TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

Financial & Actuarial Mathematics

Aggregation of Integer-Valued Risks with Copula-Induced Dependency Structure

Martin Schmidt

March 11, 2019

Table of contents

The multivariate distribution function F of a random vector $(X_1, \dots, X_d)$ contains two kinds of information:

- the univariate marginal distributions $F_1, \dots, F_d$ and
- the dependency structure among the components.

Definition

Let $d \in \mathbb{N}, d \geq 2$ and let $(U_1, \dots, U_d)$ denote a random vector on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that the random variable U_i is standard uniformly distributed for all $i = 1, \dots, d$.
A d-dimensional copula C is a multivariate distribution function on the d-dimensional unit cube with univariate standard uniform marginals,

$$\begin{aligned} C : [0,1]^d &\to [0,1] \\ (u_1, \dots, u_d) &\mapsto \mathbb{P}[U_1 \leq u_1, \dots, U_d \leq u_d]. \end{aligned}$$

Any dependency structure can be described using a copula

Sklar's theorem

Let F denote a multivariate distribution function with univariate marginals $F_1, \dots, F_d$. Then there exists a d-dimensional copula $C : [0,1]^d \to [0,1]$, such that for all $x_1, \dots, x_d \in \mathbb{R}$ it holds that

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)). \tag{1}$$

If $F_1, \dots, F_d$ are continuous, then C is unique. Conversely, if C is a copula and $F_1, \dots, F_d$ are univariate distribution functions, then the function F defined via (1) is a d-dimensional distribution function with one-dimensional margins $F_1, \dots, F_d$.

Restricted uniqueness of copulas in the case of non-continuous margins

Sklar's theorem states that for continuous marginal distributions the copula is unique on $[0,1]^d$.

$\Rightarrow$ **This is no longer true for discrete or mixed margins!**

Theorem

Let F denote the multivariate distribution function of a random vector $X = (X_1, \dots, X_d)$ on $\mathbb{R}^d$ with univariate marginals $F_1, \dots, F_d$. Then a copula C of X is uniquely determined on $\mathrm{Ran}(F_1) \times \cdots \times \mathrm{Ran}(F_d)$.

Extreme example: Consider Bernoulli-distributed random variables $X_1 \sim \mathrm{Bern}(p_1)$, $X_2 \sim \mathrm{Bern}(p_2)$ with $p_1, p_2 \in (0,1)$. Then a copula of (X_1, X_2) is only uniquely determined in the point $(1-p_1, 1-p_2)$.

Fundamental copulas

Independence copula

The independence copula $\Pi : [0,1]^d \to [0,1]$ is given by

$$\Pi(u_1, \dots, u_d) := \prod_{i=1}^{d} u_i, \quad u_1, \dots, u_d \in [0,1].$$

Comonotonicity copula

The comonotonicity copula $M : [0,1]^d \to [0,1]$ is given by

$$M(u_1, \dots, u_d) := \min\{u_1, \dots, u_d\}, \quad u_1, \dots, u_d \in [0,1].$$

Countermonotonicity copula (only in dimension $d = 2$)

The countermonotonicity copula $W : [0,1]^2 \to [0,1]$ is given by

$$W(u_1, u_2) := \max\{u_1 + u_2 - 1, 0\}, \quad u_1, u_2 \in [0,1].$$

Gaussian copulas

Definition

Let Φ denote the distribution function of a univariate standard normal distribution and let Φ_P^d be the distribution function of a d-variate normal distribution with correlation matrix P and mean 0. Then for $u_1, \ldots, u_d \in [0,1]$ the d-dimensional Gaussian copula $C_P^{\mathrm{Ga}} : [0,1]^d \to [0,1]$ is given as

$$C_P^{\mathrm{Ga}}(u_1, \ldots, u_d) := \Phi_P^d(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)).$$

- If $P = I_d$, where I_d denotes the d-dimensional identity matrix, then $C_P^{\mathrm{Ga}} \equiv \Pi$.
- If $P = J_d$, where J_d denotes a $d \times d$ matrix consisting entirely of ones, then $C_P^{\mathrm{Ga}} \equiv M$.
- If $d = 2$ and $P = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, then $C_P^{\mathrm{Ga}} \equiv W$.

t-copulas

Definition

Let t_ν be the distribution function of a univariate standard t-distribution with ν degrees of freedom, $\nu > 0$. By $t_{\nu,P}$ we denote the multivariate distribution function of a d-variate t-distribution with dispersion matrix P and $\nu > 0$ degrees of freedom. Then for $u_1, \ldots, u_d$ the d-dimensional t-copula $C^t_{\nu,P} : [0,1]^d \to [0,1]$ is given as

$$C^t_{\nu,P}(u_1, \ldots, u_d) := t_{\nu,P}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)).$$

- **Caution!** If $P = I_d$, where I_d denotes the d-dimensional identity matrix, then $C^t_{\nu,P} \not\equiv \Pi$.
- If $P = J_d$, where J_d denotes a $d \times d$ matrix consisting entirely of ones, then $C^t_{\nu,P} \equiv M$.

Setting

- Arbitrary dimension $d \in \mathbb{N}_{\geq 2}$
- $\mathbb{N}_0$-valued random variables (risks) $X_1, \ldots, X_d$
- $X_i \sim F_i$ for $i = 1, \ldots, d$
- Dependency structure of the portfolio $(X_1, \ldots, X_d)$ is given by an arbitrary copula C
- **Aggregate portfolio loss** $S = \sum_{i=1}^{d} X_i$

Notational conventions

For $n \in \mathbb{N}_0$ we define

- $\mathcal{J}_n^d = \{j = (j_1, \ldots, j_d) \in \mathbb{N}_0^d : j_1 + \cdots + j_d \leq n\}$
- $\bar{\mathcal{J}}_n^d = \{j = (j_1, \ldots, j_d) \in \mathbb{N}_0^d : j_1 + \cdots + j_d = n\}$
- $p_n = \mathbb{P}[S = n]$
- $c_n = \sum_{j \in \bar{\mathcal{J}}_n^d} C(F_1(j_1), \ldots, F_d(j_d))$
- Convention: $p_n = c_n = 0$, if $n < 0$

Further,

- $\mathcal{I}^d = \{i = (i_1, \ldots, i_d) \in \{0, 1\}^d\}$
- $\mathrm{sign}(i) = (-1)^{\sum_{k=1}^d i_k}, \quad i \in \mathcal{I}^d$

Known results

Joint probability mass function of $(X_1, \dots, X_d)$

From the properties of copulas we know that for $j_1, \dots, j_d \in \mathbb{N}_0$

$$\mathbb{P}[X_1 = j_1, \dots, X_d = j_d] = \sum_{i \in \mathcal{I}^d} \mathrm{sign}(i) C(F_1(j_1 - i_1), \dots, F_d(j_d - i_d)).$$

Distribution of the sum S

$$\mathbb{P}[S \leq n] = \sum_{j \in \mathcal{J}_n^d} \mathbb{P}[X_1 = j_1, \dots, X_d = j_d], \quad n \in \mathbb{N}_0$$

Disadvantage of the **formula above: inefficient**, many summands
$\Rightarrow$ For $d = 6$ and $n = 100$ we sum up 109.177.903.744 terms!

Theorem (Distribution of the sum S)

Let $X_1, \ldots, X_d$ denote $\mathbb{N}_0$-valued random variables with univariate distribution functions $F_1, \ldots, F_d$, respectively. Then for $n \in \mathbb{N}_0$ and every copula C of the random vector $(X_1, \ldots, X_d)$ it holds that

$$\mathbb{P}[S \leq n] = \sum_{k=0}^{\min\{d-1,n\}} (-1)^k \binom{d-1}{k} c_{n-k}.$$

$\Rightarrow$ For $d = 6$ and $n = 100$ we **reduce** the **number of summands by** approximately **99.53%**

Recursion for the probability mass function of S

Remember:

- $\bar{\mathcal{J}}_n^d = \{j = (j_1, \dots, j_d) \in \mathbb{N}_0^d : j_1 + \dots + j_d = n\}$
- $p_n = \mathbb{P}[S = n]$
- $c_n = \sum_{j \in \bar{\mathcal{J}}_n^d} C(F_1(j_1), \dots, F_d(j_d))$
- Convention: $p_n = c_n = 0$, if $n < 0$

Recursion for the probability mass function of the sum S

Starting with $p_0 = c_0$, the following recursion formula applies:

$$p_n = c_n - \sum_{k=1}^{n} \binom{k+d-1}{d-1} p_{n-k}, \quad n \in \mathbb{N}.$$

Numerical speed-up

Disadvantages of the **recursion formula**:

- **Weak efficiency** – especially for sparse univariate marginal distributions or large dimensions d combined with large n
- **High precision** is **necessary**

Theorem (Probability mass function of the sum S)

Let $X_1, \ldots, X_d$ denote $\mathbb{N}_0$-valued random variables with univariate distribution functions $F_1, \ldots, F_d$, respectively. Then for all $n \in \mathbb{N}_0$ and every copula C of the random vector $(X_1, \ldots, X_d)$ it holds that

$$\mathbb{P}[S = n] = \sum_{k=0}^{\min\{d,n\}} (-1)^k \binom{d}{k} c_{n-k}.$$

Integration over copula densities

Definition

If the probability measure associated with a copula C is absolutely continuous with respect to the Lebesgue measure on $[0,1]^d$, then by Radon–Nikodým there exists an almost everywhere unique density $c : [0,1]^d \to [0,\infty)$ such that for $u_1,\dots,u_d \in [0,1]$:

$$C(u_1,\dots,u_d) = \int_0^{u_1} \cdots \int_0^{u_d} c(v_1,\dots,v_d)\, dv_d \dots dv_1.$$

Theorem (Probability mass function of the sum S)

For $n \in \mathbb{N}_0$ we have that

$$\mathbb{P}[S=n] = \sum_{j \in \bar{\mathcal{J}}_n^d} \int_{F_1(j_1-1)}^{F_1(j_1)} \cdots \int_{F_d(j_d-1)}^{F_d(j_d)} c(v_1,\dots,v_d)\, dv_d \dots dv_1.$$

Value-at-Risk and Expected Shortfall

Value-at-Risk

At a given confidence level $\alpha \in (0,1)$, the Value-at-Risk (VaR) of a random variable S is the smallest value $s \in \mathbb{R}$ where the distribution function F_S of S reaches or exceeds the value α for the first time:

$$\mathrm{VaR}_\alpha(S) := \inf\{s \in \mathbb{R} : F_S(s) \geq \alpha\}.$$

Expected Shortfall

For a random variable S with $\mathbb{E}[|S|] < \infty$ the Expected Shortfall (ES) at a confidence level $\alpha \in (0,1)$ is given as

$$\mathrm{ES}_\alpha(S) := \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_u(S)\, du.$$

Example

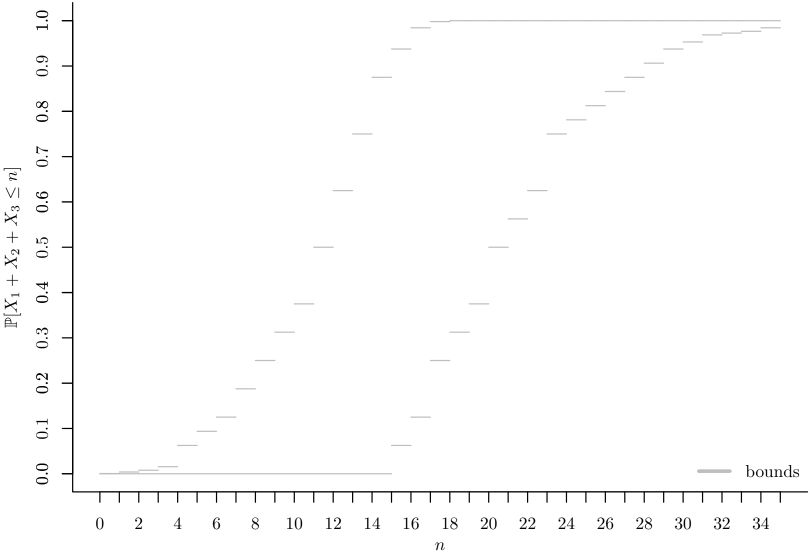
Let $d = 3$ and consider Poisson-distributed random variables

$$X_1 \sim \mathrm{Poi}(3),\ X_2 \sim \mathrm{Poi}(5) \text{ and } X_3 \sim \mathrm{Poi}(8).$$

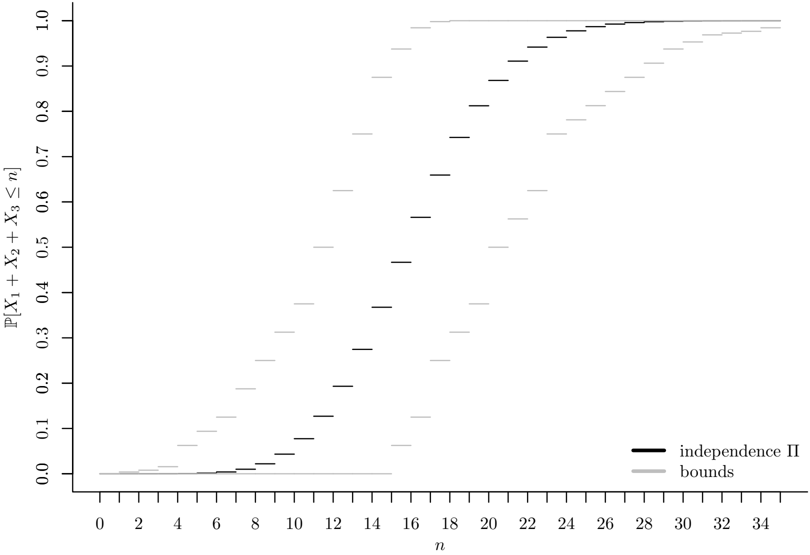
On the next slides we will present the distribution, probability mass function and risk measures for the sum $S = X_1 + X_2 + X_3$ under the following dependency scenarios:

