

Stochastische Modellierung von Katastrophen-Schäden

P. Ender

Institut für Mathematische Stochastik, Universität Karlsruhe (TH)

Zusammenfassung

Der Gesamtschaden einer Naturkatastrophe aggregiert sich aus Einzelschäden. Diese Einzelschäden sind unterschiedlichen Typs wie Gebäudeschäden, Schäden am Inventar, Folgeschäden etc. Der Zufallscharakter der Schäden führt zur Aufgabe der stochastischen Modellierung mehrdimensionaler Daten.

Ein stochastisches Modell für mehrdimensionale (Schaden-)Daten besteht aus den Randverteilungen und einer Kopula(-funktion). In diesem Modell werden die Randverteilungen und die Kopula unabhängig voneinander bestimmt. Die Kopula beschreibt die Abhängigkeit der Daten vollständig. Jede Kopula beschreibt eine Form der Abhängigkeit: Unabhängigkeit wird durch die Produkt-Kopula dargestellt. Eine positive Abhängigkeit im Extrembereich der Daten kann durch die Gumbel-Kopula modelliert werden.

Bei vorliegenden Daten über Katastrophen-Schäden stellt sich die wichtige Frage, welche Kopula die (Abhängigkeit der mehrdimensionalen) Schaden-Daten am „besten“ modelliert. Diese Frage führt zu einem statistischen Anpassungsproblem.

An einem Beispiel (volkswirtschaftliche Schäden und versicherte Schäden großer Naturkatastrophen der Versicherungsgeschichte) wird das beschriebene Modell angewendet und eine Anpassung an eine Kopula durchgeführt.

Abstract

The total damage caused by a disaster is the sum of many individual damages. The individual damages are of different types, such as building damage, inventory damage, collateral damage etc. A common feature of all damages is that they are random. This observation calls for a multivariate stochastic model for the different types of damages.

We propose a model that consists of the marginal distributions (of the individual damages) and a copula(-function). The advantage of this model is, that we can choose the marginal distributions and the copula independently of each other. The copula completely determines the dependence structure of the damage types.

Independence is modelled by the product copula. The Gumbel copula models a positive dependence in the upper tails.

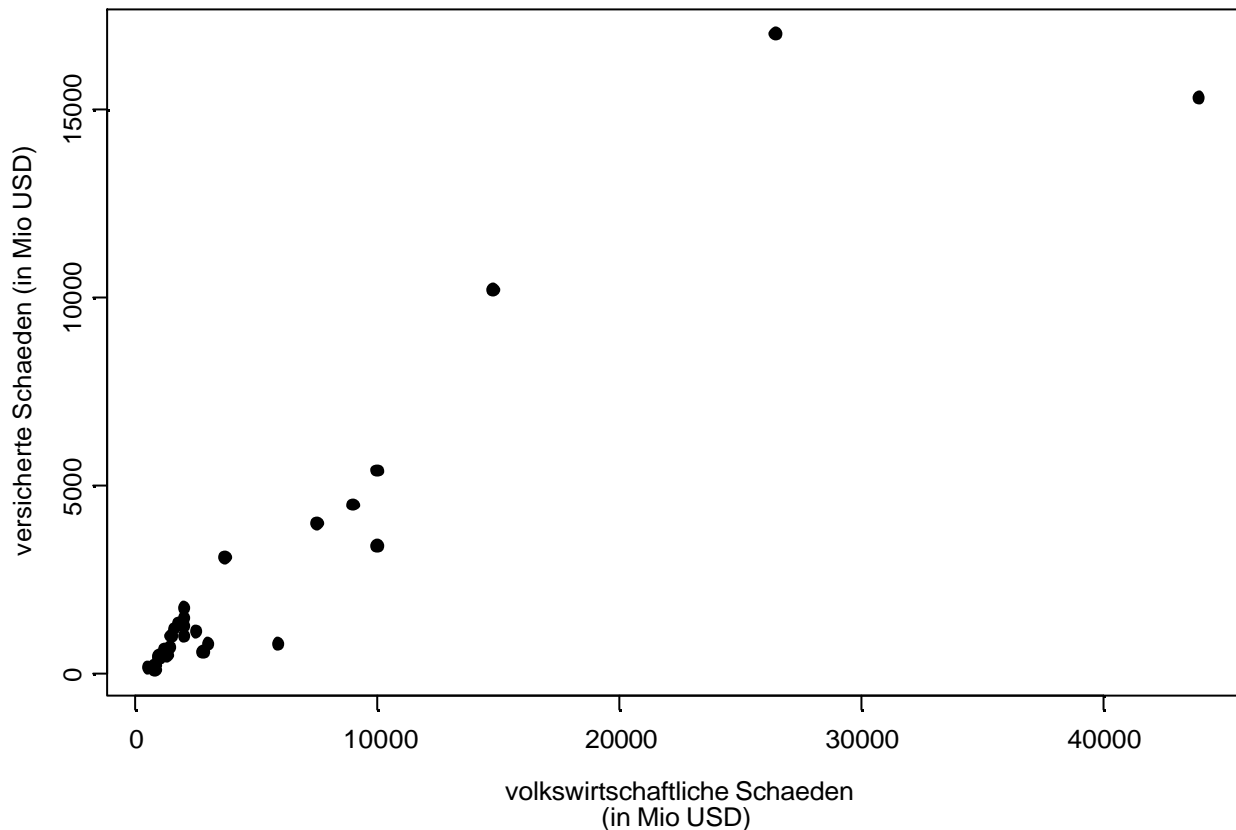
There are many more families of copula functions. So, given a data set of disaster damages, there is the natural question: Which copula is the “right” one? A statistical goodness-of-fit test can “answer” this question.

