Course 2023-2024 in Financial Risk Management
Lecture 9. Copulas and Extreme Value Theory

Thierry Roncalli*

*Amundi Asset Management[1]

*University of Paris-Saclay

September 2023

[1] The opinions expressed in this presentation are those of the authors and are not meant to represent the opinions or official positions of Amundi Asset Management.

General information

1. **Overview**
   The objective of this course is to understand the theoretical and practical aspects of risk management
2. **Prerequisites**
   M1 Finance or equivalent
3. **ECTS**
   4
4. **Keywords**
   Finance, Risk Management, Applied Mathematics, Statistics
5. **Hours**
   Lectures: 36h, Training sessions: 15h, HomeWork: 30h
6. **Evaluation**
   There will be a final three-hour exam, which is made up of questions and exercises
7. **Course website**
   `http://www.thierry-roncalli.com/RiskManagement.html`

Objective of the course

The objective of the course is twofold:

1. knowing and understanding the financial regulation (banking and others) and the international standards (especially the Basel Accords)
2. being proficient in risk measurement, including the mathematical tools and risk models

Class schedule

Course sessions

- September 15 (6 hours, AM+PM)
- September 22 (6 hours, AM+PM)
- September 19 (6 hours, AM+PM)
- October 6 (6 hours, AM+PM)
- October 13 (6 hours, AM+PM)
- October 27 (6 hours, AM+PM)

Tutorial sessions

- October 20 (3 hours, AM)
- October 20 (3 hours, PM)
- November 10 (3 hours, AM)
- November 10 (3 hours, PM)
- November 17 (3 hours, PM)

Class times: Fridays 9:00am-12:00pm, 1:00pm–4:00pm, University of Evry, Room 209 IDF

Agenda

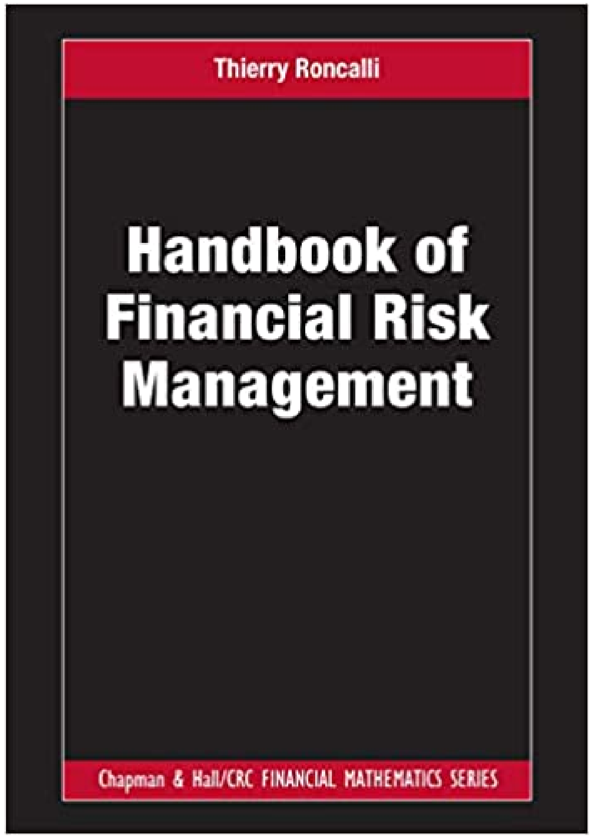
- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- Lecture 9: Copulas and Extreme Value Theory
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

Agenda

- Tutorial Session 1: Market Risk
- Tutorial Session 2: Credit Risk
- Tutorial Session 3: Counterparty Credit Risk and Collateral Risk
- Tutorial Session 4: Operational Risk & Asset Liability Management Risk
- Tutorial Session 5: Copulas, EVT & Stress Testing

Textbook

- Roncalli, T. (2020), *Handbook of Financial Risk Management*, Chapman & Hall/CRC Financial Mathematics Series.