- Fundamental copulas
- Gaussian copulas
- t-copulas with $\nu = 1$

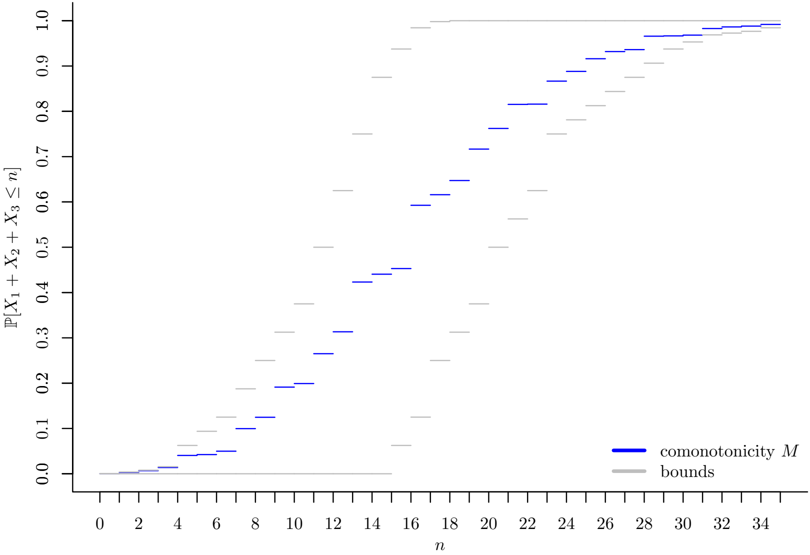
Fundamental copulas



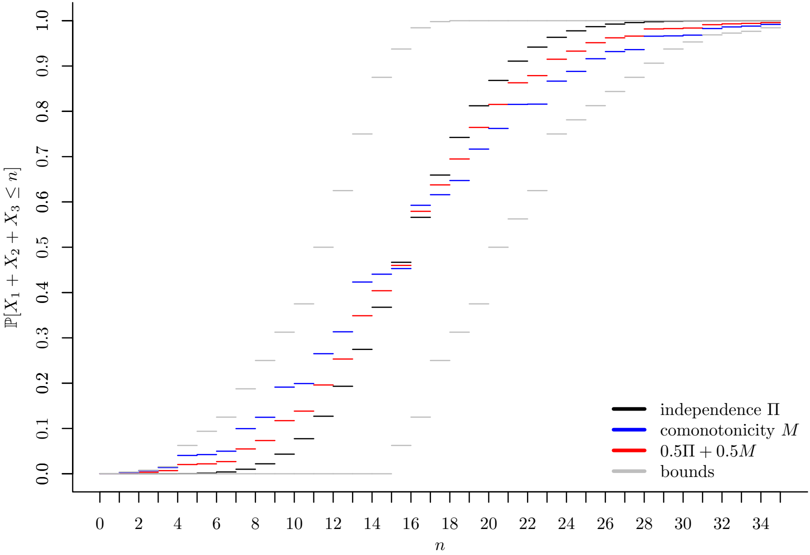
Fundamental copulas



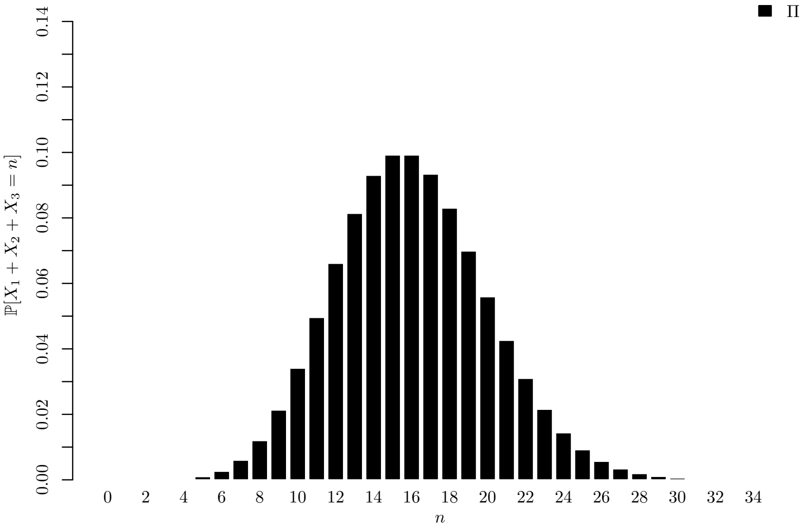
Fundamental copulas



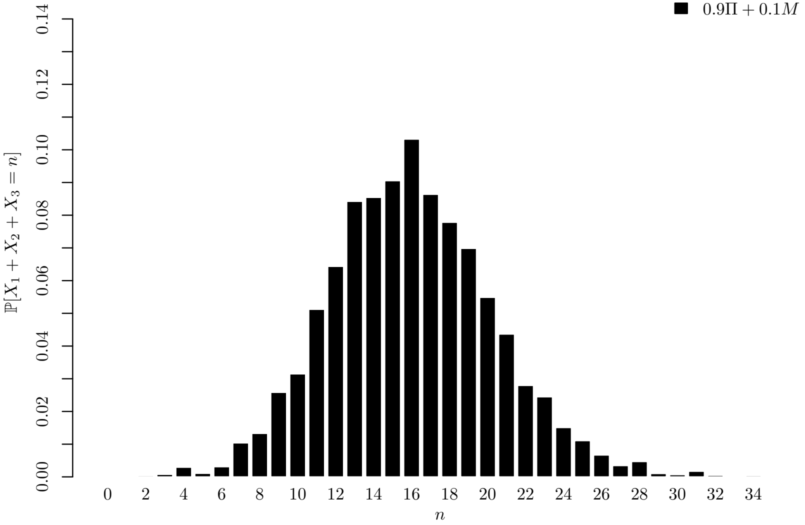
Fundamental copulas



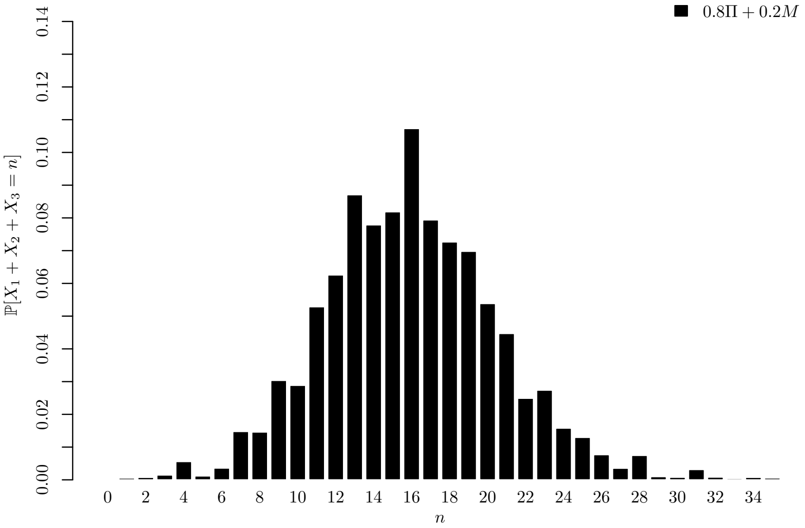
Fundamental copulas



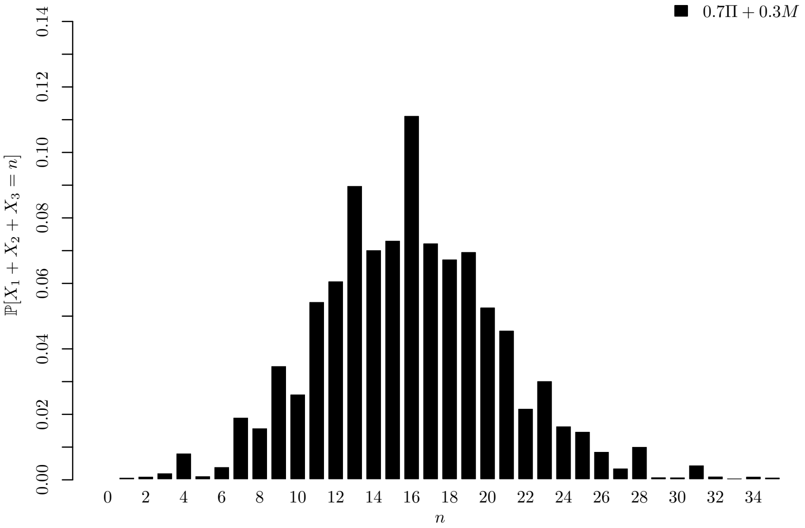
Fundamental copulas



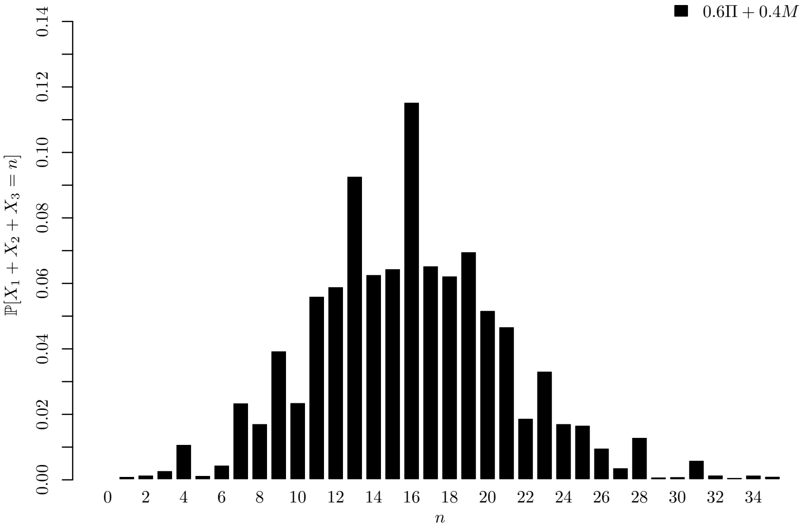
Fundamental copulas



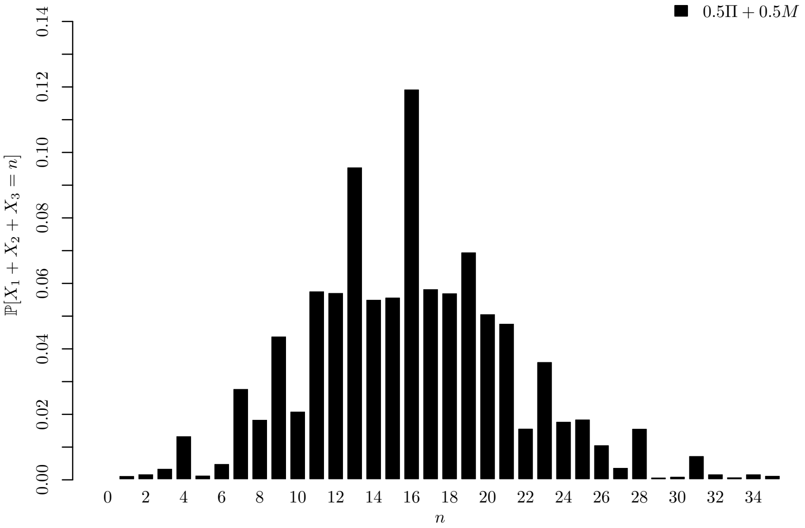
Fundamental copulas



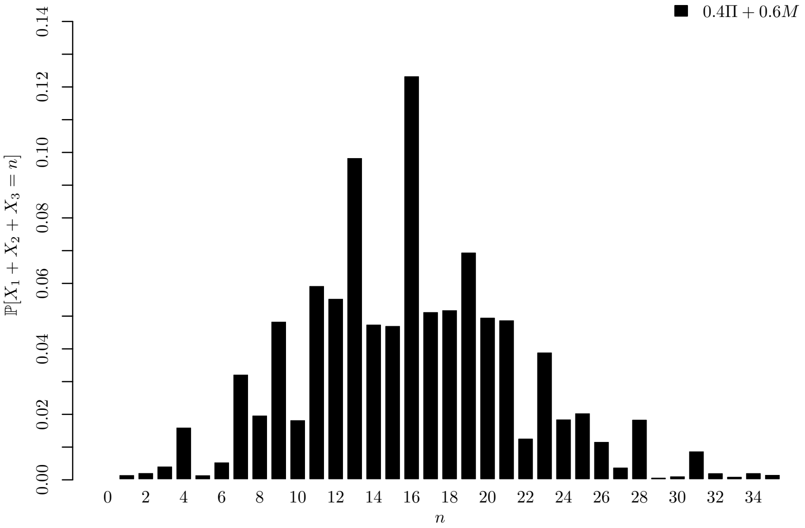
Fundamental copulas



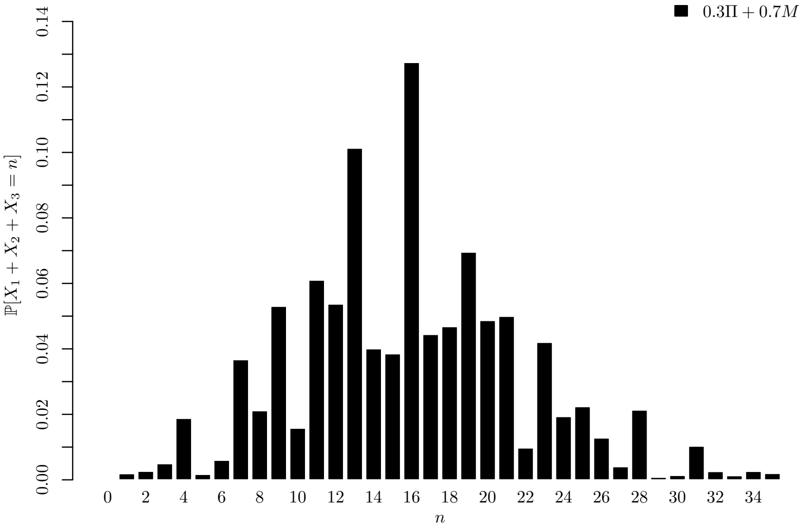
Fundamental copulas



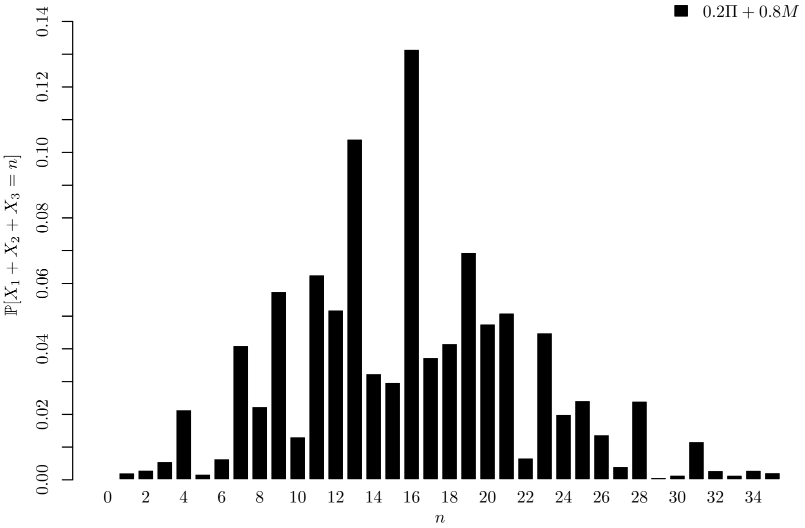
Fundamental copulas



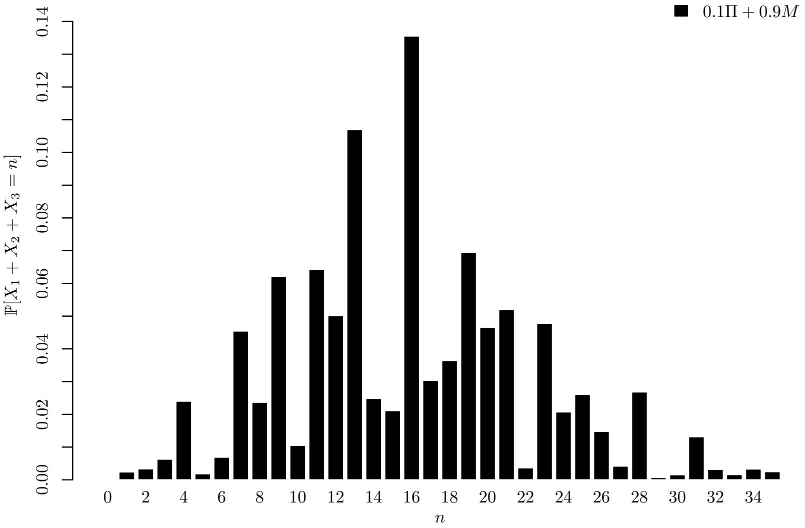
Fundamental copulas



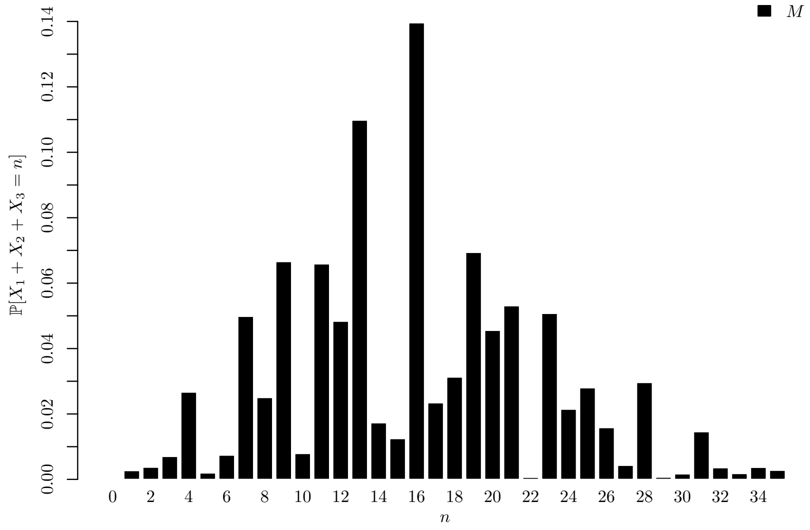
Fundamental copulas



