12/12

Arbeidsnotat
Working Paper

Recognizing and visualizing copulas: an approach using local Gaussian approximation

Geir Drage Berentsen
Bård Støve
Dag Tjøstheim
Tommy Nordbø

SNF
Et selskap i NHH-miljøet

SAMFUNNS- OG NÆRINGSLIVSFORSKNING AS
Institute for Research in Economics and Business Administration

SNF
Samfunns- og næringslivsforskning AS

- er et selskap i NHH-miljøet med oppgave å initiere, organisere og utføre eksternfinansiert forskning. Norges Handelshøyskole, Universitetet i Bergen og Stiftelsen SNF er aksjonærer. Virksomheten drives med basis i egen stab og fagmiljøene ved NHH og Institutt for økonomi (UiB).

SNF er Norges største og tyngste forskningsmiljø innen anvendt økonomisk-administrativ forskning, og har gode samarbeidsrelasjoner til andre forskningsmiljøer i Norge og utlandet. SNF utfører forskning og forskningsbaserte utredninger for sentrale beslutningstakere i privat og offentlig sektor. Forskningen organiseres i programmer og prosjekter av langsiktig og mer kortsiktig karakter. Alle publikasjoner er offentlig tilgjengelig.

SNF
Institute for Research in Economics and Business Administration

- is a company within the NHH group. Its objective is to initiate, organize and conduct externally financed research. The company shareholders are the Norwegian School of Economics and Business Administration (NHH), the University of Bergen (UiB) and the SNF Foundation. Research is carried out by SNF´s own staff as well as faculty members at NHH and the Department of Economics at UiB.

SNF is Norway´s largest and leading research environment within applied economic administrative research. It has excellent working relations with other research environments in Norway as well as abroad. SNF conducts research and prepares research-based reports for major decision-makers both in the private and the public sector. Research is organized in programmes and projects on a long-term as well as a short-term basis. All our publications are publicly available.

Working Paper No 12/12

Recognizing and visualizing copulas: an approach using local Gaussian approximation

by

Geir Drage Berentsen

Bård Støve

Dag Tjøstheim

Tommy Nordbø

SNF project no 1306 "Crisis, Restructuring and Growth"

CRISIS, RESTRUCTURING AND GROWTH

This working paper is one of a series of papers and reports published by the Institute for Research in Economics and Business Administration (SNF) as part of its research programme "Crisis, Restructuring and Growth". The aim of the programme is to map the causes of the crisis and the subsequent real economic downturn, and to identify and analyze the consequences for restructuring needs and ability as well as the consequences for the long-term economic growth in Norway and other western countries. The programme is part of a major initiative by the NHH environment and is conducted in collaboration with The Norwegian Ministry of Trade and Industry, The Research Council of Norway, The Confederation of Norwegian Enterprise/ABELIA and Sparebanken Vest/Bergen Chamber of Trade and Industry/Stavanger Chamber of Trade and Industry.

INSTITUTE FOR RESEARCH IN ECONOMICS AND BUSINESS ADMINISTRATION
BERGEN, JUNE 2012
ISSN 1503-2140

Recognizing and visualizing copulas: an approach using local Gaussian approximation

Geir Drage Berentsen, Bård Støve, Dag Tjøstheim and Tommy Nordbø

21st June 2012

Abstract

Copulas are much used to model nonlinear and non-Gaussian dependence between stochastic variables. Their functional form is determined by a few parameters, but unlike a dependence measure like the correlation, these parameters do not have a clear interpretation in terms of the dependence structure they create. In this paper we examine the relationship between a newly developed local dependence measure, the local Gaussian Correlation, and standard copula theory. We are able to describe characteristics of the dependence structure in different copula models in terms of the local Gaussian correlation. In turn, these characteristics can be effectively visualized. More formally, the characteristic dependence structure can be used to construct a goodness-of-fit test for bivariate copula models by comparing the theoretical local Gaussian correlation for a specific copula and the estimated local Gaussian correlation. A Monte Carlo study reveals that the test performs very well compared to a commonly used alternative test. We also propose two types of diagnostic plots which can be used to investigate the cause of a rejected null. Finally, our methods are used on a "classic" insurance data set.

1 Introduction

Copula theory goes back to the work of Sklar [1959]. In recent years the use of copulas has grown fast. The books of Joe [1997] and Nelsen [2006] provide an overview of copula theory,

including the most common parametric families of copulas and estimating procedures. One of the main argument for using copula theory is that non-linear dependencies between variables can be modelled. Thus copula modelling has found many useful applications, in particular in finance, where non-linear dependencies typically arise between the returns of financal assets, see e.g. Jaworski et al. [2010], Brigo et al. [2010], Cherubini et al. [2004], Chollete et al. [2009] and Okimoto [2008].

There are two interrelated issues of copula theory that we will look at in this paper: i) visualizing and quantifying the nonlinear dependence structure of a copula and ii) use this to recognize and specify a copula model from given data. Both of these issues will be explored using the new tool of local Gaussian correlation that was introduced in Tjøstheim and Hufthammer [2012]. The local Gaussian correlation is a nonlinear dependence measure, but it retains the standard correlation interpretation based on a family of local Gaussian approximations.

Typically a copula model contains a few (often only one) parameters that describes the dependence structure. A problem is that the parameters are difficult to interpret. In what way do they measure dependence? A very crude characterization of a copula model is obtained by simulating observations from it and subsequently looking at the resulting scatter diagram. For instance the Clayton copula whose scatter diagram indicates heavy tails for negative values is thought to give a possible model for dependence of financial returns, since it is a common view among finance analysts that the correlations between financial objects increase as the market is going down. But a scatter diagram is not a very precise quantification of dependence.

Tjøstheim and Hufthammer [2012] introduce a local correlation measure that is meant to give a precise mathematical description and interpretation of such phenomena. A brief survey of this concept is given in Section 2. In Section 3 it will be shown how it can be used to precisely characterize and visualize the dependence structure for a number of standard copula models.

The problem of recognizing a copula from the data is the problem of goodness-of-fit. Many proposals have been made for goodness-of-fit-testing of copula models, which dates back to Deheuvels [1979]. One has to choose the right copula from a wide range of possibilities. The most used approach is to select the copula that provides the best likelihood, e.g. by the Akaike Information Criteria (AIC), see Breymann et al. [2003]. Two recent papers which

introduce new copula selection methods are Huard et al. [2006] and Karlis and Nikoloupoulos [2008]. In the first paper the authors propose a Bayesian method to select the most probable copula family among a given set, whereas the second paper introduce a goodness-of-fit test for copula families based on Mahalanobis squared distance between original and simulated data, through parametric bootstrap techniques.

Based on the Rosenblatt transformation (see Rosenblatt [1952]), Breymann et al. [2003] propose a test procedure, further Chen et al. [2004] developed a test based on the kernel density estimator. Genest et al. [2009] reviews and performs a power study of the available goodness-of-fit tests for copulas, and a similar study is performed by Berg [2009].

In goodness-of-fit testing, when a model is rejected, a problem is to identify the cause of the rejection. This problem has been recognized by Berg [2009]; "When doing model evaluation... there is still an unsatisfied need for intuitive and informative diagnostic plots."

In this paper we introduce a new goodness-of-fit test for bivariate copula models based on the local Gaussian correlation. The test is based on calculating the difference between the local Gaussian correlation estimated nonparametrically for the data in question and estimated by using an analytical expression for the local Gaussian correlation for a specific copula. One type of diagnostic plots are obtained by plotting these estimates together along the diagonal $x_1 = x_2$. We also propose a second type of diagnostic plot which displays the results of a "local goodness-of-fit" test. Implementation issues of the goodness-of-fit test are discussed in section 4 where a simulation study is conducted to assess the power and level of the proposed test and the diagnostic plots are discussed in Section 5. Finally, a practical data example is given in Section 6.

2 Local Gaussian approximation

Let $X = (X_1, X_2)$ be a two-dimensional random variable with density $f(x) = f(x_1, x_2)$. In this section we describe how f can be approximated locally in a neighbourhood of each point $x = (x_1, x_2)$ by a Gaussian bivariate density

$$\psi(v, \mu(x), \Sigma(x)) = \frac{1}{2\pi|\Sigma(x)|^{1/2}} \exp\left[-\frac{1}{2}(v - \mu(x))^T \Sigma^{-1}(x)(v - \mu(x))\right], \tag{2.1}$$

where $v = (v_1, v_2)^T$ is the running variable, $\mu(x) = (\mu_1(x), \mu_2(x))^T$ is the local mean vector and $\Sigma(x) = (\sigma_{ij}(x))$ is the local covariance matrix. With $\sigma_i^2(x) = \sigma_{ii}(x)$, we define the local correlation at the point x by $\rho(x) = \frac{\sigma_{12}(x)}{\sigma_1(x)\sigma_2(x)}$, and in terms of the local correlation, ψ may be written as

$$\psi(v, \mu_1(x), \mu_2(x), \sigma_1^2(x), \sigma_2^2(x), \rho(x)) = \frac{1}{2\pi\sigma_1(x)\sigma_2(x)\sqrt{1-\rho^2(x)}} \exp\Bigg\{ -\frac{1}{2(1-\rho^2(x))} \times \Bigg[\left(\frac{v_1 - \mu_1(x)}{\sigma_1(x)}\right)^2 - 2\rho(x)\left(\frac{v_1 - \mu_1(x)}{\sigma_1(x)}\right)\left(\frac{v_2 - \mu_2(x)}{\sigma_2(x)}\right) + \left(\frac{v_2 - \mu_2(x)}{\sigma_2(x)}\right)^2 \Bigg] \Bigg\}. \quad (2.2)$$

First note that the representation in (2.2) is not well-defined unless extra conditiona are imposed. Actually, as is easy to check, if $f(x)$ is a global Gaussian $\mathcal{N}(\mu, \Sigma)$, infinitely many Gaussians can be chosen whose density pass through the point $(x, f(x))$. However, these Gaussian densities are not really the objects we want. We need to construct a Gaussian approximation that approximates $f(x)$ in a *neighborhood* of x and such that (2.2) holds at x. In the case of $X \sim \mathcal{N}(\mu, \Sigma)$ this is trivially obtained by taking one Gaussian; i.e., $\mu(x) = \mu$ and $\Sigma(x) = \Sigma$ for all x. In fact, these relatisonships may be taken as definitions of the local parameters for a Gaussian distribution.

In Tjøstheim and Hufthammer [2012] it was demonstrated that for a given neighbourhood characterized by a bandwidth parameter b the local population parameters $\lambda(x) = (\mu(x), \Sigma(x))$ or $\lambda(x) = (\mu_1(x), \mu_2(x), \sigma_1^2(x), \sigma_2^2(x), \rho(x))$ can be defined by minimizing a likelihood related penalty function resulting in the equations

$$\int K_b(v - x) \frac{\partial}{\partial \lambda_j} \log(\psi(v, \lambda(x))[f(v) - \psi(v, \lambda(x)] dv = 0, \quad j = 1, \ldots, 5. \quad (2.3)$$

where b is a bandwidth parameter, and $K_b(v-x) = b^{-1}K(b^{-1}(v-x))$ with K being a kernel function. We define the population value $\lambda_b(x)$ as the solutions of these set of equations. It is assumed that there is a bandwidth b_0 such that there exists a unique solution of the set of equations (2.3) for any b with $0 < b < b_0$.

It is easy to find examples where (2.3) is satisfied with a unique $\lambda_b(x)$. A trivial example is when $X \sim \mathcal{N}(\mu, \Sigma)$ is Gaussian, where $\lambda_b(x) = \lambda(x) = (\mu, \Sigma)$. The next step is defining

a step function of a Gaussian variable $Z \sim \mathcal{N}(\mu, \Sigma)$, where we will take $\mu = 0$ and $\Sigma = I_2$, the identity matrix of dimension 2. Let $R_i, i = 1, \ldots, k$ be a set of non-overlapping regions of $\mathbb{R}^2$ such that $\mathbb{R}^2 = \cup_{i=1}^k R_i$. Further, let a_i and A_i be a corresponding set of vectors and matrices in $\mathbb{R}^2$ such that A_i is non-singular and define the piecewise linear function

$$X = g_s(Z) = \sum_{i=1}^{k} (a_i + A_i Z) 1(Z \in R_i), \tag{2.4}$$

where $1(\cdot)$ is the indicator function. Let S_i be the region defined by $S_i = \{x : x = a_i + A_i z, z \in R_i\}$. It is assumed that (2.4) is one-to-one in the sense that $S_i \cap S_j = \emptyset$ for $i \neq j$ and $\cup_{i=1}^k S_i = \mathbb{R}^2$. To see that the linear step function (2.4) can be used to obtain a solution of (2.3) let x be a point in the interior of S_i and let the kernel function K have a compact support. If $v - x$ is in the support of K, then b can be made small enough so that $v - x \in S_i$. Under this restriction on b, $\lambda_b(x) = \lambda(x) \equiv \lambda_i = (\mu_i, \Sigma_i)$ where $\mu_i = a_i$ and $\Sigma_i = A_i A_i^T$ as defined in (2.4). Thus, in this sense, for a fixed but small b, there exists a local Gaussian approximation $\psi(x, \lambda_b)$ of f, with corresponding local means $\mu_{i,b}(x)$, variances $\sigma_{i,b}^2(x)$, $i = 1, 2$, and correlation $\rho_b(x)$.

It was shown in Tjøstheim and Hufthammer [2012] that once a unique population vector $\lambda_b(x)$ exists one can let $b \to 0$ to obtain a local population vector $\lambda(x)$ defined at a point x. The popolation vectors $\lambda_b(x)$ and $\lambda(x)$ are both consistent with a local log-likelihood function defined by

$$L\Big(X_1, \ldots, X_n, \lambda_b(x)\Big) = n^{-1} \sum_i K_b(X_i - x) \log \psi(X_i, \lambda_b(x)) - \int K_b(v - x) \psi(v, \lambda_b(x)) \, \mathrm{d}v. \tag{2.5}$$

for given observations $X_1, \ldots, X_n$. This likelihood is taken from Hjort and Jones [1996] where it was used for density estimation. Here, the X_i's are iid observations or more generally from an ergodic time series $\{X_t\}$. In the latter case (2.5) could be thought of as a marginal local likelihood function. The last term of (2.5) is perhaps somewhat unexpected, but it is this term that forces $\psi(x, \lambda_b(x))$ not to stray away from $f(x)$ as $b \to 0$. Indeed, using the notation $u_j(\cdot, \lambda) = \partial / \partial \lambda_j \log \psi(\cdot, \lambda)$, by the law of large numbers (or ergodic theorem),

assuming $\mathbb{E}\{K_b(X_i - x)\log\psi(x_i, \lambda_b(x))\} < \infty$, we have almost surely

$$\frac{\partial L}{\partial \lambda_j} = n^{-1}\sum_i K_b(X_i - x)u_j(X_i, \lambda_b(x)) - \int K_b(v - x)u_j(v, \lambda_b(x))\psi(v, \lambda_b(x))\,\mathrm{d}v$$
$$\to \int K_b(v - x)u_j(v, \lambda_b(x))[f(v) - \psi(v, \lambda_b(x))]\,\mathrm{d}v \quad (2.6)$$

as $n \to \infty$, where (2.6) can be identified with (2.3). Letting $b \to 0$ and requiring

$$\partial L/\partial \lambda_j = 0 \quad (2.7)$$

leads to

$$u_j\Big(x, \lambda_b(x)\Big)[f(x) - \psi(x, \lambda_b(x))] + O(b^{\mathrm{T}}b) = 0, \quad (2.8)$$

so that, ignoring solutions that yield $u_j\Big(x, \lambda_b(x)\Big) = 0$, (2.6) requires $\psi\Big(x, \lambda_b(x)\Big)$ to be close to $f(x)$, in fact with a difference of the order $O(b^{\mathrm{T}}b)$ as $b \to 0$. The numerical maximization of the local likelihood (2.5) leads to local likelihood estimates $\widehat{\lambda}_b(x)$, including estimates $\widehat{\rho}_b(x)$ of the local correlation. It is shown in Tjøstheim and Hufthammer (2012) that under relatively weak regularity conditions $\widehat{\lambda}_b(x) \to \lambda_b(x)$ for b fixed, and $\lambda_b(x) \to \lambda(x)$ for $b = b_n$ tending to zero.

