Dept. of Math. University of Oslo
Statistical Research Report No. 10
ISSN 0806–3842 December 2005

A Goodness-of-fit Test for Copulae Based on the Probability Integral Transform

Daniel Berg*
Department of Mathematics,
University of Oslo, Norway.
Henrik Bakken†
Boston Consulting Group,
Oslo, Norway.

Abstract

Copulae is a growing field of interest and application for dependency modelling. There is however no predominant way of choosing the copula model that best fits a given data set. We introduce a new goodness-of-fit test, based on the probability integral transform. The test is consistent, numerically efficient and incorporates a weighting functionality. Results show that the test performs well and that the weighting functionality is very powerful. Applied to stock portfolios the test strongly rejects the Gaussian and the Clayton copulae, while the Student's t copula provides a good fit.

Keywords: Copulae, Goodness-of-fit, Probability Integral Transform

*Address for correspondence: Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, NO-0316 Oslo, Norway. E-mail: daniel@nr.no.
†E-mail: Bakken.Henrik@bcg.com

Contents

1 Introduction

Copulae have proved to be a very useful tool in the analysis of dependency structures. The concept of copulae was introduced by Sklar (1959), but was first used for financial applications by Embrechts et al. (1999). Since then we have seen a tremendous increase of copula related research and applications. One of the most attractive properties of copulae is the decoupling of the copula and the margins, enabling us to capture the full dependency structure without considering the margins. Another very attractive feature is the invariance to strictly increasing transformations. For a thorough analysis of copulae, see Joe (1997) or Nelsen (1999).

The limitation of the copula approach is the lack of a predominant way of choosing the copula model that best fits a given data set. Prior to the use of goodness-of-fit (GOF) tests, various information criterions were employed, such as Aikaike's Information Criterion (AIC). These are suboptimal tests that do not provide us with any understanding of the size of the decision rule employed, nor its power. Hence, GOF tests are preferred.

Lately, several copula GOF tests have been proposed in literature. Chen et al. (2004) propose a test based on the probability integral transform (PIT) of Rosenblatt (1952). The PIT transforms a set of dependent variables into a set of independent $U(0,1)$ variables, given the multivariate distribution. Genest et al. (2005) propose a GOF test based on the Kendall's process, while Panchenko (Panchenko) propose a test based on positive definite bilinear forms.

Chen et al. (2004) propose two tests. The first suffers the curse of dimensionality, while the second test, which is based on the test by Breymann et al. (2003), does not have this problem. It is however not consistent. This means that the test is not strictly increasing for every deviance from the null hypothesis, there may be deviations cancelling each other. Chen et al. (2004)'s test weight the tails of the copulae, implicitly, through the squared inverse gaussian cumulative distribution function (cdf).

We introduce a new test which is consistent and numerically efficient. The test decouples the estimation of deviance from the null hypothesis and the weighting, such that any weight function or no weight at all may be applied. This flexibility in the weighting function is appropriate for instance in applications where one wishes to focus more on specific areas of the copulae, e.g. the tails.

The paper is organized as follows. In Section 2 we present some basic copula theory. Section 3 presents our new test statistic. In Section 4 we present mixing results, visualizing the power of our test in distinguishing the Gaussian copula from the Student's t and the Clayton copulae. In section 5 we analyze the dependency structure of stock portfolios, using our test statistic. Finally, section 6 summarizes our results and concludes.

2 Copulae

Consider d continuous real-valued random variables $X_1, \cdots, X_d$ with cumulative marginal distribution functions $F_1, \cdots, F_d$. Their dependence structure is described by the joint cumulative distribution function (henceforth referred to as cdf) F

$$F(x_1, \cdots, x_d) = P(X_1 \leq x_1, \cdots, X_d \leq x_d).$$

The quantile functions F_i^{-1} are defined as $F_i^{-1}(\alpha) = \inf\{x|F(x) \geq \alpha\}, \alpha \in [0,1]$, and are the inverse transforms of the univariate cdfs.

Theorem 2.1. *Assuming that F is a univariate cdf with quantile function F^{-1}:*

1. *If $U \sim U(0,1)$, then $F^{-1}(U) \sim F$,*

2. *If F is continuous and $X \sim F$, then $F(X) \sim U(0,1)$.*

Proof. See Ripley (1952). □

2.1 Definition

The copula is the dependence structure of the cdf, independent of the marginals. Thus, F is split into two components, the dependence structure and the marginal distributions. The formal definition of a copula function is (Embrechts et al., 1999):

Definition 2.1 (COPULA). *A d-copula is the distribution of a random vector in $\mathbb{R}^d$ with uniform-$(0,1)$ marginals or equivalently a d-copula is any function $C : [0,1]^d \to [0,1]$ which has the following three properties:*

1. *$C(u_1, \dots, u_d)$ is increasing in each component u_i,*
2. *$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \qquad \forall i \in [1, \dots, d], u_i \in [0,1]$,*
3. *$\forall (a_1, \dots, a_d), (b_1, \dots, b_d) \in [0,1]$ with $a_i \leq b_i$ we have*

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\ldots+i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0 \tag{2.1}$$

where $u_{j1} = a_j, u_{j2} = b_j \forall j \in [1, \dots, d]$.

Equation (2.1) is equivalent with $P(a_1 \leq U_1 \leq b_1, \dots, a_d \leq U_d \leq b_d)$.
The relation between the joint cdf and the copula is given by Sklar (1959):

Theorem 2.2. *Let F be an d-dimensional distribution function with marginals $F_1, \dots, F_d$. Then there exists an d-copula C such that $\forall \mathbf{x} \in \mathbb{R}^n$*

$$F(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d)). \tag{2.2}$$

If all F_i are continuous, then C is unique.

The copula function can be extracted from Equation (2.2):

Theorem 2.3. *If F is a continuous d-variate distribution function with univariate margins $F_1, \dots, F_d$ and quantile functions $F_1^{-1}, \dots, F_d^{-1}$, then*

$$C(\mathbf{u}) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)) \tag{2.3}$$

is the unique choice of C in Equation (2.2).