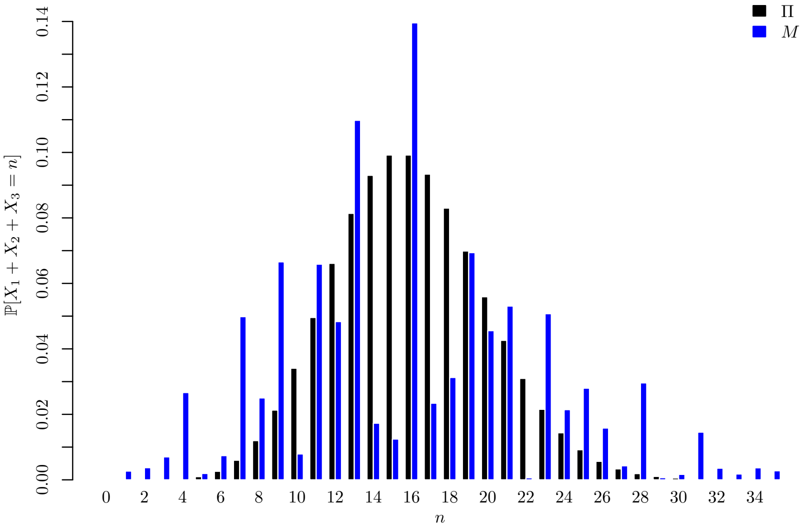
Fundamental copulas



Fundamental copulas



Fundamental copulas



Gaussian copulas

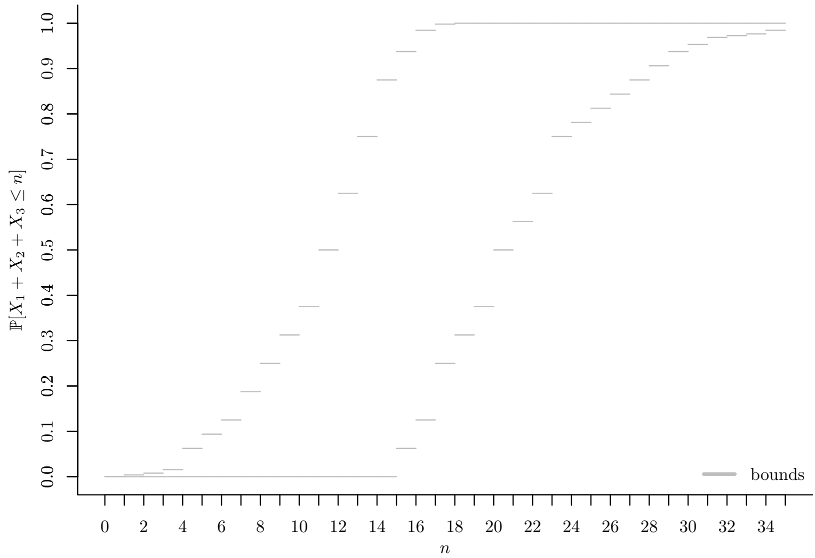
Notes for the following plots:

- By a Gaussian copula with parameter ρ, i.e. C_ρ^{Ga}, we denote a Gaussian copula with correlation matrix such that all pairwise correlations coincide to $\rho \in [-1, 1]$.
- For the Poisson-distributed random variables from our example it holds that $X_1 + X_2 \overset{d}{=} X_3$, if X_1 and X_2 are independent. So if we try to minimize the variance of the sum $X_1 + X_2 + X_3$ under a Gaussian copula we can use

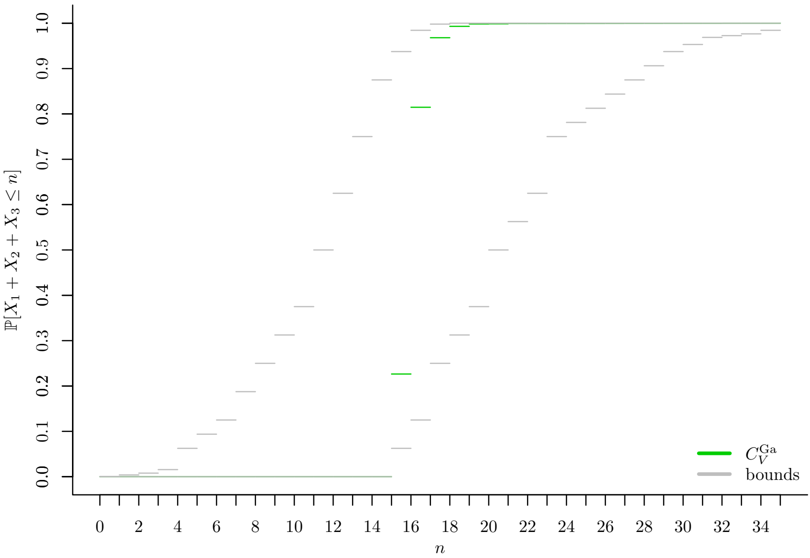
$$V = \begin{pmatrix} 1 & 0 & -\sqrt{\frac{3}{8}} \\ 0 & 1 & -\sqrt{\frac{5}{8}} \\ -\sqrt{\frac{3}{8}} & -\sqrt{\frac{5}{8}} & 1 \end{pmatrix},$$

where the entries of the correlation matrix V are obtained by simple calculation under positive semi-definite constraints.

