Insurance: Mathematics and Economics 50 (2012) 430–436

Contents lists available at SciVerse ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Dependence modeling in non-life insurance using the Bernstein copula[☆]

Dorothea Diers[a], Martin Eling[b], Sebastian D. Marek[c,*]

[a] *Provinzial NordWest Holding AG, 48143 Münster, Germany*

[b] *Institute of Insurance Economics, University of St. Gallen, 9010 St. Gallen, Switzerland*

[c] *Institute of Insurance Science, University of Ulm, 89069 Ulm, Germany*

ARTICLE INFO

Article history:
Received July 2011
Received in revised form
February 2012
Accepted 10 February 2012

Keywords:
Non-life insurance
Copulas
Bernstein copula
Goodness-of-fit
Simulation

ABSTRACT

This paper illustrates the modeling of dependence structures of non-life insurance risks using the Bernstein copula. We conduct a goodness-of-fit analysis and compare the Bernstein copula with other widely used copulas. Then, we illustrate the use of the Bernstein copula in a value-at-risk and tail-value-at-risk simulation study. For both analyses we utilize German claims data on storm, flood, and water damage insurance for calibration. Our results highlight the advantages of the Bernstein copula, including its flexibility in mapping inhomogeneous dependence structures and its easy use in a simulation context due to its representation as mixture of independent Beta densities. Practitioners and regulators working toward appropriate modeling of dependences in a risk management and solvency context can benefit from our results.

1. Introduction

Using copulas in risk management has become popular in both academia and practice recently. Copula applications are presented for modeling dependence between stock returns (Jondeau and Rockinger, 2006), CDO pricing (Hofert and Scherer, 2011), currency option pricing (Salmon and Schleicher, 2006), or internal risk models (Eling and Toplek, 2009); further areas of application can, e.g., be found in Genest et al. (2009a). However, several popular copulas, such as elliptical and Archimedean copulas, exhibit a certain degree of symmetry or are restricted to certain correlation structures, which is not always suitable for risk modeling in practice.

The Bernstein copula, which has only recently received attention in an insurance context (Pfeifer et al., 2009), has the potential to overcome these drawbacks while still being applicable in higher dimensions. It is a flexible, non-parametric copula capable of approximating any copula arbitrarily well and thus may serve as a model for an unknown underlying dependence structure (Sancetta and Satchell, 2004). We explore this flexibility in a realistic environment by fitting the Bernstein copula to empirical claims data from six lines of business and simulating the aggregate value-at-risk and tail-value-at-risk. Our data are from lines of business driven by exposure to natural perils. The resulting portfolio is difficult to model since different lines are combined, resulting in an inhomogeneous dependence structure.

Goodness-of-fit is an aspect usually not considered in the literature on copula modeling and calibration (Embrechts, 2009). Thus, this paper contributes to the literature by implementing the Bernstein copula in a higher dimension, assessing its goodness-of-fit in the modeling of dependence structures of non-life insurance risks, and illustrating its use in a simulation context. For this purpose we use the representation of the Bernstein copula as a mixture of Beta densities.[1] This representation facilitates an efficient random sampling algorithm, which has, to our knowledge, so far not been applied in the context of the Bernstein copula.

To preview our results, we show that the Bernstein copula performs especially well when multiple risk classes with inhomogeneous dependence structure are combined. We conclude that the Bernstein copula is a promising alternative for modeling dependence structures in internal risk models. Practitioners working on

☆ We are grateful to an anonymous referee for valuable suggestions. We also thank Martin Hampel, Christian Hering, Andreas Niemeyer, Thomas Parnitzke, Dietmar Pfeifer, Jan-Philipp Schmidt, and the participants of the 2010 World Risk and Insurance Economics Congress, the 2011 Annual Meeting of the German Insurance Science Association, and the 15th International Congress on Insurance: Mathematics and Economics for their comments. Furthermore, we acknowledge use of the Common Ulm Stuttgart Server project's computing facilities.