Proof. The proof is based on Theorem 2.1: If $X_i \sim F_i$ and $U_i \sim U(0,1)$, then $X_i \sim F_i^{-1}(U_i)$ and $F_i(X) \sim U_i$. By expressing X_i as a function of U_i in Equation (2.2), Equation (2.3) is obtained. □

The copula is a multivariate distribution with all univariate margins being $U(0,1)$. Hence if C is a copula, then it is the cdf of a multivariate uniform random vector. The copula can, as all cdfs, be represented by its density function $\tau(\mathbf{u})$:

$$C(\mathbf{u}) = \int_0^{u_1} \cdots \int_0^{u_d} \tau(\mathbf{u}) d\mathbf{u}. \tag{2.4}$$

This density function can also be written as

$$\tau(\mathbf{u}) = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \dots f_d(F_d^{-1}(u_d))}, \tag{2.5}$$

where f is the density of the joint distribution function and $f_1, \dots, f_d$ are the marginal densities.

2.2 Transformation Invariance

A very attractive feature of copulae, is the invariance under increasing and continuous transformations of the marginals.

Theorem 2.4. *If $x_1, \ldots, x_d$ have copula C and $T_1, \ldots, T_d$ are increasing continuous functions, then $T_1(x_1), \ldots, T_d(x_d)$ also have copula C.*

Proof. See Embrechts et al. (1999). □

Intuitively, this is due to the independence of the marginal distributions: The copula is only related to the dependence structure between the variables. Since e.g. the probability of survival, p_s, for a certain time (using the Merton (1974) default model) is a strictly increasing transformation of the stock prices, their associated copula equals the copula of the stock prices. Thus, the survival probability copula can be found, and the parameters estimated using commonly available stock data. Obviously, the corresponding default probability, p_d, is a strictly decreasing transformation of the survival probability, $p_d = 1 - p_s$, and the survival copula of the stock prices can be used for the modelling of these default probabilities, given the marginal distributions (Mashal et al., 2003).

3 Goodness-of-fit tests

One fundamental problem with copulae is to determine which copula that provides the best fit to an observed data set. Prior to the use of GOF tests, various information criterions were employed, such as Aikaike's Information Criterion (AIC). These are suboptimal tests that do not provide us with any understanding of the size of the decision rule employed nor its power, i.e. they will not give us any way of concluding whether or not one copula fits the data significantly better than another. Hence, GOF tests are preferred.

For univariate distributions, the GOF assessment can be performed by e.g. the well-known Anderson-Darling (Anderson and Darling, 1954) test, or less quantitatively using a QQ-plot. In the multivariate domain there are fewer alternatives. In addition, economic theory sheds little light on the dependence structure between financial assets, and multivariate normality is often assumed a priori. Evidence shows, however, that more appropriate dependence structures are available (Chen et al., 2004; Dobrić and Schmid, 2005).

GOF tests for copulae is basically a special case of the more general problem of testing multivariate density models, but is complicated due to the unspecified marginal distributions. Empirical margins are used since we are interested in the fit of the copula itself, not the copula and the margins together. In short, the use of empirical margins introduces infinitely many nuisance parameters. This complicates the deduction of the asymptotic distribution properties for the tests. Thus p-values are found by simulation.

Several GOF tests have been proposed, but there are no general guidelines for optimal parametric copula selection. Genest and Rivest (1993) have developed an empirical method to identify the best copula in the Archimedean case. Diebold et al. (1998), Diebold et al. (1999), Hong (2000), Berkowitz (2001), Thompson (2002) and Chen et al. (2004) focus on the probability integral transform of the data in the evaluation of density models. Genest et al. (2005) utilize the Kendall's process, while Panchenko (Panchenko) focus on positive definite bilinear forms.

Most tests project the multivariate problem to a univariate problem, then apply a univariate test. This leads to numerically efficient algorithms for problems of high dimension. Any univariate test may be used, e.g. Kolmogorov-Smirnov (KS), Anderson-Darling (AD), Cramér-von Mises (CvM) and kernel smoothing (KDE) based L2 tests. We will, in this paper, focus on the AD test, which is an unbiased cdf test.

3.1 The Probability Integral Transform

The PIT transforms a set of dependent variables into a new set of independent $U(0,1)$ variables, given the multivariate distribution. The PIT is a universally applicable way of creating a set of iid $U(0,1)$ variables from any data set with known distribution. Given a test for multivariate, independent uniformity, this transformation can be used to test whether any assumed model fits or not. The concept was first introduced by Rosenblatt (1952) and can be interpreted as the inverse of simulation.

Definition 3.1 (PROBABILITY INTEGRAL TRANSFORM)**.** *Let* $\mathbf{X} = (X_1, \ldots, X_d)$ *denote a random vector with marginal distributions* $F_i(x_i) = P(X_i \leq x_i)$ *and conditional distributions* $F(X_i \leq x_i | X_1 = x_1, \ldots, X_{i-1} = x_{i-1})$ *for* $i = [1, \ldots, d]$*. The PIT of* $\mathbf{X}$ *is defined as* $T(\mathbf{X}) = (T_1(X_1), \ldots, T_d(X_d))$ *where* $T_i(X_i)$ *is defined as follows:*

$$
\begin{aligned}
T_1(X_1) &= P(X_1 \leq x_1) = F_{X1}(x_1), \\
T_2(X_2) &= P(X_2 \leq x_2 | X_1 = x_1) = F_{X2|X_1}(x_2|x_1), \\
&\vdots \\
T_d(X_d) &= P(X_d \leq x_d | X_1 = x_1, \ldots, X_{d-1} = x_{d-1}) = F_{X_d|X_1 \ldots X_{d-1}}(x_d|x_1, \ldots, x_{d-1}).
\end{aligned}
$$

The random variables $Z_i = T_i(X_i)$*, for* $i = 1, \ldots, d$ *are uniformly and independently distributed on* $[0,1]^d$.

For more details on the PIT see e.g. Rosenblatt (1952) or Breymann et al. (2003).

A recent application of the PIT has been multivariate GOF tests. Hong and Li (2002) report Monte Carlo evidence of tests using the PIT variables outperforming tests using the original random variables. Chen et al. (2004) believe that a similar conclusion also applies to GOF tests for copulae. Hence, a PIT-based approach seems to be preferred.