2.1 Non-linear transformations of Gaussian variables

To connect the local correlation concept introduced in the preceding section with the copula concept it is advantageous to consider nonlinear transformations of Gaussian variables. A continuous one-to-one function $g : \mathbb{R}^2 \to \mathbb{R}^2$ with an inverse $h = g^{-1}$ can be approximated by a sequence of one-to-one piecewise linear functions such as in (2.4) by letting k increase and by letting the regions R_i be smaller. It will be seen below if g is continuously differentiable at z, then a Gaussian representation can be found by Taylor expansion, but unfortunately it is not unique unless x is restricted, and it cannot be identified with the representation of the previous section unless there is uniqueness.

Generally if g is continously differentiable at z, the best linear approximation of $X = g(Z)$ in a neighbourhood $\mathcal{N}(x) = \{x' : |x' - x| \le r_x\}$ of $x = g(z)$ transformed from a corresponding neighbourhood $\mathcal{N}(z) = \{z' : |z' - z| \le r_z\}$ of z is given by

$$U_z(Z) = g(z) + \frac{\partial g}{\partial z}(z)(Z - z)$$

such that $X = g(Z) = U_z(Z) + o_p(|Z - z|)$, and where $\frac{\partial g}{\partial z}$ is the Jacobi matrix. When $r_z \to 0$, then $r_x \to 0$ because of the continuity of g, and in the limit using the continuous differentiability of g, higher order terms of a Taylor expansion of $g(Z)$ can be neglected in probability, and in the limit $U_z(Z)$ gives one Gaussian representation of $X = g(Z)$ at the point $x = g(z)$.

For this representation it is tempting to define the local mean $\mu(x)$ and $\Sigma(x)$ of the density of X at the point x as the mean and covariance of the Gaussian variable $U_z(Z)$. These are expressed as functions of x using $z = h(x) = g^{-1}(x)$. Since $\mathbb{E}(Z) = 0$ and $\Sigma(z) = I_2$ this results in

$$\mu(x) = g(z) - \frac{\partial g}{\partial z}(z)z = x - \left(\frac{\partial h}{\partial x}(x)\right)^{-1} h(x) \tag{2.9}$$

and

$$\Sigma(x) = \frac{\partial g}{\partial z}(z)\left(\frac{\partial g}{\partial z}(z)\right)^T = \left(\frac{\partial h}{\partial x}(x)\right)^{-1}\left(\left(\frac{\partial h}{\partial x}(x)\right)^{-1}\right)^T. \tag{2.10}$$

It is an easy matter to verify that $f_{U_z(Z)} = \psi(v, \mu(x), \Sigma(x))$ yields a representation of type (2.1).

The representations (2.9) and (2.10) are unique for a given X and g, But for a given density $f(x)$, it can be generated in several ways, leading to non-uniqueness. This raises two questions: When can a stochastic variable X be represented as a function of a Gaussian variable Z and to what degree is the representations in (2.9) and (2.10) unique? The first question is essentially answered in Rosenblatt [1952]. We state it as a lemma.

Lemma 2.1. *Let X have a density $f_X(x)$ on $\mathbb{R}^2$ with cumulative distribution function $F_X(x) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f_X(w_1, w_2)\,\mathrm{d}w_1\,\mathrm{d}w_2$. Then there exists a one-to-one function g such that $X = g(Z)$, where $Z \sim \mathcal{N}(0, I_2)$.*

Proof. We have $f_X(x) = f_{X_1}(x_1)f_{X_2|X_1}(x_2|x_1)$. Then $U_1 = F_{X_1}(X_1)$ is uniform. There also exists a standard normal variable Z_1 such that $U_1 = \Phi(Z_1)$, where Φ is the cumulative distribution of the standard normal density. Hence, $X_1 = F_{X_1}^{-1}(\Phi(Z_1))$. In the same manner, there exists a uniform variable U_2 independent of U_1 (see Rosenblatt [1952]) such that $U_2 = F_{X_2|X_1}(X_2|X_1)$, and there exists a $Z_2 \sim \mathcal{N}(0, 1)$ independent of Z_1 such that $U_2 = \Phi(Z_2)$,

and hence

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_{X_1}^{-1}(\Phi(Z_1)) \\ F_{X_2|X_1}^{-1}(\Phi(Z_2)|F_{X_1}^{-1}(\Phi(Z_1)) \end{bmatrix} \doteq g(Z), \tag{2.11}$$

where $F_{X_2|X_1}^{-1}$ is interpreted as the inverse of $F_{X_2|X_1}$ with X_1 fixed (i.e., with U_1, Z_1 fixed). Here g is one-to-one due to the strict monotonicity of F_X. □

As pointed out in Rosenblatt [1952] this representation is non-unique, since we also have

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_{X_1|X_2}^{-1}(\Phi(Z_1)|F_{X_2}^{-1}(\Phi(Z_2)) \\ F_{X_2}^{-1}(\Phi(Z_2)) \end{bmatrix} \doteq g'(Z'), \tag{2.12}$$

where in general $g \neq g'$ and $Z \neq Z'$. This also means that $\mu'(x) \neq \mu(x)$ and $\Sigma'(x) \neq \Sigma(x)$.

However, there may be a set of points (x_1, x_2) for which the two Rosenblatt representations (2.11) and (2.12), with ρ and ρ', respectively, coincide. It is shown in Tjøstheim and Hufthammer [2012] that $\rho(x_1, x_2) = \rho'(x_2, x_1)$ if X_1 and X_2 are exchangable, i.e. $F_{X_1,X_2}(x_1, x_2) = F_{X_1,X_2}(x_2, x_1)$ for all pairs (x_1, x_2), in which case they coincide along the diagonal $x = (s, s)$. More generally they would coincide along the curve defined by $F_{X_1}(x_1) = F_{X_2}(x_2)$ (see section 3). Since the two Rosenblatt representations are bases for any representation of $f_X(x)$, (including a density generated by a general functional relationship $X = g(Z)$), we have uniqueness at points where they coincide. The local parameters along such curves are consistent with the local parameters derived from the local penalty function (2.3). Indeed, for a point x where the Rosenblatt representations give a unique $\lambda(x) = (\mu(x), \Sigma(x))$ such that $f(x) = \psi(x, \lambda(x))$, a local Gaussian approximation with $\lambda_b(x)$ can be found that satisfies the local penalty equation (2.3) and that converges to $\lambda(x)$. Simply choose a linear stepwise representation (2.4), such that $x \in S_i$ for some i, and take $A_i = \Sigma^{1/2}(x)$ and $a_i = \mu(x)$. Then with a small enough bandwidth, $\lambda_b(x) = \lambda_i = (a_i, A_i A_i^T)$ $= (\mu(x), \Sigma(x))$, and $\lambda_b(x) \to \lambda(x)$ trivially as $b \to 0$. If for a point x there is not a unique Rosenblatt representation, i.e. off-diagonal terms in the above example, then such an approach is not possible since there is not a unique $\lambda(x)$ that could serve a starting point for the construction. Nevertheless, for such points x, under the regularity conditions mentioned at the end of the previous sub-section the existence of a unique $\lambda(x)$ can be determined by the local penalty function resulting in (2.5) and the local likelihood estimate $\widehat{\lambda}(x)$ converges towards $\lambda(x)$ (see Tjøstheim and Hufthammer [2012]), but unlike the points along the curve

defined by $F_{X_1}(x_1) = F_{X_2}(x_2)$ we have not managed to find an explicit expression for $\lambda(x)$ for a general x. A simulation experiment confirming these facts are given in section 4, but first we will derive explicit formulas for $\rho(x)$ along the curve $F_1(x_1) = F_2(x_2)$ for several copulas with $F_i = F_{X_i}$.

3 Local Gaussian correlation for copula models

We start by rephrasing the local parameters given by (2.9) and (2.10) in the previous section and making it explicit for the local Gaussian correlation in the case when g is the Rosenblatt transformation (2.11). We will subsequently look at (2.12) and then examine under what conditions these two transformations will give rise to a unique local Gaussian correlation to be used in the rest of the paper. For the transformation (2.10) the matrix

$$\frac{\partial h}{\partial x}(x) = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} \end{bmatrix} \tag{3.1}$$

is lower triangular and

$$\left(\frac{\partial h}{\partial x}(x)\right)^{-1} = \left(\frac{\partial h_1}{\partial x_1}\frac{\partial h_2}{\partial x_2}\right)^{-1} \begin{bmatrix} \frac{\partial h_2}{\partial x_2} & 0 \\ \frac{-\partial h_2}{\partial x_1} & \frac{\partial h_1}{\partial x_1} \end{bmatrix},$$

which, by (2.10), results in the following local covariance matrix

$$\Sigma(x) = \left(\frac{\partial h_1}{\partial x_1}\frac{\partial h_2}{\partial x_2}\right)^{-2} \begin{bmatrix} \left(\frac{\partial h_2}{\partial x_2}\right)^2 & -\frac{\partial h_2}{\partial x_1}\frac{\partial h_2}{\partial x_2} \\ -\frac{\partial h_2}{\partial x_1}\frac{\partial h_2}{\partial x_2} & \left(\frac{\partial h_1}{\partial x_1}\right)^2 + \left(\frac{\partial h_2}{\partial x_1}\right)^2 \end{bmatrix}.$$

The local Gaussian correlation is then given by

$$\rho(x) = \rho(x_1, x_2) = \frac{\Sigma_{12}(x)}{\sqrt{\Sigma_{11}(x)\Sigma_{22}(x)}} = \frac{-\frac{\partial h_2}{\partial x_1}}{\sqrt{\left(\frac{\partial h_1}{\partial x_1}\right)^2 + \left(\frac{\partial h_2}{\partial x_1}\right)^2}}, \tag{3.2}$$

where we return to its validity and uniqueness below. Next, consider a continuous random variable $X = (X_1, X_2)$ with joint cumulative distribution function F and margins $F_1(x_1)$

and $F_2(x_2)$. Due to the reprentation theorem of Sklar [1959], F can be written as

$$F(x_1, x_2) = C((F_1(x_1), F_2(x_2)), \tag{3.3}$$

where the copula $C : [0,1]^2 \to [0,1]$ is a unique bivariate distribution function with uniform margins. In the case when F is given by (3.3) we may re-express (2.11) and thus $\rho(x_1, x_2)$ in terms of the copula C and the margins F_1 and F_2. Let U_1 and U_2 be distributed according to C and let

$$\begin{aligned} C_1(u_1, u_2) &= Pr(U_2 \leq u_2 | U_1 = u_1) \\ &= \lim_{\Delta u_1 \to 0} \frac{C(u_1 + \Delta u_1, u_2) - C(u_1, u_2)}{\Delta u_1} = \frac{\partial}{\partial u_1} C(u_1, u_2) \end{aligned} \tag{3.4}$$

Then, using the notation $F_{2|1}$ for for the distribution function of X_2 given X_1, we may write $F_{2|1}(x_2|x_1)$ as

$$F_{2|1}(x_2|x_1) = P(U_2 \leq F_2(x_2) | U_1 = F_1(x_1)) = C_1(F_1(x_1), F_2(x_2)),$$

and consequently $F_{2|1}^{-1}(x_2|x_1)$ may be written as

$$F_{2|1}^{-1}(x_2|x_1) = F_2^{-1}\left(C_1^{-1}(F_1(x_1), x_2)\right),$$

where $C_1^{-1}(u, v)$ is interpreted as the inverse of $C_1(u, v)$ with u fixed. It follows that (2.11) may be written as

$$g(Z) = \begin{bmatrix} F_1^{-1}(\Phi(Z_1)) \\ F_2^{-1}\left(C_1^{-1}\left(\Phi(Z_1), \Phi(Z_2)\right)\right) \end{bmatrix} \tag{3.5}$$

Note that this transformation (only with $\Phi(Z_1)$ and $\Phi(Z_2)$ replaced by two independent uniform $[0,1]$ variables) is a standard way of sampling from the distribution $C(F_1(x_1), F_2(x_2))$ (See e.g. Nelsen, 2006, page 35-37). In the continuous case, g is one-to-one if the copula density $c(u_1, u_2)$ satisfies $c(u_1, u_2) > 0$ for all points $(u_1, u_2) \in [0,1]^2$ (This guarantees the invertibility of $C_1(u_1, u_2)$ with respect to u_2). The inverse $h = g^{-1}$ is then given by

$$h(X) = \begin{bmatrix} h_1(X_1, X_2) \\ h_2(X_1, X_2) \end{bmatrix} = \begin{bmatrix} \Phi^{-1}\left(F_1(X_1)\right) \\ \Phi^{-1}\left(C_1\left(F_1(X_1), F_2(X_2)\right)\right) \end{bmatrix} \tag{3.6}$$

Towards finding an expression for $\rho(x_1, x_2)$ using (3.2) let ϕ denote the standard normal density function and let

$$C_{11}(u_1, u_2) = \frac{\partial^2}{\partial u_1^2} C(u_1, u_2).$$

Then the two partial derivatives of h involved in (3.2) is given by

$$\frac{\partial h_1}{\partial x_1} = \frac{f_1(x_1)}{\phi\left(\Phi^{(-1)}\left(F_1(x_1)\right)\right)}, \tag{3.7}$$

$$\frac{\partial h_2}{\partial x_1} = \frac{C_{11}\left(F_1(x_1), F_2(x_2)\right) f_1(x_1)}{\phi\left(\Phi^{(-1)}\left(C_1\left(F_1(x_1), F_2(x_2)\right)\right)\right)}, \tag{3.8}$$

where f_1 is the marginal density function of X_1. Inserting equation (3.7) and (3.8) into (3.2) the local Gaussian correlation given by (3.2) and model (3.3) may be written as

$$\rho(x_1, x_2) = \frac{-C_{11}(F_1(x_1), F_2(x_2))\phi\left(\Phi^{-1}\left(F_1(x_1)\right)\right)}{\sqrt{\phi^2\left(\Phi^{-1}\left(C_1(F_1(x_1), F_2(x_2))\right)\right) + C_{11}^2(F_1(x_1), F_2(x_2))\phi^2\left(\Phi^{-1}\left(F_1(x_1)\right)\right)}} \tag{3.9}$$

