European Journal of Operational Research 238 (2014) 476–485

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

Construction of asymmetric copulas and its application in two-dimensional reliability modelling

Shaomin Wu *

Kent Business School, University of Kent, Canterbury, Kent CT2 7PE, United Kingdom

ARTICLE INFO

Article history:
Received 18 August 2013
Accepted 10 March 2014
Available online 18 March 2014

Keywords:
Copula
Tail dependence
Warranty
Two-dimensional reliability data
Asymmetric

ABSTRACT

Copulas offer a useful tool in modelling the dependence among random variables. In the literature, most of the existing copulas are symmetric while data collected from the real world may exhibit asymmetric nature. This necessitates developing asymmetric copulas that can model such data. In the meantime, existing methods of modelling two-dimensional reliability data are not able to capture the tail dependence that exists between the pair of age and usage, which are the two dimensions designated to describe product life. This paper proposes a new method of constructing asymmetric copulas, discusses the properties of the new copulas, and applies the method to fit two-dimensional reliability data that are collected from the real world.

1. Introduction

1.1. Motivation

Copulas are a tool for constructing multivariate distributions and formalising the dependence structures between random variables. The notion of copula was first introduced by Abe Sklar in 1959 when he responded to a question with respect to the relationship between a multidimensional probability function and its lower dimensional margins (Sklar, 1959). It has attracted considerable attention in recent years in both theoretical and application aspects.

Sklar's Theorem states that any cumulative distribution function of a random vector can be written in terms of marginal distribution functions and a copula that describes the dependence structure between the variables (Sklar, 1959). That is, given a vector of random variables $(X_1, \ldots, X_d)$, its cumulative distribution function $H(x_1, \ldots, x_d)(= P(X_1 \leqslant x_1, \ldots, X_d \leqslant x_d))$, and marginals $F_k(x_k)(= P(X_k \leqslant x_k)$, where $k = 1, \ldots, d$, Sklar proved that $H(x_1, \ldots, x_d)$ can be written as $H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$ and named $C(\cdot)$ as a *copula* (Sklar, 1959). Copulas are useful in statistical applications because they allow one to easily model and estimate the distribution of a random vector through estimating the marginals and the copula separately.

In the literature, many parametric copula families have been proposed (Joe, 1996; Nelsen, 1999). Most of them are symmetric in the sense that the variables in a copula are exchangeable, $C(v_1, v_2) = C(v_2, v_1)$ $(v_1, v_2 \in [0, 1])$ for the bivariate copula case, for example. However, there are many natural processes that possess asymmetric dependence structures. Using symmetric copulas to model such processes may not be able to capture the nature of the data. This necessitates developing asymmetric copulas. This paper serves this necessity by developing a method of constructing asymmetric copulas.

1.2. Related work

1.2.1. Construction of asymmetric copulas

There are many copula families. The reader is referred to the monographs by Joe (1996) and Nelsen (1999) for detailed accounts of the theory and surveys of commonly used copulas and to the review papers (Embrechts, 2009; Genest, Remillard, & Beaudoin, 2009; Kolev, Dos Anjos, & Mendes, 2006), work on tail dependence (Hua & Joe, 2011; Zhang, 2008), and papers on applications (Grundke & Polle, 2012; Ye, Liu, & Miao, 2012).

Some work has been done for constructing asymmetric copulas ((Alfonsi & Brigo, 2005; Liebscher, 2008), for example). Alfonsi and Brigo (2005) described a method based on periodic functions. Liebscher (2008) introduced two methods to construct asymmetric multivariate copulas: the first is connected with products of copulas and the second generalises the Archimedean copulas, and the resulting copulas are asymmetric. However, their research did

* Tel.: +44 1227 827 940.
E-mail address: s.m.wu@kent.ac.uk

http://dx.doi.org/10.1016/j.ejor.2014.03.016
0377-2217/

not show how those methods can be used to construct asymmetric copulas with tail dependence in a given direction.

1.2.2. Modelling two-dimensional reliability data

Estimating reliability functions is important for both product manufacturers and asset managers. System reliability is usually estimated based on one dimension, either usage (for example, the number of door openings of an elevator, the number of pages copied by a copy machine) or age. In reality, however, failures or other adverse events in systems may depend on both the age and the usage history of the systems. Such examples can be found in warranty claim data analysis, in which the warranty policy for certain types of products specifies the limits of coverage in terms of both age and usage. For example, a sold car may be covered by a 3 year and 30,000 mile warranty, which implies that the warranty supplier—which is usually the manufacturer—will repair the car if it fails within the warranty coverage. It is important for the manufacturer to estimate the reliability of their products for the purpose of product improvement and fiscal planning. There is considerable research on warranty data analysis. For more detailed discussion, the reader is referred to recently published review papers (Wu, 2012, 2013) and papers on reliability data analysis (see (Anastasiadis, Anderson, & Chukova, 2013; Wu & Akbarov, 2011, 2012), for example).