3.2 New Test

The new test B, proposed in this paper, was developed with three purposes in mind. We want it to be consistent, numerically efficient and unbiased. The test is somewhat similar to the test G by Breymann et al. (2003) and Chen et al. (2004). However, their test is inconsistent, meaning that some deviations from the null hypothesis may be neglected. Our new test solves this problem by transforming the data before projecting the multivariate problem to a univariate problem.

Let $\mathbf{Z}$ be the uniformly and independently distributed variables on $[0,1]^d$, obtained from applying the PIT to a multivariate data set $\mathbf{X}$. Define a new vector $\mathbf{Z}^*$ as

$$
Z_i^* = P(r_i \leq \widetilde{Z}_i | r_1, \ldots, r_{i-1}) = \left(1 - \left(\frac{1 - \widetilde{Z}_i}{1 - r_{i-1}}\right)^{d-(i-1)}\right) \cdot I(\widetilde{Z}_i \geq r_{i-1}), \tag{3.1}
$$

for $i = 1, \ldots, d$, where $\widetilde{\mathbf{Z}} = (\widetilde{Z}_1, \ldots, \widetilde{Z}_d)$ is the sorted counterpart of $\mathbf{Z}$, r_i is rank variable i[1] from $\mathbf{Z}$ and $I(x)$ is the indicator function ($I(x) = 0$ for $x < 0$, $I(x) = 1$ for $x \geq 0$). Let

$$
Y = \sum_{i=1}^{d} \gamma(Z_i; \boldsymbol{\alpha}) \cdot \Phi^{-1}(Z_i^*)^2, \tag{3.2}
$$

where γ is a weight function used for weighting $\Phi^{-1}(Z_i^*)^2$ depending on its corresponding value Z_i, and $\boldsymbol{\alpha}$ is the set of weight parameters. Further let $F_Y(\cdot)$ be the cdf of Y, i.e. the cdf of a linear combination of squared normal variables. The new test B is then defined as the cdf of $F_Y(Y)$:

$$
B(w) = P\left[F_Y(Y) \leq w\right], \qquad w \in [0,1]. \tag{3.3}
$$

[1] Rank variables are the observed variables, ordered ascendingly.

Under the null hypothesis $B(w) = w$. The density function of $B(w)$ is $b(w) = 1$.

Given n observations of the d-dimensional vector $\mathbf{Z}$, the empirical version of $B(w)$, $\hat{B}(w)$, equals:

$$\hat{B}(w) = \frac{1}{n+1} \sum_{j=1}^{n} I\left(F_Y(Y) \leq w\right), \qquad w = \frac{1}{n+1}, \ldots, \frac{n}{n+1}. \tag{3.4}$$

The rationale behind equation (3.1) can be explained as follows. The test G by Breymann et al. (2003) and Chen et al. (2004) projects the multivariate problem to a univariate problem by computing $Y = \sum_{i=1}^{d} \Phi^{-1}(Z_i)^2$. This test is inconsistent, as the authors themselves point out. To avoid this problem we transform the data $\mathbf{Z}$ to $\mathbf{Z}^*$. We wish to find $P(r_i < \widetilde{Z}_i | r_1, \ldots, r_{i-1}) = 1 - P(r_i \geq \widetilde{Z}_i | r_1, \ldots, r_{i-1})$. The only way r_i can be greater than or equal to $\widetilde{Z}_i$ is if all remaining $d - (i-1)$ variables are greater than or equal to $\widetilde{Z}_i$. Since the remaining $d - (i-1)$ variables are independent the probability of all being greater than or equal to $\widetilde{Z}_i$ is the product of each r_k being greater than or equal to $\widetilde{Z}_i$:

$$P(\widetilde{Z}_i \leq r_k < 1 | r_k > r_{i-1}) = \frac{P(r_k \geq \widetilde{Z}_i \cap r_k > r_{i-1})}{P(r_k > r_{i-1})} = \frac{P(r_k \geq \widetilde{Z}_i)}{P(r_k > r_{i-1})} = \frac{1 - \widetilde{Z}_i}{1 - r_{i-1}}, \quad k \in [i, d].$$

The indicator function in equation (3.1) is included since $\widetilde{Z}_i$ must be greater than or equal to r_{i-1}.

The problem of obtaining the distribution of a linear combination of squared normal variables has been addressed by many authors. Hence, several representations for the cdf and density can be found in the literature, including i.a. power series expansions (Shah and Khatri, 1961), χ^2 series (Ruben, 1962) and Laguerre series (Shah, 1963; Kotz et al., 1967a,b). For computational purposes we use simulation to obtain F_Y.

The use of the squared inverse gaussian cdf, $\Phi^{-1}(\mathbf{Z}^*)^2$, in equation (3.2), weights large deviances from the null hypothesis more than small deviances. We may also wish to weight certain regions of the original copula, e.g. the tails like Breymann et al. (2003) and Chen et al. (2004) implicitly do. The weighting function γ is introduced for this purpose.

3.2.1 Weighting Functionality

The weighting functionality, incorporated in the test through γ, adds extra flexibility compared to the G test of Breymann et al. (2003) and Chen et al. (2004). This test weights the tails of the copula implicitly through the use of $\Phi^{-1}(Z_i)^2$. Our test opens for a much more general weighting procedure. The weight function can be of any form, for example:

- Power tail weighting: $\gamma(Z_i; \alpha) = (Z_i - \frac{1}{2})^{\alpha}, \quad \alpha \in (2, 4, \ldots)$.
- Left/Right power tail weighting:
  1. Left power tail: $\gamma(Z_i; \alpha) = 1 - Z_i^{1/\alpha}$,
  2. Right power tail: $\gamma(Z_i; \alpha) = 1 - (1 - Z_i)^{1/\alpha}$.
- Inverse Student's t tail weighting: $t_{\nu}^{-1}(Z_i)^2$.

We may also choose to not weight any specific region at all. Figure 3.1 shows the effect of power- and Student's t tail weighting as well as left- and right power tail weighting. We see that as we increase α (or decrease ν) the weight is increasingly pushed into the tails. Tail weighting means that we weight the tails of the copula, i.e. in the bivariate case the upper right corner and the bottom left corner of the copula. For some applications these regions may be of special interest to us. For simplicity we only consider power tail weighting in this paper.