However, repeating the above steps with the Rosenblatt representation (2.12) as a starting point in stead of (2.11) leads to another local Gaussian correlation $\rho'(x_1, x_2)$ given by

$$\rho'(x_1, x_2) = \frac{-C_{22}(F_1(x_1), F_2(x_2))\phi\left(\Phi^{-1}\left(F_2(x_2)\right)\right)}{\sqrt{\phi^2\left(\Phi^{-1}\left(C_2(F_1(x_1), F_2(x_2))\right)\right) + C_{22}^2(F_1(x_1), F_2(x_2))\phi^2\left(\Phi^{-1}\left(F_2(x_2)\right)\right)}} \tag{3.10}$$

where $C_2(u_1, u_2) = \frac{\partial}{\partial u_2} C(u_1, u_2)$ and $C_{22}(u_1, u_2) = \frac{\partial^2}{\partial u_2^2} C(u_1, u_2)$. As pointed out in section 2, the local correlation given by (3.9) and (3.10) is only consistent with the local correlation derived from the penalty function (2.3) at points (x_1, x_2) where $\rho'(x_1, x_2) = \rho(x_1, x_2)$. In the copula case, when the copula is exchangeable (i.e. $C(u_1, u_2) = C(u_2, u_1)$), these points are found along the curve defined by $F_1(x_1) = F_2(x_2)$. To see this let (x_1, x_2) be a point along this curve so that $F_1(x_1) = F_2(x_2) = u$. Then

$$\rho(x_1, x_2) = \frac{-C_{11}(u, u)\phi\left(\Phi^{-1}\left(u\right)\right)}{\sqrt{\phi^2\left(\Phi^{-1}\left(C_1(u, u)\right)\right) + C_{11}^2(u, u)\phi^2\left(\Phi^{-1}\left(u\right)\right)}} \tag{3.11}$$

$$\rho'(x_1, x_2) = \frac{-C_{22}(u, u)\phi\left(\Phi^{-1}\left(u\right)\right)}{\sqrt{\phi^2\left(\Phi^{-1}\left(C_2(u, u)\right)\right) + C_{22}^2(u, u)\phi^2\left(\Phi^{-1}\left(u\right)\right)}} \tag{3.12}$$

It then follows from echangeability of C that $C_1(u,u) = C_2(u,u)$ and $C_{11}(u,u) = C_{22}(u,u)$ and thus $\rho(x_1,x_2) = \rho'(x_1,x_2)$. In case the margins are identical which they are if X_1 and X_2 are exchangeable, we have equality along the diagonal $x_1 = x_2$ and we are recovering the result mentioned in Tjøstheim and Hufthammer [2012]. In the rest of the paper we take (3.9) as the local Gaussian correlation along such curves. It will be used for characterizing the dependence properties of the copulas and for testing goodness of fit.

Note that since $\phi\left(\Phi^{-1}\left(F_1(\cdot)\right)\right) > 0$, the sign of $\rho(x_1,x_2)$ is determined by the sign of $-C_{11}(F_1(x_1),F_2(x_2))$. To see that this is reasonable, consider a random variable X_1 positively related to the variable X_2 in the neighbourhood of (x_1,x_2), in the sense that $m(s) := P(X_2 \leq x_2 | X_1 = s) = C_1(F_1(s),F_2(x_2))$ is decreasing as s increases (in a neighbourhood of x_1). Then since $m'(s) < 0$, we have that $-C_{11}(F_1(x_1),F_2(x_2)) > 0$ and thus $\rho(x_1,x_2) > 0$ in the neighbourhood of (x_1,x_2).

When X_1 and X_2 are independent their copula is the independence copula $C(u_1,u_2) = u_1u_2$. Then $C_{11}(u_1,u_2) = 0$ which implies that $\rho(x_1,x_2) = 0$ along the curve $F_1(x_1) = F_2(x_2)$. In Tjøstheim and Hufthammer [2012] it is shown that independence implies $\rho(x) = 0$ everywhere and that a necessary and sufficient condition for independence is that $\rho(x) \equiv 0$, $\mu_i(x) \equiv \mu_i(x_i)$, $\sigma_i^2(x) \equiv \sigma_i^2(x_i)$, $i = 1,2$. We have not been able to find examples where $\rho(x) \equiv 0$ and where we do not have independence.

Tjøstheim and Hufthammer [2012] consider the connection between $\rho(x_1,x_2)$ and the upper and lower tail coefficient given by

$$\lambda_u = \lim_{q \to 1^-} P(F_2(X_2) > q | F_1(X_1) > q) \quad \text{and} \quad \lambda_l = \lim_{q \to 0^+} P(F_2(X_2) \leq q | F_1(X_1) \leq q).$$

Due to the local Gaussian representation it can be shown that under a weak monotonicity condition the lower tail coefficient can be expressed as

$$\lambda_l = 2 \lim_{s \to -\infty} \Phi\left(s\sqrt{\frac{1-\rho(s,s)}{1+\rho(s,s)}}\right),$$

It is seen that if there is lower tail dependence, we must have $\rho(s,s) \to 1$ as $s \to -\infty$. Thus

by (3.9), $\rho(x_1, x_2)$ for copula models with lower tail dependence should satisfy

$$\lim_{s\to-\infty}\rho(s,s)=\lim_{q\to0^+}\frac{-C_{11}(q,q)\phi\left(\Phi^{-1}\left(q\right)\right)}{\sqrt{\phi^2\left(\Phi^{-1}\left(C_1(q,q)\right)\right)+C_{11}^2(q,q)\phi^2\left(\Phi^{-1}\left(q\right)\right)}}=1. \tag{3.13}$$

For exchangeable copulas with lower tail dependence λ_l, it can be shown that $\lim_{q\to0^+}\phi^2\left(\Phi^{-1}\left(C_1(q,q)\right)\right)=\phi^2\left(\Phi^{-1}(\lambda_l/2)\right)\neq0$. So for (3.13) to hold when $\lambda_l\neq0$ we must have that $\lim_{q\to0^+}-C_{11}(q,q)\phi\left(\Phi^{-1}\left(q\right)\right)=\infty$. This can for example be verified for the Clayton copula. For the speed at which $\rho(s,s)\to1$ for the Clayton copula we refer to figure 1 in the case of standard Gaussian margins.

3.1 Examples

In practice, given a copula, the formula (3.9) often becomes quite complicated. As a consequence, for the examples in section 3.1.1 and section 3.1.2, we only formulate the functions C_1 and C_{11} and refer to figure 1 - 4 for the characteristics of $\rho(x_1, x_2)$ for each copula. In the examples we have simply used standard normal margins for both X_1 and X_2 and all copulas considered are exchangeable so that the local correlation given by (3.9) is well defined along the diagonal $x_1 = x_2$. In figures 1 - 4 a) we have plotted $\rho(s, s)$ against s. The copula parameters in these plots are chosen so that they correspond to a specific value of Kendall's tau ($\tau = 0.2, 0.4, 0.6, 0.8$), which in general is uniquely related to the (one-parameter) copula C by the formula

$$\tau=m(\theta)=4\int\int_{[0,1]^2}C(u,v)\,\mathrm{d}C(u,v)-1. \tag{3.14}$$

We have not been able to find an analytic expression for $\rho(x_1, x_2)$ for general (x_1, x_2), but using the local likelihood algoritm $\rho(x_1, x_2)$ can be estimated for all (x_1, x_2) for which there is data. Figures 1 - 4 b) display the estimated local correlation based on one realisation of $n = 500$ samples from each of the copula models considered, with copula parameter i corresponding to $\tau = 0.4$. There is some boundary bias in the estimation by the estimated dependence pattern revealed in figures 1 - 4 b) are consistent with the theoretical ones along the diagonal in figures 1 - 4 a). See also the comparison made in figure 7.

3.1.1 Archimedean copulas

An important class of copulas is the class of Archimedean copulas, which have been extensively studied. Archimedean copulas are popular, because they allow dependence modeling with only one parameter governing the strength of dependence. These copulas are completely defined by their so-called generator function φ, with the following properties. Let $\varphi : [0,1] \to [0,\infty]$ be a continuous and strictly decreasing function with $\varphi(1) = 0$ and $\varphi(0) \leq \infty$. Define the pseudo-inverse of φ with domain $[0,\infty]$ by

$$\varphi^{[-1]} = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0), \\ 0, & \varphi(0) < t \leq \infty. \end{cases} \tag{3.15}$$

A bivariate Archimedean copula is a copula on the form

$$C(u_1, u_2) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2)), \tag{3.16}$$

where φ satisfies the above assumptions and is convex. For simplicity we will only consider Archemedean copulas where $\varphi(0) = \infty$. The generator function is then said to be strict and we may replace the pseudo-inverse $\varphi^{[-1]}$ by the ordinary functional inverse φ^{-1}. The functions C_1 and C_{11} needed to compute $\rho(x_1, x_2)$ are then easily obtain by differentiation of equation (3.16) with respect to u_1

$$C_1(u_1, u_2) = \frac{\varphi'(u_1)}{\varphi'\left(C(u_1, u_2)\right)}. \tag{3.17}$$

$$C_{11}(u_1, u_2) = \frac{\varphi''(u_1)\varphi'\left(C(u_1, u_2)\right)^2 - \varphi'(u_1)^2\varphi''\left(C(u_1, u_2)\right)}{\varphi'\left(C(u_1, u_2)\right)^3}. \tag{3.18}$$

In the following three examples we consider the commonly used Archimedean copulas Clayton, Gumbel and Frank.

Example 3.1 (Clayton copula)**.** The Clayton copula is an asymmetric copula, exhibiting greater dependence in the negative tail than in the positive (i.e. lower tail dependence). The generator for the Clayton copula is $\varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$ for $\theta \geq -1$. The Clayton copula can

thus be written as

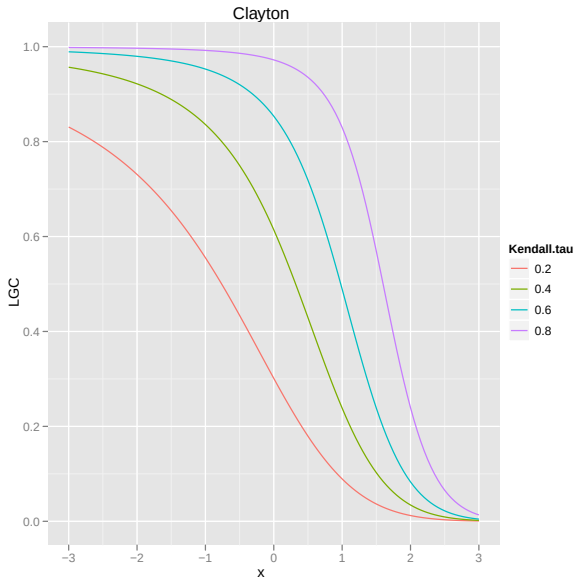
$$C_\theta^{Cl}(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta},$$

with derivatives

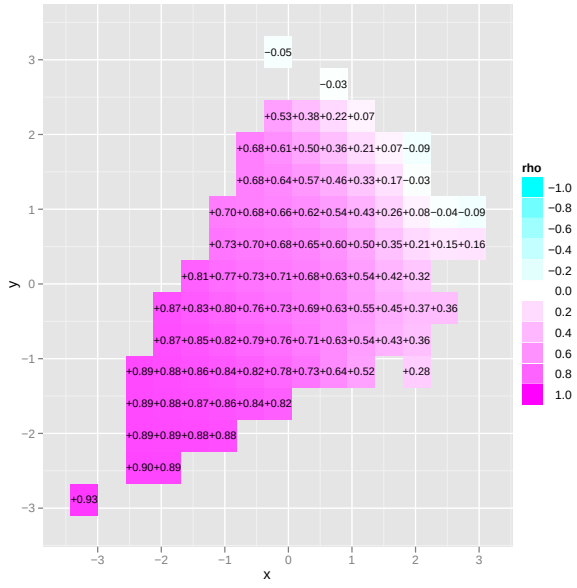
$$C_1(u_1, u_2) = \left(1 + u_1^\theta(u_2^{-\theta} - 1)\right)^{-\frac{\theta+1}{\theta}}$$
$$C_{11}(u_1, u_2) = (\theta + 1)u_2^{\theta-1}(1 - u_2^{-\theta})(1 + u_1^\theta(u_2^{-\theta} - 1))^{-1/\theta-2}$$

Since $C_{11}(q, q) = (\theta + 1)q^{-1}(1 - q^\theta)(2 - q^\theta)^{-1/\theta-2} \to 0$ as $q \to 1^-$ we have that $\rho(s, s) \to 0$ as $s \to \infty$. On the other hand, it can be shown that $-C_{11}(q, q)\phi\left(\Phi^{-1}(q)\right) \to \infty$ as $q \to 0^+$ so that $\rho(s, s) \to 1$ as $s \to -\infty$. These features can be seen in figure 1. These plots give a directly interpretable local dependence structure in terms of local correlation.

Figure 1: Local gaussian correlation for the clayton copula: (a) along the diagonal $x_1 = x_2$; (b) estimated based on $n = 500$ observations.



(a)



(b)

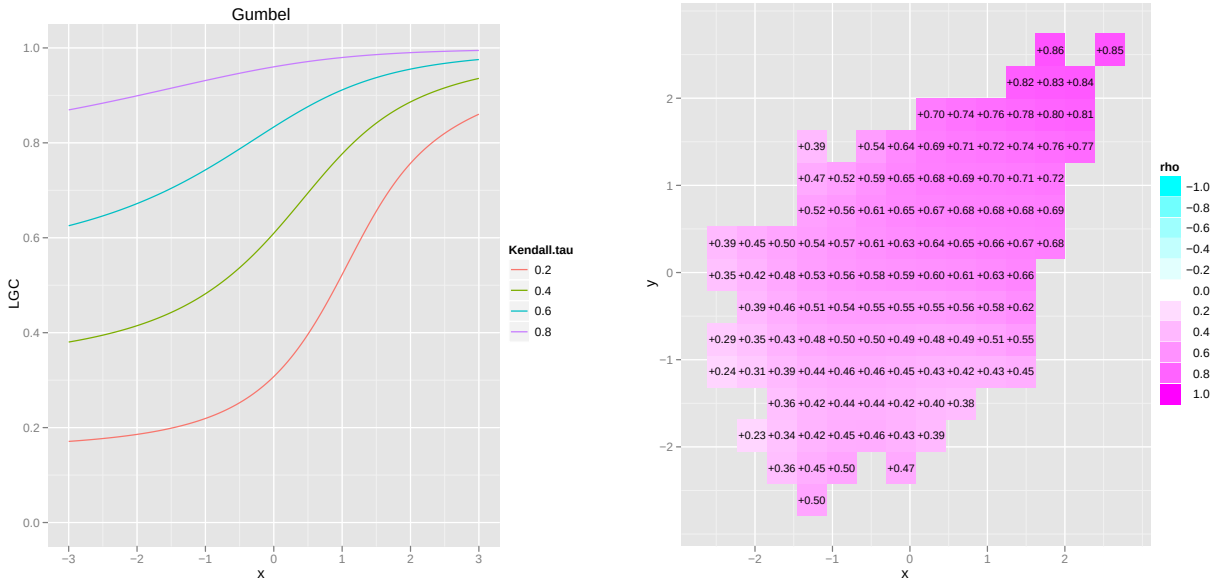
Example 3.2 (Gumbel copula)**.** The Gumbel copula is also an asymmetric copula, exhibiting greater dependence in the positive tail than in the negative (i.e. upper tail dependence).

