Extremes (2006) 9: 55–62
DOI 10.1007/s10687-006-0024-9

Copulas: Tales and facts—rejoinder

Thomas Mikosch

Published online: 28 October 2006

I am grateful for the contributions to the discussion. Independently of whether the discussants and I do or do not agree on some of the points in my paper I appreciate that they gave serious thoughts to the problems raised. This discussion is only the beginning of an honest analysis of the pros and cons of the use of multivariate distributions with non-symmetric marginals. In the course of this discussion, I have learned more about multivariate distributions. But I am as uncertain as before as to how one would answer the question: do we better understand stochastic dependence and risks by using copulas?

The contributions to this discussion, the literature cited therein and recently published papers show that there has been progress in understanding the probabilistic and statistical problems of copula modeling. *Johan Segers* points out that the pseudo likelihood inference approach (using the empirical distributions for the marginals and assuming a parametric copula) is in general not efficient. *Liang Peng* mentions that more efficient estimators can be found e.g., for elliptical copulas. *Laurens de Haan* mentions that there is recent progress on non-parametric estimation of the probability of multivariate failure sets far out in the tails.

I would like to thank *Paul Embrechts* for his contribution which is diplomatic, pedagogic and insightful, as always. He mentions correctly that risk measures (such as VaR) can never be estimated precisely. There is no doubt about this because the estimation depends on the data and on the model

T. Mikosch (✉)
Laboratory of Actuarial Mathematics, University of Copenhagen,
Universitetsparken 5, 2100 Copenhagen, Denmark
e-mail: mikosch@math.ku.dk

chosen. Every model has its limitations, but some models contain more grains of truth than others. What I have been missing in the copula discussion so far is this particular modeling aspect, besides problems with statistical estimation. In financial practice, whether we like it or not, risk managers do what they are told by people like us. They do not automatically use a two-stage procedure: fit the marginals, fit the copula. They do it because we told them. As much as we taught them before that the multivariate Gaussian distribution is a great model and therefore, correlations are all you need to know.

Many thanks to *Harry Joe* for his contribution. I like the expressed thoughts which are motivated by statistical experience and the handling of real-life multivariate data. Many readers will benefit from the ideas and references on statistical problems mentioned in this contribution. In particular, the statement, *it is mainly for multivariate extreme value inference that I find copulas most useful*, is what most people in extreme value theory will agree with. There is one point where I want to comment. I do not think that the suggestion of simulating from a multivariate distribution by using conditional distributions is a good approach in general. For example, one would not simulate a multivariate Gaussian or elliptically distributed vector in this way. There are only a few situations where we know the marginal distributions and the conditional probabilities. We cannot assume this when dealing with most stochastic processes of interest, e.g., solutions to stochastic differential equations (SDEs). (Of course, Professor Joe says that copulas and stochastic processes do not fit....) By now, there exists a wide range of literature on the numerical solution of SDEs, see e.g., Kloeden and Platen (1992). Further, the proposed method is inefficient in particular when large dimensions are involved and when rare events are concerned. In these cases, special simulation methods have been developed among them variance reduction methods such as importance sampling; see e.g., Glasserman (2004). There exists a vast literature on the simulation of particular stochastic processes.

I am grateful to *Christian Genest and Bruno Rémillard* that they care about my reputation. I am known as a quarrelsome person who sometimes asks questions when mathematical fashion is at its top and, of course, not everybody likes this in particular when he or she is swimming with the stream with new clothes.

It took the discussants several pages of fierce fighting against my pamphlet to come to the conclusion that I am guilty in all points they raise. Their approach looks a bit religious to me. Even some of the discussants of my paper have mercy and agree that I might be right in one point or the other. I do not want to react to all the statements (accusations?) presented by Professors Genest and Rémillard, but I want to make clear where we differ. This contribution is typical for people who are partly blind since they are used to look through copula glasses. Therefore it is useful to comment on some of the points raised.

The main point of disagreement seems the notion of *stochastic dependence*. In my opinion (and, I guess, in the opinion of those who use stochastic processes), the dependence structure of a stochastic model is given by its distribution. This means e.g. for an $\mathbb{R}^d$-valued process $(X_t)_{t\in T}$, continuous or discrete in time on a nice subset $T \subset \mathbb{R}$, that it is determined by its finite-dimensional distributions. Those include the one-dimensional marginal distributions, i.e., the distributions of X_t for every fixed $t \in T$. The distribution of a stochastic process is a synonym for its dependence structure. In this sense, there is *no dependence structure without marginal distributions.* A vector $(Y_1, \ldots, Y_n)$ of mutually independent components has a dependence structure which is determined by the independence of its components and the marginal distributions. This notion of dependence structure makes sense even when dealing with random elements assuming values in an abstract space such as random sets, random measures, e.g., point processes. In these cases we might not have the notion of distribution function or copula due to the lack of a partial ordering in the space. But we still have distributions. The reader may guess what I think about the statement, *virtually all modern concepts, measures and orderings of dependence are copula-based.*

