An Introduction to Copulae

Daniel Berg

University of Oslo & Norwegian Computing Center

Statistics seminar.

NTNU – March-14-2006.

Outline

1. Introduction
2. Definitions and Theorems
3. Dependence Concepts
4. Copula Families
5. Estimation
6. Simulation
7. Higher Dimensional Copulae
8. Application
9. Goodness-of-fit
10. Concluding remarks and further reading

1. Introduction

- Dependency modelling
- Linear correlation coefficient - a measure of linear dependence
- In e.g. financial markets we often see non-linear dependency structures
- Elliptical distributions - linear dependence structure - correlation coefficient meaningful
- Non-elliptical distributions - alternative measures of dependence needed $\Rightarrow$ Copulae
- Any multivariate distribution function can serve as a copula

1.1. Brief historical background:

- 1940's: Hoeffding studies properties of multivariate distributions
- 1959: The word **copula** appears for the first time (Sklar)
- 1999: Introduced to financial applications (Embrechts, McNeil, Straumann)
- 2006: Several insurance companies, banks and other financial institutions apply copulae as a risk management tool

2. Definitions and Theorems

Definition (Copula)

A d-dimensional copula is a multivariate distribution, $\mathcal{C}$, with standard uniform marginal distributions.

Theorem (Sklar)

Every multivariate distribution F, with margins, $F_1, F_2, \ldots, F_d$ can be written as

$$F(x_1, \ldots, x_d) = \mathcal{C}(F_1(x_1), \ldots, F_d(x_d)), \tag{2.1}$$

for some copula $\mathcal{C}$.

2. 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):

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

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

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

$$\mathcal{C}(\mathbf{u}) = P(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_d \leq u_d) = \int_0^{u_1} \cdots \int_0^{u_d} c(\mathbf{u}) \mathrm{d}\mathbf{u},$$

2. 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

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

- Hence,

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

- 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. Attractive features of copulae:

- 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

2.2. Examples

Example 1: Independence copula

If $U \sim U(0,1)$ and $V \sim U(0,1)$ are independent, then

$$C(u,v) = uv = \Pi = \mathbb{P}(U \leq u)\mathbb{P}(V \leq v) = \mathbb{P}(U \leq u, V \leq v) = H(u,v),$$

where $H(u,v)$ is the distribution function of (U,V). C is called the independence copula.

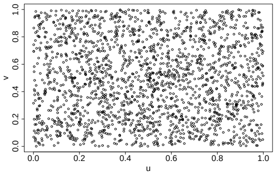


Figure: Simulations from the bivariate independence copula.

2.2 Examples

Example 2: Gaussian copula (implicit)

$$C_{\rho}^{Ga}(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left\{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\right\} \mathrm{d}x\mathrm{d}y,$$

where ρ is the linear correlation coefficient.
Example 3: Student's t copula (implicit)

$$C_{\rho,\nu}^{t}(u,v) = \int_{-\infty}^{t^{-1}(u)} \int_{-\infty}^{t^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left\{1+\frac{x^2-2\rho xy+y^2}{\nu(1-\rho^2)}\right\}^{-(\nu+2)/2} \mathrm{d}x\mathrm{d}y,$$

where ν is the degrees of freedom and ρ is the linear correlation coefficient.

2.2 Examples 2-3: Illustration

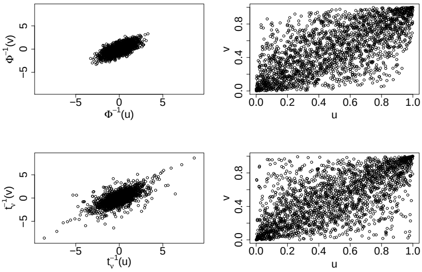


Figure: Simulations from the bivariate Gaussian- and Student's t distribution, and the associated copulae ($\rho = 0.7,\ \nu = 4$).

2.2 Examples

Example 4: Clayton copula (explicit)

$$C_{\delta}^{Cl}(u,v) = (u^{-\delta} + v^{-\delta} - 1)^{-1/\delta},$$

where $0 < \delta < \infty$ is the parameter controlling the dependence. Perfect dependence is obtained if $\delta \to \infty$, while $\delta \to 0$ implies independence.

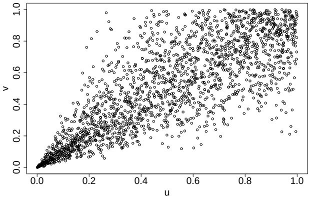


Figure: Simulations from the bivariate Clayton copula ($\delta = 3$).