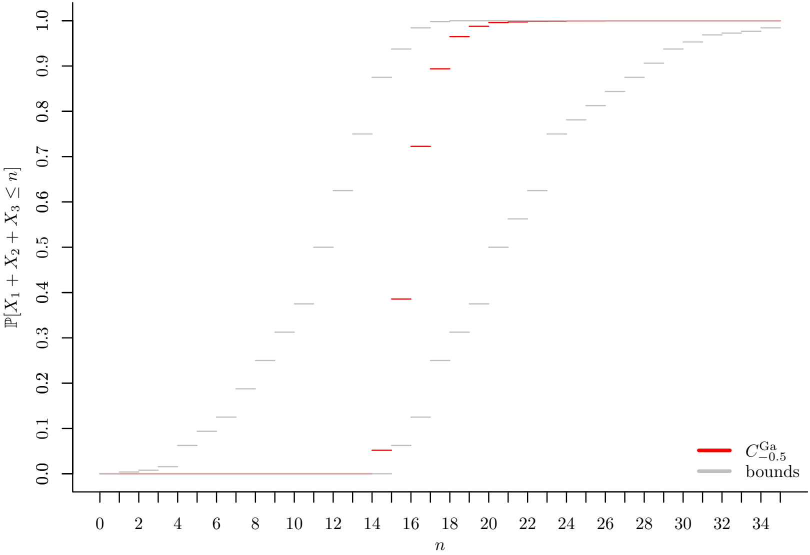
Gaussian copulas



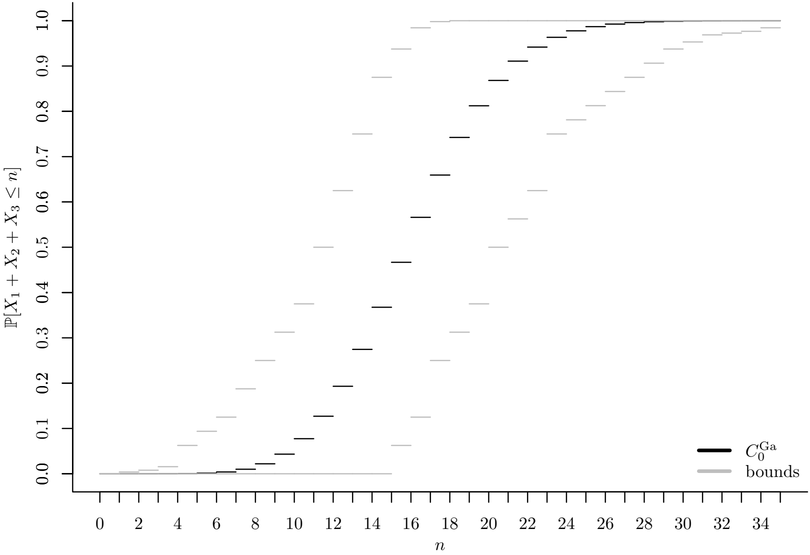
Gaussian copulas



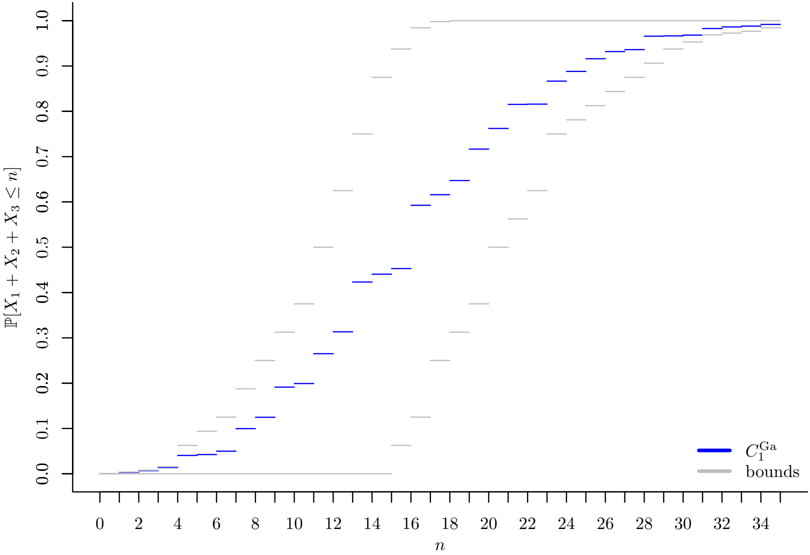
Gaussian copulas



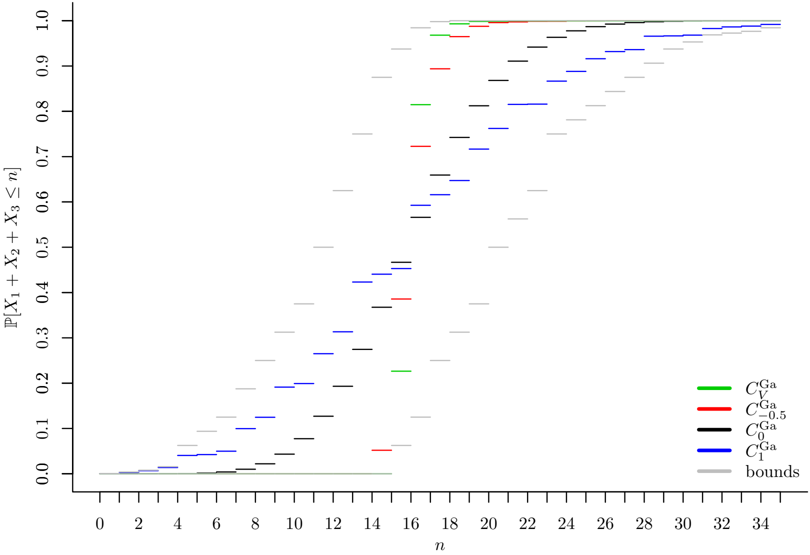
Gaussian copulas



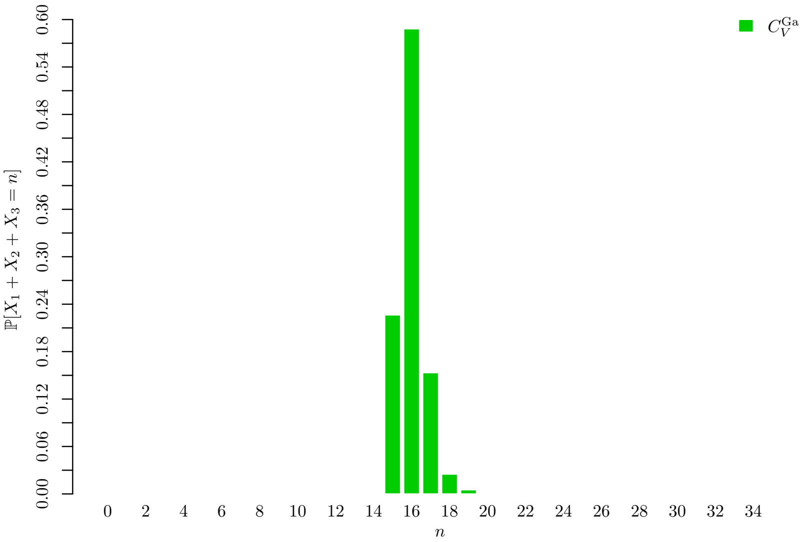
Gaussian copulas



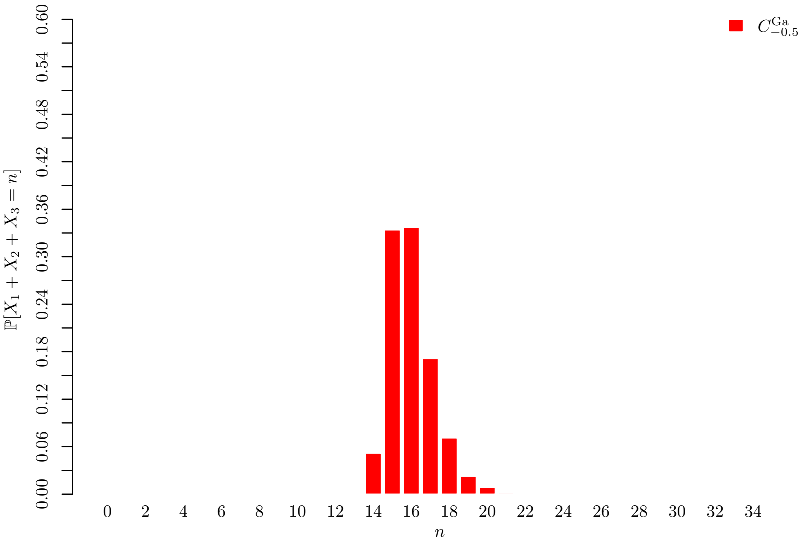
Gaussian copulas



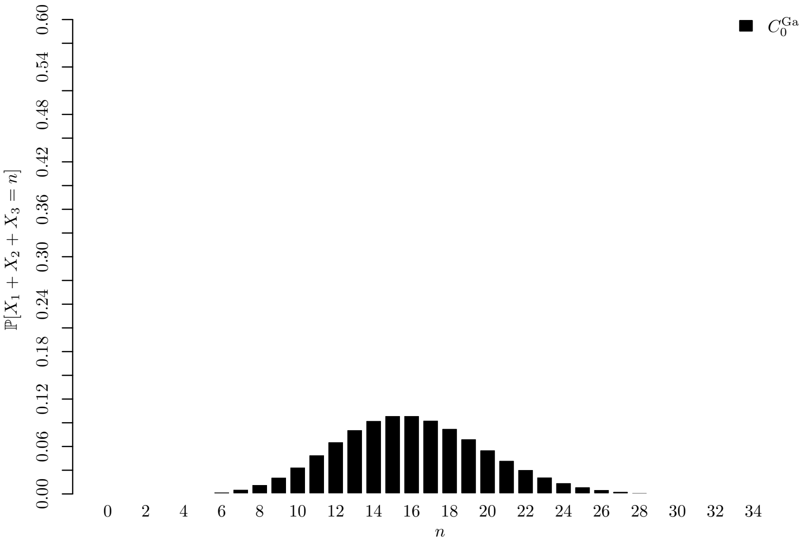
Gaussian copulas



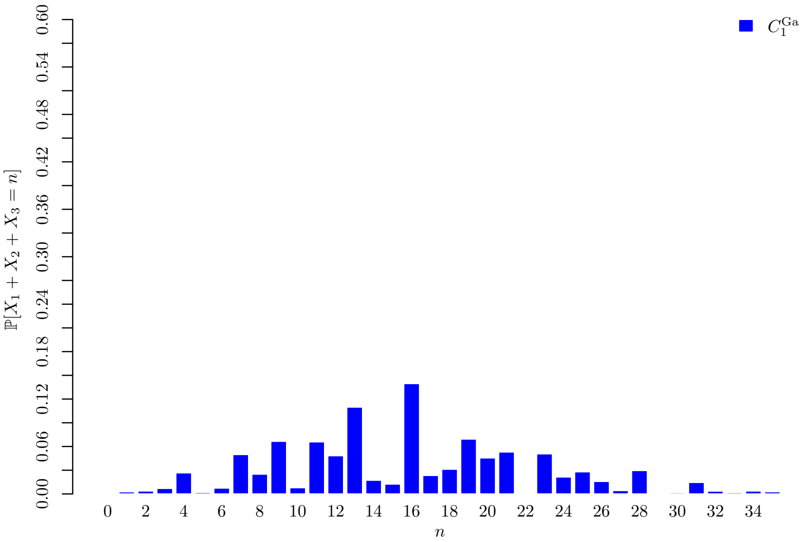
Gaussian copulas



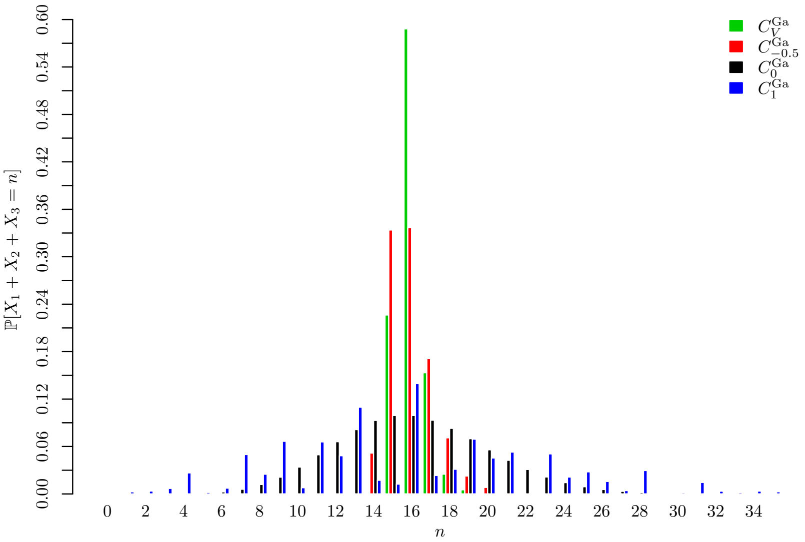
Gaussian copulas



Gaussian copulas



Gaussian copulas

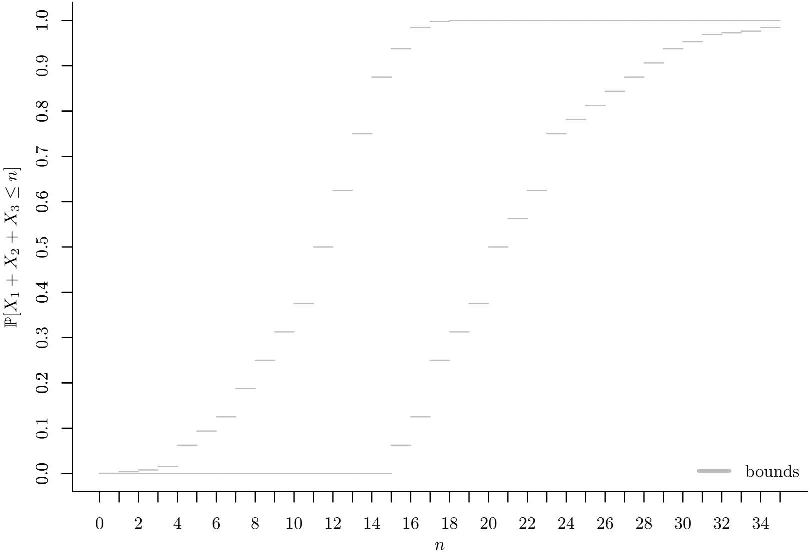


t-copulas

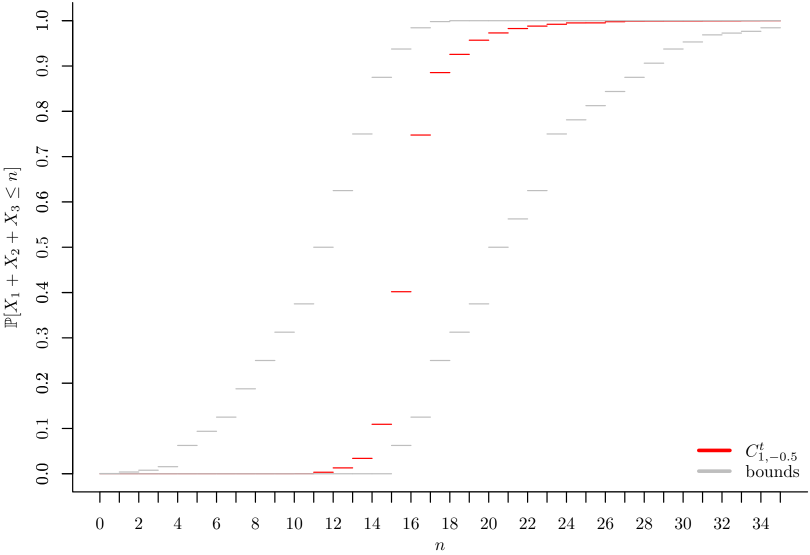
Note for the following plots:

- By a t-copula with parameters ν and ρ, i.e. $C^t_{\nu,\rho}$, we denote a t-copula with ν degrees of freedom and dispersion matrix with 1 in the diagonal and all other entries coinciding to $\rho \in [-1, 1]$.

