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

Daniel Berg
daniel@danielberg.no

University of Oslo / Norwegian Computing Center

21st Nordic Conference on Mathematical Statistics
Rebild, Denmark, 14th June 2006
Outline

1. Introduction
2. Copula - Definitions and Theorems
3. Copula goodness-of-fit testing
   - 3.1. Probability integral transform
   - 3.2. Breymann, Dias and Embrecht’s approach \( G \)
   - 3.3. New approach \( B \)
4. Power results from simulation
5. Application to daily return data
6. Summary
Introduction

- Copulae - a popular and flexible way of modelling dependence
- Copula choice may have huge impacts on e.g. capital allocation
- Is the data appropriately modelled by a given parametric copula?
- We propose a new copula goodness-of-fit approach.
2. Copula - Definitions and Theorems

**Definition (Copula)**

*A d-dimensional copula is a multivariate distribution, *C*, with standard uniform marginal distributions.*

**Theorem (Sklar)**

*Every multivariate distribution* *F*, with margins, *F*, , ..., *F*, can be written as*

\[
F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)),
\]

*for some copula* *C*. 
2. Copula - Definitions and Theorems

- Given a random vector $\mathbf{X} = (X_1, \ldots, X_d)$ the copula of their joint distribution function may be extracted from equation (2.1):

$$C(u_1, \ldots, u_d) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)),$$

where the $F_i^{-1}$'s are the quantile functions of the margins.

- The copula is often represented by its density function $c(u)$:

$$C(u) = P(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_d \leq u_d) = \int_0^{u_1} \cdots \int_0^{u_d} c(u)\,du,$$
2. Copula - Definitions and Theorems

- For the implicit copula of an absolutely continuous joint df $F$ with strictly continuous marginal df’s $F_1, \ldots, F_d$, the copula density is given by

$$c(u) = \frac{f(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \cdots f_d(F_d^{-1}(u_1))}.$$

- Hence,

$$c(F_1(x_1), \ldots, F_d(x_d)) = \frac{f(x_1, \ldots, x_d)}{f_1(x_1) \cdots f_d(x_d)}.$$

- This means that a general $d$-dimensional density can be written as

$$f(x_1, \ldots, x_d) = c(F_1(x_1), \ldots, F_d(x_d)) \cdot f_1(x_1) \cdots f_d(x_d)$$

for some copula density $c(\cdot)$. 
2.1. Copula - Attractive features

- A copula describes how the marginals are tied together in the joint distribution
- The joint df is decomposed into the marginal dfs and a copula
- The marginal dfs and the copula can be modelled and estimated separately, independent of each other
- Given a copula, we can obtain many multivariate distributions by selecting different marginal dfs
- The copula is invariant under increasing and continuous transformations
3. Copula Goodness-of-fit Testing

- Determine whether a copula appropriately fits the data.
- Univariate distributions \(\Rightarrow\) e.g. Anderson-Darling test or less quantitatively using QQ-plot.
- Multivariate domain \(\Rightarrow\) fewer alternatives.
- Copula GOF is a special case of the more general problem of testing multivariate density models.
- Complicated due to the use of empirical margins. Hence, \(P\)-values are usually found by simulation.
3. Copula Goodness-of-fit Testing

- Several approaches proposed lately, e.g.
  - Breymann et al. (2003) - based on the probability integral transform (PIT)
  - Genest et al. (2006) - based on the empirical copula and Kendall’s process

- Dimension reduction techniques reduce the multivariate problem to a univariate problem.
3.1. Probability Integral Transform

- Transforms a set of dependent variables into a new set of independent $U(0, 1)$ variables, given the multivariate distribution.
- 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 the fit of any assumed model.
- The concept was first introduced by Rosenblatt (1952) and can be interpreted as the inverse of simulation.
3.1. Probability Integral Transform

▷ The idea is to PIT the observed copula, assuming a $\mathcal{H}_0$ copula, and then test for independence. The null hypothesis may be a parametric copula family.

▷ An advantage with the PIT in this setting is that the null- and alternative hypotheses are the same, regardless of the distribution before the PIT.

▷ The PIT also enables weighting in a simple way since the data, under $\mathcal{H}_0$, is always iid $U(0, 1)$. 
3.2. Breymann et al. (2003)’s approach: G

- Let $Z$ be an iid $U(0, 1)^d$ vector under $\mathcal{H}_0$. Now define

$$Y_G = \sum_{i=1}^{d} \Phi^{-1}(z_i)^2,$$

$$W_G = F_{\chi^2_d}(Y_G),$$

$$F_G(w) = P(W_G \leq w), \quad w \in [0, 1].$$

Under $\mathcal{H}_0$ $F_G(w) = w$ and its density function $f_g(w) = 1$.