We illustrate this stochastic model and the fitting of a copula by an example of the economic and insured losses caused by great natural disasters.

Große Naturkatastrophen der Versicherungsgeschichte

Auf der CD-ROM „Welt der Naturgefahren“ der Münchener-Rück können die großen Naturkatastrophen der Versicherungsgeschichte abgefragt werden. Ein Plot der versicherten Schäden über den volkswirtschaftlichen Schäden ist in folgender Abbildung dargestellt.

Grosse Naturkatastrophen



In den volkswirtschaftlichen Schäden sind die versicherten Schäden inbegriffen, deshalb seien

- X = volkswirtschaftliche Schäden ohne versicherte Schäden
- = nichtversicherte Schäden,
- Y = versicherte Schäden.

Die Korrelation zwischen X und Y untersuchen wir mit dem Korrelationskoeffizienten nach Kendall t . Für die empirische Version von t erhalten wir den Wert 0.365 . Ein Test auf Unabhängigkeit ergibt den p-Wert 0.006 . Deshalb können die (Zufalls-)Größen X und Y als signifikant positiv korreliert angesehen werden.

Bevor wir die Daten über die Katastrophen-Schäden weiter untersuchen, geben wir ein stochastisches Modell für den Zufallsvektor (X, Y) der nichtversicherten und versicherten Schäden an.

Stochastisches Modell

Eine *Kopula* ist eine bivariate Verteilungsfunktion auf dem Einheitsquadrat, welche gleichverteilte Randverteilungen besitzt. Es seien X und Y Zufallsvariablen mit stetiger Verteilungsfunktion $F(x) = P(X \leq x)$ ($P(X \leq x)$ ist Wahrscheinlichkeit, dass die Zufallsvariable X eine Realisierung unterhalb des Wertes x annimmt.) bzw. $G(y) = P(Y \leq y)$ und gemeinsamer Verteilungsfunktion $H(x, y) = P(X \leq x, Y \leq y)$. Dann existiert nach Nelson (1998), Theorem 2.3.3 eine (eindeutige) Kopula C mit der Eigenschaft

$$H(x, y) = C(F(x), G(y)).$$

Bei der Auswahl der Kopula für unser statistisches Anpassungsproblem beschränken wir uns auf zwei Kopula-Familien, die *Clayton*- oder *Pareto*-Kopula mit Parameter $J \geq 0$ (welche aus einer bivariaten Pareto-Verteilung abgeleitet wird), und der *Gumbel*-Kopula mit Parameter $J \geq 1$ (welche aus einer bivariaten Extremwert-Verteilung abgeleitet wird).

Anpassung einer Kopula

Welche der beiden Kopula-Familien ist ein „besseres“ Modell für unsere Daten der großen Naturkatastrophen?

Beide Kopula-Familien haben den Vorteil, dass man den jeweiligen Parameter aus der empirischen Version von t unter Zuhilfenahme eines eindeutigen Zusammenhangs zwischen dem jeweiligen Parameter und t schätzen kann. Damit ist dieser Schätzer unabhängig von den Randverteilungen.