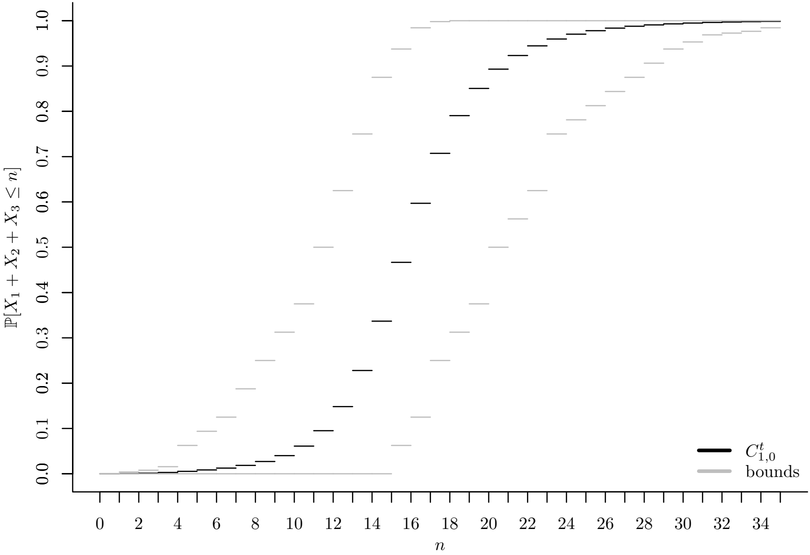
t-copulas


$\mathbb{P}[X_1 + X_2 + X_3 \leq n]$
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
n
bounds

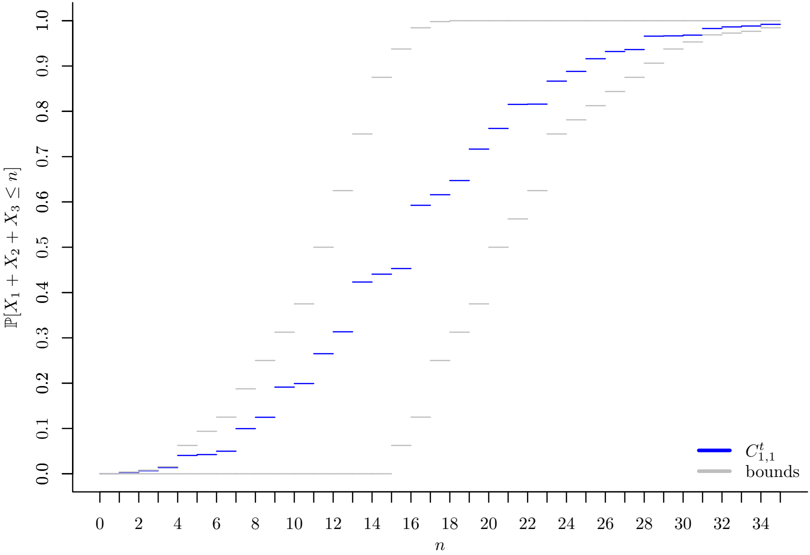
t-copulas


$\mathbb{P}[X_1 + X_2 + X_3 \leq n]$
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
n
$C^t_{1,-0.5}$
bounds

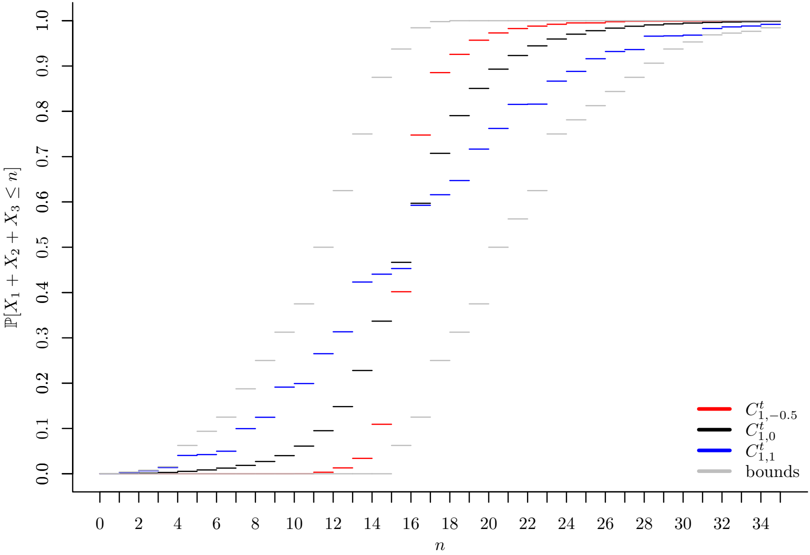
t-copulas


$\mathbb{P}[X_1 + X_2 + X_3 \leq n]$
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
n
$C^t_{1,0}$
bounds

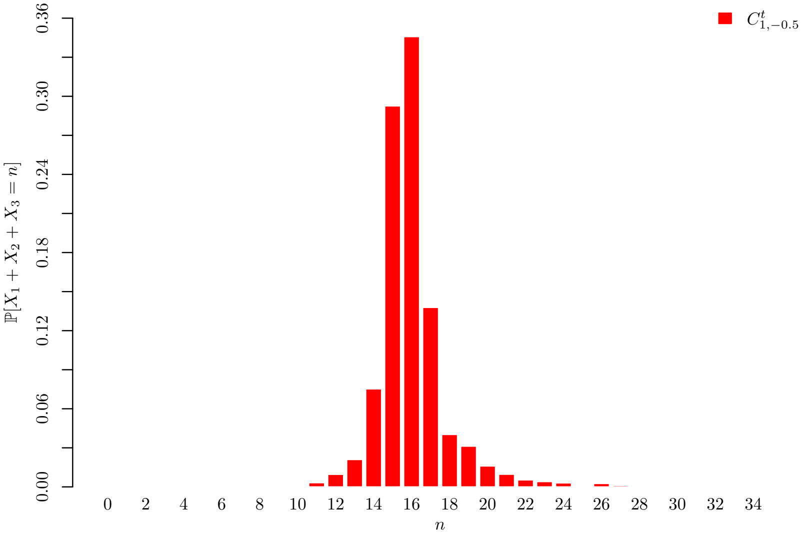
t-copulas


$\mathbb{P}[X_1 + X_2 + X_3 \leq n]$
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
n
$C^t_{1,1}$
bounds

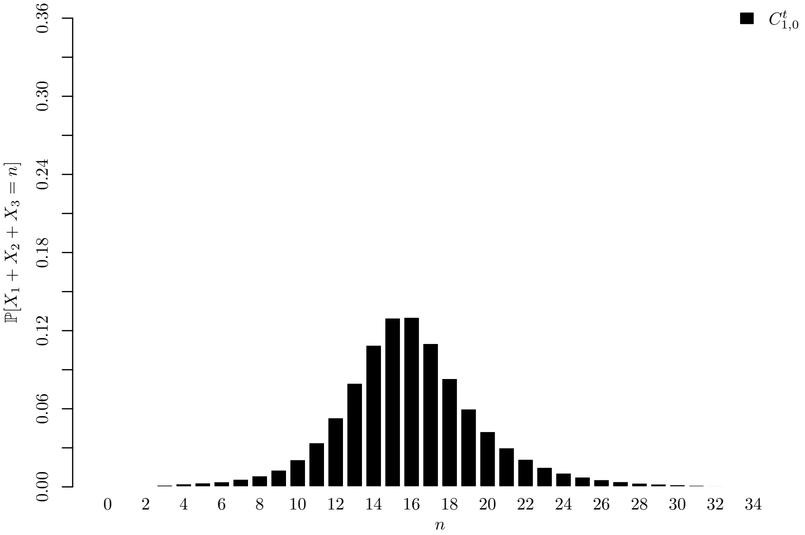
t-copulas


$\mathbb{P}[X_1 + X_2 + X_3 \leq n]$
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34
n
$C^t_{1,-0.5}$
$C^t_{1,0}$
$C^t_{1,1}$
bounds

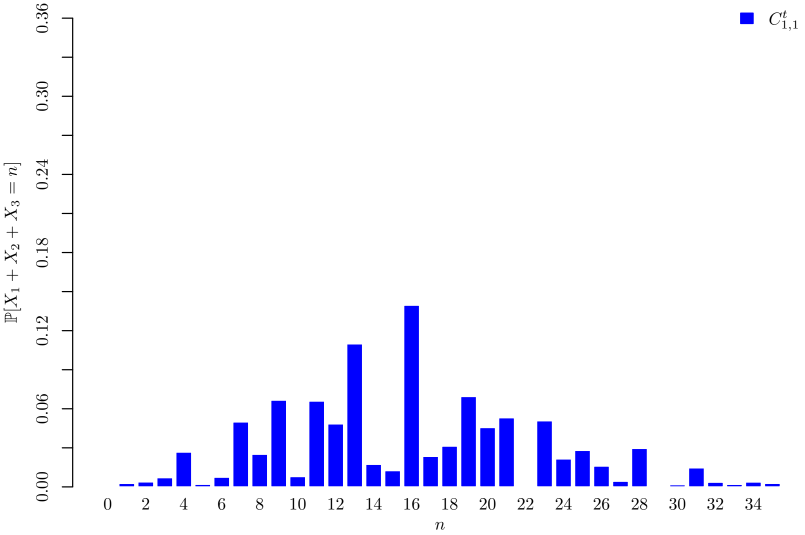
t-copulas


$C^t_{1,-0.5}$
$\mathbb{P}[X_1 + X_2 + X_3 = n]$
0.00
0.06
0.12
0.18
0.24
0.30
0.36
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34
n

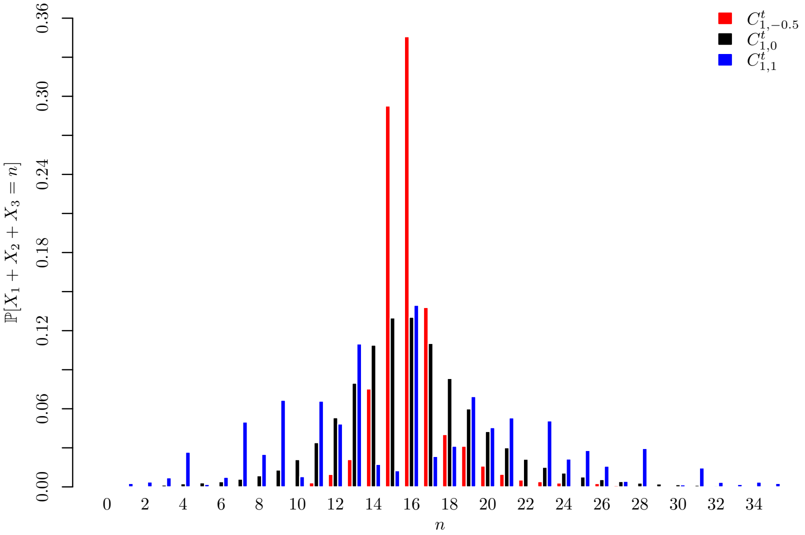
t-copulas


$C^t_{1,0}$
$\mathbb{P}[X_1 + X_2 + X_3 = n]$
0.00
0.06
0.12
0.18
0.24
0.30
0.36
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
n

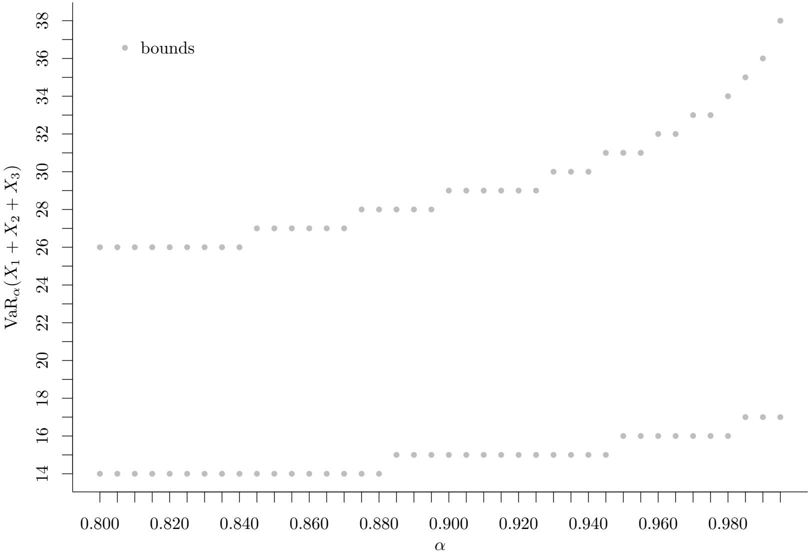
t-copulas


$C^t_{1,1}$
$\mathbb{P}[X_1 + X_2 + X_3 = n]$
0.00
0.06
0.12
0.18
0.24
0.30
0.36
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
n