Its generator function is $\varphi(t) = (-\ln t)^\theta$ for $\theta \geq 1$, thus the Gumbel copula can be written as

$$C_\theta^{Gu}(u_1, u_2) = \exp\Big[-((-\ln u_1)^\theta + (-\ln u_2)^\theta)^{1/\theta}\Big].$$

The functions C_1 and C_{11} are quite complicated and are therefore not given here. The characteristics of $\rho(x_1, x_2)$ for the Gumbel copula can be seen in figure 2 where we clearly see the upper tail dependence numerically quantified in terms of the local correlation.

Figure 2: Local gaussian correlation for the gumbel copula: (a) along the diagonal $x_1 = x_2$; (b) estimated based on $n = 500$ observations.



(a) (b)

Example 3.3 (Frank copula)**.** Define $q_z = e^{-\theta z} - 1$. The generator for the Frank copula is $\varphi(t) = -\ln(q_t/q_1)$. Then the Frank copula may be written as

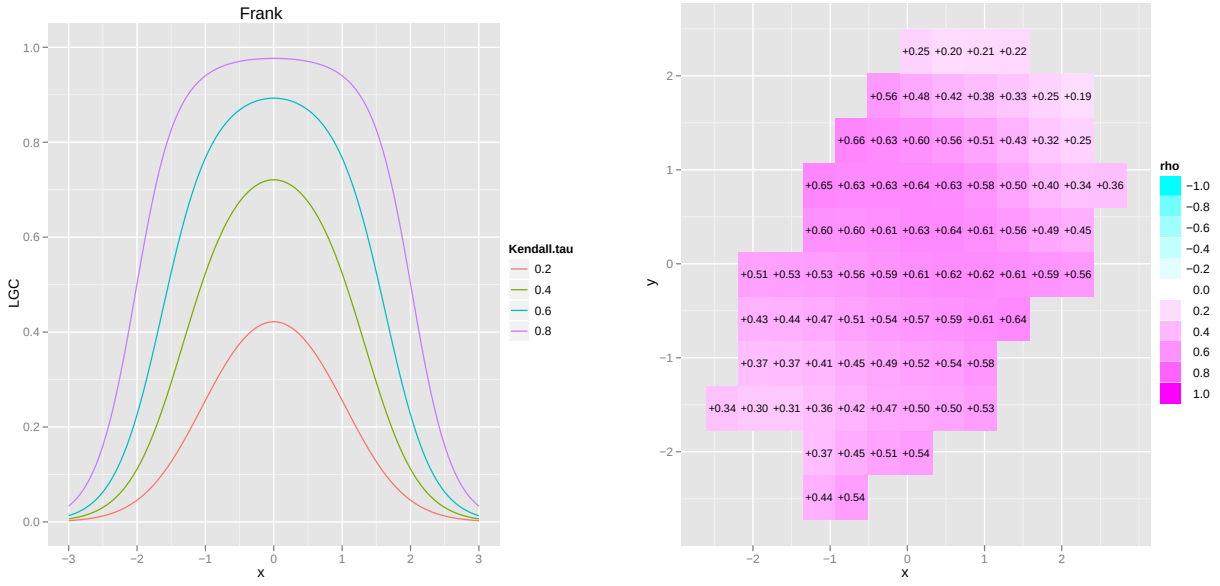
$$C_\theta^{Fr}(u_1, u_2) = -\theta^{-1} \ln\{1 + q_{u_1} q_{u_2}/q_1\}$$

The derivatives C_1 and C_1 are

$$C_1(u_1, u_2) = \frac{q_{u_1} q_{u_2} + q_{u_2}}{q_{u_1} q_{u_2} + q_1}$$
$$C_{11}(u_1, u_2) = \frac{q'_{u_1} q_{u_2} (q_1 - q_{u_2})}{(q_{u_1} q_{u_2} + q_1)^2}$$

It is easily seen that $C_{11}(q, q) \to 0$ when $q \to 0^+$ and when $q \to 1^-$. Thus the $\rho(x_1, x_2)$ goes to zero in both the upper and lower tail. This feature is reflected in figure 3; close to constant dependence in the center which vanish in the tails.

Figure 3: Local gaussian correlation for the gumbel copula: (a) along the diagonal $x_1 = x_2$; (b) estimated based on $n = 500$ observations.



(a) (b)

3.1.2 Elliptical copulas

Elliptical copulas are simply the copulas of elliptical distributions. The key advantage of elliptical copula is that one can specify different levels of correlation between the marginals,

but a disadvantage is that elliptical copulas typically do not have closed form expressions. The most commonly used elliptical distributions are the Gaussian and Student-t distributions.

Example 3.4 (Gaussian copula)**.** For a given correlation matrix $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ the Gaussian copula with correlation matrix Σ can be written as

$$C_{\Sigma}^{Gauss}(u_1, u_2) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \tag{3.19}$$

where Φ_{Σ} is the joint bivariate distribution function of a Gaussian variable with mean vector zero and correlation matrix Σ. In general, when (X_1, X_2) is Gaussian with mean vector zero and correlation matrix Σ, then $X_1|X_2 = x_2 \sim N(\rho x_2, 1-\rho^2)$. It follows that for the Gaussian copula

$$\begin{aligned} C_1(u_1, u_2) &= P(U_2 \leq u_2 | U_1 = u_1) = P(\Phi^{-1}(U_2) \leq \Phi^{-1}(u_2) | \Phi^{-1}(U_1) = \Phi^{-1}(u_1)) \\ &= \Phi\left(\frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(u_1)}{\sqrt{1-\rho^2}}\right) \end{aligned}$$

Letting $R = \frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(u_1)}{\sqrt{1-\rho^2}}$ and differentiating this expression once more with respect to u_1 we get

$$C_{11}(u_1, u_2) = \frac{-\rho}{\sqrt{1-\rho^2}\phi(\Phi^{-1}(u_1))}\phi(R)$$

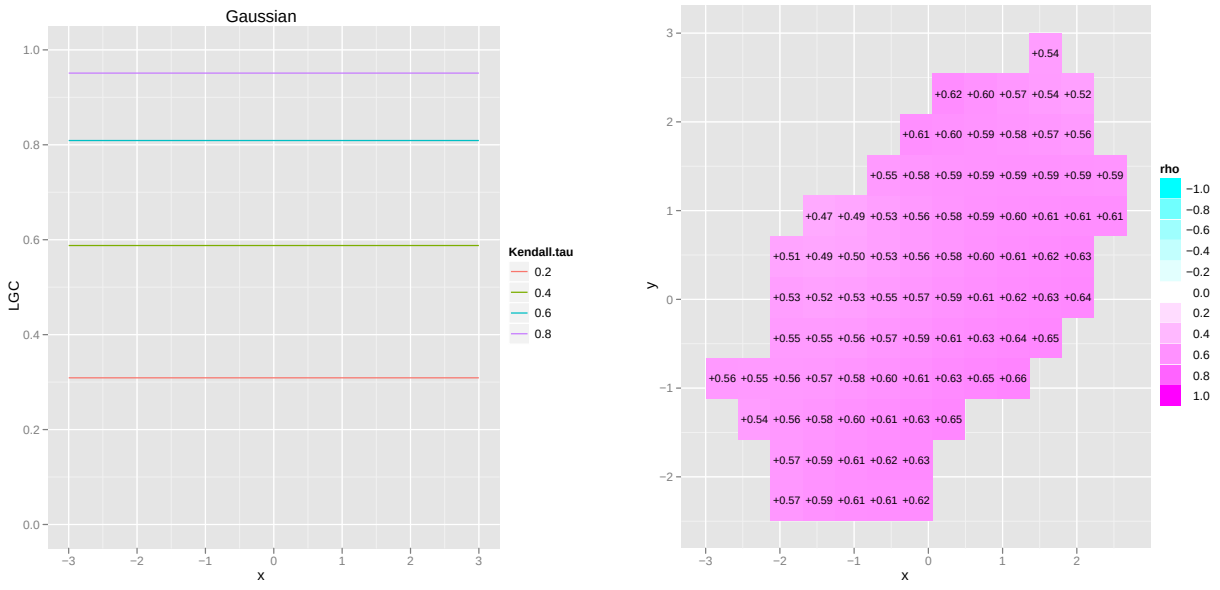
Thus for a Gaussian copula model $C^{Gauss}(F_1(x_1), F_2(x_2))$ with arbitrary margins F_1 and F_2, $\rho(x_1, x_2)$ is

$$\begin{aligned} \rho(x_1, x_2) &= \frac{-C_{11}(F_1(x_1), F_2(x_2))\phi\left(\Phi^{-1}\left(F_1(x_1)\right)\right)}{\sqrt{\phi^2\left(\Phi^{-1}\left(C_1(F_1(x_1), F_2(x_2))\right)\right) + C_{11}^2(F_1(x_1), F_2(x_2))\phi^2\left(\Phi^{-1}\left(F_1(x_1)\right)\right)}} \\ &= \frac{\phi(R)\frac{\rho}{\sqrt{1-\rho^2}}\frac{\phi\left(\Phi^{-1}(F_1(x_1))\right)}{\phi(\Phi^{-1}(F_1(x_1)))}}{\sqrt{\phi(R)^2 + \phi^2(R)\frac{\rho^2}{1-\rho^2}\frac{\phi^2(\Phi^{-1}(F_1(x_1)))}{\phi^2(\Phi^{-1}(F_1(x_1)))}}} = \frac{\rho}{\sqrt{1-\rho^2+\rho^2}} = \rho \end{aligned} \tag{3.20}$$

This is of course valid for all (x_1, x_2), not only on a curve $F_1(x_1) = F_2(x_2)$, and it shows that a constant local Gaussian correlation is a feature of the Gaussian copula rather than the bivariate Gaussian distribution. Note that the local mean and local variance are not in

general constant for non-Gaussian marginals. For a non-Gaussian copula, $\rho(x_1, x_2)$ will in general depend on the margins. It remains to prove the converse statement that $\rho(x) = c$, $(-1 < c < 1, c \neq 0)$ implies the Gaussian copula. We do not know of any other than the Gaussian that has this property.

Figure 4: Local gaussian correlation for the gaussian copula: (a) along the diagonal $x_1 = x_2$; (b) estimated based on $n = 500$ observations (where the copula parameter is $\rho = 0.5877$).



(a) (b)

Example 3.5 (T-copula)**.** In the case that (X_1, X_2) is t-distributed with ν degrees of freedom and correlation coefficient ρ, we have that $X_1|X_2 = x_2$ is t-distributed with $\nu + 1$ degrees of freedom, expected value ρx_2 and variance $\left(\frac{\nu+x_2^2}{\nu+1}\right)(1-\rho^2)$. With t_ν as the standard t-distribution function, a similar argument as for the Gaussian copula leads to

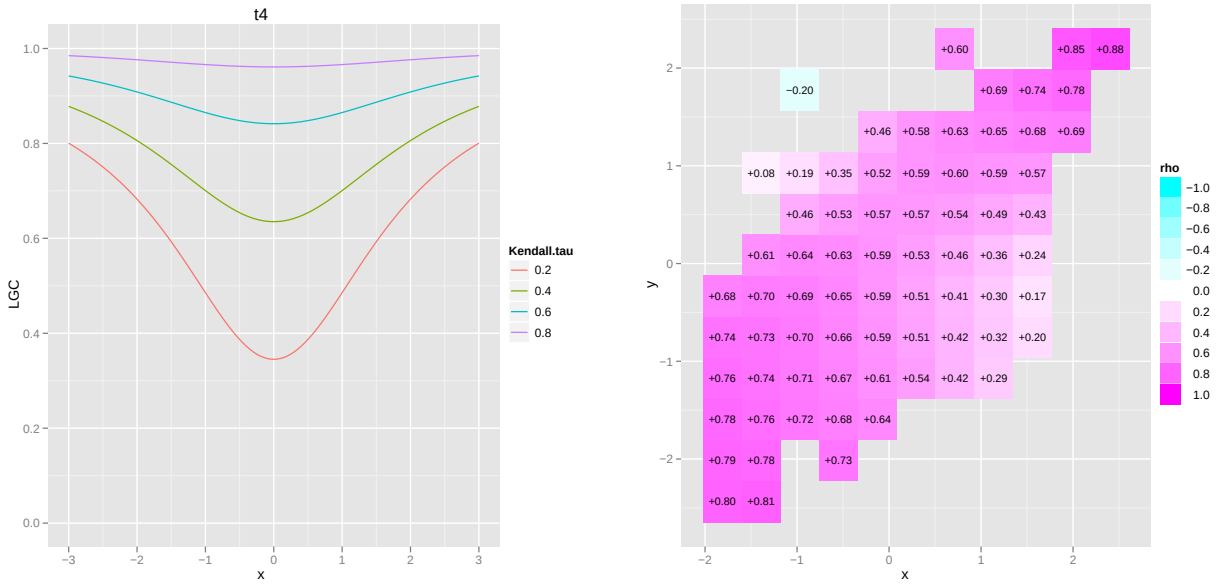
$$C_1(u_1, u_2) = t_{\nu+1}\left(\frac{t_\nu^{-1}(u_2) - \rho t_\nu^{-1}(u_1)}{\sqrt{\frac{(\nu+t_\nu^{-1}(u_1)^2)(1-\rho^2)}{\nu+1}}}\right) := t_{\nu+1}(R),$$

and with f_{t_v} as the standard t-density function and $b = \sqrt{\frac{(\nu + t_\nu^{-1}(u_1)^2)(1-\rho^2)}{\nu+1}}$

$$C_{11}(u_1, u_2) = \frac{\partial R}{\partial u_1} f_{t_{\nu+1}}(R) = \frac{-f_{t_{\nu+1}}(R)}{f_{t_\nu}(t_\nu^{-1}(u_1))b^2} \left(\rho b + \frac{1-\rho^2}{\nu+1} t_\nu^{-1}(u_1) R \right),$$

No simple formula for $\rho(x_1, x_2)$ comes as a result of this. In figure 5 a) we see that $\rho(s, s)$ increase towards each tail which is consistent with the t-copula having both upper and lower tail-dependence.