Additional materials

- Slides, tutorial exercises and past exams can be downloaded at the following address:

  `http://www.thierry-roncalli.com/RiskManagement.html`

- Solutions of exercises can be found in the companion book, which can be downloaded at the following address:

  `http://www.thierry-roncalli.com/RiskManagementBook.html`

Agenda

- Lecture 1: Introduction to Financial Risk Management
- Lecture 2: Market Risk
- Lecture 3: Credit Risk
- Lecture 4: Counterparty Credit Risk and Collateral Risk
- Lecture 5: Operational Risk
- Lecture 6: Liquidity Risk
- Lecture 7: Asset Liability Management Risk
- Lecture 8: Model Risk
- **Lecture 9: Copulas and Extreme Value Theory**
- Lecture 10: Monte Carlo Simulation Methods
- Lecture 11: Stress Testing and Scenario Analysis
- Lecture 12: Credit Scoring Models

Sklar's theorem

A bi-dimensional copula is a function $\mathbf{C}$ which satisfies the following properties:

1. $\operatorname{Dom} \mathbf{C} = [0,1] \times [0,1]$
2. $\mathbf{C}(0,u) = \mathbf{C}(u,0) = 0$ and $\mathbf{C}(1,u) = \mathbf{C}(u,1) = u$ for all u in $[0,1]$
3. $\mathbf{C}$ is 2-increasing:

$$\mathbf{C}(v_1,v_2) - \mathbf{C}(v_1,u_2) - \mathbf{C}(u_1,v_2) + \mathbf{C}(u_1,u_2) \geq 0$$

for all $(u_1,u_2) \in [0,1]^2$, $(v_1,v_2) \in [0,1]^2$ such that $0 \leq u_1 \leq v_1 \leq 1$ and $0 \leq u_2 \leq v_2 \leq 1$

Remark

This means that $\mathbf{C}$ is a cumulative distribution function with uniform marginals:

$$\mathbf{C}(u_1,u_2) = \Pr\{U_1 \leq u_1, U_2 \leq u_2\}$$

where U_1 and U_2 are two uniform random variables

Sklar's theorem

We consider the function $\mathbf{C}^{\perp}(u_1, u_2) = u_1 u_2$. We have:

- $\mathbf{C}^{\perp}(0, u) = \mathbf{C}^{\perp}(u, 0) = 0$
- $\mathbf{C}^{\perp}(1, u) = \mathbf{C}^{\perp}(u, 1) = u$
- Since we have $v_2 - u_2 \geq 0$ and $v_1 \geq u_1$, it follows that $v_1 (v_2 - u_2) \geq u_1 (v_2 - u_2)$ and :

$$v_1 v_2 + u_1 u_2 - u_1 v_2 - v_1 u_2 \geq 0$$

$\Rightarrow$ $\mathbf{C}^{\perp}$ is a copula function and is called the product copula

Multivariate probability distribution with given marginals

Let $\mathbf{F}_1$ and $\mathbf{F}_2$ be two univariate distributions.
$\mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$ is a probability distribution with marginals $\mathbf{F}_1$ and $\mathbf{F}_2$:

- $u_i = \mathbf{F}_i(x_i)$ defines a uniform transformation ($u_i \in [0, 1]$)
- $\mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(\infty)) = \mathbf{C}(\mathbf{F}_1(x_1), 1) = \mathbf{F}_1(x_1)$

Sklar also shows that:

- Any bivariate distribution $\mathbf{F}$ admits a copula representation:

$$\mathbf{F}(x_1, x_2) = \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$$

- The copula $\mathbf{C}$ is unique if the marginals are continuous

Multivariate probability distribution with given marginals

The Gumbel logistic distribution is equal to:

$$\mathbf{F}(x_1, x_2) = \left(1 + e^{-x_1} + e^{-x_2}\right)^{-1}$$

We have:

$$\mathbf{F}_1(x_1) \equiv \mathbf{F}(x_1, \infty) = \left(1 + e^{-x_1}\right)^{-1}$$

and $\mathbf{F}_2(x_2) \equiv \left(1 + e^{-x_2}\right)^{-1}$. The quantile functions are then:

$$\mathbf{F}_1^{-1}(u_1) = \ln u_1 - \ln(1 - u_1)$$

and $\mathbf{F}_2^{-1}(u_2) = \ln u_2 - \ln(1 - u_2)$. We finally deduce that:

$$\mathbf{C}(u_1, u_2) = \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right) = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}$$