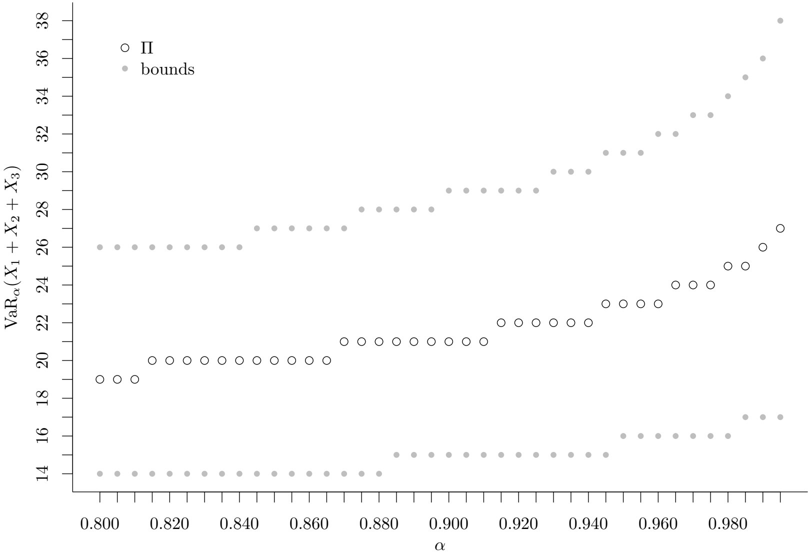
t-copulas


$C^t_{1,-0.5}$
$C^t_{1,0}$
$C^t_{1,1}$
$\mathbb{P}[X_1 + X_2 + X_3 = n]$
0.00
0.06
0.12
0.18
0.24
0.30
0.36
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
n

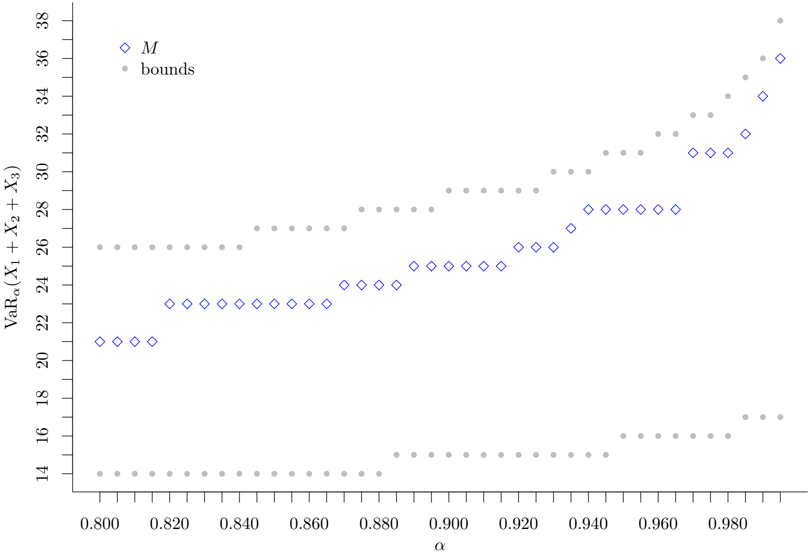
Fundamental copulas – Value-at-Risk



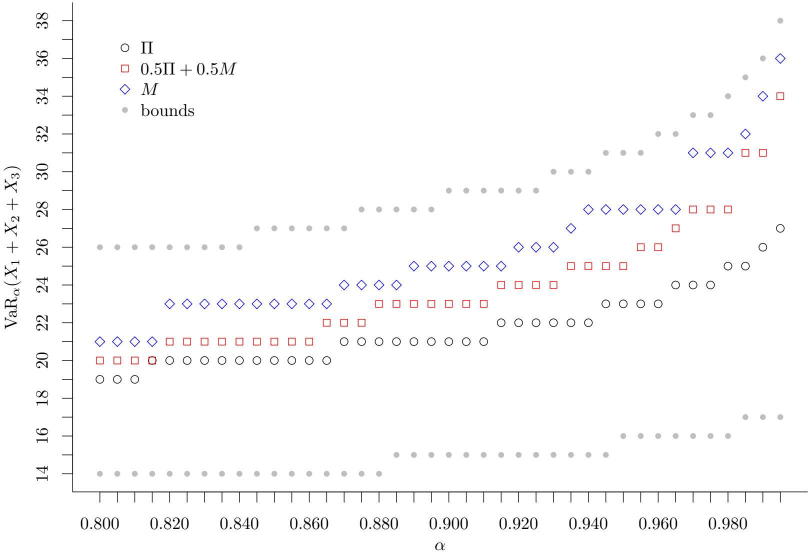
Fundamental copulas – Value-at-Risk



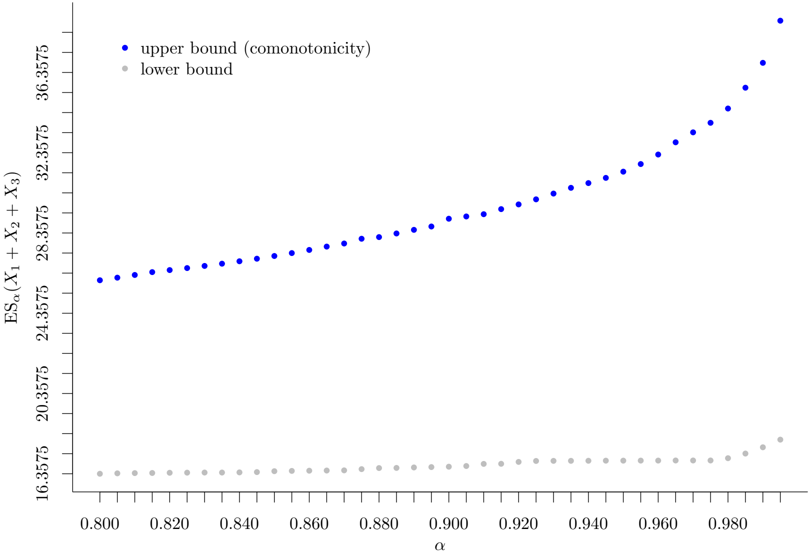
Fundamental copulas – Value-at-Risk



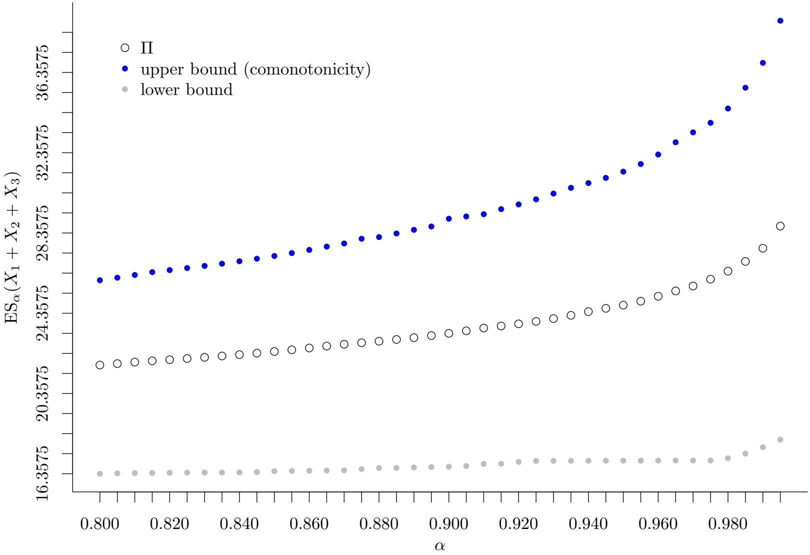
Fundamental copulas – Value-at-Risk



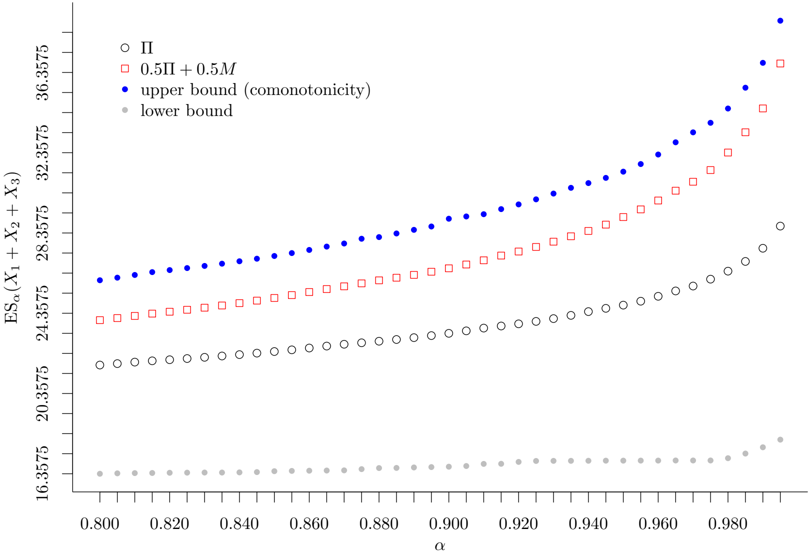
Fundamental copulas – Expected Shortfall



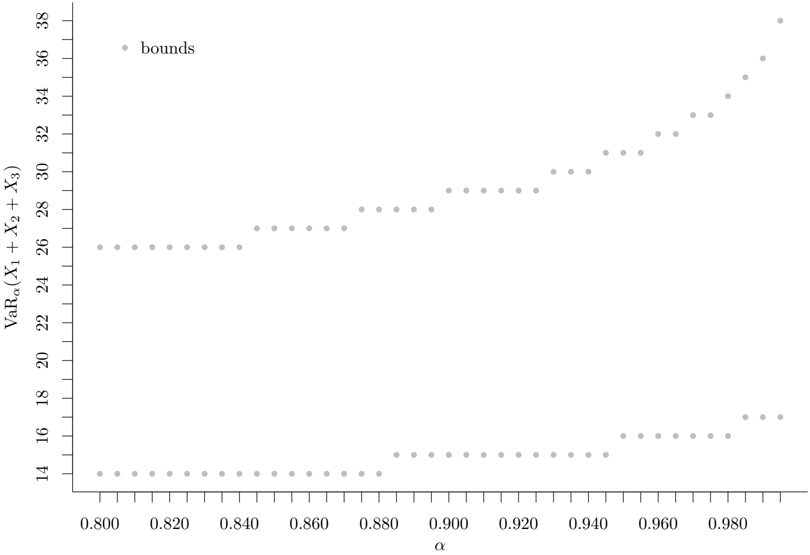
Fundamental copulas – Expected Shortfall



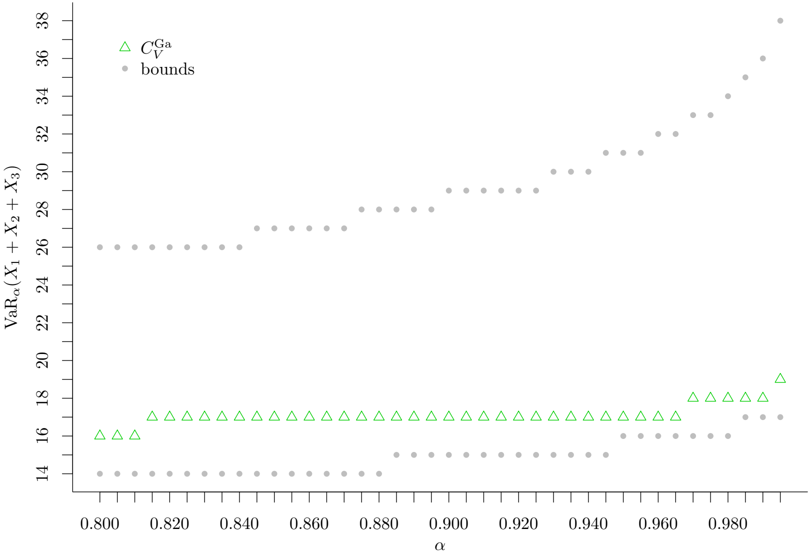
Fundamental copulas – Expected Shortfall