The notion of dependence structure as defined by Professors Genest and Rémillard is restricted to $\mathbb{R}^d$-valued vectors. It is completely based on copulas. For example, the discussants do not make a difference between the dependence structure of an iid sequence and a sequence of independent non-identically distributed random variables. In their world, a strictly stationary real-valued process has the same dependence structure as a non-stationary process as long as the finite-dimensional distributions of the two processes have the same copulas. This is confusing and does not correspond to standard practice in probability theory. Using only copulas, one misses out on some of the most important classes of applied stochastic processes and random fields: some of these are defined only via their moment characteristics. These include covariance stationary processes and processes exhibiting higher order stationarity. Such classes of processes are used most frequently in space-time analyses. One cannot characterize these processes by using copulas. For example, one cannot define a white noise process in this way. One needs the marginal distributions as well.

The contribution gives the impression that it is enough for stochastic process theory if one can deal with Markov processes because then one can *formally* apply the copula toolbox, see Nelsen (1999), Section 6.3. However, the world of stochastic processes is much larger than this class, and even in this particular case, I doubt that copulas are of much use. With a very few exceptions, we do not know the marginal distributions and the transition probabilities of a Markov process. Even if we knew these quantities, why on earth would one want to use copulas to describe transition probabilities?

The authors try to give the impression that copulas can be flexibly used for stochastic modeling purposes. I have tried to explain both in my paper and in this response (and most of the discussants have agreed on this point) that the copula concept fails whenever genuine stochastic process theory is involved. The copula concept does not use any modern probabilistic techniques (by this I mean any theory which was developed after the 1920s). It is a pre-historic concept. Modern probability and stochastic process theory is about modeling with a particular goal in mind. One uses particular stochastic models for special purposes in physics, queuing theory, telecommunications, finance, insurance, bioinformatics, the atmospheric sciences, economics, etc. These models determine the dependence structure in time and space, i.e., the underlying distributions which may have some (usually unknown) copula structure, but it is anachronistic to start with a fixed (Archimedean, t-, Gaussian, elliptical,...) copula structure with the aim of stochastic modeling.

The use of monotone transformations for one-dimensional marginals. Professors Genest and Rémillard blame me to be a person who wants to stop the progress in modern science because of the negligence of monotone marginal transformations. I have never doubted the use of monotone marginal transformations in probability theory and statistics, the quantile transform being just one of them. As proof of their truthful statement that monotone transformations are useful, they give a historical account of papers where distinguished researchers (members of the French Academy and Nobel Prize winners among them) have used monotone marginal transformations as a *technical means* or in order to get *unified* or *simpler representations*. Although not being as important as the cited people, I have used monotone transformation many times, following among others Hoeffding in several papers on U-statistics. There exists very fine work on dependence with fixed marginals without a religious copula context, see e.g., Rüschendorf (2005) or Zolotarev (1997).

The post-modern approach: classification of copulas. In contrast to the classical use of monotone marginal transformations as a technical, simplifying or unifying means a new quality of copula research started in the beginning/middle of the 1990s. The texts by Joe (1997) and Nelsen (1999) were timely. They are pedagogically written introductions to the field in a mathematically rigorous way. It was then when the notion of 'copula' slowly started becoming a fashion beyond the academic treatment. These books and other publications introduced particular classes of copulas, the Archimedean, Gaussian and t-copula classes being the favorites. These are textbook examples because they allow one to calculate many things explicitly; they are not motivated by the goal of realistic stochastic modeling. There is a stark contrast between the earlier probabilistic and theoretical statistical work on copulas, which did not make use of a particular form of these objects, and the newly introduced classes of copulas. The classification of copulas, in particular the introduction of parametric classes, opened the door for statistical analyses.

A favorite class of copulas is the Archimedean copula which is closely related to a mixture distribution. It has disadvantages because of its exchangeable structure, see Professor Joe's response. Professors Genest and Rémillard mention Marshall and Olkin who had thoughts about using Archimedean copulas in a statistical context. They are fiercely defending this specific model, but, for example, would one recommend it for prices in a joint market, implying that they are conditionally independent? The Archimedean copula is very popular among practitioners in finance, perhaps because it is so (too?) simple. Copulas and in particular the Archimedean one are recommended for modeling dependence in actuarial applications in the Solvency II framework.