is the Gumbel logistic copula

Expression of the copula density function

If the joint distribution function $\mathbf{F}(x_1, x_2)$ is absolutely continuous, we obtain:

$$\begin{aligned} f(x_1, x_2) &= \partial_{1,2}\, \mathbf{F}(x_1, x_2) \\ &= \partial_{1,2}\, \mathbf{C}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \\ &= c(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \end{aligned}$$

where $f(x_1, x_2)$ is the joint probability density function, f_1 and f_2 are the marginal densities and c is the copula density:

$$c(u_1, u_2) = \partial_{1,2}\, \mathbf{C}(u_1, u_2)$$

Remark

The condition $\mathbf{C}(v_1, v_2) - \mathbf{C}(v_1, u_2) - \mathbf{C}(u_1, v_2) + \mathbf{C}(u_1, u_2) \geq 0$ is equivalent to $\partial_{1,2}\, \mathbf{C}(u_1, u_2) \geq 0$ when the copula density exists.

Expression of the copula density function

In the case of the Gumbel logistic copula, we have:

$$\mathbf{C}(u_1, u_2) = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}$$

and:

$$c(u_1, u_2) = \frac{2 u_1 u_2}{(u_1 + u_2 - u_1 u_2)^3}$$

Expression of the copula density function

We deduce that:

$$c(u_1, u_2) = \frac{f\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)}{f_1\left(\mathbf{F}_1^{-1}(u_1)\right) \cdot f_2\left(\mathbf{F}_2^{-1}(u_2)\right)}$$

If we consider the Normal copula, we have:

$$\mathbf{C}(u_1, u_2; \rho) = \Phi\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho\right)$$

and:

$$\begin{aligned} c(u_1, u_2; \rho) &= \frac{2\pi\left(1-\rho^2\right)^{-1/2} \exp\left(-\frac{1}{2(1-\rho^2)}\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right)\right)}{(2\pi)^{-1/2} \exp\left(-\frac{1}{2}x_1^2\right) \cdot (2\pi)^{-1/2} \exp\left(-\frac{1}{2}x_2^2\right)} \\ &= \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\frac{\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right)}{(1-\rho^2)} + \frac{1}{2}\left(x_1^2 + x_2^2\right)\right) \end{aligned}$$