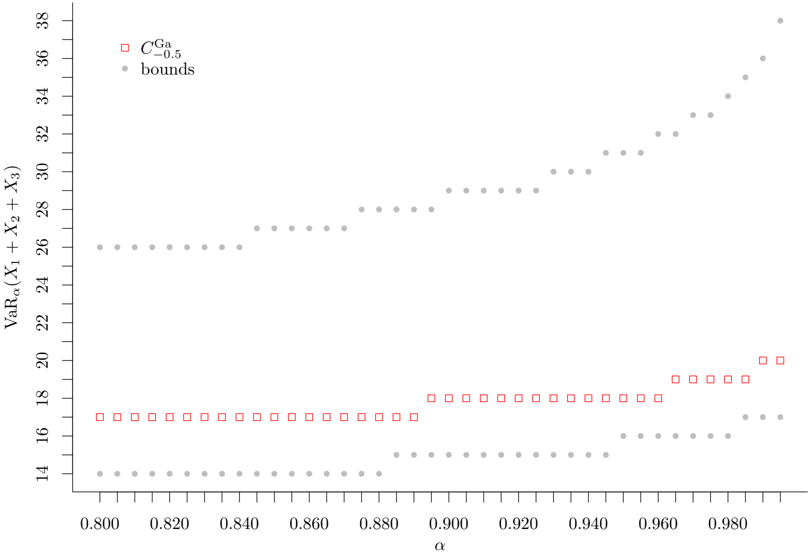
Gaussian copulas – Value-at-Risk



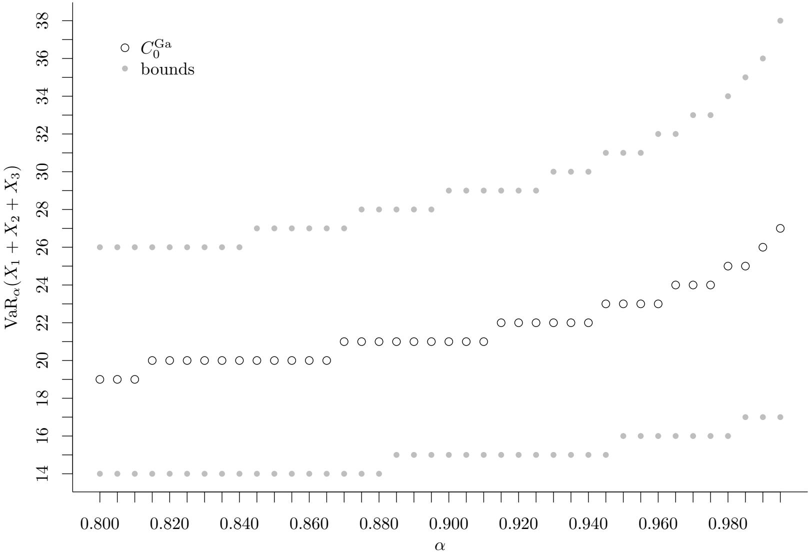
Gaussian copulas – Value-at-Risk



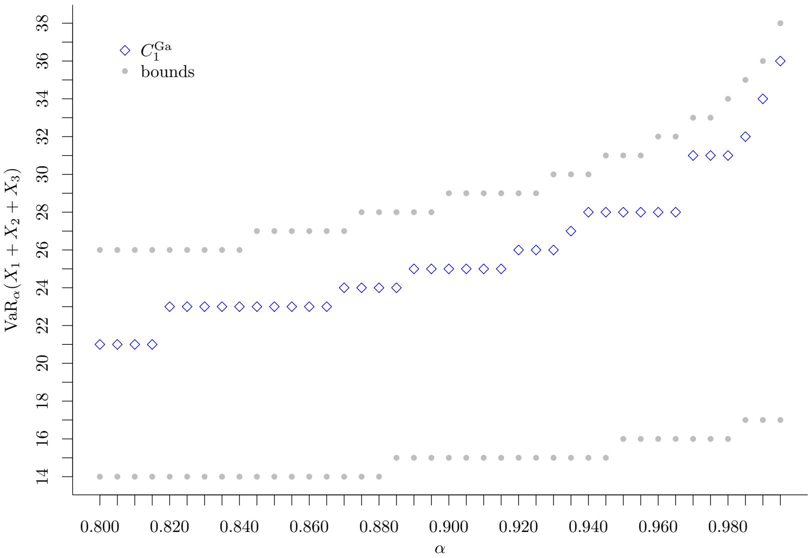
Gaussian copulas – Value-at-Risk



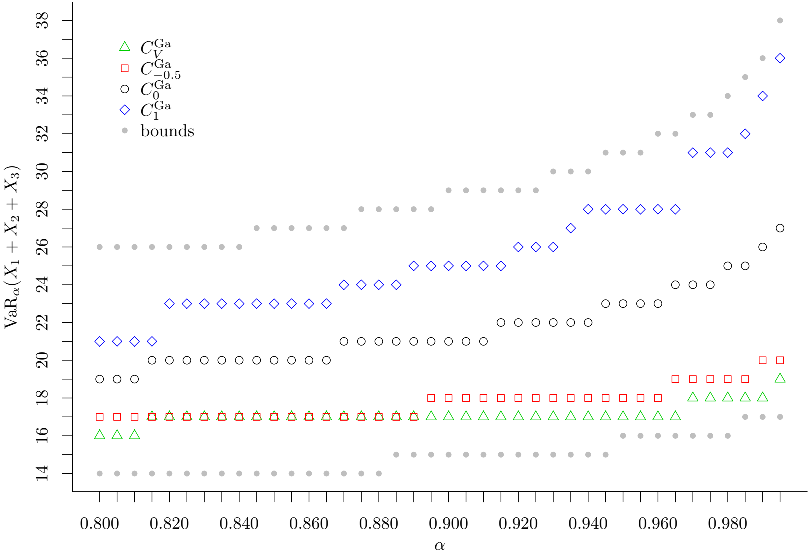
Gaussian copulas – Value-at-Risk



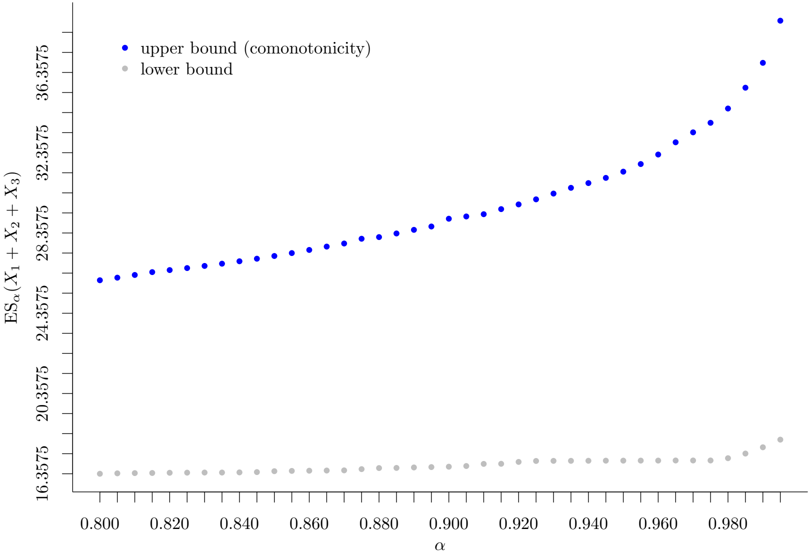
Gaussian copulas – Value-at-Risk



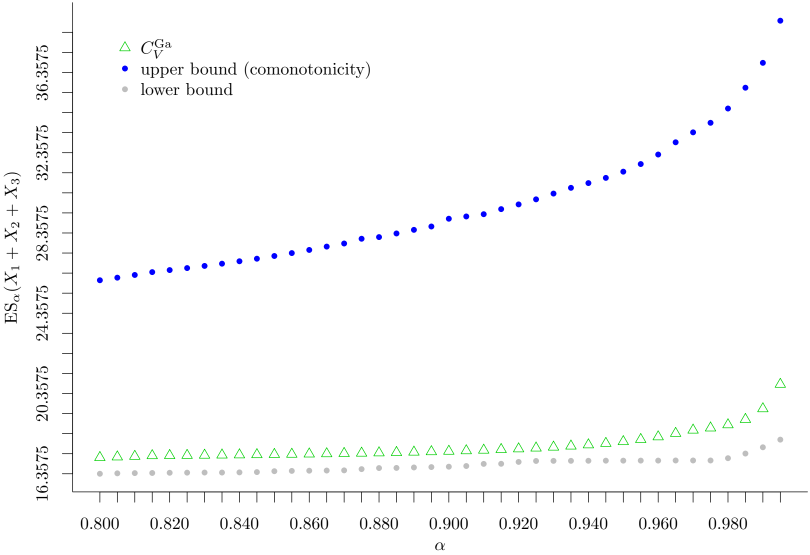
Gaussian copulas – Value-at-Risk



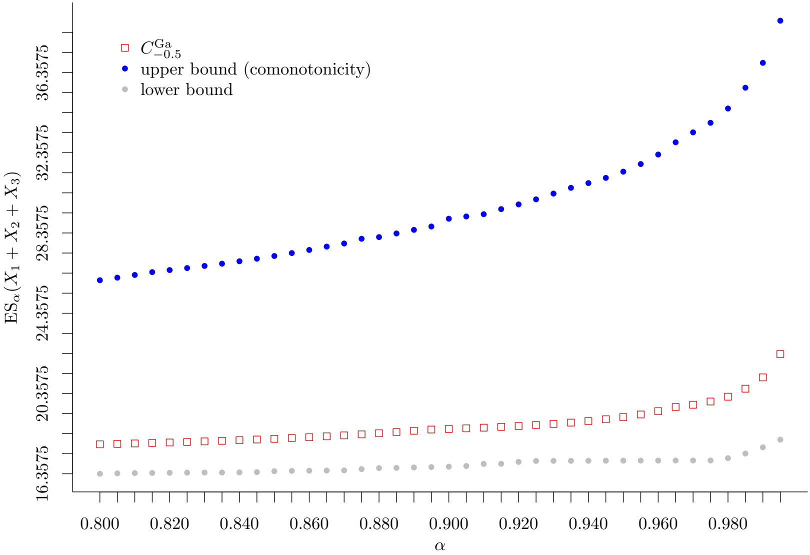
Gaussian copulas – Expected Shortfall



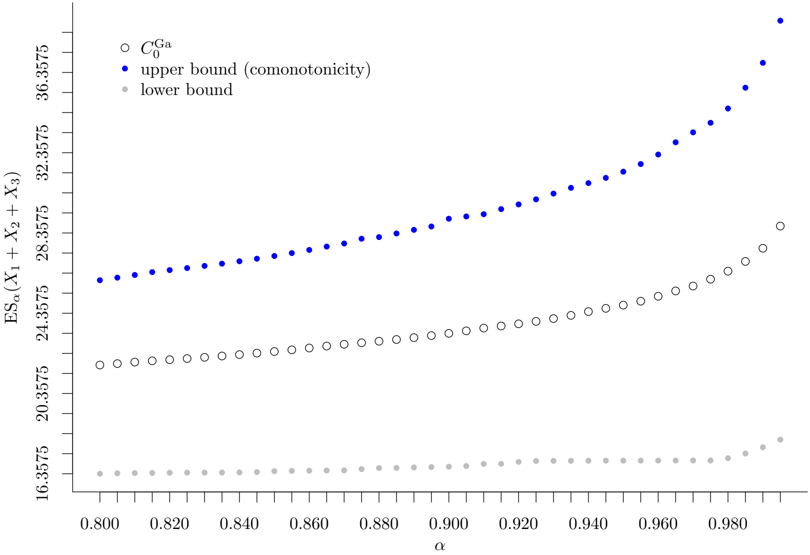
Gaussian copulas – Expected Shortfall



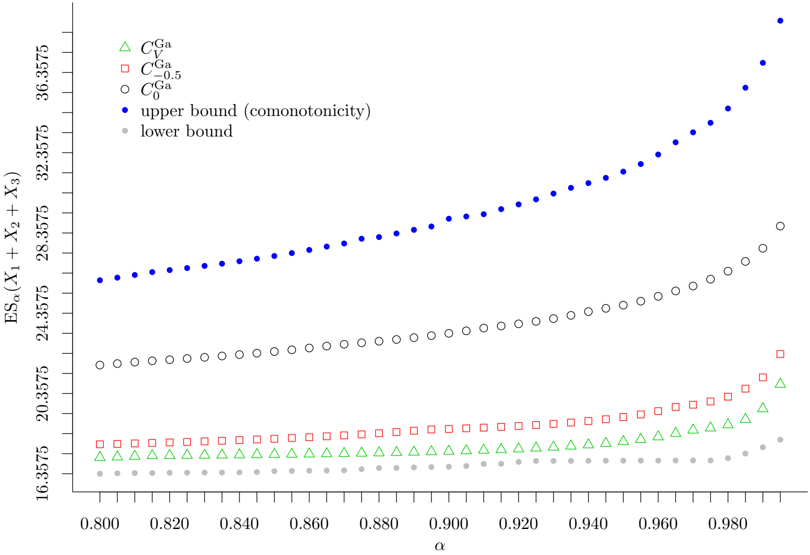
Gaussian copulas – Expected Shortfall



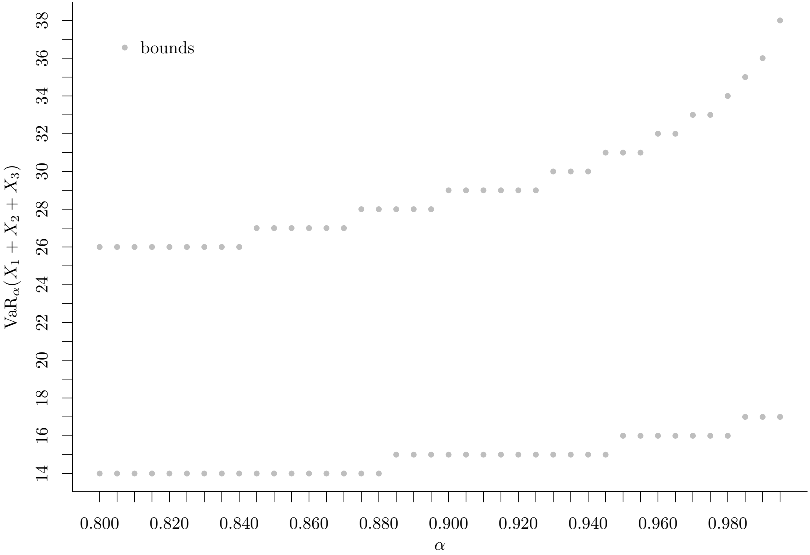
Gaussian copulas – Expected Shortfall



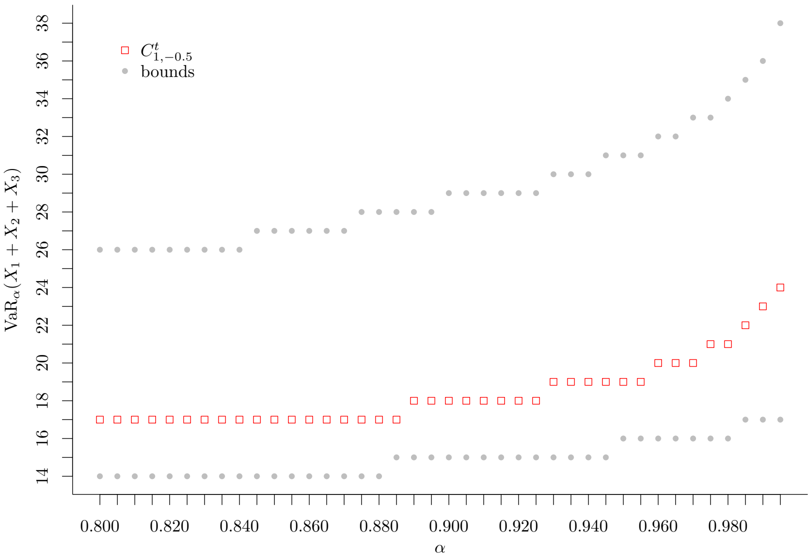
Gaussian copulas – Expected Shortfall



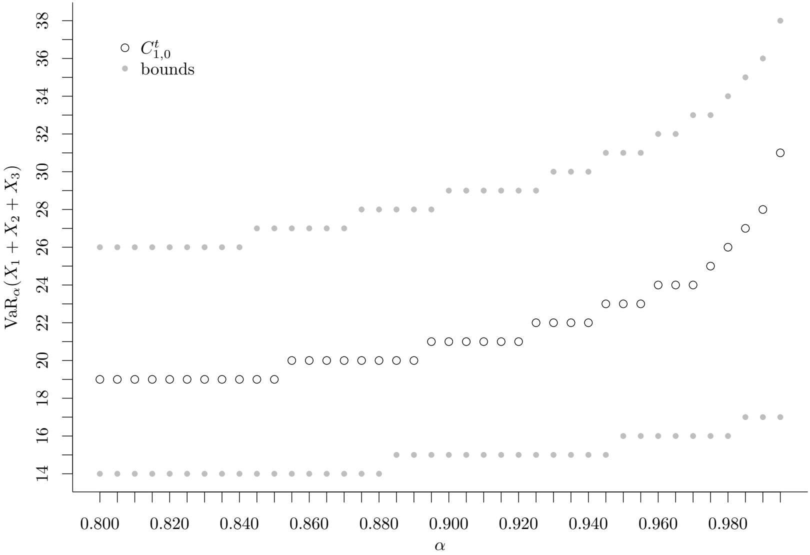
t-copulas – Value-at-Risk



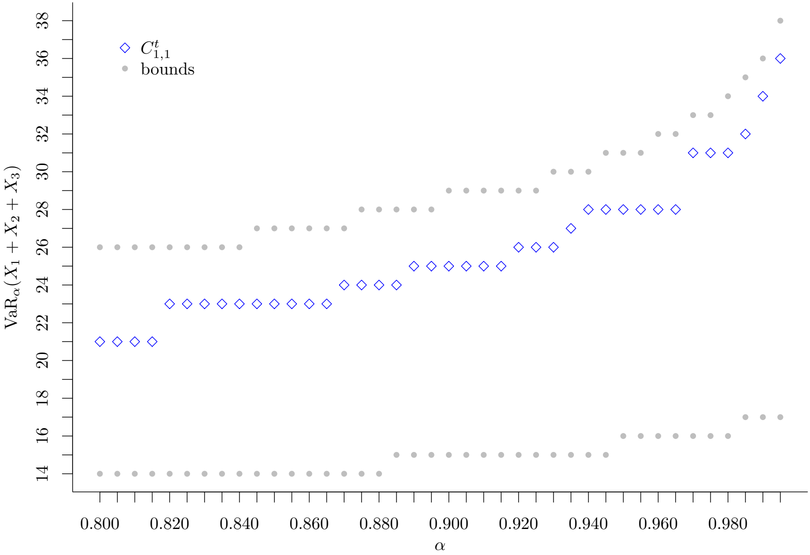
t-copulas – Value-at-Risk



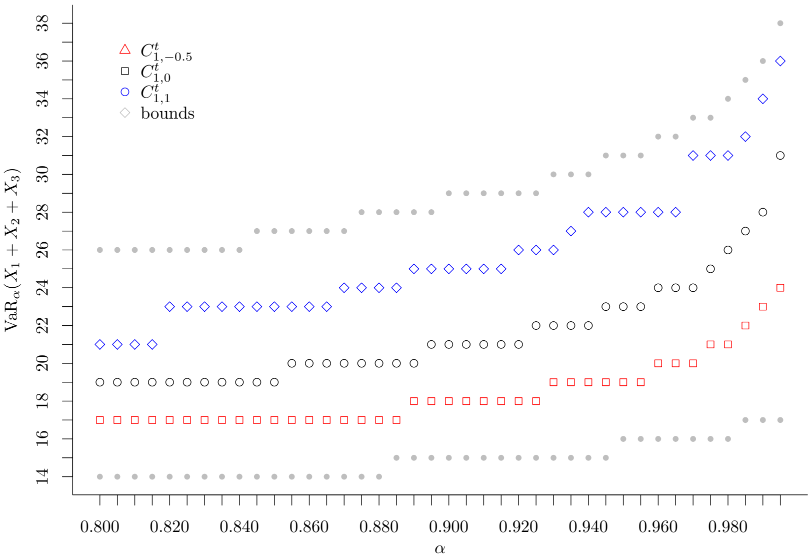