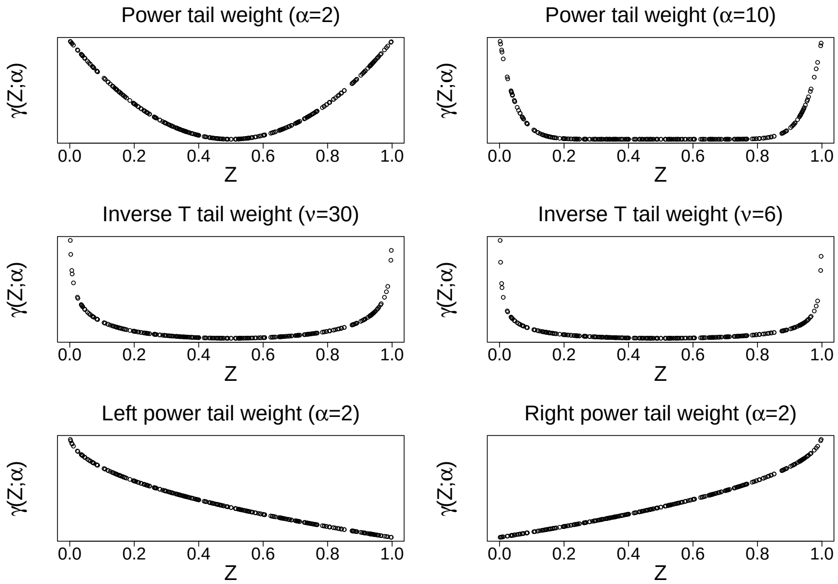


Figure 3.1: The effect of tail weighting.

3.2.2 Testing Procedure

Suppose we have n independent observations from a d-dimensional copula $\mathbf{X}$. The testing procedure would then be as follows: First, PIT $\mathbf{X}$ under a $\mathcal{H}_0$ copula. The resulting copula, $\mathbf{Z}$, should be the independent copula if $\mathcal{H}_0$ is true. Then, for each $j = 1, \ldots, n$ do:

- From $\mathbf{Z}_j = (z_{j1}, \ldots, z_{jd})$, compute the weights $\gamma(z_{ji}; \boldsymbol{\alpha})$, $i = 1, \ldots, d$, for a given weight function γ and weight parameters $\boldsymbol{\alpha}$.
- Compute $\mathbf{Z}_j^*$ according to equation (3.1). These variables are iid $U(0,1)^d$ under $\mathcal{H}_0$.
- Compute the univariate variable Y_j according to equation (3.2).

To find F_Y we repeatedly (10000 times) simulate an independent $U(0,1)^d$ vector $\mathbf{Z}$ on which we perform the three steps above. Given F_Y we can compute $\hat{B}(w)$ according to equation (3.4). The variables $\hat{B}(w)$ are now, under $\mathcal{H}_0$, iid $U(0,1)^n$. The next step is to perform a univariate GOF test on $\hat{B}$: $\mathcal{T}(\hat{B})$, where $\mathcal{T}$ is any univariate GOF test. We use the Anderson-Darling test:

$$\mathcal{T}^{AD} = n \int \frac{\left(\hat{F}(\gamma) - \gamma\right)^2}{\gamma(1-\gamma)} \mathrm{d}F(\gamma). \tag{3.5}$$

The resulting test value for our observed copula is $\mathcal{T}^{AD}(\hat{B})$. Since the cdf's are discrete, the integral in equation (3.5) becomes a sum. The discrete AD test statistic can be shown to be (Marsaglia and Marsaglia, 2004):

$$\mathcal{T}^{AD} = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\left[\ln\left(\hat{F}\left(\frac{j}{n+1}\right)\right) + \ln\left(1 - \hat{F}\left(\frac{n+1-j}{n+1}\right)\right)\right],$$

and straightforward insertion of densities into this equation gives:

$$\mathcal{T}^{AD}(\hat{B}) = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)\left[\ln\left(\hat{B}(\frac{j}{n+1})\right) + \ln\left(1 - \hat{B}(\frac{n+1-j}{n+1})\right)\right]. \tag{3.6}$$

To obtain p-values we simulate a copula of dimension $n \times d$ under $\mathcal{H}_0$. We perform the entire testing procedure, with the same weight parameters as we used for our observed copula. This gives us a value for $\mathcal{T}^{AD}_{\mathcal{H}_0}(\hat{B})$ when we know $\mathcal{H}_0$ is true. We repeat this 10000 times and obtain a distribution for $\mathcal{T}^{AD}_{\mathcal{H}_0}$. This distribution can be used to find the p-value of the $\mathcal{H}_0$ copula for our observed copula. For a discussion of the number of simulations required to obtain the desired confidence, see Appendix A.

4 Results

In this section we assess the power of our new test statistic. By performing mixing tests we get an impression of the tests ability to detect tail heaviness and skewness properties. The tests ability to distinguish the Gaussian from the Student's t copula indicates the power at detecting tail heaviness, while the ability to distinguish the Gaussian from the Clayton copula indicates the power at detecting skewness. Similar mixing tests are performed in Chen et al. (2004).

The mixing tests are performed by mixing a Gaussian copula with a Student's t or a Clayton copula to construct a mixed copula $\mathcal{C}_{mix}$:

$$\mathcal{C}_{mix} = (1-p) \cdot \mathcal{C}_g + p \cdot \mathcal{C}_a, \quad p \in [0,1],$$

where $\mathcal{C}_g$ denotes the Gaussian copula and $\mathcal{C}_a$ denotes the alternative copula. In this paper the alternative copulae considered are the Student's t copula, $\mathcal{C}_t$, and the Clayton copula, $\mathcal{C}_c$. For $p = 0$, $\mathcal{C}_{mix}$ is a Gaussian copula while for $p = 1$, $\mathcal{C}_{mix}$ is a Student's t copula or a Clayton copula. For $0 < p < 1$ we sample from the Gaussian copula with probability $(1-p)$ and from the alternative copula with probability p. In a financial setting, for a portfolio of d stocks, this can be interpreted as follows. Some days the dependency structure follow a Gaussian copula and other days a Student's t copula.