Figure 5: Local gaussian correlation for the student t-copula: (a) along the diagonal $x_1 = x_2$; (b) Estimated based on $n = 500$ observations.

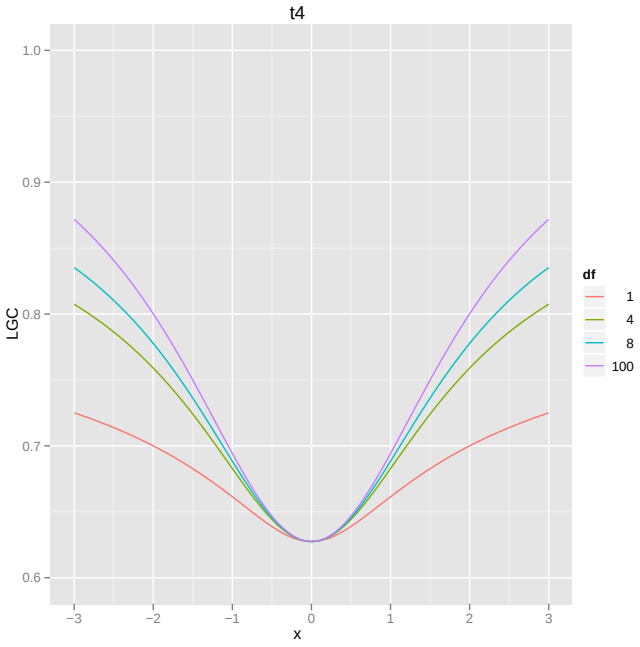


(a) (b)

Example 3.6. (Role of margins) It is of great interest to investigate how the choice of margins affects the local Gaussian correlation. For many combinations of marginal distributions $\rho(x_1, x_2)$ typically reveals the same pattern as long as the copula is kept fixed. In figure 6 we have plotted $\rho(x_1, x_2)$ along the diagonal for the t-copula (4 degrees of freedom, $\rho = 0.58$), but with margins $F_1 = F_2 = t_v$ where t_v is the t-distribution function with v degrees of freedom. We have used four different degrees of freedom v. In this case we see that the the

choice of margins only affects the speed at which $\rho(s,s)$ goes to 1 in the tails. This is quite clear since the degrees of freedom determines the rate at which $t_v(x) \to 0, 1$ as $x \to \pm\infty$, which in turn influences the rate at which the numerator of (3.2) goes to ∞.

Figure 6: $\rho(x_1, x_2)$ for the t-copula with 4 degrees of freedom and $\rho = 0.587$ ($\tau = 0.4$). Both margins are t-distributed with v degrees of freedom, $v = 1, 4, 8, 100$.



4 Evaluating copula models

Given iid observations $X_1, \ldots, X_n$ from $F(x_1, x_2) = C(F_1(x_1), F_2(x_1))$ consider the issue of using local Gaussian correlation to test the null hypothesis

$$H_0 : C \in \mathcal{C}, \quad \mathcal{C} = \{C_\theta : \theta \in \Theta\}, \tag{4.1}$$

where Θ is the parameter space. Let $\rho_\theta(\cdot)$ denote the local Gaussian correlation of the distribution function $C_\theta(F_1(x_1), F_2(x_1))$ (given by (3.9)). A natural approach is to compare the function $\rho_\theta(\cdot)$ estimated under H_0 with the nonparametric estimate described in section 2. In the following section we discuss how a plug-in estimator of ρ_θ can be obtained by

replacing θ, F_1 and F_2 in the analytical expression (3.9) with corresponding estimates under H_0.

4.1 Parametric estimation of local Gaussian correlation

Since $\rho_\theta(\cdot)$ depends on the margins F_1 and F_2 we would, for a full parametric approach, need to make the additional parametric assumption that

$$H_0^{'} : F_1 \in \mathcal{F}_1,\ F_2 \in \mathcal{F}_2,$$

and thus restricting ourselves to the more narrow null hypothesis $H_0 \cap H_0^{'}$. This problem may be overcome by estimating F_j by the empirical distribution function

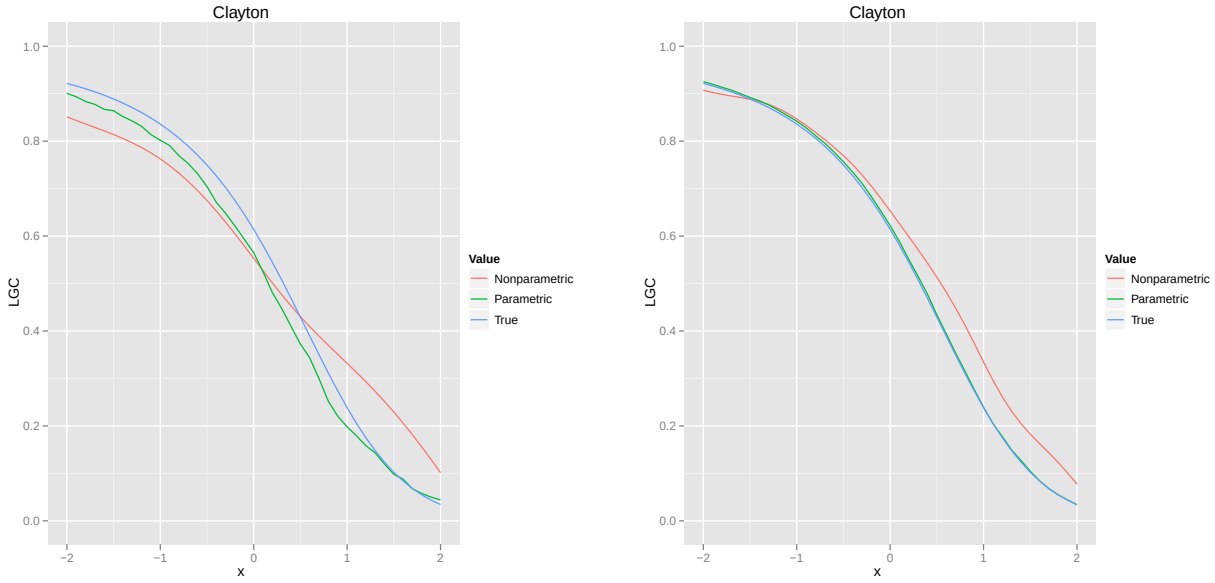
$$\hat{F}_j(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_{ij} \leq x), \quad j = 1, 2. \tag{4.2}$$

The same issue is also encountered when estimating the copula parameter θ under H_0, where a full maximum likelihood approach or the "Inference Functions for Margins" (IMF) approach [see Joe, 1997] requires the additional assumption $H_0^{'}$. This assumption can be avoided by replacing F_j in the likelihood by the empirical distribution function $\hat{F}_j(x)$ (4.2). This method is denoted the pseudo-likelihood [Demarta and McNeil, 2005] or the canonical maximum likelihood [Romano, 2002]. To avoid that the copula density blows up at the boundary of $[0,1]^2$ one typically base the pseudo-likelihood estimation on the scaled ranks $U_1 = (U_{11}, U_{12}) \ldots, U_n = (U_{n1}, U_{n2})$ where $U_{ij} = n\hat{F}_j(X_{ij})/(n+1)$. These values are called the pseudo-observations, and given independent observations $X_1, ..., X_n$ from $C(F_1(x_1), F_2(x_1))$ they can be interpreted as a sample from the underlying copula C. However, one should note that the pseudo-observations are not mutually independent. By using the pseudo-observations one could also estimate the copula parameter using the relation to Kendall-s tau given by equation (3.14). For other rank-based estimators see Tsukahara [2005] and Chen et al. [2006].

Let $\theta_n = \theta_n(U_1, \ldots, U_n)$ denote an estimate of the copula parameter under H_0 based on the pseudo observations $(U_1, \ldots, U_n)$. Then a plug-in estimator $\rho_{\theta_n}(\cdot)$ of $\rho_\theta(\cdot)$ is given by (3.9) with θ replaced by θ_n and with the marginal distribution functions F_j replaced by $\hat{F}_j$, $j = 1, 2$. Let $\rho_{n,b}(\cdot)$ denote the estimate obtained by using the local likelihood method

described in section 2. In general, $\rho_{\theta_n}(\cdot)$ will converge considerably faster towards $\rho_\theta(\cdot)$ than $\rho_{n,b}(\cdot)$ since it is based on $\hat{F}_j$ and θ_n which has the ordinary parametric convergence rate, whereas (Tjøstheim and Hufthammer [2012]) $\rho_{n,b}(\cdot)$ has a substantially slower rate. Figure 7 displays the estimates $\rho_{\theta_n}(\cdot)$ and $\rho_{n,b}(\cdot)$ along the diagonal when F_1 and F_2 are standard gaussian and C is the Clayton copula. In figure 7 (a) the estimates are based on $n = 500$ observations and in figure 7 (b) $n = 5000$. Note that $\rho_{n,b}(\cdot)$ is biased. This bias can be adjusted for, but this problem will be studied generally in a separate publication.

Figure 7: Plot of the true local Gaussian correlation $\rho_\theta(s, s)$, the parametric estimate $\rho_{\theta_n}(s, s)$ and the nonparametric estimate $\rho_{n,b}(s, s)$ against s for the Clayton copula with standard normal margins: (a) $n = 500$, $b_1 = b_2 = 1$; (b) $n = 5000$, $b_1 = b_2 = 0.5$



(a) (b)

4.2 A bootstrap based Goodness-of-fit test

We now turn to the problem of constructing a goodness-of-fit test for H_0. Having established a parametric estimate ρ_{θ_n} of the local Gaussian correlation under H_0 we propose to base a

goodness-of-fit test on the process

$$P_n(\cdot) = \rho_{n,b}(\cdot) - \rho_{\theta_n}(\cdot), \tag{4.3}$$

where $\rho_{n,b}(\cdot)$ is the estimate obtained by using the local likelihood method described in section 2. Aggregation of P_n^2 on $\mathbb{R}^2$ is done over a prespecified grid $(x_1, \ldots, x_p)$ by

$$T_n = \sum_{i=1}^{p} P_n(x_i)^2, \tag{4.4}$$

where large values of T_n lead to the rejection of H_0. By the construction of T_n it is clear that its asymptotic distribution (when scaled properly by some function $\delta(n, b)$) in general depends on the underlying copula and the parameter θ, which in turns means that critical values can not be tabulated by means of the asymptotic properties. Moreover, it is known (see e.g. Terasvirta et al. [2010] chapter 7.7) that in general the asymptotics of functional tests like T_n are not very accurate. We therefore use a parametric bootstrap similar to that of Genest et al. [2009] (see also Stute et al. [1993]) to obtain approximate P-values. The parametric bootstrap procedure is as follows:

Parametric bootstrap

1. Estimate θ, F_1 and F_2 by $\theta_n = \theta_n(U_1, \ldots, U_n)$, $\hat{F}_1$ and $\hat{F}_2$.
2. Obtain $\rho_{n,b}(\cdot)$ by the local likelihood method and $\rho_{\theta_n}(\cdot)$ by replacing θ, F_1 and F_2 in (3.9) by θ_n, $\hat{F}_1$ and $\hat{F}_2$. Compute the value of T_n.
3. For some large integer R, repeat the following steps for every $k \in \{1, \ldots, R\}$:
   (a) Generate a random sample $X_{1k}^*, \ldots, X_{nk}^*$ from the distribution $F^*(x) = C_{\theta_n}(\hat{F}_1(x_1), \hat{F}_2(x_2))$.
   (b) Compute $T_{n,k}^*$ by repeating step 1 and 2 for this sample.

The P-value for this test can then be approximated by $R^{-1} \sum_{k=1}^{R} \mathbf{1}(T_{n,k}^* > T_n)$.

In the Monte Carlo study in section 4.4 we only consider one-parameter copulas so

we have chosen to estimate θ by $\theta_n = m^{-1}(\hat{\tau})$ where $\hat{\tau}$ is the sample Kendall's tau and m is defined by 3.14. The margins F_1 and F_2 are typically estimated by their empirical counterpart given by (4.2), but for n small we suggest using smoothed nonparametric estimates.

Remember that the local likelihood estimate $\rho_{n,b}(\cdot)$ is only consistent with the analytical expression (3.9) along the curve $F_1(x_1) = F_2(x_2)$. This means that each gridpoint $x_i = (x_{i1}, x_{i2})$, $i = 1, \dots p$ should be chosen so that $\hat{F}_1(x_{i1}) \approx \hat{F}_2(x_{i2})$. For example, given suitable values x_{i1}, $i = 1, \dots, p$ one can take $x_{i2} = \hat{F}_2^{-1}(\hat{F}_1(x_{i1}))$, where $\hat{F}_j$ is given by (4.2). Such a curve is illustrated in figure 8. However, this is somewhat restrictive and we therefore provide a second bootstrap procedure which does not rely on the analytical expression (3.9) but where we instead estimate ρ_θ by monte carlo approximation:

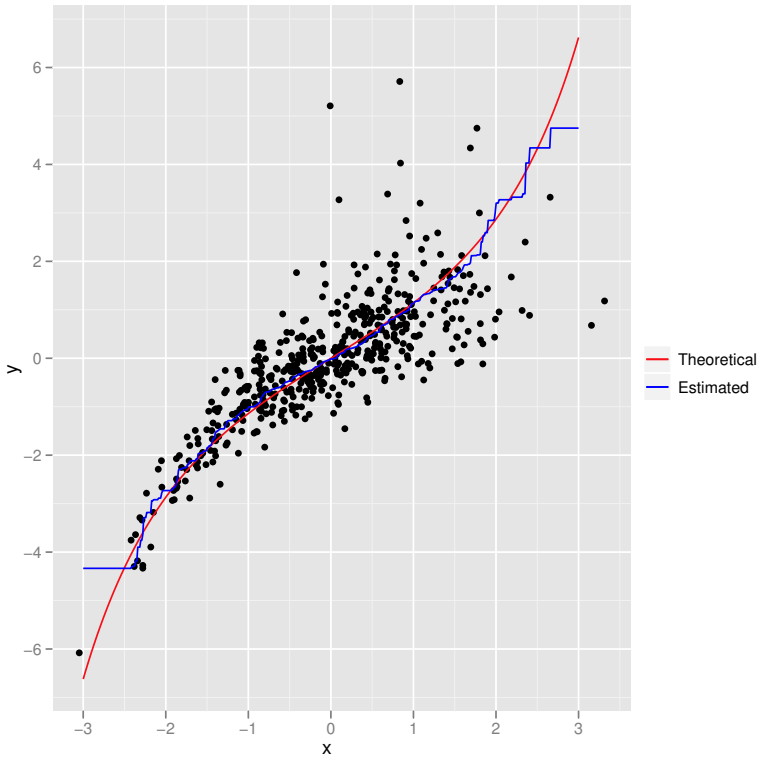
Double parametric bootstrap

1. Estimate θ, F_1 and F_2 by $\theta_n = \theta_n(U_1, \dots, U_n)$, $\hat{F}_1$ and $\hat{F}_2$.
2. Obtain $\rho_{n,b}(\cdot)$ by the local likelihood method and $\rho_{\theta_n}(\cdot)$ by monte carlo approximation:
   (a) For some (preferable large) integer $m \geq n$ generate a random samle $V_1^*, \dots, V_m^*$ from the distribution $F^*(x) = C_{\theta_n}(\hat{F}_{X_1}(x_1), \hat{F}_{X_2}(x_2))$.
   (b) Approximate $\rho_{\theta_n}(\cdot)$ by $\rho_{n,b}(\cdot)$ based on $V_1^*, \dots, V_m^*$.
   (c) Compute the corresponding value of T_n.
3. For some large integer R, repeat the following steps for every $k \in \{1, \dots, R\}$:
   (a) Generate a random samle $X_{1k}^*, \dots, X_{nk}^*$ from the distribution $F^*(x) = C_{\theta_n}(\hat{F}_1(x_1), \hat{F}_2(x_2))$.
   (b) Compute $T_{n,k}^*$ by repeating step 1 and 2 for this sample.