where $x_1 = \Phi_1^{-1}(u_1)$ and $x_2 = \Phi_2^{-1}(u_2)$

Expression of the copula density function

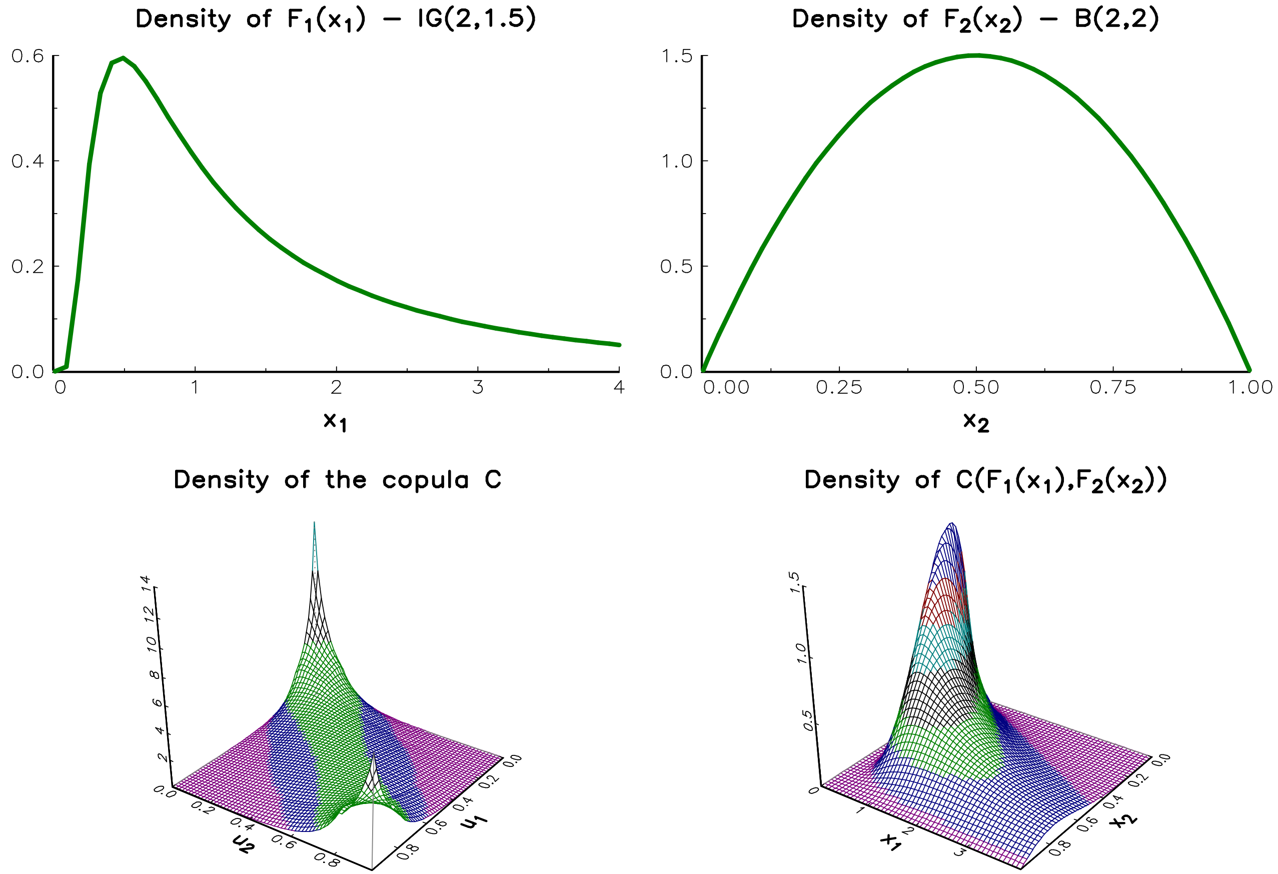


Figure: Construction of a bivariate probability distribution with given marginals and the Normal copula

Concordance ordering

Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be two copula functions. We say that the copula $\mathbf{C}_1$ is smaller than the copula $\mathbf{C}_2$ and we note $\mathbf{C}_1 \prec \mathbf{C}_2$ if we have:

$$\mathbf{C}_1(u_1, u_2) \leq \mathbf{C}_2(u_1, u_2)$$

for all $(u_1, u_2) \in [0, 1]^2$

Let $\mathbf{C}_\theta(u_1, u_2) = \mathbf{C}(u_1, u_2; \theta)$ be a family of copula functions that depends on the parameter θ. The copula family $\{\mathbf{C}_\theta\}$ is totally ordered if, for all $\theta_2 \geq \theta_1$, $\mathbf{C}_{\theta_2} \succ \mathbf{C}_{\theta_1}$ (positively ordered) or $\mathbf{C}_{\theta_2} \prec \mathbf{C}_{\theta_1}$ (negatively ordered)

Remark

The Normal copula family is positively ordered

Fréchet bounds

We have:

$$\mathbf{C}^{-} \prec \mathbf{C} \prec \mathbf{C}^{+}$$

where:

$$\mathbf{C}^{-}(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$$

and:

$$\mathbf{C}^{+}(u_1, u_2) = \min(u_1, u_2)$$

The multivariate case

The canonical decomposition of a multivariate distribution function is:

$$\mathbf{F}(x_1, \ldots, x_n) = \mathbf{C}(\mathbf{F}_1(x_1), \ldots, \mathbf{F}_n(x_n))$$

We have:

$$\mathbf{C}^- \prec \mathbf{C} \prec \mathbf{C}^+$$

where:

$$\mathbf{C}^-(u_1, \ldots, u_n) = \max\left(\sum_{i=1}^{n} u_i - n + 1, 0\right)$$

and:

$$\mathbf{C}^+(u_1, \ldots, u_n) = \min(u_1, \ldots, u_n)$$

Remark

$\mathbf{C}^-$ is not a copula when $n \geq 3$

Countermonotonicity and comonotonicity

Let $X = (X_1, X_2)$ be a random vector with distribution $\mathbf{F}$. We define the copula of (X_1, X_2) by the copula of $\mathbf{F}$:

$$\mathbf{F}(x_1, x_2) = \mathbf{C}\langle X_1, X_2\rangle (\mathbf{F}_1(x_1), \mathbf{F}_2(x_2))$$