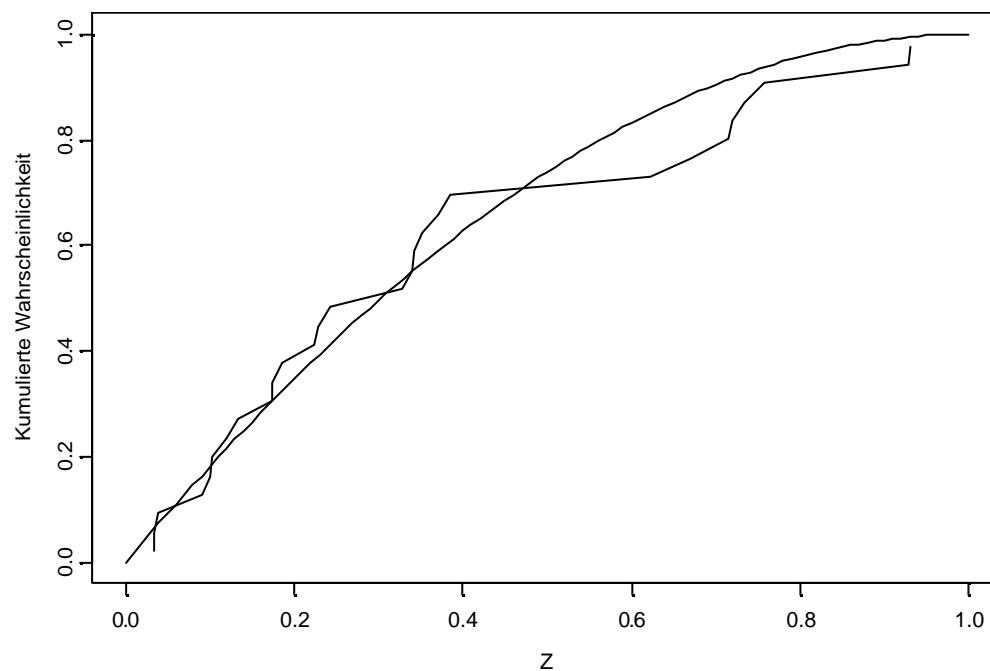
Bei der Clayton-Kopula ergibt sich ein Schätzwert von 1.15 für den Parameter J und bei der Gumbel-Kopula ein Schätzwert von 1.575 für den Parameter J .

Es gibt einige grafische Methoden zur Bestimmung der Güte einer angepassten Kopula, wovon wir eine jetzt vorstellen:

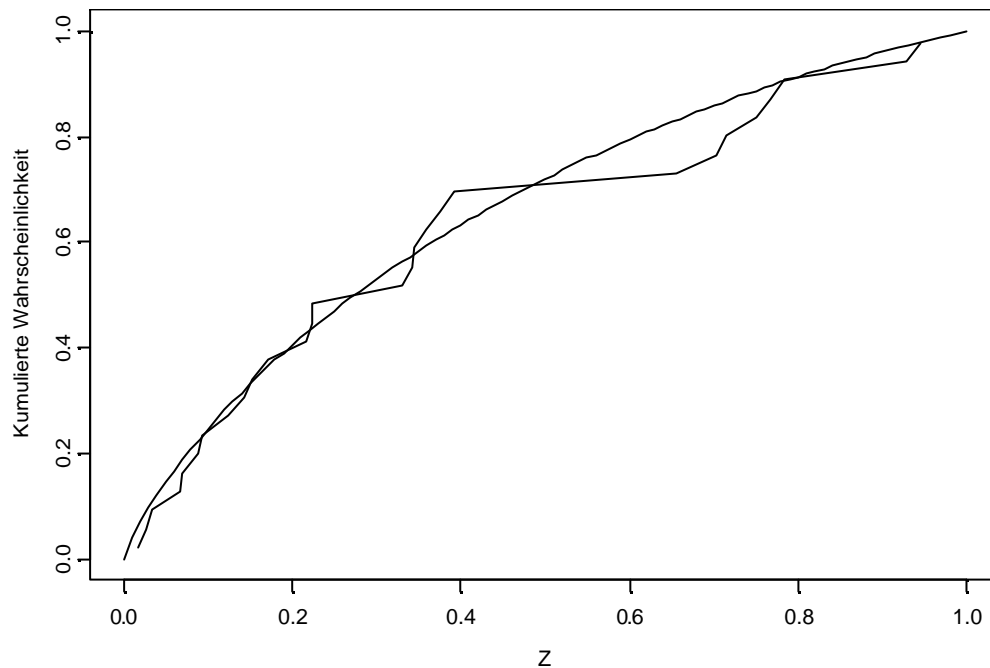
Der bivariaten Zufallsvektor (X, Y) wird durch die „Transformation“ $Z = C(F(X), G(Y))$ auf eine (eindimensionale) Zufallsgröße umgewandelt, welche die Verteilungsfunktion $K(z) = P(Z = z)$ besitzen soll. Diese spezielle Verteilungsfunktion kann zum einen