- Properties:
  - Coincides with the approaches proposed by Malevergne and Sornette (2003) when the latter is based on PIT. Also coincides with the second approach proposed by Chen et al. (2004).
  - Implicitly weights the tails of the copula through $\Phi^{-1}(\cdot)^2$
  - **NOT** consistent, some deviations may cancel out
3.3. New approach: $B$

- Extends $G$, solving the consistency issue by transforming the vector $Z$. Decouples deviance measure from weighting functionality.
- Let $Z$ be an iid $U(0, 1)^d$ vector under $\mathcal{H}_0$. Define a new vector $Z^*$ as

$$Z_i^* = \left(1 - \left(\frac{1 - \tilde{Z}_i}{1 - r_{i-1}}\right)^{d-(i-1)}\right),$$

for $i = 1, \ldots, d$, where $\tilde{Z} = (\tilde{Z}_1, \ldots, \tilde{Z}_d)$ is the sorted counterpart of $Z$ and $r_i$ is rank variable $i$ from $Z$. 
3.3. New approach: \( B \)

Next, let

\[
Y_B = \sum_{i=1}^{d} \gamma(z_i; \alpha) \cdot \Phi^{-1}(Z_i^*)^2,
\]

where \( \gamma \) is a weight function used for weighting \( \Phi^{-1}(Z_i^*)^2 \) depending on its corresponding value \( z_i \), and \( \alpha \) is the set of weight parameters.

Further let \( F_{Y_B}(\cdot) \) be the cdf of \( Y_B \), i.e. the cdf of a linear combination of squared normal variables. Then

\[
W_B = F_{Y_B}(Y_B),
\]

\[
F_B(w) = P(W_B \leq w), \quad w \in [0, 1].
\]

Under \( \mathcal{H}_0 \) \( F_B(w) = w \) and \( f_b(w) = 1 \).
3.3.1. Weighting functionality

The weight function may be of any form, for example:

- **Power tail weighting:** \( \gamma(z_i; \alpha) = (z_i - 0.5)^\alpha \)

- **Left/Right power tail weighting:**
  - Left power tail: \( \gamma(z_i; \alpha) = 1 - z_i^{1/\alpha} \)
  - Right power tail: \( \gamma(z_i; \alpha) = 1 - (1 - z_i)^{1/\alpha} \)

- **Inverse Student’s t tail weighting:** \( \gamma(z_i; \alpha) = t_{\nu}^{-1}(z_i)^2 \)
3.3.1. Weighting functionality

Figure: The effect of tail weighting.
3.3.2. Testing Procedure

Suppose we have \( n \) independent observations from a \( d \)-dimensional copula \( X \). The testing procedure would then be as follows:

1. PIT \( X \) under a \( \mathcal{H}_0 \) copula. This procedure usually involves estimating the parameters of the \( \mathcal{H}_0 \) copula, \( \hat{\theta} \). The resulting copula, \( Z \), should be the independent copula if \( \mathcal{H}_0 \) is true.

2. Then, for each \( j = 1, \ldots, n \), do:
   - From \( Z_j \), compute weights \( \gamma(z_{ji}; \alpha), i = 1, \ldots, d \).
   - Compute \( Z_j^* \). These variables are iid \( U(0, 1)^d \) under \( \mathcal{H}_0 \).
   - Compute the univariate variable \( Y_{Bj} \).
   - Given \( F_{Y_B} \) (e.g. from simulations), compute \( W_{Bj} \).
   - Given \( W_{Bj} \) compute \( F_{Bj}(w) \), an iid \( U(0, 1) \) vector under \( \mathcal{H}_0 \).

3. Compute some univariate test \( \hat{T} \) using \( F_B(w) \) or \( f_B(w) \).

4. Repeatedly (\( N \) times) perform step 1-3 using a simulated observed data set \( X^* \), simulated from the \( \mathcal{H}_0 \) distribution with parameter \( \hat{\theta} \). The resulting \( N \) values of \( \hat{T}^* \) form the distribution of \( T \).