Definition

- X_1 and X_2 are countermonotonic – or $\mathbf{C}\langle X_1, X_2\rangle = \mathbf{C}^-$ – if there exists a random variable X such that $X_1 = f_1(X)$ and $X_2 = f_2(X)$ where f_1 and f_2 are respectively decreasing and increasing functions. In this case, $X_2 = f(X_1)$ where $f = f_2 \circ f_1^{-1}$ is a decreasing function
- X_1 and X_2 are independent if the dependence function is the product copula $\mathbf{C}^\perp$
- X_1 are X_2 are comonotonic – or $\mathbf{C}\langle X_1, X_2\rangle = \mathbf{C}^+$ – if there exists a random variable X such that $X_1 = f_1(X)$ and $X_2 = f_2(X)$ where f_1 and f_2 are both increasing functions. In this case, $X_2 = f(X_1)$ where $f = f_2 \circ f_1^{-1}$ is an increasing function

Countermonotonicity and comonotonicity

- We consider a uniform random vector (U_1, U_2):

$$\begin{aligned} \mathbf{C}\langle U_1, U_2\rangle &= \mathbf{C}^- \Leftrightarrow U_2 = 1 - U_1 \\ \mathbf{C}\langle U_1, U_2\rangle &= \mathbf{C}^+ \Leftrightarrow U_2 = U_1 \end{aligned}$$

- We consider a standardized Gaussian random vector (X_1, X_2). We have $U_1 = \Phi(X_1)$ and $U_2 = \Phi(X_2)$. We deduce that:

$$\begin{aligned} \mathbf{C}\langle X_1, X_2\rangle &= \mathbf{C}^- \Leftrightarrow \Phi(X_2) = 1 - \Phi(X_1) \Leftrightarrow X_2 = -X_1 \\ \mathbf{C}\langle X_1, X_2\rangle &= \mathbf{C}^+ \Leftrightarrow \Phi(X_2) = \Phi(X_1) \Leftrightarrow X_2 = X_1 \end{aligned}$$

Countermonotonicity and comonotonicity

- We consider a random vector (X_1, X_2) where $X_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right)$. We have

$$U_i = \Phi\left(\frac{X_i - \mu_i}{\sigma_i}\right)$$

We deduce that:

$$\begin{aligned}
\mathbf{C}\left\langle X_1, X_2\right\rangle = \mathbf{C}^- \quad & \Leftrightarrow \quad \Phi\left(\frac{X_2 - \mu_2}{\sigma_2}\right) = 1 - \Phi\left(\frac{X_1 - \mu_1}{\sigma_1}\right) \\
& \Leftrightarrow \quad \Phi\left(\frac{X_2 - \mu_2}{\sigma_2}\right) = \Phi\left(-\frac{X_1 - \mu_1}{\sigma_1}\right) \\
& \Leftrightarrow \quad X_2 = \left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right) - \frac{\sigma_2}{\sigma_1}X_1
\end{aligned}$$

and:

$$\mathbf{C}\left\langle X_1, X_2\right\rangle = \mathbf{C}^+ \Leftrightarrow X_2 = \left(\mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1\right) + \frac{\sigma_2}{\sigma_1}X_1$$

Countermonotonicity and comonotonicity

- We consider a random vector (X_1, X_2) where $X_i \sim \mathcal{LN}\left(\mu_i, \sigma_i^2\right)$. We have:

$$U_i = \Phi\left(\frac{\ln X_i - \mu_i}{\sigma_i}\right)$$

We deduce that:

$$\begin{aligned}
\mathbf{C}\left\langle X_1, X_2\right\rangle = \mathbf{C}^- \quad & \Leftrightarrow \quad \ln X_2 = \left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right) - \frac{\sigma_2}{\sigma_1}\ln X_1 \\
& \Leftrightarrow \quad X_2 = e^{\left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right)} e^{-\frac{\sigma_2}{\sigma_1}\ln X_1} \\
& \Leftrightarrow \quad X_2 = e^{\left(\mu_2 + \frac{\sigma_2}{\sigma_1}\mu_1\right)} X_1^{-\frac{\sigma_2}{\sigma_1}}
\end{aligned}$$