There are three methods that have been developed for analysing two-dimensional reliability data. Those methods are briefed below.

The univariate method. It indirectly estimates $F(x_1, x_2)$ through the following two steps: to estimate $F_2(x_2|x_1)$ and $F_1(x_1)$, separately, then to obtain $F(x_1, x_2) = F_2(x_2|x_1)F_1(x_1)$ (where x_1 and x_2 represent age and usage, respectively). This method treats the usage as a function of the age (Alam & Suzuki, 2009; Lawless, Crowder, & Lee, 2009).

The bivariate method. It directly estimates $F(x_1, x_2)$ from data. For example, Jung and Bai (2007) developed a bivariate method and assumed that the bivariate distribution can describe the positive correlation between the age and the usage.

The time scale method. In addition to the above two methods, Gertsbakh and Kordonsky (1998) proposed a method that integrates the two scales (age and usage) to create a single composite scale and failures are modelled as a counting process. For example, in Gertsbakh and Kordonsky (1998), a new variable $Z = \alpha X_1 + (1-\alpha)X_2$ is introduced, where $\alpha \in (0,1), X_1$ and X_2 are the random variables representing the age and the usage, respectively.

However, the above methods have the following drawbacks.

- Both the univariate and the time scale methods assume a relationship, for example, a linear relationship, between the usage and the age, and then derive a bivariate joint lifetime distribution. A drawback of those two methods is apparent because such an assumption may be violated.
- Compared with the univariate method and the time scale method, the bivariate method is simpler and more straightforward, in the sense that the former two methods need two steps: first to estimate the relationship between age and usage and then estimate a joint distribution, but the bivariate method needs only one step that directly estimates a joint distribution. When the bivariate method is utilised, however, one needs to select a bivariate distribution in which the two variables are positively correlated. Furthermore, when the usage is measured as a discrete random variable (the number of pages copied by a copy machine, for example) and the age is described as a continuous random variable, to find a bivariate distribution that combines a discrete marginal with a continuous marginal may be difficult.
- In addition to the above two drawbacks, an important fact has never been addressed in the existing literature so far. This fact is the existence of the *tail dependence* between age and usage, as explained below.
  - (a) If the age of the product is small, its usage should be small. This is because the age is the calendar time and it is not possible to develop large cumulative usage within a short period of the calendar time. Another reason is due to the operating limit, for example, a car usually cannot be driven faster than 100 miles per hour, hence the usage within a time interval is limited.
  - (b) If the age is large, on the other hand, the usage can be small. For example, some cars are not frequently used. Hence, although they are very old, their mileage can be very small.

Those two points (a) and (b) can also be observed from Fig. 1, which is a scatterplot with the age on the *X*-axis and the usage on the *Y*-axis of the warranty claims of a particular model of car. The data are collected from a car manufacturer. As can be seen, there is no observation in the left-upper region in the figure (marked with the ellipse), which agrees with point (a). But there are some observations in the right-lower region in the figure (marked with the star), which agrees with point (b).

The above two points (a) and (b) can also be observed in those warranty data illustrated in Fig. 2 of Alam and Suzuki (2009) and in Fig. 2 of Rai and Singh (2004).

1.3. Contribution and importance of this work

This paper proposes a new method to construct asymmetric copulas and then applies the method to model two-dimensional reliability data. Its contribution is summarised in the following.

- It proposes a new method of constructing asymmetric copulas and a convex-combination of asymmetric copulas that can exhibit different tail dependence along different directions.
- It is the first paper to use asymmetric copulas to model the lifetime distribution on two-dimensional reliability data.

The proposed method has the following merits.

Ability to handle tail dependence. The proposed method of construction of asymmetric copulas differs from other existing methods in the sense that it can directly construct a copula with tail dependence in a given direction.

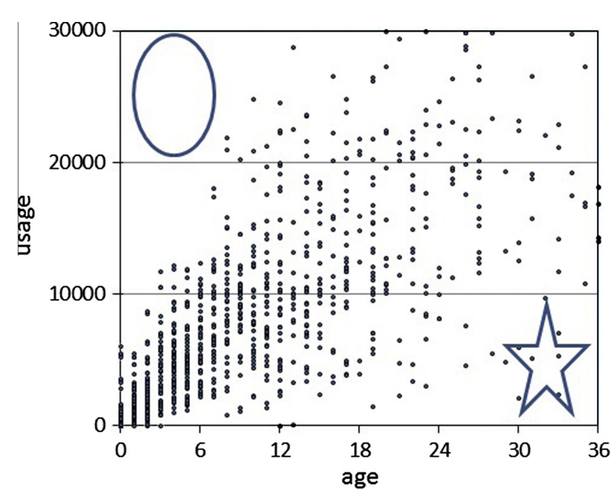

Fig. 1. Warranty claim data.