The discussants say completely correctly that one should not depend on one model but what if one does not talk about alternatives? They mention the authors in IME and CJS. Do they contrast their approach via copulas with other approaches? I checked the last five years of IME and also *Scandinavian Actuarial Journal*, but did not find a discussion. Shouldn't one think that insurance and finance applications require stochastic models tailored for their purposes and that one should have an honest discussion about why one uses this or that class of models in these areas? Does statistical convenience justify using any model/distribution?

About dogmas and false logic. In the attempts of Professors Genest and Rémillard to discard my paper, they too frequently use the same rather obvious psychological tricks. They like referring to authorities (the time-honored Kendall's tau, the Nobel prize winning Kantorovich, the Academy member Deheuvels,...), to anonymous people represented by journals (such as CJS, IME), publications by themselves, their coworkers, the coworkers of their coworkers. With a few exceptions, they mention those papers as the relevant papers on the statistics of copulas. They ask too often questions of the form, *How does Dr. Mikosch dare to raise doubts about this time-honored concept?* Most of these historical references are out of context: if anybody in the past used the quantile transform this is no argument in favor of introducing a rather arbitrary bunch of copula families and for fitting them to any kind of data. Moreover, since I have not been invited by the editors of CJS and IME to comment on some individual papers, I am not in the position to say anything about them.

About 50% of the statements of the discussants contain rhetoric questions about what Dr. Mikosch *might think* e.g., about the non-differentiability of a Brownian sample path, the use of an ARCH process or a diffusion process, or the definition of a Brownian bridge. I did not raise these problems: they are pointless with regard to the copula topic, but for the sake of completeness here are some comments to these quiz-like questions.

(1) I may be mistaken but the comparison of a non-Gaussian distribution with Gaussian marginals and the non-differentiability of a Brownian sample path is a bit out of context. If the discussants want to tell that mathematical progress does not know any limitations, I agree with them. But they will admit that Brownian motion appears as a *very natural model* in different contexts (physics, biology, finance,...) and as a good approximation to reality in contrast to various rather artificial copula models. Brownian motion has non-differentiable paths because it is an approximation. With this purpose in mind it was introduced, not in order to make those mathematicians happy who like non-differentiable functions. (2) I do not understand what the discussants intend to say about the relation between continuous and discrete distributions, on the one hand, and the fitting of extreme value distributions, on the other hand. The comparison of continuous and discrete distributions is a question of convenience of modeling and in statistical analyses indeed. But it is not reasonable to fit an extreme value distribution in a situation when there is nothing extreme in the data. Extreme value distributions are limit distributions

for maxima. There must be a reason that the data are believed to be extreme themselves, such as annual maxima of the height of sea waves. This means, that there should be a weak convergence continuity argument (maximum domain of attraction condition) in the background for using extreme value distributions. (3) Most copula models used in the literature are artificial. Professor Joe gives a good statistical reason for choosing these distributions: statistical convenience. In my discussion, I have been asking for natural stochastic models. The ARCH in the financial time series context partly satisfies this condition. Engle's ARCH (as solution to a stochastic *difference* (not differential, sorry) equation) *is a toy model* ; its inventor would not question this fact. But he got the Nobel prize because the model captures some of the essential properties of many financial return series, among them the fat tails of the marginal distributions. The model initiated the research for a whole class of similar, more realistic, models, see e.g., Andersen et al (2007). (4) The class of all diffusions is as useful as the class of all distributions. In applications, one does not choose *any* diffusion, but *some specific* ones, e.g., in finance one chooses some particular SDEs in order to describe the evolution of a price process. (5) A mixture model is not by definition *natural* (not even because Marshall and Olkin used it). It depends on the context. If one mixes in insurance mathematics over the parameter of the Poisson claim number, one does this because of the empirical observation that the variance is larger than the mean. If one mixes over the variance of the normal distribution one gets heavier tails; see e.g., the approach by Barndorff-Nielsen and Shephard (2001) in financial econometrics. (6) The discussants give the reader the impression that it is obligatory to first transform to uniform marginals before one may transform to another marginal distribution, e.g., the exponential distribution. One can always directly transform to the desired marginal distribution. (7) I do not understand why the discussants mention the Brownian bridge and the empirical process of an iid sequence in the copula context. The empirical process is based on an iid sequence, the copulas for the finite-dimensional distributions of such an iid sequence are trivial. It is convenient to use the quantile transform of a continuous distribution for the empirical process because the resulting uniform empirical process lives on a compact time interval. This makes the proof of weak convergence results easier. Moreover, the limits of the Kolmogorov–Smirnov test statistic and related goodness-of-fit test statistics are then reduced to functionals acting on the Brownian bridge. But there is no must to define a bridge process on $[0, 1]$ and one could define a limiting process of the empirical process of an iid beta-distributed sequence as a bridge process as well. Moreover, the transformation of the empirical process of an arbitrary continuous distribution to the uniform empirical process is not always useful, e.g., when studying local properties of the empirical process. (8) The figures in the contribution. As an *exploratory tool,* it is convenient to look at scatter plots with uniform marginals. But scatter plots do not give proof of a uniform distribution on $(0, 1)^2$ with independent marginals. If you take too many points in the left top scatter plot you only see a black box... What would this show? (9) I do not follow the argument that one has to use copulas in credit risk. People have chosen a copula model for credit risk because they found it convenient. Nobody forced them to do it except fashion perhaps. (10) Section 3.5 by Professors Genest and Rémillard is another fine example of seeing the world only through the dark glasses of copulas. Suppose one wants to fit a multivariate normal or t-distribution or any parametric distribution. Then one would not fit the distribution by fitting copulas and marginal distributions at all, but use maximum likelihood, for example, to estimate the parameters of the model. I do not doubt that consistency of the marginal and copula estimators yields consistency of the estimator of the multivariate distribution, but one does not have to estimate the marginal distributions and the copula if one wants to fit a parametric multivariate distribution. (11) The discussants use the word *association* in the context of stochastic dependence quite freely. This word is reserved for a specific kind of dependence, not for *any* kind of dependence; see Resnick (1987) for a guide to association. (12) The discussants exaggerate the theoretical value of copulas by comparing this concept with characteristic functions and distribution functions. The copula of a distribution does not determine the distribution without knowing the marginal distributions. Therefore it is much less useful than characteristic functions or distributions. For example, copulas are of restricted use for proving weak convergence results because the marginal distributions are not part of the concept.