The P-value for this test can then be approximated by $R^{-1} \sum_{k=1}^{R} \mathbf{1}(T_{n,k}^* > T_n)$.

In Genest et al. [2009] a similar bootstrap procedure is used for a number of test statistics in the context of copula goodness-of-fit testing. There it is concluded that for the double bootstrap to be efficient the number m of repetitions must be substantially larger

Figure 8: Plot of the curves $x_2 = F_2^{-1}(F_1(x_1))$ and $x_2 = \hat{F}_2^{-1}(\hat{F}_1(x_1))$ when the data comes from $C_\theta(\phi(x_1), t_4(x_2))$, where C_θ is the Clayton copula with $\theta = 3$, ϕ is the standard normal distribution function and t_4 is the Student's t-distribution function with 4 degrees of freedom.



than the sample size n (minimum $m = 2500$ when $n = 150$). In our case we can expect that even larger values of m is required since a larger m is balanced out by a smaller bandwidth b. This makes the double-boostrap computational demanding and, consequently, we only considered the one-level parametric bootstrap in the simulation study in section 4.4. The advantage of this boostrap procedure is that we can choose the grid $(x_1, \ldots, x_p)$ freely and that it can be used whenever the analytical expression (3.9) is not available. The selection of gridpoints can be done as in Jones and Koch [2003] and Berentsen and Tjøstheim [2012]: First place a regular grid over the area of interest and then select the gridpoints satisfying $\hat{f}(x_i) > C$ for some constant C and a density estimator $\hat{f}$. Alternatively, if the user is interested in a good fit in a particular region the grid can be specified manually, for example in the context of risk management where the fit of the tails are specially important.

4.3 Choice of bandwidth

Choosing the correct bandwidth for the estimate $\rho_{n,b}$ is in general a difficult task and an important topic for future research. A practical bandwidth algorithm is proposed by Tjøstheim and Hufthammer [2012], but when testing $H_0 : C \in \mathcal{C}$ we may choose the bandwidth such that it is optimal if H_0 is true. In general, when the distribution function of X is given by $F(x) = C_\theta(F_1(x_1), F_2(x_2))$ the mean sum of squared error over a grid $(x_1, \ldots, x_p)$ is given by

$$\mathrm{MSSE}(\rho_{n,b}(\cdot)) = E\left(\sum_{i=1}^{p}\left(\rho_{n,b}(x_i) - \rho_\theta(x_i)\right)^2\right) \tag{4.5}$$

Since $\rho_{\theta_n}(\cdot)$ converges faster than $\rho_{n,b}(\cdot)$ (under H_0) it is reasonable to choose b as the minimizer of

$$\widehat{\mathrm{MSSE}}(\rho_{n,b}(\cdot)) = E^*\left(\sum_{i=1}^{n}\left(\rho_{n,b}(x_i) - \rho_{\theta_n}(x_i)\right)^2\right) \tag{4.6}$$

where the expectation E^* is with respect to the distribution function $F^*(x) = C_{\theta_n}(\hat{F}_1(x_1), \hat{F}_2(x_2))$ estimated under H_0. If the grid is the same as for the test statistic T_n, this amounts to minimising the bootstrap mean of the statistic T_n, i.e. $\widehat{\mathrm{MSSE}}(\rho_{n,b}(\cdot)) = E^*(T_n)$.

Table 1 reports bandwidth estimates based on minimizing (4.6). For simplicity $b_1 = b_2 = b$. The resampling distribution $F^*(x) = C_{\theta_n}(\hat{F}_1(x_1), \hat{F}_2(x_2))$ is calculated from a single sample from the copula model $F(x) = C_\theta(F_1(x_1), F_2(x_2))$, with the variations $n = 250, 500$

and $\tau = 0.2, 0.4$. Standard normal margins where used, i.e. $F_1 = F_2 = \Phi$. For comparison, the minimiser of (4.5) (which is computed by monte carlo integration) is given in parentheses. Figure 9 displays $\widehat{\text{MSSE}}(\rho_{n,b}(\cdot))$ as a function of b when C is the Clayton copula.

Not supricingly, neither (4.5) nor (4.6) has a minimum for the Gaussian copula (both decrease as b increases). This is a result of the local Gaussian likelihood beeing equivalent with the global Gaussian likelihood when $b \to \infty$. However, it is not recommended to use a very large bandwidth when testing for the Gaussian copula, since too much smoothing results in poor power when the null hypothesis is false.

Figure 9: $\widehat{\text{MSSE}}(\rho_{n,b}(\cdot))$ versus b for the Clayton copula.

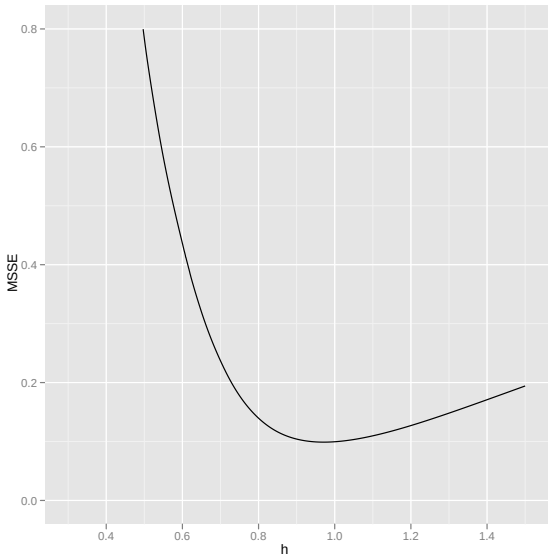


Table 1: Estimated bandwidth based on $\widehat{\text{MSSE}}(\rho_{n,b}(\cdot))$ calculated from a single sample from each copula model for $n = 250, 500$ and $\tau = 0.2, 0.4$. The minimizer of $\text{MSSE}(\rho_{n,b}(\cdot))$ is given in the parentheses.

Copula	$\tau = 0.2$		$\tau = 0.4$	
	$n = 250$	$n = 500$	$n = 250$	$n = 500$
Clayton	0.9731(1.0536)	0.9710(0.9515)	0.8465(0.8525)	0.8116(0.7955)
Gumbel	1.1030(1.0674)	0.9198(0.9350)	1.0084(0.9602)	0.9116(0.8896)
Frank	1.4065(1.7846)	1.0166(1.0423)	0.8545(0.7824)	0.7165(0.7122)
Gaussian	∞	∞	∞	∞
t4	0.9176(0.8728)	0.8290(0.7778)	0.8786(0.7977)	0.7675(0.7055)

4.4 Simulation study

A Monte Carlo study is performed to assess the finite-sample properties of the proposed goodness-of-fit test (4.4) (based on the one-level parametric bootstrap). In order to examine its performance, we compare it with a much used test proposed by Genest and Rémillard [2008].

This particular test is chosen because of its good overall performance in the simulation studies of Genest et al. [2009] and Berg [2009]. The test is based on the empirical copula process

$$C_n(u) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(U_{i1} \leq u_1, U_{i2} \leq u_2), \tag{4.7}$$

where $U_i = (U_{i1}, U_{i2})$ $i = 1, \ldots, n$ are the pseudo-observations and $u = (u_1, u_2) \in [0, 1]^2$. A natural test consist in comparing a distance between C_n and an estimate C_{θ_n} of C obtained under H_0. Then a goodness-of-fit test may be based on the Cramer-von-Mise type statistic

$$A_n = n \int_{[0,1]^2} \{C_n(u) - C_{\theta_n}(u)\}^2 \mathrm{d}C_n(u), \tag{4.8}$$

which in turn may be estimated by $\hat{A}_n = \sum_{i=1}^{n} \{C_n(U_i) - C_{\theta_n}(U_i)\}^2$. Further one proceed by parametric bootstrap analogue to the first procedure described in section 4.2 to find an approximate P-value for the test. If an analytical expression for C_θ is not available one may resort to a double parametric bootstrap analogue to the second bootstrap procedure described in section 4.2. For a more detailed description of these test procedures we refer to

Genest et al. [2009] or Berg [2009].

We are interested in examining the nominal level (arbitrarily fixed at 5 % throughout the study) and the power against fixed alternatives. The simulation design is as follows:

- Five H_0 copulas: Clayton, Gumbel, Frank, Gaussian and Student with 4 degrees of freedom
- Five H_1 copulas: Clayton, Gumbel, Frank, Gaussian and Student with 4 degrees of freedom
- Two degrees of global dependence: Kendall's tau $\tau = \{0.2, 0.4\}$
- Two sample sizes: $n = \{250, 500\}$

For every combination of the above setup, a sample of size n is drawn from $C(F_1(x_1), F_2(x_2))$ where C is the copula under H_1 with dependence parameter corresponding to τ. As margins we have used $F_1 = F_2 = \Phi$, but when testing H_0 : Clayton against H_1 : Gaussian we have also considered $F_1 = F_2 = t_4$ (labelled Gaussian*) and $F_1 = \Phi$, $F_2 = t_4$ (labelled Gaussian**). The test statistics for our proposed test (4.4) and for the alternative test (4.8) is then computed under H_0, and P-values are estimated using the parametric bootstrap procedure described in section 4.2. In the estimation of (4.4) the bandwidth $b = b_1 = b_2$ is taken from table 1, except when H_0 is Gaussian in which case we put $b_1 = b_2 = 1$. The prespecified grid $(x_1, \ldots, x_p)$ are 10 fixed points along the diagonal, except in the case when $F_1 = \Phi$, $F_2 = t_4$, in which case we have used 10 fixed point along the curve $x_2 = t_4^{-1}(\Phi(x_1))$ (corresponding to the curve $F_1(x_1) = F_2(x_2)$). Choosing the grid in all cases according to $x_1 = \hat{F}_1^{-1}\left(\hat{F}_2(x_2)\right)$ would lead to very similar results. The procedure is repeated over 500 independent runs to estimate the nominal level and power for the test for each case. The number of bootstrap samples was fixed at $R = 500$.

Table 2 reports the level and power of the test (4.4) and the test (4.8) (in parenthesis). Each line of the table shows the percentage of rejection of H_0 associated with the two tests for different combinations of sample size and dependence, and given a choice of the copula under H_0 and a true underlying copula. The nominal levels match relatively well the prescribed size of 5%. However, in some cases the level is slightly underestimated. The power of our proposed test is very good compared to the A_n based test, and the power of the A_n based test in our simulation study corresponds very well with the power found in

similar studies by Genest et al. [2009] and Berg [2009]. Note that for testing the Gaussian and Student hypothesis, powers are in general lower than for testing Clayton, Gumbel and Frank hypothesis. This is also in line with the previous mentioned studies.

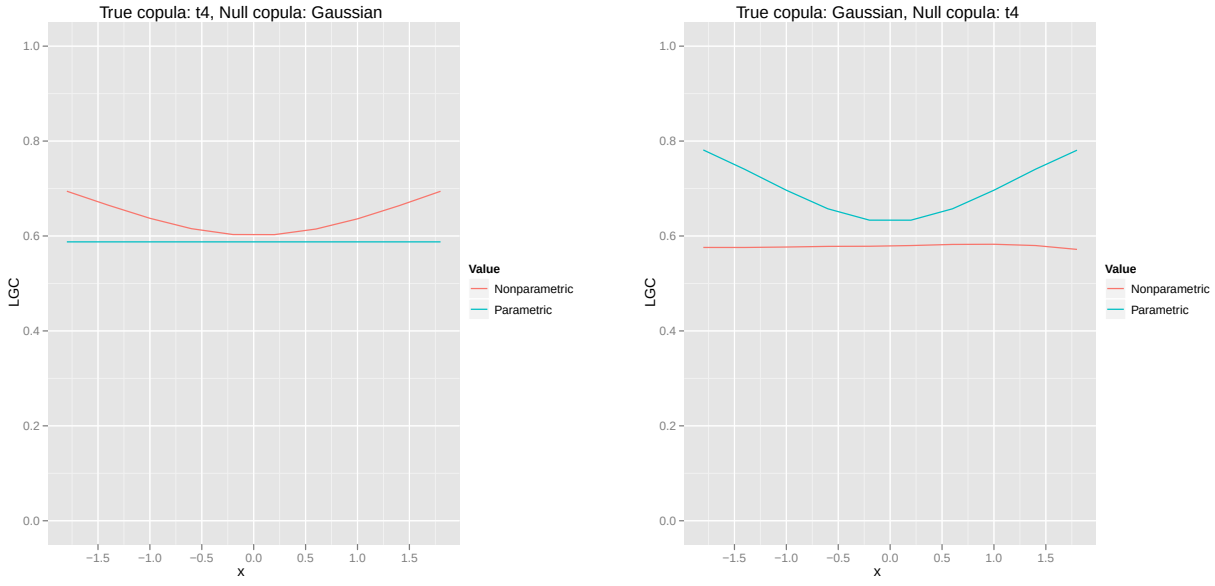
Note that there is a rather large asymmetry in power when testing case (a); H_0 : Gaussian copula, when H_1 is the t-copula with 4 degrees of freedom and case (b); H_0 : t-copula with 4 degrees of freedom, when H_1 is the Gaussian copula. We investigated this feature further in the case $n = 250$ and $\tau = 0.4$ by considering the average estimate of $\rho_{n,b}(x_i)$ and $\rho_{\theta_n}(x_i)$ (under the H_0) in the selected gridpoints of the 500 independent samples used in the simulation study. The result can be seen in figure 10. We see that even though $\rho_{n,b}(x_i)$ and $\rho_{\theta_n}(x_i)$ are changing roles, $\rho_{n,b}(x_i)$ seems to be underestimating in the tails of the student t-copula in case a). Thus $\rho_{n,b}(x_i)$ is closer to $\rho_{\theta_n}(x_i)$, which in turns means that T_n does not capture the difference between $\rho_{n,b}(x_i)$ and $\rho_{\theta_n}(x_i)$ in case (a) to the same degree as in case (b). Case (a) and (b) are also the only cases where the power did not increase with the level of dependence. Indeed, as was seen in figure 4 the local Gaussian correlation for the student t-copula becomes more constant as the level of dependence increase and thus resembling more the Gaussian structure.