* Corresponding author.
E-mail address: sebastian.marek@uni-ulm.de (S.D. Marek).
URL: http://www.uni-ulm.de/mawi/ivw (S.D. Marek).

[1] The representation of densities based on Bernstein polynomials as mixture of Beta distributions is an already established result; for instance, Sancetta and Satchell (2004) use it in their representation of the Bernstein copula's Spearman's rho. We are grateful to an anonymous referee for making us aware of this representation.

0167-6687/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.insmatheco.2012.02.007

calibration and implementation of such models, as well as regulators responsible for the validation of internal risk models, can benefit from our results.

The remainder of this paper is organized as follows. In Section 2, we present the analyzed copulas, parameter estimation, and sampling algorithms. In Section 3, we introduce the data. Section 4 reports the results of the goodness-of-fit analysis. In Section 5, we show the value-at-risk and tail-value-at-risk simulation results. We conclude in Section 6.

2. Analyzed copulas

2.1. Motivation for analyzing the Bernstein copula

We use the non-parametric Bernstein copula to analyze its potential for solving selected problems in the application of standard copula approaches (symmetry, parameter restrictions, applicability in high dimensions). The type of problem considered here is of high importance in light of Solvency II, which requires adequate modeling of the dependences between different types of risks in an insurance company. In this context, Eling and Toplek (2009) discuss various elliptical and nested Archimedean copulas. We extend that analysis with a flexible modeling alternative that is easy to use and readily calibrated to empirical data.

There are many other non-parametric copulas, such as the grid-type copula (Pfeifer et al., 2009; Kulpa, 1999), the box copula (Hummel, 2009), or the Fourier copula (Lowin, 2010). We focus on the Bernstein copula for four reasons. First, it recently has attracted attention in insurance modeling and is thus a natural candidate for further analysis. In this context, Pfeifer et al. (2009) focus on the mathematical properties and calculation of the Bernstein copula and apply it to a two-dimensional dataset of insurance claims. We use the Bernstein copula in a higher-dimensional framework and analyze its statistical fit. Second, the Bernstein copula is attractive from a modeling perspective. Standard copula approaches from the elliptical and Archimedean class provide a certain degree of symmetry or are restricted to certain correlation structures, which may not always be desirable. The Bernstein copula is not bound to these limitations and thus can provide a more adequate estimate of the underlying dependence structure. Third, because calibration and simulation efforts do not increase exponentially with the dimension, the Bernstein copula is also suitable in higher dimensions, which is a major advantage compared to other parametric and non-parametric estimators. Fourth, its mathematical properties are interesting as the Bernstein estimator converges to the underlying dependence structure, provides a higher rate of consistency than other common non-parametric estimators (Sancetta and Satchell, 2004; Kulpa, 1999), and does not suffer from boundary bias as do kernel-based copulas.

However, the Bernstein copula also has the same disadvantages like other non-parametric estimators (e.g. the bias-variance tradeoff). Even though it can approximate any behavior in the tail, it cannot model asymptotic tail dependence (Sancetta and Satchell, 2004). Approximation quality may thus vary.

2.2. Considered copulas

In this analysis, we consider the independence copula as well as various parametric and non-parametric copulas. Specifically, we analyze the parametric elliptical (Gauss and Student) and Archimedean (Gumbel and Clayton) classes and the Bernstein, grid-type, and kernel copulas as non-parametric alternatives.

The *independence copula*, in which all risk classes are considered as independent, serves as the benchmark. The implementation of dependence structures and copulas in internal risk models increases model complexity and thus costs. Therefore, more complex copulas should at least perform better than this benchmark.

From the class of *elliptical copulas* we choose the Gauss and the Student copulas. Stemming from elliptical distributions, both induce symmetric dependence structures, which may not always be suitable in an insurance context with possibly erratic or clustered claim realizations.