2.2 Examples

Example 5: Gumbel copula (explicit)

$$C_\theta^{Gu}(u,v) = \exp\{-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}\},$$

where $1 < \theta < \infty$ is the parameter controlling the dependence. Perfect dependence is obtained if $\theta \to \infty$, while $\theta \to 1$ implies independence.

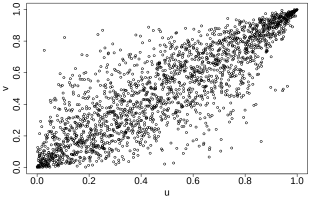


Figure: Simulations from the bivariate Gumbel copula ($\theta = 3$).

3. Dependence Concepts

We will consider the following dependence measures:

- Linear correlation
- Concordance
  - Kendall's tau
  - Spearman's rho
- Tail dependence

3.1. Linear correlation

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

- Sensitive to outliers
- Measures the "average dependence" between X and Y
- Invariant under strictly increasing **linear** transformations
- May be misleading in situations where multivariate df is not elliptical

3.1. Linear correlation

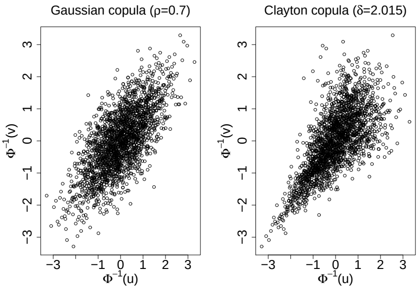


Figure: Illustration of the potential pitfalls of the linear correlation coefficient. Both distributions have linear correlation coefficient equal to 0.7.

3.2. Concordance

Let (x_i, y_i) and (x_j, y_j) be two observations from a random vector (X, Y) of continuous random variables.

- Concordance: $(x_i - x_j)(y_i - y_j) > 0$
- Discordance: $(x_i - x_j)(y_i - y_j) < 0$

Let (X_1, Y_1) and (X_2, Y_2) be independent vectors of cont. random variables with joint df's H_1 and H_2 and copulae C_1 and C_2, respectively. Let Q define the difference between the prob. of concordance and discordance of (X_1, Y_1) and (X_2, Y_2):

$$
\begin{aligned}
Q &= \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) > 0\right) - \mathbb{P}\left((X_1 - X_2)(Y_1 - Y_2) < 0\right) \\
&= Q(C_1, C_2) = 4 \int_0^1 \int_0^1 C_2(u, v) \mathrm{d}C_1(u, v) - 1.
\end{aligned}
$$

3.2.1. Kendall's tau

$$\begin{aligned}\rho_\tau(X,Y) = Q(C,C) &= 4\int_0^1\int_0^1 C(u,v)\mathrm{d}C(u,v) - 1\\ &= 4\mathbb{E}(C(U,V)) - 1.\end{aligned}$$

- Less sensitive to outliers
- Measures the "average dependence" between X and Y
- Invariant under strictly increasing transformations
- Depends only on the copula of (X,Y)
- For elliptical copulae: $\mathrm{cor}(X,Y) = \sin\left(\frac{\pi}{2}\rho_\tau\right)$

3.2.2. Spearmans's rho

$$\begin{aligned}\rho_S(X,Y) &= 3Q(C,\Pi)\\ &= 12\int_0^1\int_0^1 uv\,\mathrm{d}C(u,v) - 3\\ &= 12\int_0^1\int_0^1 C(u,v)\,\mathrm{d}u\mathrm{d}v - 3.\end{aligned}$$

- Less sensitive to outliers
- Measures the "average dependence" between X and Y
- Invariant under strictly increasing transformations
- Depends only on the copula of (X,Y)
- $\rho_S(X,Y) = \rho(F_X(X), F_Y(Y))$
- For elliptical copulae: $\mathrm{cor}(X,Y) = 2\sin\left(\frac{\pi}{6}\rho_S\right)$

3.3. Tail dependence

Let (X, Y) be a r.v. with marginal df's F_X and F_Y. The coefficient of upper and lower tail dependence of (X, Y) is defined as:

$$\lambda_u(X, Y) = \lim_{\alpha \to 1} \mathbb{P}(Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)),$$

$$\lambda_l(X, Y) = \lim_{\alpha \to 0} \mathbb{P}(Y \leq F_Y^{-1}(\alpha) | X \leq F_X^{-1}(\alpha)).$$

I.e. the tail dependence is the prob. of observing a large(small) Y, given that X is large(small). If $\lambda_u > 0$ ($\lambda_l > 0$), then we say that (X, Y) has upper (lower) tail dependence.