As regards the statistics literature, I am happy that various papers on statistical issues have been written in the last few years. But judging by the contributions of Professors Joe, Segers, Peng, de Haan, who are certainly better informed than me, a couple of problems seem to be left, not only as regards goodness-of-fit tests. What about robustness results, e.g., when plugging some parametric or empirical copula into a functional of the multivariate distribution, such as VaR? VaR is about estimation far out in the tails. Does the empirical copula or a parametric copula combined with the empirical marginal distributions give a good fit for such an extremal event? Is it recommendable for an insurance company to use the empirical copula or empirical marginals for evaluating the dependence structure of the assets in a portfolio and for calculating reserves for the next year in order to avoid ruin in the case of a 9/11 event? I hope not.

The discussion on copulas is certainly not finished. As experience shows, fashions come and go, but if researchers make serious contributions to the field then some theoretical and practical wisdom will also remain for the future. It seems that statisticians and probabilists have started thinking about copulas more seriously. It is not enough to introduce new copula families, give them a name and fit them to any kind of data. Living in the 21st century, we stand on the shoulders of giants such as Kolmogorov, Lévy, Wiener and Cramér who did things not just because they could or because it was convenient. We should ask about the why's and wherefore's of the recent copula activities, and given the responses to my paper I am optimistic that more, in particular young, people will ask the same questions as I did, even if our answers will be very different.

References

Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, T. (eds.): Handbook of Financial Time Series. Springer, Berlin Heidelberg New York, (2007) (in press)

Barndorff-Nielsen, O.E., Shephard, N.: Modelling by Lévy processes for financial econometrics. In: Barndorf-Nielsen, O.E., Mikosch T., Resnick, S.I. (eds.) Lévy Processes: Theory and Applications, pp. 283–318. Birkhauser, Boston (2001)

Glasserman, P.: Monte-Carlo Methods in Financial Engineering. Springer, Berlin Heidelberg New York (2004)

Joe, H.: Multivariate Models and Dependence Concepts. Chapman & Hall, London (1997)

Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin Heidelberg New York (1992)

Nelsen, R.B.: An Introduction to Copulas. Lecture Notes in Statistic, vol. 139. Springer, Berlin Heidelberg New York (1999)

Resnick, S.I.: Extreme Values, Regular Variation, and Point Processes. Springer, Berlin Heidelberg New York (1987)

Rüschendorf, L.: Stochastic ordering of risks, influence of dependence and a.s. constructions. In: Advances on Models, Characterizations and Applications. Stat. Textb. Monogr, vol. 180, pp. 19–56. Chapman & Hall, London, UK (2005)

Zolotarev, V.: Modern Theory of Summation of Random Variables. VSP, Utrecht, The Netherlands (1997)