It is quite clear that the power of the test dependends strongly on the choice of gridpoints $(x_1, \ldots, x_p)$ and for a arbitrary grid (not restricted by $\hat{F}_1(x_{1i}) = \hat{F}_2(x_{2i})$) the double bootstrap procedure must be employed. However, when testing for the Gaussian copula we do not face the uniqueness problems discussed in section 2.1 since $\rho(x_1, x_2) = \rho'(x_1, x_2) = \rho$ so for this test we can choose the grid freely. An obvious case where we can improve power by expanding the grid is when the alternative hypothesis is the student t-copula. This is because the t-copula also exhibits negative dependence along the off-diagonal when the level of dependence is small. To illustrate this we added two points along the off-diagonal to the grid and repeated the simulation study for this particular setup (H_0: Gaussian, H_1: student t-copula $v = 4$ degrees of freedom). The result can be seen in table 2 in the row labelled Student 4 df*. Adding just these two points leads to a very significant increase of power.

An obvious drawback of this simulation study is that we have focused on standard normal margins. It is therefore difficult to make too categorical conclusions about the power of the test since the result may vary according to which margins we choose. For example, a smaller experiment with t-distributed margins revealed that larger sample sizes was needed to obtain reasonable levels and we also had to choose gridpoints further out in the tails to obtain good

power. This can partially be explained by figure 6 where it is seen that $\rho(s,s) \to 1$ slower as the degrees of freedom of the t distributed margins decrease. This feature is also the case for the other copulas considered in this paper. Thus, to discriminate between two copulas with different types of tail-dependence $\rho_{\theta_n}(x_1, x_2)$ must be compared with $\rho_{n,b}(x_1, x_2)$ further out in the tails when the margins themselves are heavy-tailed.

Figure 10: Average estimate of $\rho_{n,b}(x_i)$ and $\rho_{\theta_n}(x_i)$ in case (a) and case (b) when $n = 250$ and $\tau = 0.4$.



(a) (b)

Copula under H_0	True copula	$\tau = 0.2$		$\tau = 0.4$	
		$n = 250$	$n = 500$	$n = 250$	$n = 500$
Clayton	Clayton	**5.8(4.6)**	**5.2(5.0)**	**5.2(5.2)**	**4.4(5.0)**
	Gumbel	97.6(90.0)	100.0(99.8)	100.0(100.0)	100.0(100.0)
	Frank	77.8(60.3)	96.8(88.8)	97.2(97.6)	99.8(100.0)
	Gaussian	76.6(58.2)	92.6(77.8)	97.0(95.4)	100.0(100.0)
	Gaussian*	61.8	71.0	81.4	95.4
	Gaussian**	66.8	77.6	83.8	93.0
	Student 4 df	82.6(69.2)	97.6(84.6)	99.6(97.8)	100.0(100.0)
Gumbel	Clayton	99.0(81.2)	100.0(99.0)	100.0(99.8)	100.0(100.0)
	Gumbel	**5.4(5.8)**	**6.4(5.2)**	**3.0(5.2)**	**5.2(4.8)**
	Frank	66.0(20.2)	90.6(49.0)	89.4(50.2)	99.6(92.2)
	Gaussian	52.6(11.4)	83.6(36.0)	73.8(27.8)	96.2(69.0)
	Student 4 df	37.0(20.2)	76.8(54.6)	71.8(36.6)	98.0(80.4)
Frank	Clayton	68.0(50.6)	94.6(85.6)	97.6(95.8)	100.0(100.0)
	Gumbel	50.4(40.4)	87.8(62.8)	86.2(76.0)	100.0(97.4)
	Frank	**3.2(4.8)**	**4.0(4.6)**	**2.4(4.6)**	**4.2(5.0)**
	Gaussian	9.4(8.0)	20.6(15.6)	35.4(19.2)	69.8(49.6)
	Student 4 df	82.2(27.8)	99.6(52.0)	95.0(46.0)	100.0(87.0)
Gaussian	Clayton	59.4(44.2)	90.8(73.6)	98.4(93.4)	100.0(100.0)
	Gumbel	24.6(33.2)	64.6(42.8)	51.6(58.2)	95.0(82.4)
	Frank	7.8(7.6)	14.0(7.0)	28.2(21.4)	65.0(35.4)
	Gaussian	**4.6(5.2)**	**4.0(4.8)**	**3.8(5.0)**	**4.0(5.4)**
	Student 4 df	24.0(20.6)	78.4(26.6)	11.0(21.2)	60.0(23.8)
	Student 4 df†	35.4	82.8	34.2	79.4
Student 4 df	Clayton	77.4(35.8)	97.0(69.8)	97.6(88.4)	100.0(99.6)
	Gumbel	51.2(23.0)	82.6(34.0)	64.6(45.8)	91.2(63.4)
	Frank	78.4(9.2)	99.0(16.8)	91.8(26.8)	100.0(48.4)
	Gaussian	64.6(5.2)	92.8(7.8)	61.8(4.0)	92.0(2.8)
	Student 4 df	**4.4(5.4)**	**4.8(5.0)**	**2.8(4.8)**	**4.8(4.8)**

Table 2: Percentage of rejection of H_0 by the T_n-based test and the A_n-based test (in paranthesis) for data sets of different sizes arising from different copula models with dependence $\tau = 0.2$ or $\tau = 0.4$. * Power when both margins are t-distributed. ** Power when one margin is standard normal and one is t-distributed. † Power when the grid is expanded.

5 Visualizing departures from H_0

If the null is rejected, the cause of the rejection can be investigated by visually comparing the non-parametric estimate $\rho_{n,b}(\cdot)$ with the estimate under the null hypothesis $\rho_{\theta_n}(\cdot)$. One possibility is to plot $\rho_{n,b}(\cdot)$ and $\rho_{\theta_n}(\cdot)$ seperately on a grid on $\mathbb{R}^2$. However, this requires monte carlo estimation of ρ_{θ_n} since (3.9) is not valid on all of $\mathbb{R}^2$ and it can also be argued that such plots are too detailed and, perhaps, difficult to interpret. We therefore believe that such a direct comparison is best suited along curves as in figure 11-12 (a).

To obtain diagnostic plots on $\mathbb{R}^2$ we suggest to plot the results of "local goodness-of-fit" tests performed over a grid on $\mathbb{R}^2$, not limited to a curve. That is, for every point x_j of a pre-specified grid $(x_1, \ldots, x_p)$ we first test the null hypothesis that $\rho(x_j) = \rho_\theta(x_j)$, where $\rho_\theta(x_j)$ is the local Gaussian correlation at x_j under the original H_0. Our test statistic for this purpose is simply $\rho_{n,b}(x_j)$, and the following bootstrap procedure is used to construct critical values:

Local goodness-of-fit test

1. Estimate θ, F_1 and F_2 by $\theta_n = \theta_n(U_1, \ldots, U_n)$, $\hat{F}_1$ and $\hat{F}_2$.
2. Compute the value of $\rho_{n,b}(x_j)$ by the local likelihood method.
3. For some large integer R, repeat the following steps for every $k \in \{1, \ldots, R\}$:
   (a) Generate a random sample $X^*_{1k}, \ldots, X^*_{nk}$ from the distribution $F^*(x) = C_{\theta_n}(\hat{F}_1(x_1), \hat{F}_2(x_2))$.
   (b) Compute $\rho^{*,k}_{n,b}(x_j)$ by repeating step 2 for this sample.

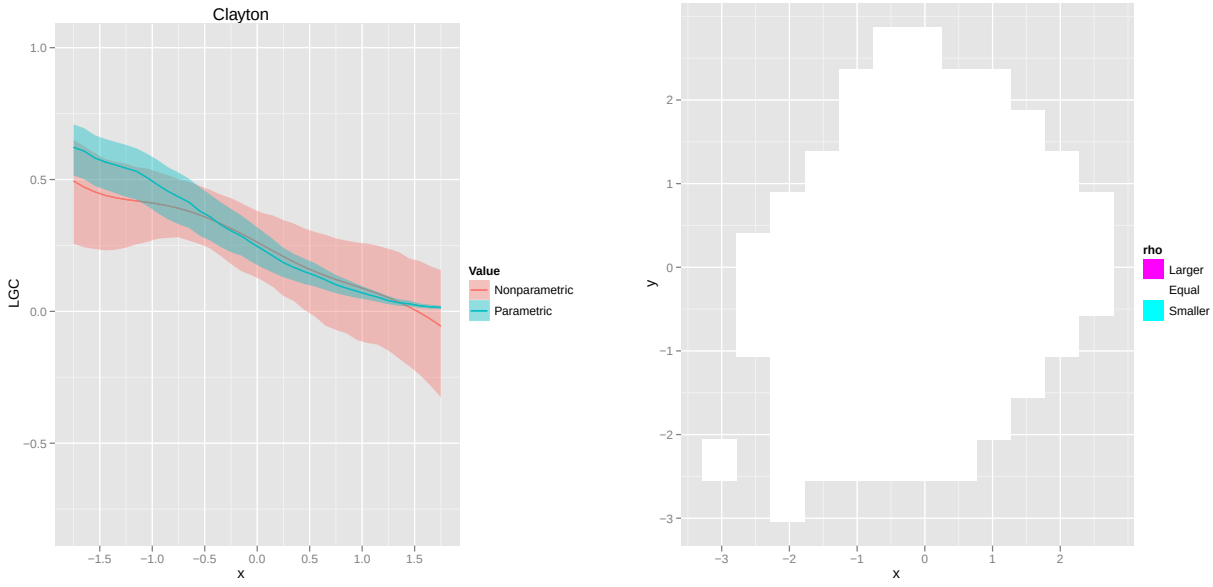
For a given significance level α, we reject the hypothesis that $\rho(x_j) = \rho_\theta(x_j)$ if $\rho_{n,b}(x_j)$ is respectively smaller or larger than the $(\alpha/2)\%$ or $(1-\alpha/2)\%$ quantile of $\rho^{*,1}_{n,b}(x_j), \ldots, \rho^{*,R}_{n,b}(x_j)$.

To present the test results we adopt the idea of Jones and Koch [2003] used to construct so-called "dependence maps": If $\rho_{n,b}(x_j)$ is significantly larger, x_j is assigned the colour magenta; if significantly smaller, the colour cyan; if the null hypothesis is not rejected, the colour white.

Figures 11 (a)-(b) illustrates how the diagnostic plots look when the null hypothesis that C is the Clayton copula is true. The $n = 500$ real data comes from $C(\Phi(x_1), \Phi(x_2))$, where C is the Clayton copula with parameter $\theta = 0.5$. The T_n test does not reject Clayton as the null copula (P-value=0.788). In figure 11 (a) we have plotted $\rho_{n,b}(\cdot)$ and $\rho_{\theta_n}(\cdot)$ along the curve $x_1 = x_2$ (since the true margins are both standard normal this curve corresponds to the curve $F_1(x_1) = F_2(x_2)$). Using an estimated curve with $x_1 = \hat{F}_1^{-1}\left(\hat{F}_2(x_2)\right)$ would give very similar results. For this plot we have also added standard 95% bootstrap confidence intervals. In 11 (b) we have plottet the results of the local goodness-of-fit test, where we have used $\alpha = 0.05$ and $R = 1000$ bootstrap samples.

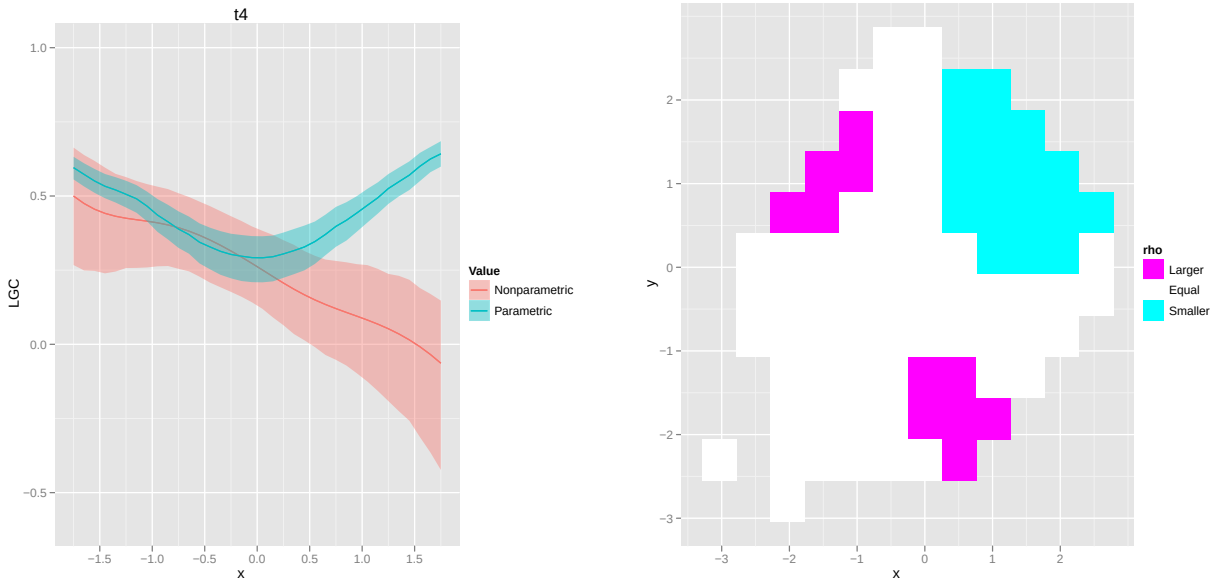
Figures 12 (a)-(b) is based on the same data as above, but now the null hypothesis is that C is the t-copula with 4 degrees of freedom. With a 5% significance level this hypothesis is rejected by the T_n test (P-value=0.002). Figure 12 (a) clearly shows that $\rho_{n,b}(\cdot)$ is smaller than ρ_{θ_n} in the upper tail. This is an indication that the estimated t-copula assigns to much dependence in the upper tail compared to the data. This is indeed confirmed by the local goodness-of-fit tests displayed in 12 (b). Also notice that 12 (b) reveals that $\rho(x_j)$ is significally larger than ρ_θ in points of the 2nd and 4th quadrant. This is a result of the t-copula having negative dependence in these regions when the copula parameter is close to zero (See Berentsen and Tjøstheim [2012] for a detailed study of the t-distribution.)