and:

$$\begin{aligned}
\mathbf{C}\left\langle X_1, X_2\right\rangle = \mathbf{C}^+ \quad & \Leftrightarrow \quad \ln X_2 = \left(\mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1\right) + \frac{\sigma_2}{\sigma_1}\ln X_1 \\
& \Leftrightarrow \quad X_2 = e^{\left(\mu_2 - \frac{\sigma_2}{\sigma_1}\mu_1\right)} X_1^{\frac{\sigma_2}{\sigma_1}}
\end{aligned}$$

Countermonotonicity and comonotonicity

- If $X_1 \sim \mathcal{LN}(0,1)$ and $X_2 \sim \mathcal{LN}(0,1)$, we have:

$$\mathbf{C}\langle X_1, X_2\rangle = \mathbf{C}^- \Leftrightarrow X_2 = \frac{1}{X_1}$$

- If $X_1 \sim \mathcal{LN}\left(0,2^2\right)$ and $X_2 \sim \mathcal{LN}(0,1)$, we have:

$$\mathbf{C}\langle X_1, X_2\rangle = \mathbf{C}^+ \Leftrightarrow X_2 = \sqrt{X_1}$$

Linear dependence vs non-linear dependence

The concepts of counter- and comonotonicity concepts generalize the cases where the linear correlation of a Gaussian vector is equal to -1 or $+1$

Non-linear stochastic dependence

Scale invariance property

If h_1 and h_2 are two increasing functions on $\operatorname{Im} X_1$ and $\operatorname{Im} X_2$, then we have:

$$\mathbf{C}\left\langle h_1\left(X_1\right), h_2\left(X_2\right)\right\rangle = \mathbf{C}\left\langle X_1, X_2\right\rangle$$

Non-linear stochastic dependence

Proof (marginals)

We note $\mathbf{F}$ and $\mathbf{G}$ the probability distributions of the random vectors (X_1, X_2) and $(Y_1, Y_2) = (h_1(X_1), h_2(X_2))$. The marginals of $\mathbf{G}$ are:

$$\begin{aligned}
\mathbf{G}_1(y_1) &= \Pr\{Y_1 \leq y_1\} \\
&= \Pr\{h_1(X_1) \leq y_1\} \\
&= \Pr\left\{X_1 \leq h_1^{-1}(y_1)\right\} \quad \text{(because } h_1 \text{ is strictly increasing)} \\
&= \mathbf{F}_1\left(h_1^{-1}(y_1)\right)
\end{aligned}$$

and $\mathbf{G}_2(y_2) = \mathbf{F}_2\left(h_2^{-1}(y_2)\right)$. We deduce that $\mathbf{G}_1^{-1}(u_1) = h_1\left(\mathbf{F}_1^{-1}(u_1)\right)$ and $\mathbf{G}_2^{-1}(u_2) = h_2\left(\mathbf{F}_2^{-1}(u_2)\right)$

Non-linear stochastic dependence

Proof (copula)

By definition, we have:

$$\mathbf{C}\langle Y_1, Y_2\rangle (u_1, u_2) = \mathbf{G}\left(\mathbf{G}_1^{-1}(u_1), \mathbf{G}_2^{-1}(u_2)\right)$$

Moreover, it follows that:

$$\begin{aligned}
\mathbf{G}\left(\mathbf{G}_1^{-1}(u_1), \mathbf{G}_2^{-1}(u_2)\right) &= \Pr\left\{Y_1 \leq \mathbf{G}_1^{-1}(u_1), Y_2 \leq \mathbf{G}_2^{-1}(u_2)\right\} \\
&= \Pr\left\{h_1(X_1) \leq \mathbf{G}_1^{-1}(u_1), h_2(X_2) \leq \mathbf{G}_2^{-1}(u_2)\right\} \\
&= \Pr\left\{X_1 \leq h_1^{-1}\left(\mathbf{G}_1^{-1}(u_1)\right), X_2 \leq h_2^{-1}\left(\mathbf{G}_2^{-1}(u_2)\right)\right\} \\
&= \Pr\left\{X_1 \leq \mathbf{F}_1^{-1}(u_1), X_2 \leq \mathbf{F}_2^{-1}(u_2)\right\} \\
&= \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)
\end{aligned}$$

