*. Concepts, Techniques, Tools. Princeton University Press, Princeton.
- [8] R.B. NELSEN (2006): *An Introduction to Copulas*. 2<sup>nd</sup> Ed., Springer, N.Y.R.D. REISS, M. THOMAS (2001): *Statistical Analysis of Extreme Values*. With Applications to Insurance, Finance, Hydrology and Other Fields. 2<sup>nd</sup> Ed., Birkhäuser, Basel.
- [9] O. OKHRIN, Y. OKHRIN AND W. SCHMID (2011): *Determining the structure and estimation of hierarchical Archimedean copulas*. Under revision in Journal of Econometrics.
- [10] D. PFEIFER, D. STRABBURGER, J. PHILIPPS (2011): *Modelling and simulation of dependence structures in nonlife insurance with Bernstein copulas*. Revised Preprint.