- Gaussian copula: $\lambda_u = \lambda_l = 2 \lim_{x \to \infty} \Phi \left(x \sqrt{1-\rho} / \sqrt{1+\rho} \right) = 0$
- Student-t copula: $\lambda_u = \lambda_l = 2 t_{\nu+1} \left(-\sqrt{\nu+1} \sqrt{(1-\rho)/(1+\rho)} \right)$. Asymptotic tail dependence, even when $\rho = 0$.
- Clayton copula: $\lambda_u = 0,\ \lambda_l = 2^{-1/\delta}$.
- Gumbel copula: $\lambda_l = 0,\ \lambda_u = 2 - 2^{1/\theta}$.

4. Copula Families

We will consider the two most important families of copulae:

- ▷ Elliptical copulae
- ▷ Archimedean copulae

4.1. Elliptical Copulae

- Implied by well-known multivariate df's, derived through Sklar's theorem
- Extends the multivariate normal $\mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$.
- Extend to arbitrary dimensions and are rich in parameters. A d-dim elliptical copula has at least $d(d-1)/2$ parameters
- Easy to simulate
- Drawback: Do not have closed form expressions and are restricted to have radial symmetry

Examples: Gaussian copula, Student's t copula

4.2. Archimedean Copulae

An Archimedean copula is defined as follows:

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)).$$

The function φ is called the generator of the copula.

- Allow for a great variety of dependence structures
- Closed form expressions
- **Not** derived from mv df's using Sklar's theorem
- Drawback: Higher dimensional extensions difficult

Examples: Clayton copula, Gumbel copula

4.2. Archimedean Copulae

Example 1: Clayton copula
The generator function for the Clayton copula is given by $\varphi(t) = (t^{-\delta} - 1)/\delta$, where $\delta \in (0, \infty)$. This gives the Clayton copula:

$$C_\delta(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v)) = (u^{-\delta} + v^{-\delta} - 1)^{-1/\delta}.$$

The Clayton copula has lower tail dependence.
Example 2: Gumbel copula
The generator function for the Gumbel copula is given by $\varphi(t) = (-\ln t)^\theta$, where $\theta \geq 1$. This gives the Gumbel copula:

$$C_\theta(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v)) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}).$$

The Gumbel copula has upper tail dependence.

5. Estimating Copula Parameters

Fully parametric method:

- Denoted Inference functions for margins (IFM) method.
- Assumes parametric univariate marginal distributions.
- Parameters of margins are first estimated, then each parametric margin is plugged into the copula likelihood, and this full likelihood is maximized.
- Success depends upon finding appropriate parametric models for the margins, which is not always straightforward

Semi-parametric method:

- Denoted the pseudo-likelihood or canonical maximum likelihood (CML) method
- No parametric assumptions for the margins, use empirical cdf's, then plug into likelihood

5.1 Estimation - Elliptical copulae

Gaussian copula:

- Correlation matrix $\boldsymbol{R}$ ($d(d-1)/2$ parameters)
- ML estimator: $\widehat{\boldsymbol{R}} = \arg\max_{\boldsymbol{R}\in\mathcal{P}} \sum_{j=1}^{n} \log c(\boldsymbol{U}_j; \boldsymbol{R})$, where the pseudo samples $\boldsymbol{U}_j$ are generated using either the IFM or the CML method.

Student's t copula:

- Correlation matrix $\boldsymbol{R}$ and degree-of-freedom ν ($1 + d(d-1)/2$ parameters)
- ML wrt $\boldsymbol{R}$ and ν simultaneously difficult
- Simpler: two-stage approach in which $\boldsymbol{R}$ is estimated first using Kendall's tau, and then the pseudo-likelihood function is maximized wrt ν.

5.2 Estimation - Archimedean copulae

Clayton and Gumbel copulae:

- One parameter, δ and θ respectively
- Numerical optimization of likelihood
- Bivariate - utilize the following relationships to Kendall's tau:

$$\widehat{\delta} = \frac{2\widehat{\rho}_\tau}{1-\widehat{\rho}_\tau}, \qquad \widehat{\theta} = \frac{1}{1-\widehat{\rho}_\tau}.$$

6. Simulating from Copulae

Gaussian copula:

- Simulate $\boldsymbol{X} \sim \mathcal{N}_d(\mathbf{0}, \boldsymbol{R})$
- Set $\boldsymbol{U} = (\Phi(X_1), \ldots, \Phi(X_d))$ or $\boldsymbol{U} = (F(X_1), \ldots, F(X_d))$ where the F's are the quantile functions