5. Compute the \( p \)-value, \( p = \frac{1 + \sum_{k=1}^{N} I(\hat{T}^* \geq \hat{T})}{N+1} \).
4. Results

To assess the power of the test we performed so called 'Mixing' tests:

- \( C^{\text{Mix}} = (1 - \beta) \cdot C^{\text{Ga}} + \beta \cdot C^{\text{Alt}}, \quad \beta \in [0, 1], \quad C^{\text{Alt}} \in [C^{\text{St}}, C^{\text{Cl}}]. \)
- \( \mathcal{H}_0: \) Gaussian copula
- PIT under \( \mathcal{H}_0 \) and compute \( p \)-value.
- Repeat 500 times to obtain rejection rates as a function of the mixing parameter \( \beta \) and the alternative copula.
4. Results

Figure: The effect of \( n \) - the number of observations. G/T mixing, power tail weighting, \( d = 2, \alpha = 4, \rho = 0.5, \nu = 4 \), 5\% significance level.
4. Results

Figure: The effect of $d$ - the dimension. G/T mixing, power tail weighting, $n = 500$, $\alpha = 4$, $\rho = 0.5$, $\nu = 4$, 5% significance level.
4. Results

Figure: The effect of $\alpha$ - the power tail weighting parameter. Gaussian-Student-t mixing, power tail weighting, $d = 5, n = 500, \rho = 0.5, \nu = 4$, 5% significance level.
4. Results

Figure: $G$ 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
4. Results

Figure: $G$ 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
5. Application

- Portfolio of 50 large cap stocks. Daily log-returns from September 26th 2001 to September 16th 2005, i.e. $d = 50$ and $n = 1000$.
- Randomly select collections of 2 assets.
- PIT under Gaussian, Student-t and Clayton (one-parameter) $\mathcal{H}_0$ respectively.
- Compute $P$-value.
- Repeat 100 times $\Rightarrow$ rejection rates.
- Repeat for collections of 5 and 10 assets.
5. Application

<table>
<thead>
<tr>
<th>Dimension</th>
<th>No Weight</th>
<th>( \alpha = 2 )</th>
<th>( \alpha = 4 )</th>
<th>( \alpha = 10 )</th>
<th>( \alpha = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian copula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No weight</td>
<td>0.076</td>
<td>0.132</td>
<td>0.176</td>
<td>0.466</td>
<td>0.512</td>
</tr>
<tr>
<td>( \alpha = 2 )</td>
<td>0.700</td>
<td>0.930</td>
<td>0.930</td>
<td>0.920</td>
<td>0.910</td>
</tr>
<tr>
<td>( \alpha = 4 )</td>
<td>0.740</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>( \alpha = 10 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha = 20 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dimension</th>
<th>No Weight</th>
<th>( \alpha = 2 )</th>
<th>( \alpha = 4 )</th>
<th>( \alpha = 10 )</th>
<th>( \alpha = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student-t copula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No weight</td>
<td>0.042</td>
<td>0.022</td>
<td>0.032</td>
<td>0.044</td>
<td>0.034</td>
</tr>
<tr>
<td>( \alpha = 2 )</td>
<td>0.120</td>
<td>0.090</td>
<td>0.060</td>
<td>0.050</td>
<td>0.070</td>
</tr>
<tr>
<td>( \alpha = 4 )</td>
<td>0.260</td>
<td>0.040</td>
<td>0.150</td>
<td>0.130</td>
<td>0.190</td>
</tr>
<tr>
<td>( \alpha = 10 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha = 20 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dimension</th>
<th>No Weight</th>
<th>( \alpha = 2 )</th>
<th>( \alpha = 4 )</th>
<th>( \alpha = 10 )</th>
<th>( \alpha = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton copula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No weight</td>
<td>0.622</td>
<td>0.354</td>
<td>0.792</td>
<td>0.434</td>
<td>0.396</td>
</tr>
<tr>
<td>( \alpha = 2 )</td>
<td>0.980</td>
<td>0.990</td>
<td>0.980</td>
<td>0.970</td>
<td>0.950</td>
</tr>
<tr>
<td>( \alpha = 4 )</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>( \alpha = 10 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha = 20 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: Rejection rates for the fit of the Gaussian, Student-t and Clayton copulae.
6. Summary

- New approach B merges the efficiency of one-dimensional tests with the consistency of multi-dimensional tests.
- The weighting functionality adds valuable flexibilities to the analyst.
- Mixing tests show that the approach has good power for tail heaviness and skewness. The weighting functionality also seems to be very powerful.
- Applied to daily log-returns of stock portfolios the Student-t copula outperforms the Gaussian and Clayton copulae, as expected and in accordance with the findings of other studies.
References