t-copulas – Value-at-Risk



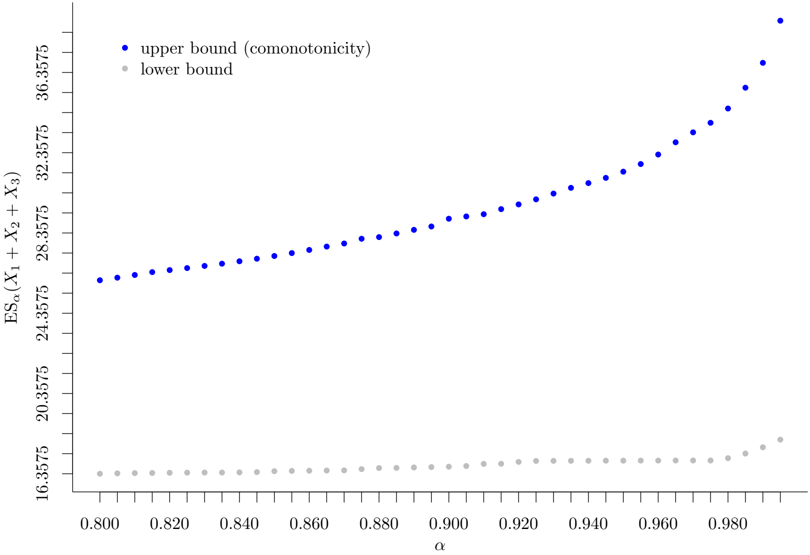
t-copulas – Value-at-Risk



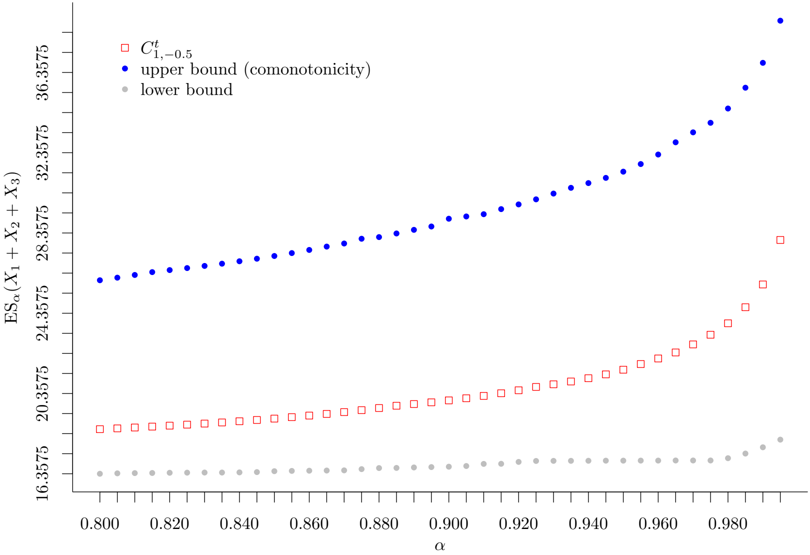
t-copulas – Value-at-Risk



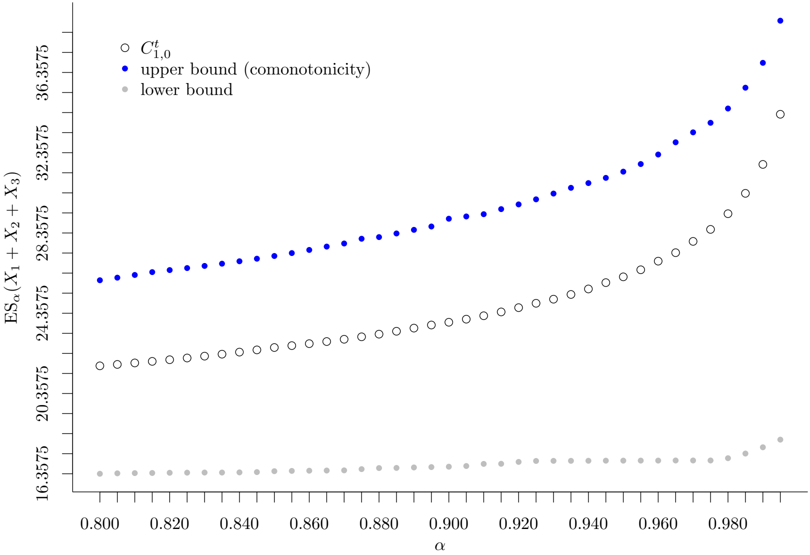
t-copulas – Expected Shortfall



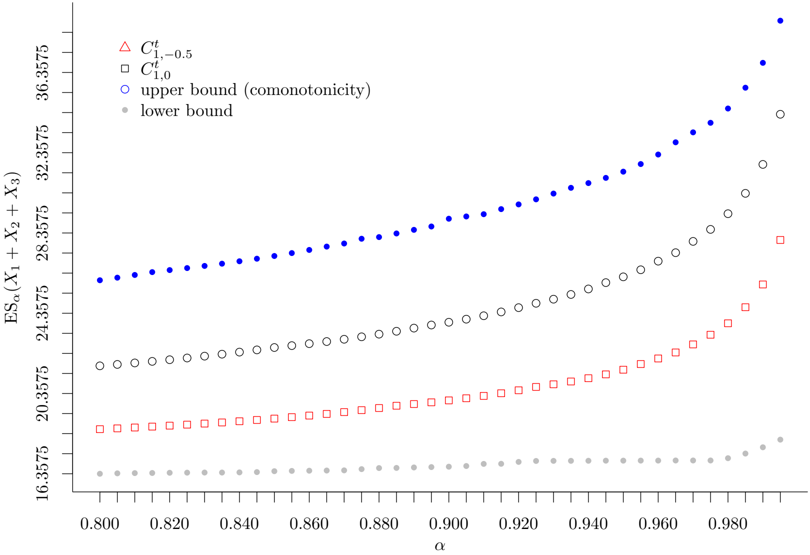
t-copulas – Expected Shortfall



t-copulas – Expected Shortfall



t-copulas – Expected Shortfall



References

- Durante, F., and Sempi, C., *Principes of Copula Theory.* CRC Press, 2015.
- Genest, C., and Nešlehová, J., *A primer on copulas for count data.* ASTIN Bulletin 37, 2 (2007), 475–515.
- Gijbels, I., and Herrmann, K., On the distribution of sums of random variables with copula-induced dependence. IME 59, C (2014), 27–44.
- Joe, H., *Dependence Modeling with Copulas.* CRC Press, 2014.
- McNeil, A. J., Frey, R., and Embrechts, P., *Quantitative Risk Management.* Princeton University Press, 2005.
- Puccetti, G., and Rüschendorf, L., *Computation of sharp bounds on the distribution of a function of dependent risks.* J. Comput. Appl. Math 236, 7 (2012), 1833–1840.