Student's t copula:

- Simulate $\boldsymbol{X} \sim t_d(\mathbf{0}, \boldsymbol{R}, \nu)$
- Set $\boldsymbol{U} = (t_\nu(X_1), \ldots, t_\nu(X_d))$ or $\boldsymbol{U} = (F(X_1), \ldots, F(X_d))$ where the F's are the quantile functions

6. Simulating from Copulae

Clayton copula:

By noting that the inverse of the generator is equal to the Laplace transform of a Gamma variate $X \sim Ga(1/\delta, 1)$, the simulation algorithm becomes:

- Simulate a gamma variate $X \sim Ga(1/\delta, 1)$
- Simulate d iid $U(0,1)$ variables $V_1, \ldots, V_d$
- Return $\boldsymbol{U} = \left((1 - \frac{\log V_1}{X})^{-1/\delta}, \ldots, (1 - \frac{\log V_d}{X})^{-1/\delta}\right)$

Gumbel copula:

By noting that the inverse of the generator function is equal to the Laplace transform of a positive stable variate $X \sim St(1/\theta, 1, \gamma, 0)$, where $\gamma = \left(\cos\left(\frac{\pi}{2\theta}\right)\right)^{\theta}$ and $\theta > 1$, the simulation algorithm becomes:

- Simulate a positive stable variate $X \sim St(1/\theta, 1, \gamma, 0)$
- Simulate d iid $U(0,1)$ variables $V_1, \ldots, V_d$
- Return $\boldsymbol{U} = \left(\exp\left(-\left(-\frac{\log V_1}{X}\right)^{1/\theta}\right), \ldots, \exp\left(-\left(-\frac{\log V_d}{X}\right)^{1/\theta}\right)\right)$

6. Simulating from Copulae

In general we could apply the conditional marginal cdf's:

$$F_{i|1,\dots,i-1}(u_i|u_1,\dots,u_{i-1}) = \frac{\partial^{i-1}C(u_1,\dots,u_i)}{\partial u_1 \cdots \partial u_{i-1}} \Big/ \frac{\partial^{i-1}C(u_1,\dots,u_{i-1})}{\partial u_1 \cdots \partial u_{i-1}}.$$

The simulation algorithm then becomes:

- Simulate a rv u_1 from $U(0,1)$,
- Simulate a rv u_2 from $F_{2|1}(\cdot|u_1)$,

  $\vdots$

- Simulate a rv u_d from $F_{d|1,\dots,d-1}(\cdot|u_1,\dots,u_{d-1})$.
- Generally means simulating a rv V_i from $U(0,1)$ from which $u_i = F^{-1}_{i|1,\dots,i-1}(V_i|u_1,\dots,u_{i-1})$ can be obtained, if necessary by numerical root finding.

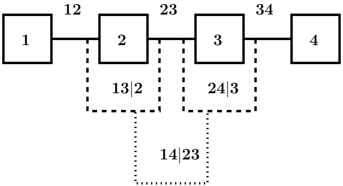
7. Higher Dimensional Copulae

I. Copulae with at least $d(d-1)/2$ bivariate dependence parameters:

- Build multivariate copulae from bivariate copula
- Based on iteratively mixing conditional copulae
- Very flexible tool for dependency modelling
- Does not require any assumption of conditional independence
- Also referred to as 'Vines' (Cooke and Bedford, 2002)
- Drawback: difficult, slow, depends heavily on permutation

7. Higher Dimensional Copulae

I. Copulae with at least $d(d-1)/2$ bivariate dependence parameters:



Example:

$$
\begin{aligned}
c_{1234}(u_1,u_2,u_3,u_4) = c_{12}(u_1,u_2) \cdot c_{23}(u_2,u_3) \cdot c_{34}(u_3,u_4) \\
\cdot\, c_{13|2}(F_{1|2}(u_1|u_2), F_{3|2}(u_3|u_2)) \cdot c_{24|3}(F_{2|3}(u_2|u_3), F_{4|3}(u_4|u_3)) \\
\cdot\, c_{14|23}(F_{1|23}(u_1|u_2,u_3), F_{4|23}(u_4|u_2,u_3)),
\end{aligned}
$$

where $F_{i|1\ldots i-1}(u_i|u_1,\ldots,u_{i-1}) = \frac{\partial^{i-1}C(u_1,\ldots,u_i)}{\partial u_1\cdots\partial u_{i-1}} \Big/ \frac{\partial^{i-1}C(u_1,\ldots,u_{i-1})}{\partial u_1\cdots\partial u_{i-1}}$.