From the class of *Archimedean copulas* we consider the Clayton and Gumbel copulas. For this class there are two possible setups: the exchangeable case and the non-exchangeable case. The exchangeable, single-parametric case is less adequate in our context because it results in identical margins and dependence among all risk classes. In the non-exchangeable case, nested Archimedean copulas (NACs) couple multiple Archimedean copulas with different generators or, as a simplification, with the same generator. This construction better reflects the dependence situation in our data and will be described in more detail in Section 3.

As non-parametric copulas we consider the Bernstein copula, the closely related grid-type copula, and a kernel-based copula. For the *Bernstein* and *grid-type copula*, we use the notation from Pfeifer et al. (2009) with $d \in N$ denoting the dimension of the copula (i.e., the number of considered risk classes), $n \in \mathbb{N}$ the sample size used for calibration of the copula, and $m_i \in \mathbb{N}$ the grid size for $i = 1, \ldots, d$. Let $T_i = \{0, 1, \ldots, m_i - 1\}$ and $I_{k_1,\ldots,k_d} := \times_{i=1}^{d} \left(\frac{k_i}{m_i}, \frac{k_i+1}{m_i} \right]$ for all possible choices of $(k_1, \ldots, k_d) \in \times_{i=1}^{d} T_i$. Thus $I_{k_1,\ldots,k_d}$ describes a grid over the $[0, 1]^d$ hypercube with $\prod_{i=1}^{d} m_i$ cells. Let $U = (U_1, \ldots, U_d)$ be some discrete random vector with uniform margins over T_i. With $p(k_1, \ldots, k_d) := P\left(\bigcap_{i=1}^{d} \{U_i = k_i\} \right)$ and $(u_1, \ldots, u_d) \in [0, 1]^d$, the general Bernstein and grid-type copula can be given as

$$c(u_1, \ldots, u_d) := \sum_{k_1=0}^{m_1-1} \cdots \sum_{k_d=0}^{m_d-1} p(k_1, \ldots, k_d) \prod_{i=1}^{d} m_i \phi(m_i, k_i, u_i).$$

For $\phi(m, k, u) = B(m-1, k, u) := \binom{m-1}{k} u^k (1-u)^{m-1-k}$, i.e., the Bernstein polynomials, we receive the Bernstein density, and for $\phi(m, k, u) = \mathbb{1}_{\left(\frac{k}{m}, \frac{k+1}{m} \right]}(u)$, i.e., the indicator function, we receive the grid-type density. For this analysis, we choose $m_i \equiv m$ constant.

For each of the cells $I_{k_1,\ldots,k_d}$, a probability estimate can be obtained, which is summarized in a d-dimensional contingency table p. The Bernstein polynomials can be interpreted as smoothing functions that disperse parts of the probability mass to surrounding cells. The degree of the polynomials determines the intensity of smoothing.

For the *kernel copula* we use a setup adopted from Fermanian and Scaillet (2003) with a Gaussian kernel. We evaluate the kernel copula by $C(u_1, \ldots, u_d) = \hat{F}(\hat{\xi}_1, \ldots, \hat{\xi}_d)$, with $\hat{F}$ as multivariate Gaussian kernel distribution function and $\hat{\xi}_i$ as a kernel-based estimate of the quantile with probability level u_i of the i-th data vector. This construction corrects to some extent the boundary bias inherent in kernel-based copulas. There are alternatives that may provide better performance; for example, Chen and Huang (2007) propose local-linear kernels and Bouezmarni and Rombouts (2009) suggest a semi-parametric setup with non-parametric margins and parametric copula functions. For our higher-dimensional analysis we need, however, an efficient sampling algorithm which is not yet available for these modified estimators. For this reason we stay with the Fermanian and Scaillet (2003) setup.

2.3. Copula estimation

For parameter estimation of elliptical copulas we rely on the canonical maximum likelihood (CML) method. We disregard the