- mit der aus den „neuen Daten“ Z_1, \dots, Z_n gewonnenen empirischen Verteilungsfunktion
- oder
- mit der unter der Annahme einer speziellen Kopula-Familie geschätzten Verteilungsfunktion geschätzt werden. Liegen beide Schätzungen für K „nahe beieinander“, so ist die Anpassung „gut“, ansonsten „schlecht“.

Empirische und geschätzte Verteilungsfunktion für K bei einer angepassten Clayton-Kopula:



Empirische und geschätzte Verteilungsfunktion für K bei einer angepassten Gumbel-Kopula:



Die Gumbel-Kopula passt die Daten im „oberen Bereich“ von Z besser an als die Clayton-Kopula. Durch einen Anpassungstest können wir diese „Vermutung“ weiter untersuchen.

Anpassungstest für Kopula-Funktionen

Es sei die bivariate Stichprobe $(X_1, Y_1), \dots, (X_n, Y_n)$ als Folge unabhängiger Zufallsvektoren mit **unbekannter** gemeinsamer Verteilungsfunktion H und **unbekannter** Kopula C gegeben. Weiter sei $\{C(\cdot; \mathbf{J}) : \mathbf{J} \in \Theta\}$ eine Kopula-Familie (z.B. Clayton oder Gumbel) mit einem reellen Parameterbereich Θ .

Zu testen ist die **(Null-)Hypothese**

$$H_0 : C \in \{C(\cdot; \mathbf{J}) : \mathbf{J} \in \Theta\}$$

(d.h. es existiert ein $\mathbf{J}_0 \in \Theta$ mit $C = C(\cdot; \mathbf{J}_0)$) gegen die **Alternative**

$$H_1 : C \notin \{C(\cdot; \mathbf{J}) : \mathbf{J} \in \Theta\}.$$

Ein allgemeines Test-Verfahren für einen Test auf Kopula-Funktionen besteht aus den folgenden vier Schritten:

1. Bilde die empirische Kopula C_n als nichtparametrischen Schätzer für die unbekannte Kopula.
2. Mit einem Schätzer $\hat{\mathbf{J}}$ für den Parameter \mathbf{J} bilde einen parametrischen Schätzer $C(\cdot; \hat{\mathbf{J}})$ aus der Kopula-Familie $\{C(\cdot; \mathbf{J}) : \mathbf{J} \in \Theta\}$.
3. Bestimme das Abstandsmaß $T_n = \iint (C_n(u, v) - C(u, v; \hat{\mathbf{J}}))^2 dudv$ und benutze T_n als Teststatistik.
4. Lehne H_0 ab, falls T_n einen „kritischen“ Wert, welcher vom Stichprobenumfang n und einem Signifikanzniveau α (z.B. $\alpha = 5\%$) abhängt, überschreitet.

Eine Anwendung auf unser Beispiel der großen Naturkatastrophen ergibt bei einem Stichprobenumfang von $n=28$ die Werte für das Abstandsmaß nach folgender Tabelle.

Kopula	Abstandsmaß T_{28}
Clayton	0.0326
Gumbel	0.0068

Die Kopula-Familie, welche den kleineren Wert für das Abstandsmaß liefert, ist eine „bessere“ Anpassung. In unserem Beispiel ist demzufolge die Gumbel-Kopula eine „bessere“ Anpassung als die Clayton-Kopula.

Weitere Untersuchungen

1. Bestimmung der (Grenz-)Verteilung der Teststatistik.
2. Untersuchungen der Güte des Tests (z.B. bei festen Alternativen).
3. Entwicklung eines Test auf Tail-Abhängigkeit.

Literatur

- Klugman, S. A. und R. Parsa, 1999: Fitting bivariate loss distributions with copulas. Insurance Mathematics & Economics 24 (1999), 139 - 148.
- Nelson, R. B., 1998: An Introduction to Copulas. Springer, New York.