7. Higher Dimensional Copulae

II. Archimedean copulae with $d-1$ bivariate dependence parameters:

- Build multivariate copulae from bivariate copula
- Based on iteratively mixing conditional copulae
- Less flexible but more intuitive and faster than 'vines'
- Only applicable to Archimedean copulae with strict generator functions

$$\begin{aligned} C^3(u_1,u_2,u_3) &= \varphi^{-1}[\varphi(u_1)+\varphi(u_2)+\varphi(u_d)] \\ &= \varphi^{-1}[\varphi(\varphi^{-1}[\varphi(u_1)+\varphi(u_2)])+\varphi(u_3)] \\ &= C^2(C^2(u_1,u_2),u_3), \\ \Rightarrow C^d(u_1,\ldots,u_d) &= C^2(C^{d-1}(u_1,\ldots,u_{d-1}),u_d). \end{aligned}$$

Example:

$$C^3(u_1,u_2,u_3) = \varphi_2^{-1}[\varphi_2 \circ \varphi_1^{-1}[\varphi_1(u_1)+\varphi_1(u_2)]+\varphi_2(u_3)].$$

8. Application

Using some simplifying assumptions we simulated losses for a portfolio of 30 American firms, assuming Gaussian, Student's t and Vine (Student's t C-vine) dependency structures.

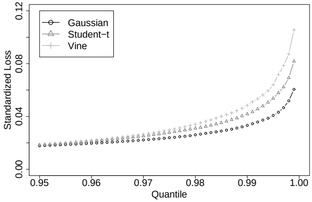


Figure: Value-at-risk for a portfolio of 30 american firms, assuming Gaussian, Student's t and Vine dependency structure.

9. Copula Goodness-of-fit Testing

- Special case of testing for multivariate density models
- Complicated due to the unspecified marginal df's. Asymptotic distributional properties becomes very difficult to derive $\Rightarrow$ p-values obtained through simulation.
- χ^2- and other tests based on binning the probability space will not be feasible in higher dimensions as the need for data would be too great.
- Some tests focus on multivariate smoothing procedures. These are computationally very demanding in high dimensions.
- A more promising class of tests project the multivariate problem to a univariate problem, then apply a univariate GOF statistic, e.g. Anderson-Darling (AD).
- We may base the testing on the probability integral transform (PIT).

9.1 Probability Integral Transform (PIT)

- The PIT 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
- First introduced by Rosenblatt (1952)
- Inverse of simulation
- GOF: the observed copula is PIT assuming a $\mathcal{H}_0$ copula. Then a test of independence is performed.

9.1 Probability Integral Transform (PIT)

DEFINITION: Probability Integral Transform

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

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

The random variables $Z_i = T_i(X_i)$, for $i = 1, \dots, d$ are uniformly and independently distributed on $[0, 1]^d$. $F(x_i | x_1, \dots, x_{i-1})$ is found by

$$
F_{i|1 \dots i-1}(u_i | u_1, \dots, u_{i-1}) = \frac{\partial^{i-1} C(u_1, \dots, u_i)}{\partial u_1 \cdots \partial u_{i-1}} \Big/ \frac{\partial^{i-1} C(u_1, \dots, u_{i-1})}{\partial u_1 \cdots \partial u_{i-1}}.
$$

9.2 Proposed tests

G: Breymann et al. (2003)

$$Y_j^G = \sum_{i=1}^{d} \Phi^{-1}(z_{ji})^2, \quad j = 1, \ldots, n,$$
$$G(w) = P\left(F_{\chi_d^2}(Y^G \leq w)\right), \quad w \in [0,1].$$

- Coincides with the tests proposed by Malevergne and Sornette (2003) when the latter is based on PIT. Also coincides with the test proposed by Chen et al. (2004).
- Very fast
- Tail weight
- **NOT** consistent

9.2 Proposed tests

B: Berg and Bakken (2005)

$$z^*_{ji} = P(r_i \leq \widetilde{z}_{ji} | r_1, \ldots, r_{i-1}) = \left(1 - \left(\frac{1-\widetilde{z}_{ji}}{1-r_{i-1}}\right)^{d-(i-1)}\right),$$

$$Y^B_j = \sum_{i=1}^{d} \gamma(z_{ji}; \boldsymbol{\alpha}) \cdot \Phi^{-1}(z^*_{ji})^2, \quad j = 1, \ldots, n,$$

where $\gamma(\cdot)$ is a weight function and $\boldsymbol{\alpha}$ are weight parameters. Then

$$B(w) = P(F_B(Y^B) \leq w), \quad w \in [0,1].$$

- Similar to G-test but based on transformed data Z^*.
- Fast
- Any weight
- Consistent