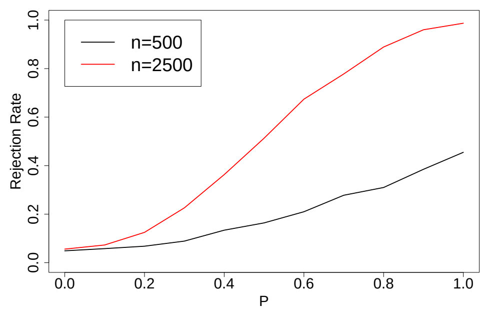
Our null hypothesis is that the mixed copula is a Gaussian copula. We PIT $\mathcal{C}_{mix}$ under this null hypothesis and compute $\mathcal{T}^{AD}(\hat{B})$ and the corresponding p-value, given some weighting type and parameter. This is repeated 1000 times in order to obtain rejection rates and corresponding power curves.

For both the Gaussian-Student's t and the Gaussian-Clayton mixing we examined the general case of no weighting and the case of power tail weighting with $\alpha \in (2, 4, 10, 20)$.

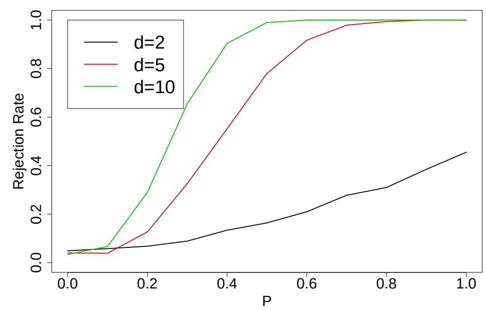
Figure 1(a) shows the effect the number of observations has on the power of the test for Gaussian-Student's t mixing. We see that the power increases dramatically as the number of observations increase from 500 to 2500, which is to be expected. The effect of the dimension on Gaussian-Student's t mixing is shown in Figure 1(b). The dimension seems to have an even greater effect on the power of the test than the number of observations, at least when we move from bivariate copulae to higher dimensions. This is also as expected, since we used the same degrees-of-freedom for the Student's t copula for all dimensions, and the distance between the Gaussian and the Student's t copula increases with the dimension(see Appendix B for a discussion of distance between distributions). In Figure 1(c) we have used power tail weighting for the Gaussian-Student's t mixing and we see that it is very powerful. Figure 4.2 compares our new test B with the G test and we see that by applying heavy power tail weighting our test performs almost as good as the G test for tail heaviness. It seems like the B test performs very well at determining skewness for the chosen d and n, and that the performance increases as we add some tail weight. However, as we increase the tail weight too much the performance decreases dramatically. This may be why the G test breaks down here, because this test implicitly adds heavy tail weight. We see similar results for other combinations of d and n.

5 Application

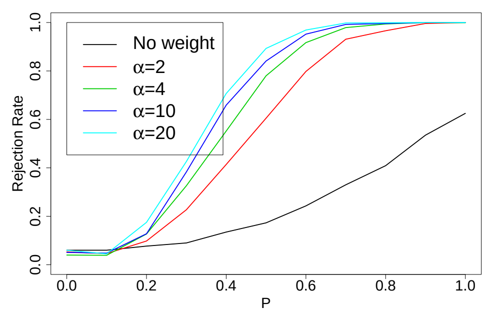
The dependency structure can have big impacts in several applications, e.g. capital allocation or the pricing of credit derivatives, such as basket default swaps.



(a) Effect of n - the number of observations. Gaussian-Student's t mixing, power tail weighting, $d = 2,\ \alpha = 4,\ \rho = 0.5,\ \nu = 4,$ 5% significance level.

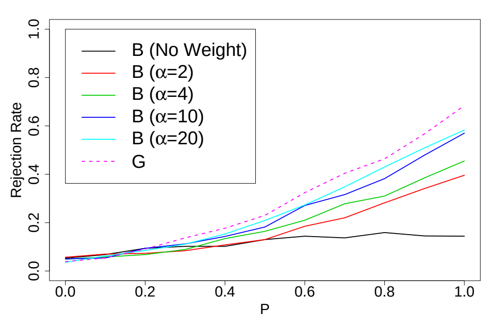


(b) Effect of d - the dimension. Gaussian-Student's t mixing, power tail weighting, $n = 500,\ \alpha = 4,\ \rho = 0.5,\ \nu = 4,$ 5% significance level.

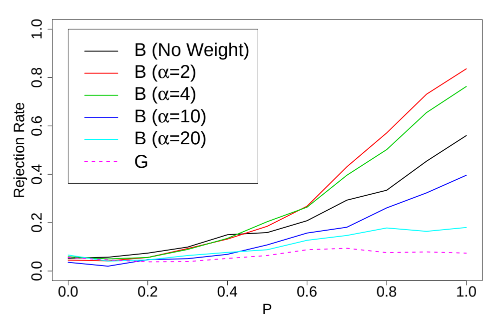


(c) Effect of α - the power tail weighting parameter. Gaussian-Student's t mixing, power tail weighting, $d = 5,\ n = 500,\ \rho = 0.5,\ \nu = 4,$ 5% significance level.

Figure 4.1: Power curves for the B test, for varying parameters. On the x-axis we see the mixing parameter while on the y-axis we see the portion of times the Gaussian copula is rejected.



(a) G test versus B test for $d = 2$ and $n = 500$. No weight and various power tail weights for the B test. Gaussian-Student's t mixing, $\rho = 0.5,\ \nu = 4,$ 5% significance level.



(b) G test versus B test for $d = 5$ and $n = 500$. No weight and various power tail weights for the B test. Gaussian-Clayton mixing, $\rho = 0.5,\ \delta = 0.5,$ 5% significance level.

Figure 4.2: Power comparison for the G and B tests. On the x-axis we see the mixing parameter while on the y-axis we see the portion of times the Gaussian copula is rejected.

Table 5.1: Rejection rates of the Gaussian, Student's t and Clayton copulae, applied to the raw returns. 5% significance level.