Figure 11: Evaluation of H_0: Clayton when Clayton is the true copula: (a) Diagonal plot of $\rho_{n,b}$ and ρ_{θ_n} where ρ_{θ_n} is estimated under H_0: Clayton; (b) Pointwise test, H_0 : Clayton



(a) (b)

Figure 12: Evaluation of H_0: t-copula when Clayton is the true copula: (a) Diagonal plot of $\rho_{n,b}$ and ρ_{θ_n} where ρ_{θ_n} is estimated under H_0: t-copula; (b) Pointwise test, H_0 : Clayton

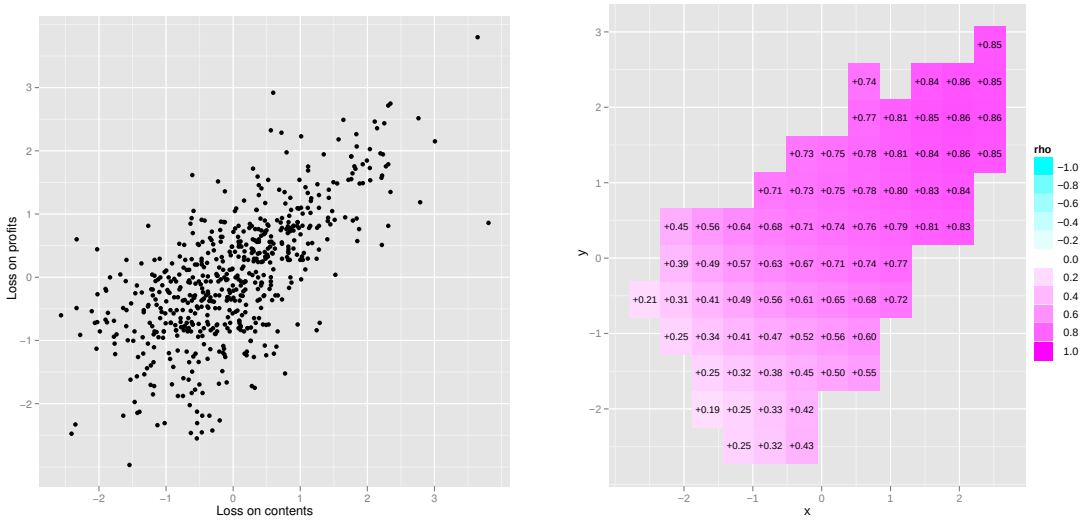


(a) (b)

6 A real data study

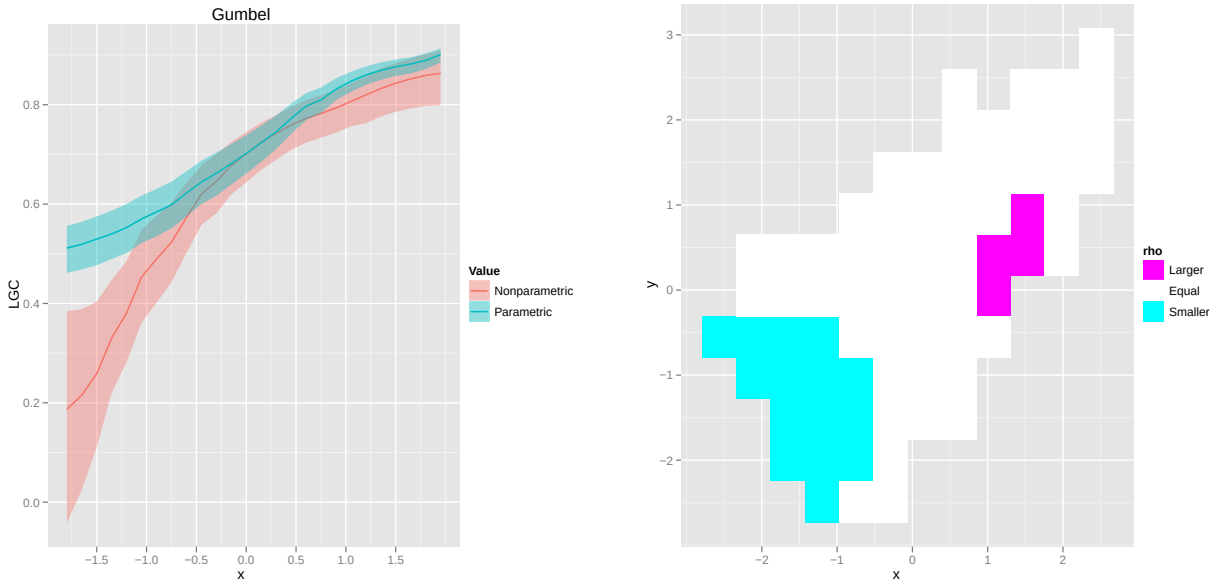
The Danish fire insurance claims data has been shown much interest in actuarial science and extreme value theory [see e.g. Mcneil, 1997]. The data consists of 2167 losses over one million DKK from the years 1980 to 1990 inclusive. There were registered in total 604 cases where a loss in both contents and profits occured, and the standardized log-transformed values of these claims can be seen in figure 13 (a) along with the nonparametric estimate $\rho_{n,b}(\cdot)$ in figure 13 (b). From the latter figure we see that the dependence increase towards the upper tail. Indeed, in terms of AIC (Akaike Information Criteria), the best ranked copula amongst Clayton, Frank, Gumbel, Gaussian and student-t was the Gumbel copula (-379.6), followed by the Gaussian copula (-327.9). Nevertheless, the null hypothesis that the copula of the data is the Gumbel copula was rejected by the goodness-of-fit test proposed in section 4.2 (P-value ≈ 0). However, by using the diagnostic plots proposed in section 5 we now have the possibility to investigate the characteristics of the discrepancy between the data and the null hypothesis. In figure 14 (a) the parametric estimate of the local gaussian correlation is plotted together with the nonparametric estimate along the curve $x_2 = \hat{F}_2^{-1}(\hat{F}_1(x_1))$ (which is close to the diagonal curve). We see that the (fitted) Gumbel copula assigns too large correlation in the lower tail compared to the data. This is also supported by figure 14 (b) which displays the result of the local goodness-of-fit test. It is perhaps not so suprising that the Gumbel copula is rejected since the data set is fairly complicated, and it is unlikely that the dependence relationship should be adequately described by just one parameter as is the case for the Gumbel copula. Screening some of the outliers did increase the p-value slightly, but not enough to accept the null hypothesis.

Figure 13: Standardized log-transformed values of loss on contents and loss on profits: (a) Scatterplot; (b) Estimated local Gaussian correlation



(a) (b)

Figure 14: Diagnostic plots: (a) Parametric versus nonparametric local Gaussian correlation; (b) Local goodness-of-fit test



(a) (b)

7 Concluding remarks

In this paper we have developed the theoretical relationship between the local Gaussian correlation and the dependence structure in different bivariate copula models, i.e. the dependence of a exchangeable copula can be expressed by an explicit formula for the $\rho(x_1, x_2)$ along the curve $F_1(x_1) = F_2(x_2)$. Further, by plotting $\rho(x_1, x_2)$ along such curves we are able to shed some new light over the dependence structure of some of the most useful copulas. In particular, because of the easy expression of the Archimedean copulas we can in many cases easily calculate the population value of the local Gaussian correlation when marginal distribution has been chosen. This gives us the possibility to visualize the dependence arising from different copula models in the framework of the standard correlation coefficient, but interpreted locally.

The local Gaussian correlation can be used for selecting the most appropriate copula in a situation where bivariate data are given and one wishes to fit a copula model. As a first exploratory step we suggest to look at two-dimensional plots of the (nonparametric) estimate of $\rho(x_1, x_2)$. This gives the user a crude description of the dependence structure in the data.

Given a copula candidate we have proposed a goodness-of-fit test, where we compare the difference between the nonparametric and parametric estimate of local Gaussian correlation estimated for the data in question. To construct the null distribution of the proposed test statistic we have used two different parametric bootstrap procedure. One which computes the parametric estimate of $\rho(x_1, x_2)$ by the analytical formula derived in section 2 while the other procedure does this estimation by monte carlo approximation. A Monte Carlo simulation experiment confirms that the former bootstrap test performs very well in the case of standard normal margins, also compared to a bona fide test based on the empirical copula. However, these result are not general since other choices of margins has been seen to influence both the power and the level of the test. For example, in the case of t-distributed margins with $v = 4$ degrees of freedom, larger sample sizes was needed to obtain reasonable levels and the grid points of the test statistic had to be changed in order to obtain comparable power. Also, when the level of dependence in the data is very high the level of the one-level bootstrap test does not meet the prescribed level. We believe this to be the result of bias in the nonparametric estimate of $\rho(x_1, x_2)$, since the same feature is not observed for the double parametric bootstrap test.

As a supplement to the goodness-of-fit test we propose two types of diagnostic plots which gives the user an overview of the copula fit on a local level. Besides information about the fit (or inadequacies) of the current model, these plot can in many cases also point the user in the direction of a better model.

At this point a limitation to the proposed methods is that they require input about bandwidths and choice of gridpoints. Though some guidelines for selecting such inputs has been proposed in this paper there is a need to develope more efficient and automatic methods.

References

G. D. Berentsen and D. Tjøstheim. Recognizing and visualizing departures from independence in bivariate data using local gaussian correlation. *Working paper*, 2012.

D. Berg. Copula goodness-of-fit testing: an overview and power comparison. *The European Journal of Finance*, 15:675–701, 2009.

W. Breymann, A. Dias, and P. Embrechts. Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, 3(1):1–14, 2003.

D. Brigo, A. Pallavicini, and R. Torresetti. *Credit Models and the Crisis: A Journey Into CDOs, Copulas, Correlations and Dynamic Models*. The Wiley Finance Series. John Wiley & Sons, 2010. ISBN 9780470665664. URL `http://books.google.no/books?id=Mt5Zxt0rVBkC`.

X. Chen, Y. Fan, and A. Patton. Simple tests for models of dependence between multiple financial time series, with applications to u.s. equity returns and exchange rates. Working Paper no 483, London Economics Financial Markets Group, 2004.

X. Chen, Y. Fan, and V. Tsyrennikov. Efficient estimation of semiparametric multivariate copula models. *Journal of the American Statistical Association*, 101(475):1228–1240, 2006. doi: 10.1198/016214506000000311. URL `http://pubs.amstat.org/doi/abs/10.1198/016214506000000311`.

Umberto Cherubini, Elisa Luciano, and Walter Vecchiato. *Copula methods in finance*. John Wiley & Sons, Hoboken, NJ, 2004. URL `http://www.loc.gov/catdir/bios/wiley047/2004002624.html`.

Loran Chollete, Andreas Heinen, and Alfonso Valdesogo. Modeling international financial returns with a multivariate regime switching copula. *Journal of Financial Econometrics*, 7:437–480, 2009.

P. Deheuvels. La fonction de dependance empirique et ses proprietes: Un test non parametric d'independence. *Bulletin de la Classe des Sciences, 5e Series*, pages 274–292, 1979.

Stefano Demarta and Alexander J. McNeil. The t copula and related copulas. *International Statistical Review*, 73(1):111–129, 2005. ISSN 1751-5823. doi: 10.1111/j.1751-5823.2005.tb00254.x. URL http://dx.doi.org/10.1111/j.1751-5823.2005.tb00254.x.

C. Genest and B. Rémillard. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales Henri Poincaré*, 44:1096—-1127, 2008.

C. Genest, B. Remillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance Mathematics and Economics*, 44:199–213, 2009.

N. L. Hjort and M. C. Jones. Locally parametric nonparametric density estimation. *Ann. Statist.*, 24(4):1619–1647, 1996. ISSN 0090-5364. doi: 10.1214/aos/1032298288. URL http://dx.doi.org/10.1214/aos/1032298288.

D. Huard, G. Evin, and A.C. Favre. Bayesian copula selection. *Journal of Computational Statistics and Data Analysis*, 51:809–822, 2006.

P. Jaworski, F. Durante, W. Härdle, and T. Rychlik. *Copula theory and its applications: proceedings of the workshop held in Warsaw, 25-26 September 2009*. Lecture notes in statistics. Springer, 2010. ISBN 9783642124648. URL http://books.google.no/books?id=vX233feaA6MC.

H. Joe. *Multivariate models and dependence concepts*. Chapman and Hall, London, 1997.

M. C. Jones and I. Koch. Dependence maps: local dependence in practice. *Statistics and Computing*, 13(3):241–255, August 2003. URL http://oro.open.ac.uk/22636/.

D. Karlis and A.K. Nikoloupoulos. Copula model evaluation based on parametric boostrap. *Journal of Computational Statistics and Data Analysis*, 52:3342–3353, 2008.

A. J. Mcneil. Estimating the tails of loss severity distributions using extreme value theory. *ASTIN bulletin*, 27:117–132, 1997.

R. B. Nelsen. *An introduction to copulas*. Springer, New York, 2006.

Tatsuyoshi Okimoto. New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis*, 43:787–816, 2008.

Claudio Romano. Calibrating and simulating copula functions: An application to the italian stock market. *Risk Management*, page 23, 2002.

M. Rosenblatt. Remarks on a multivariate transformation. *The Annals of Mathematical Statistics*, 23:470–472, 1952.

A. Sklar. Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institute de Statistique de l'Universite de Paris 8*, pages 229–231, 1959.

Winfried Stute, Wenceslao Manteiga, and Manuel Quindimil. Bootstrap based goodness-of-fit-tests. *Metrika*, 40:243–256, 1993. ISSN 0026-1335. URL `http://dx.doi.org/10.1007/BF02613687`. 10.1007/BF02613687.

Timo Terasvirta, Dag Tjøstheim, and Clive W. J. Granger. *Modelling Nonlinear Economic Time Series*. Oxford University Press, 2010.

D. Tjøstheim and K.O. Hufthammer. Local gaussian correlation: A new measure of dependence. Manuscript revised for Journal of the American Statistical Assosiation, 2012. URL `http://folk.uib.no/st11188/local-gaussian-correlation/local-gaussian-correlation.pdf`.

Hideatsu Tsukahara. Semiparametric estimation in copula models. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 33(3):pp. 357–375, 2005. ISSN 03195724. URL `http://www.jstor.org/stable/25046185`.

Copulas are much used to model nonlinear and non-Gaussian dependence between stochastic variables. Their functional form is determined by a few parameters, but unlike a dependence measure like the correlation, these parameters do not have a clear interpretation in terms of the dependence structure they create. In this paper we examine the relationship between a newly developed local dependence measure, the local Gaussian Correlation, and standard copula theory. We are able to describe characteristics of the dependence structure in different copula models in terms of the local Gaussian correlation. In turn, these characteristics can be effectively visualized. More formally, the characteristic dependence structure can be used to construct a goodness-of-fit test for bivariate copula models by comparing the theoretical local Gaussian correlation for a specific copula and the estimated local Gaussian correlation. A Monte Carlo study reveals that the test performs very well compared to a commonly used alternative test. We also propose two types of diagnostic plots which can be used to investigate the cause of a rejected null. Finally, our methods are used on a ”classic” insurance data set.

SNF
Et selskap i NHH-miljøet

SAMFUNNS- OG NÆRINGSLIVSFORSKNING AS

Institute for Research in Economics and Business Administration

Breiviksveien 40
N-5045 Bergen
Norway
Phone: (+47) 55 95 95 00
Fax: (+47) 55 95 94 39
E-mail: publikasjon@snf.no
Internet: http://www.snf.no/

Trykk: Allkopi Bergen