9.2 Proposed tests

Q: Panchenko (2005)

$$Q = \langle f_1 - f_2 | \kappa_d | f_1 - f_2 \rangle = Q_{11} - 2Q_{12} + Q_{22},$$

$$\widehat{Q}_{kp} = \frac{1}{n^2} \sum_{j_k=1}^{n} \sum_{j_p=1}^{n} \kappa_d(\boldsymbol{X}_k^{j_k}, \boldsymbol{X}_p^{j_p}),$$

where

$$\kappa_d(\boldsymbol{x}_1, \boldsymbol{x}_2) = \exp\left\{ -\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2 / (2dh^2) \right\}.$$

- Based on positive bilinear forms
- Very slow
- No weight
- Consistent

9.2 Proposed tests

K: Genest et al. (2006)

$$K(w) = P(C(\boldsymbol{Z}) \leq w), \quad w \in [0,1],$$

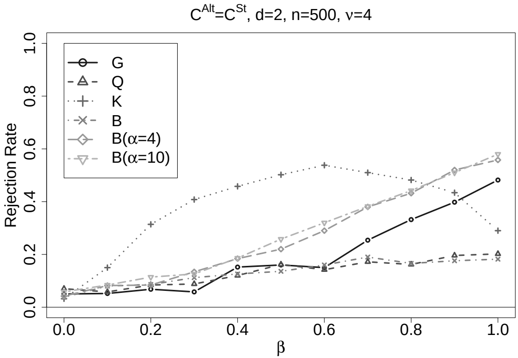
$$\widehat{K}(w) = \frac{1}{n+1} \sum_{j=1}^{n} I\left(\widehat{C}(\boldsymbol{z}_j) \leq w\right), \quad w = \frac{1}{n+1}, \ldots, \frac{n}{n+1}.$$

- Based on the empirical copula and Kendall's process
- Slow
- Left tail weight
- Consistent

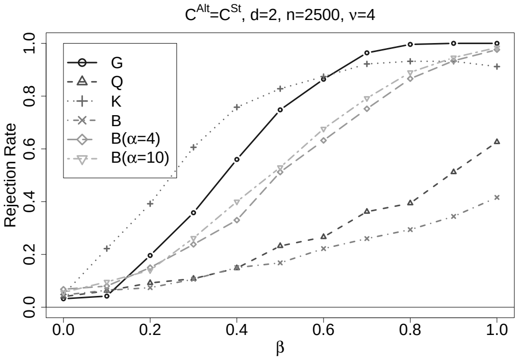
9.3 Mixing results

- Mix a Gaussian copula with an alternative copula to construct a mixed copula $\mathcal{C}^{mix} = (1-\beta)\cdot\mathcal{C}^{Ga} + \beta\cdot\mathcal{C}^{Alt}, \quad \beta \in [0,1]$
- $\mathcal{C}^{Alt}$: Student's t ($\mathcal{C}^{St},\ \nu = 4$), Clayton($\mathcal{C}^{Cl},\ \delta = 1.0$) and survival Clayton($\mathcal{C}^{sCl},\ \delta = 1.0$)
- $\mathcal{H}_0:\ \mathcal{C}^{Ga}$
- Simulate from $\mathcal{C}^{Ga}$ and $\mathcal{C}^{Alt}$ and mix. Then PIT $\mathcal{C}^{mix}$ under $\mathcal{H}_0$. Finally compute test statistic and corresponding p-value
- Repeat 500 times to obtain rejection rates
- Consider G, Q, K and B test. For the B-test we consider no weight and power tail weighting: $\gamma(Z_i;\alpha) = (Z_i - \frac{1}{2})^{\alpha}, \quad \alpha = [4, 10]$.