Gaussian copula					
	No Weight / Power tail weight (parameter α)				
Dimension	No weight	$\alpha = 2$	$\alpha = 4$	$\alpha = 10$	$\alpha = 20$
2	0.076	0.132	0.176	0.466	0.512
5	0.700	0.930	0.930	0.920	0.910
10	0.740	1.000	1.000	1.000	1.000
Student's t copula					
	No Weight / Power tail weight (parameter α)				
Dimension	No weight	$\alpha = 2$	$\alpha = 4$	$\alpha = 10$	$\alpha = 20$
2	0.042	0.022	0.032	0.044	0.034
5	0.120	0.090	0.060	0.050	0.070
10	0.260	0.040	0.150	0.130	0.190
Clayton copula					
	No Weight / Power tail weight (parameter α)				
Dimension	No weight	$\alpha = 2$	$\alpha = 4$	$\alpha = 10$	$\alpha = 20$
2	0.622	0.354	0.792	0.434	0.396
5	0.980	0.990	0.980	0.970	0.950
10	1.000	1.000	1.000	1.000	1.000

We analyze the dependency structure of stock portfolios by looking at their daily log returns. The total portfolio consists of 1000 observations of 50 large cap stocks from the New York Stock Exchange, spanning the period September 26th, 2001 to September 16th, 2005.

Asset collections of dimension 2, 5 and 10 were randomly selected 100 times from the full data set. As in Chen et al. (2004) and Panchenko (Panchenko) we examine the raw returns and the $GARCH(1,1)$ filtered returns, i.e. each individual assets return is filtered through a standard $GARCH(1,1)$ process. This filtering is done to remove serial dependence in each individual time series. For details of $GARCH$ processes, see e.g. Bollerslev (1986). Next, we fit a Gaussian, Student's t and Clayton copula to the portfolios and apply our B test to investigate which copula that provides the best fit. When fitting copulae to the data the parameters of the copulae are estimated by numerically optimizing the likelihood. For the Student's t copula a semi-parametric approach is followed. This method is denoted the pseudo-likelihood (Demarta and McNeil, 2005) or the canonical maximum likelihood (CML) method (Romano, Romano), and is described in Genest et al. (1995).

Tables 5.1 and 5.2 show the rejection rates for the raw and $GARCH(1,1)$ filtered returns, respectively. The Clayton copula seems to provide the worst fit, as expected for stock data. In addition we have only considered Clayton copula with one parameter, hence the poor performance for higher dimensions is not surprising. The Gaussian copula is not that easily rejected for the bivariate case, even though we see an increasing rejection rate for the raw returns as we increase the tail weight. For higher dimensions we see that the Gaussian copula is strongly rejected, for both raw and $GARCH$ filtered returns. The Student's t copula seems to provide a very good fit for all dimensions and for both raw and $GARCH$ filtered returns. It is not surprising that the Student's t copula outperforms the Gaussian copula since it has one extra parameter. However, the low rejection rates for the Student's t copula are interesting.

6 Conclusion

We have introduced a new copula goodness-of-fit test B, which merges the efficiency of one-dimensional tests with the consistency of multi-dimensional tests. The test is consistent and can

Table 5.2: Rejection rates for the Gaussian, Student's t and Clayton copulae, applied to the $GARCH(1,1)$ filtered returns. 5% significance level.

Gaussian copula					
	No Weight / Power tail weight (parameter α)				
Dimension	No weight	$\alpha = 2$	$\alpha = 4$	$\alpha = 10$	$\alpha = 20$
2	0.216	0.176	0.062	0.058	0.156
5	0.170	0.470	0.570	0.470	0.500
10	0.620	0.530	0.840	0.620	0.770
Student's t copula					
	No Weight / Power tail weight (parameter α)				
Dimension	No weight	$\alpha = 2$	$\alpha = 4$	$\alpha = 10$	$\alpha = 20$
2	0.048	0.114	0.008	0.012	0.006
5	0.100	0.010	0.000	0.010	0.010
10	0.220	0.000	0.010	0.070	0.030
Clayton copula					
	No Weight / Power tail weight (parameter α)				
Dimension	No weight	$\alpha = 2$	$\alpha = 4$	$\alpha = 10$	$\alpha = 20$
2	0.436	0.282	0.100	0.248	0.076
5	0.680	0.770	0.970	0.870	0.920
10	0.980	1.000	1.000	0.970	0.990

be considered a modification of the tests by Breymann et al. (2003) and Chen et al. (2004). The novelty of the test is the transformation $\mathbf{Z}^*$, making the projection from a multivariate problem to a univariate problem consistent, and the weighting functionality. The test enables the user to weight any region of the copula in any way desirable. We believe that this weighting functionality adds a very attractive flexibility to the user.

Mixing results show that the test has good power and that the weighting functionality is very powerful. They also show that by applying heavy power tail weighting we can achieve almost the same power as Breymann et al. (2003) and Chen et al. (2004) at distinguishing the Gaussian copula from the Student's t copula. For distinguishing the Gaussian copula from the Clayton copula, the test by Breymann et al. (2003) and Chen et al. (2004) breaks down for lower values of n whereas our new test performs well.

Application to stock portfolios show that the Student's t copula provide a fairly good fit to the data while the Gaussian copula is strongly rejected for higher dimensions. This is in accordance with the findings of i.a. Dobrić and Schmid (2005) and Chen et al. (2004).

Further work involve comparison of our new test with the tests of Genest et al. (2005) and Panchenko (Panchenko), which both seem to be promising tests with sound theoretical foundation. We also believe that the transformation $\mathbf{Z}^*$ may be utilized to improve the latter. Further tests of various weight functions will also be of interest as will studies of the impact on the p-values of the order in which we PIT the data.

Acknowledgements

This work is part of a Strategic Department Program called Statistical Analysis of Risk at the Department of Mathematics, University of Oslo and the Norwegian Computing Center. The programme is funded by the Research Council of Norway. The authors acknowledge the support and guidance of colleagues at the Norwegian Computing Center, in particular Assistant Research Director Kjersti Aas and Chief Research Scientist Xeni Kristine Dimakos. This work was carried out while Henrik Bakken was a MSc. student at the Norwegian University of Science and Technology, working on his diploma thesis at the Norwegian Computing Center.

A Iterations required

Given α, the level of significance for the test, it is imperative to have an acceptable resolution for the empirical distribution in the vicinity of this critical limit.

For a number of simulations N, the resolution will improve for values of α closer to 50% since there then is a higher probability for an even distribution of simulations above and below the critical limit. We suggest a method for determining the amount of iterations required that is based on confidence intervals.

The actual probability of rejection of $\mathcal{H}_0$ when it is true (Type I error) is α, the probability of accepting it is $1-\alpha$. For N simulations, this becomes a regular binomial distribution, $\mathrm{Bin}(N,\alpha)$, with expected value $\mu = N\alpha$ and variance $\sigma^2 = N\alpha(1-\alpha)$. In addition, as N becomes large, this distribution is well approximated by a univariate Gaussian distribution. Define a as the number of simulations that are rejected, even though $\mathcal{H}_0$ is true. Then,

$$a \sim \mathcal{N}(N\alpha, N\alpha(1-\alpha)),$$

given that N is large. The choice of N is then determined by the allowed deviation that we allow a to have from α. Given the significance level γ and fractional deviation β, the following confidence intervals can be constructed:

$$\begin{aligned} P\left(|a - N\alpha| \leq N\alpha \cdot \beta\right) &\leq 1-\gamma, \qquad \alpha \leq 0.5, \\ P\left(|a - N\alpha| \leq N(1-\alpha) \cdot \beta\right) &\leq 1-\gamma, \qquad \alpha > 0.5. \end{aligned}$$

The reason why the problem is twofold, is that when α is larger than 0.5, the allowed fractional deviation determined by β is smaller for $(1-\alpha)$ than α and the former becomes the dominating limit. This implies that $N(\alpha)$ is symmetric around 0.5, $N(\alpha) = N(1-\alpha)$. Hence, the $\alpha \leq 0.5$ case is considered and applied similarly to $\alpha > 0.5$. The probability that the empirical significance level a/N deviates with more than a fraction β from the actual level α is required to be less than $1-\gamma$. This can be rewritten as

$$P\left(N\alpha \cdot (1-\beta) \leq a \leq N\alpha \cdot (1+\beta)\right) \leq 1-\gamma, \qquad \alpha \leq 0.5.$$

Since a is normally distributed and due to distribution symmetry,

$$P\left(Z \leq \frac{N\alpha(1+\beta) - N\alpha}{\sqrt{N\alpha(1-\alpha)}}\right) \leq 1 - \frac{\gamma}{2}, \qquad \alpha \leq 0.5.$$

The corresponding Z-value $z_{1-\gamma/2}$ can then be expressed as

$$z_{1-\gamma/2} = \beta\sqrt{\frac{N\alpha}{1-\alpha}}, \qquad \alpha \leq 0.5.$$

By rewriting this expression with respect to N:

Proposition A.1. *To find the distribution of the test statistic $\mathcal{T}$ of a GOF test using Monte Carlo simulation, given the significance level α, the recommended number of simulations N' of $\mathcal{T}$ is*

$$N'(\alpha) = \left(\frac{z_{1-\gamma/2}}{\beta}\right)^2 \frac{1-\alpha}{\alpha}, \qquad \alpha \leq 0.5, \tag{A.1}$$

$$N'(\alpha) = N'(1-\alpha), \qquad \alpha > 0.5. \tag{A.2}$$

Remark A.1. Notice that only one parameter besides α is required for determination of N', namely the fraction $\frac{z_{1-\gamma/2}}{\beta}$. There is thus one redundant parameter, one of the two parameters β and γ can be locked at a constant value (e.g. $\beta = 5\%$) without losing model flexibility.

Remark A.2. Note that the number of times the copula must be simulated is $n \cdot N$ since each test statistic measure is based upon n simulations.

Setting $\beta = 10\%$, table A.1 displays the number of iterations recommended for varying α and γ.

It is evident that α is the main driver for computational complexity. The number of calculations increase exponentially as α reduces, and explodes as $\alpha \to 0^+$.

Table A.1: Recommended number of simulations for $\beta = 10\%$.

$\beta = 10\%$	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 1\%$	$\alpha = 0.1\%$
$\gamma = 10\%$	2435	5141	26785	270284
$\gamma = 5\%$	3457	7299	38030	383762
$\gamma = 1\%$	5971	12606	65685	662826
$\gamma = 0.01\%$	9745	20572	107193	1081674

B The Kullbach-Leibler Information Criterion

When applying a GOF test for determining the best copula, the test's usefulness is determined by its ability to show how one alternative is clearly better than the other. It is likely that a test statistic comparing the bivariate Gaussian copula with a bivariate t-copula with 50 degrees of freedom would return approximately the same result for both copulae. This is due to the fact that they are very similar and that the student t-copula converges towards the Gaussian copula as $\nu \to \infty$. The difference can be illustrated by a measure of the relative distance, or information entropy, between two sets of probability densities. From Cooke and Bedford (2002):

Definition B.1 (RELATIVE INFORMATION). *Let ν and μ be probability measures on a probability space such that ν is absolutely continuous with respect to μ ($\nu \ll \mu$) with Radon-Nikodym derivative $\frac{d\nu}{d\mu}$, then the relative information or Kullbach Leibler divergence $I(\nu|\mu)$ of ν with respect to μ is*

$$I(\nu|\mu) = \int \ln\left(\frac{d\nu}{d\mu}(x)\right) d\nu(x). \tag{B.1}$$

The Radon-Nikodym-derivative is a probability density, specifically the density that transforms the probability measure ν to μ, given that $\nu \ll \mu$. The relative information equals 0 if and only if $\nu = \mu$. For independent probability measures, ν is not absolutely continuous with respect to μ, and we define $I(\nu|\mu) = \infty$. The measure can be interpreted as measuring the degree of uniformness of ν with respect to μ and it is always non-negative.

In Nielsen and Chuang (2003), the relative information measure is used to calculate the Shannon entropy of a probability distribution: $I(\nu|\mu)$ where μ is uniform. Shannon's noiseless coding theorem says that the minimal physical requirements needed to store an information source (a probability density) equals the Shannon entropy. This can be used to calculate the optimal compression rate of information, see Nielsen and Chuang (2003) Chapter 12 for proof.