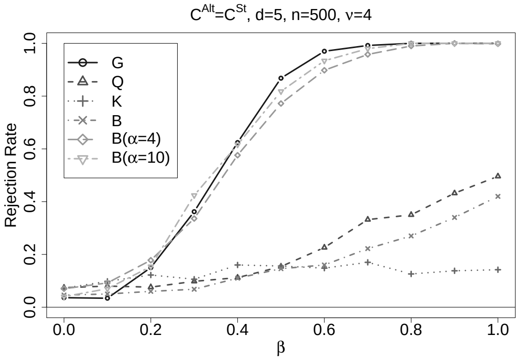
9.3 Mixing results



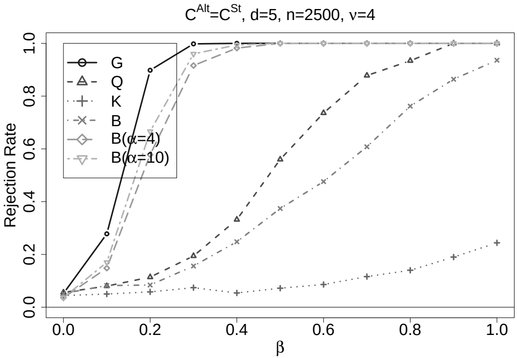
9.3 Mixing results



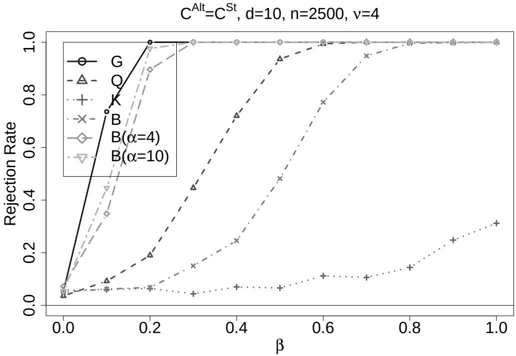
9.3 Mixing results



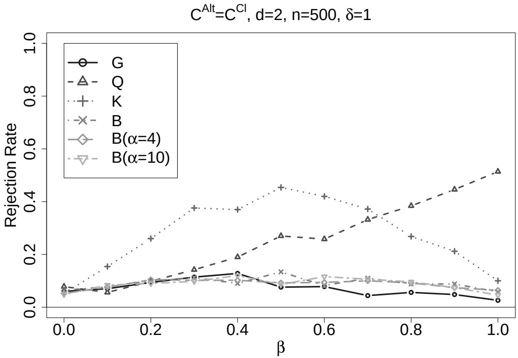
9.3 Mixing results



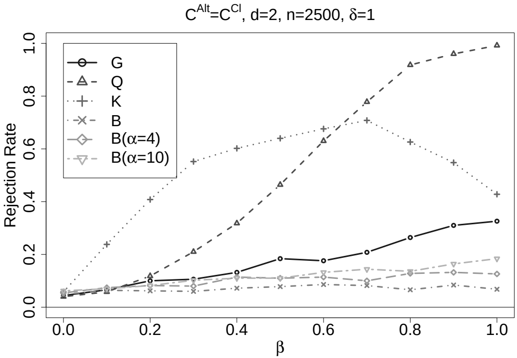
9.3 Mixing results



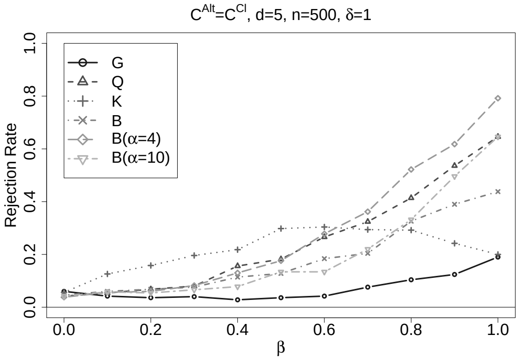
9.3 Mixing results



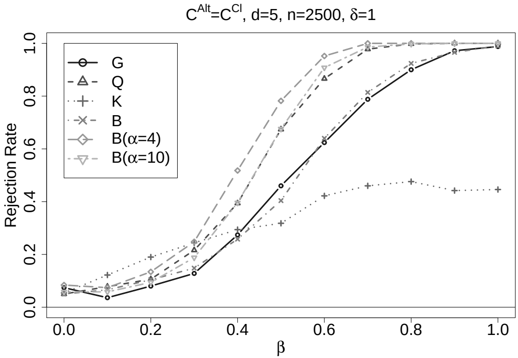
9.3 Mixing results



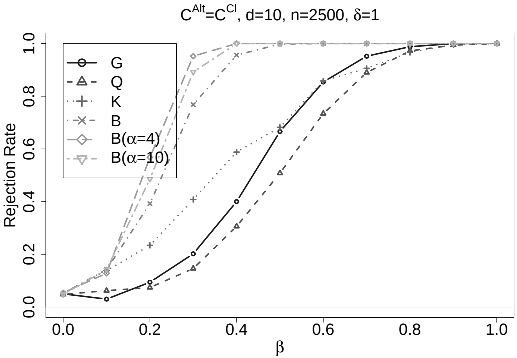
9.3 Mixing results



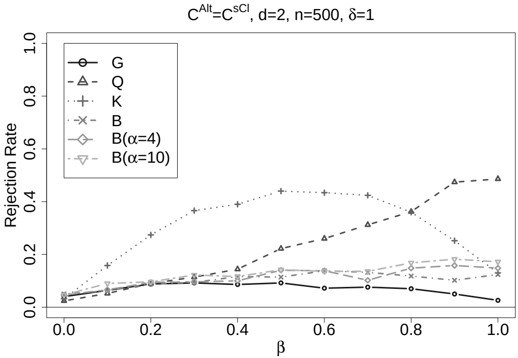
9.3 Mixing results



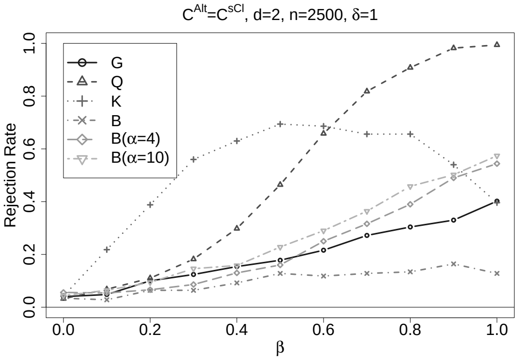
9.3 Mixing results



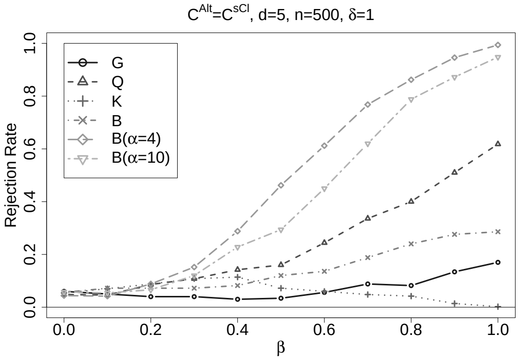
9.3 Mixing results



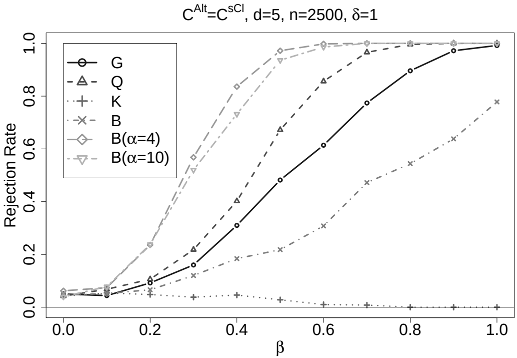
9.3 Mixing results



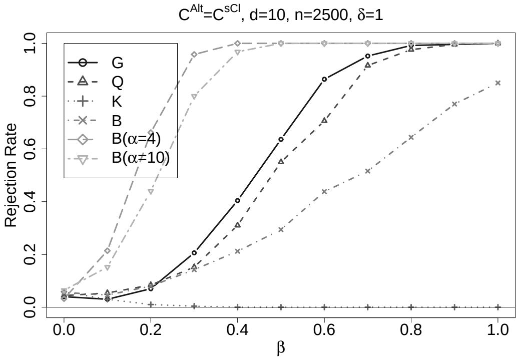
9.3 Mixing results



9.3 Mixing results



9.3 Mixing results



10. Summary

- Linear correlation coefficient not sufficient outside the world of elliptical distributions $\Rightarrow$ alternative dependence measures
- Copula families: Elliptical, Archimedean
- Estimation and simulation
- Complex multivariate highly dependent models can be built, based on bivariate copulae
- Significant impacts, i.e. on portfolio VaR
- Goodness-of-fit:
  - Bivariate: several candidates
  - Dimension > 2: B-test

References

Berg, D. and H. Bakken (2005, December). A goodness-of-fit test for copulae based on the probability integral transform. Note, Norwegian Computing Centre, SAMBA/41/05.

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.

Genest, C., J.-F. Quessy, and B. Rémillard (2006). Goodness-of-fit procedures for copula models based on the probability integral transform. *Scandinavian Journal of Statistics 33*.

Malevergne, Y. and D. Sornette (2003). Testing the gaussian copula hypothesis for financial assets dependence. *Quantitative Finance 3*, 231–250.

Panchenko, V. (2005). Goodness-of-fit test for copulas. *Physica A 355(1)*, 176–182.