From the measure of relative information, a distance measure between two copulae can derived:

Definition B.2 (KLIC DISTANCE). *The Kullbach-Leibler Information Criterion (KLIC) between two copulae $C_a(\mathbf{u})$ and $C_b(\mathbf{u})$ with densities $\tau_a(\mathbf{u})$ and $\tau_b(\mathbf{u})$ is defined as:*

$$KLIC(C_a : C_b) = \int \ln\left(\frac{\tau_a(\mathbf{u})}{\tau_b(\mathbf{u})}\right) \tau_a(\mathbf{u}) d\mathbf{u}. \tag{B.2}$$

This measure is always greater than or equal to zero. Also, $KLIC(C_a : C_b) = KLIC(C_b : C_a)$.

We show by simulation that the KLIC distance between the Gaussian and Student-t copula increases with dimension d, see Table B.1. Even though the distance between a $\nu = 50$ Student-t copula and a Gaussian copula is negligible for $d = 2$, for $d = 30$ it is twice as large as the bivariate distance at $\nu = 4$. Thus, the approximation that a Student-t copula with $\nu > 30$ is very similar to a Gaussian copula is valid only for few dimensions. The GOF tests will therefore produce sharper contrasts when the number of assets are increased.

Table B.1: KLIC distance between Gaussian and Student-t copulae

	Degrees of freedom ν					
Dimension	4	6	10	20	30	50
2	0.022	0.011	0.005	0.001	0.001	0.000
3	0.059	0.027	0.013	0.004	0.001	0.001
5	0.168	0.089	0.036	0.008	0.004	0.002
10	0.481	0.262	0.112	0.036	0.020	0.005
20	1.123	0.611	0.293	0.100	0.061	0.023
30	1.795	0.955	0.468	0.184	0.100	0.047

References

Anderson, T. W. and D. A. Darling (1954). A test of goodness of fit. *Journal of the American Statistical Association 49*, 765–769.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics 19*, 465–474.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

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

Chen, X., Y. Fan, and A. Patton (2004). Simple tests for models of dependence between multiple financial time series, with applications to u.s. equity returns and exchange rates. Financial Markets Group, London School of Economics, Discussion Paper 483. Revised July 2004.

Cooke, R. and T. Bedford (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*, 1031–1068.

Demarta, S. and A. J. McNeil (To appear in 2005). The t copula and related copulas. *International Statistical Review*.

Diebold, F. X., T. Gunther, and A. S. Tay (1998). Evaluating density forecasts, with applications to financial risk management. *International Economic Review 39*, 863–883.

Diebold, F. X., J. Hahn, and A. S. Tay (1999). Multivariate density forecast and calibration in financial risk management: High-frequency returns on foreign exchange. *The Review of Economics and Statistics 81(4)*, 661–673.

Dobrić, J. and F. Schmid (2005). Testing goodness of fit for parametric families of copulas - application to financial data. *Communications in Statistics: Simulation and Computation 34*.

Embrechts, P., A. McNeil, and D. Straumann (1999, May). Correlation and dependence in risk management: Properties and pitfalls. *RISK*, 69–71.

Genest, C., K. Ghoudi, and L. Rivest (1995). A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika 82*, 543–552.

Genest, C., J.-F. Quessy, and B. Rémillard (2005). Goodness-of-fit procedures for copula models based on the probability integral transform. *Scandinavian Journal of Statistics*.

Genest, C. and L.-P. Rivest (1993). Statistical inference procedures for bivariate archimedean copulas. *Journal of the American Statistical Association*, 1034–1043.

Hong, Y. (2000). Evaluation of out-of-sample density forecasts with applications to s&p 500 stock prices. Working Paper, Department of Economics and Department of Statistical Science, Cornell University.

Hong, Y. and H. Li (2002). Nonparametric specification testing for continuous-time models with application to spot interest rates. Working Paper, Department of Economics, Cornell University.

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

Kotz, S., N. L. Johnson, and D. W. Boyd (1967a). Series representations of distributions of quadratic forms in normal variables. i. central case. *Annals of Mathematical Statistics 38*, 823–837.

Kotz, S., N. L. Johnson, and D. W. Boyd (1967b). Series representations of distributions of quadratic forms in normal variables. ii. non-central case. *Annals of Mathematical Statistics 38*, 838–848.

Marsaglia, G. and J. Marsaglia (2004). Evaluating the anderson-darling distribution. *Journal of Statistical Software 9*.

Mashal, R., M. Naldi, and A. Zeevi (2003). Extreme events and multi-name credit derivatives. In *Credit Derivatives The Definitive Guide*, pp. 313–338. Great Britain: Risk Waters Group.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance 29*, 449–470.

Nelsen, R. B. (1999). *An Introduction to Copulas*. New York: Springer Verlag.

Nielsen, M. and I. Chuang (2003). *Quantum Computation and Quantum Information*. Cambridge, UK: Cambridge University Press.

Panchenko, V. Goodness-of-fit test for copulas. *Physica A 355(1)*.

Ripley, B. (1952). *Stochastic simulation*. New York: John Wiley & Sons.

Romano, C. Calibrating and simulating copula functions: An application to the italian stock market. Working paper n. 12, CIDEM, Universit'a degli Studi di Roma.

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

Ruben, H. (1962). Probability content of regions under spherical normal distributions. iv. the distribution of homogeneous and non-homogeneous quadratic functions of normal variables. *Annals of Mathematical Statistics 33*, 542–570.

Shah, B. K. (1963). Distribution of definite and indefinite quadratic forms from a non-central normal distribution. *Annals of Mathematical Statistics 34*, 186–190.

Shah, B. K. and C. G. Khatri (1961). Distribution of a definite quadratic form for non-central normal variates. *Annals of Mathematical Statistics 32*, 883–887.

Sklar, A. (1959). Fonctions dé repartition á n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris 8*, 299–231.

Thompson, S. B. (2002). Evaluating the goodness of fit of conditional distributions, with an application to affine term structure models. Working paper, Department of Economics, Harvard University.