Because we have $\mathbf{C}\langle X_1, X_2\rangle (u_1, u_2) = \mathbf{F}\left(\mathbf{F}_1^{-1}(u_1), \mathbf{F}_2^{-1}(u_2)\right)$, we deduce that:

$$\mathbf{C}\langle Y_1, Y_2\rangle = \mathbf{C}\langle X_1, X_2\rangle$$

Non-linear stochastic dependence

We have:

$$\begin{aligned}\mathbf{G}\left(y_1, y_2\right) &= \mathbf{C}\left\langle X_1, X_2\right\rangle \left(\mathbf{G}_1\left(y_1\right), \mathbf{G}_2\left(y_1\right)\right) \\ &= \mathbf{C}\left\langle X_1, X_2\right\rangle \left(\mathbf{F}_1\left(h_1^{-1}\left(y_1\right)\right), \mathbf{F}_2\left(h_2^{-1}\left(y_2\right)\right)\right)\end{aligned}$$

Applying an increasing transformation does not change the copula function, only the marginals

The copula function is the minimum exhaustive statistic of the dependence

Non-linear stochastic dependence

If X_1 and X_2 are two positive random variables, the previous theorem implies that:

$$\begin{aligned}
\mathbf{C}\left\langle X_1, X_2 \right\rangle &= \mathbf{C}\left\langle \ln X_1, X_2 \right\rangle \\
&= \mathbf{C}\left\langle \ln X_1, \ln X_2 \right\rangle \\
&= \mathbf{C}\left\langle X_1, \exp X_2 \right\rangle \\
&= \mathbf{C}\left\langle \sqrt{X_1}, \exp X_2 \right\rangle
\end{aligned}$$

Concordance measures

A numeric measure m of association between X_1 and X_2 is a measure of concordance if it satisfies the following properties:

1. $-1 = m \langle X, -X \rangle \leq m \langle \mathbf{C} \rangle \leq m \langle X, X \rangle = 1$;
2. $m \left\langle \mathbf{C}^{\perp} \right\rangle = 0$;
3. $m \langle -X_1, X_2 \rangle = m \langle X_1, -X_2 \rangle = -m \langle X_1, X_2 \rangle$;
4. if $\mathbf{C}_1 \prec \mathbf{C}_2$, then $m \langle \mathbf{C}_1 \rangle \leq m \langle \mathbf{C}_2 \rangle$;

We have:

$$\mathbf{C} \prec \mathbf{C}^{\perp} \Rightarrow m \langle \mathbf{C} \rangle < 0$$

and:

$$\mathbf{C} \succ \mathbf{C}^{\perp} \Rightarrow m \langle \mathbf{C} \rangle > 0$$

Kendall's tau and Spearman's rho

- Kendall's tau is the probability of concordance minus the probability of discordance:

$$\begin{aligned}\tau &= \Pr\left\{(X_i - X_j)\cdot(Y_i - Y_j) > 0\right\} - \Pr\left\{(X_i - X_j)\cdot(Y_i - Y_j) < 0\right\} \\ &= 4 \iint_{[0,1]^2} \mathbf{C}(u_1, u_2)\,\mathrm{d}\mathbf{C}(u_1, u_2) - 1\end{aligned}$$

- Spearman's rho is the linear correlation of the rank statistics:

$$\begin{aligned}\varrho &= \frac{\operatorname{cov}\left(\mathbf{F}_X(X), \mathbf{F}_Y(Y)\right)}{\sigma\left(\mathbf{F}_X(X)\right)\cdot\sigma\left(\mathbf{F}_Y(Y)\right)} \\ &= 12 \iint_{[0,1]^2} u_1 u_2\,\mathrm{d}\mathbf{C}(u_1, u_2) - 3\end{aligned}$$

- For the normal copula, we have:

$$\tau = \frac{2}{\pi}\arcsin\rho \qquad \text{and} \qquad \varrho = \frac{6}{\pi}\arcsin\frac{\rho}{2}$$

Exhaustive vs non-exhaustive statistics of stochastic dependence

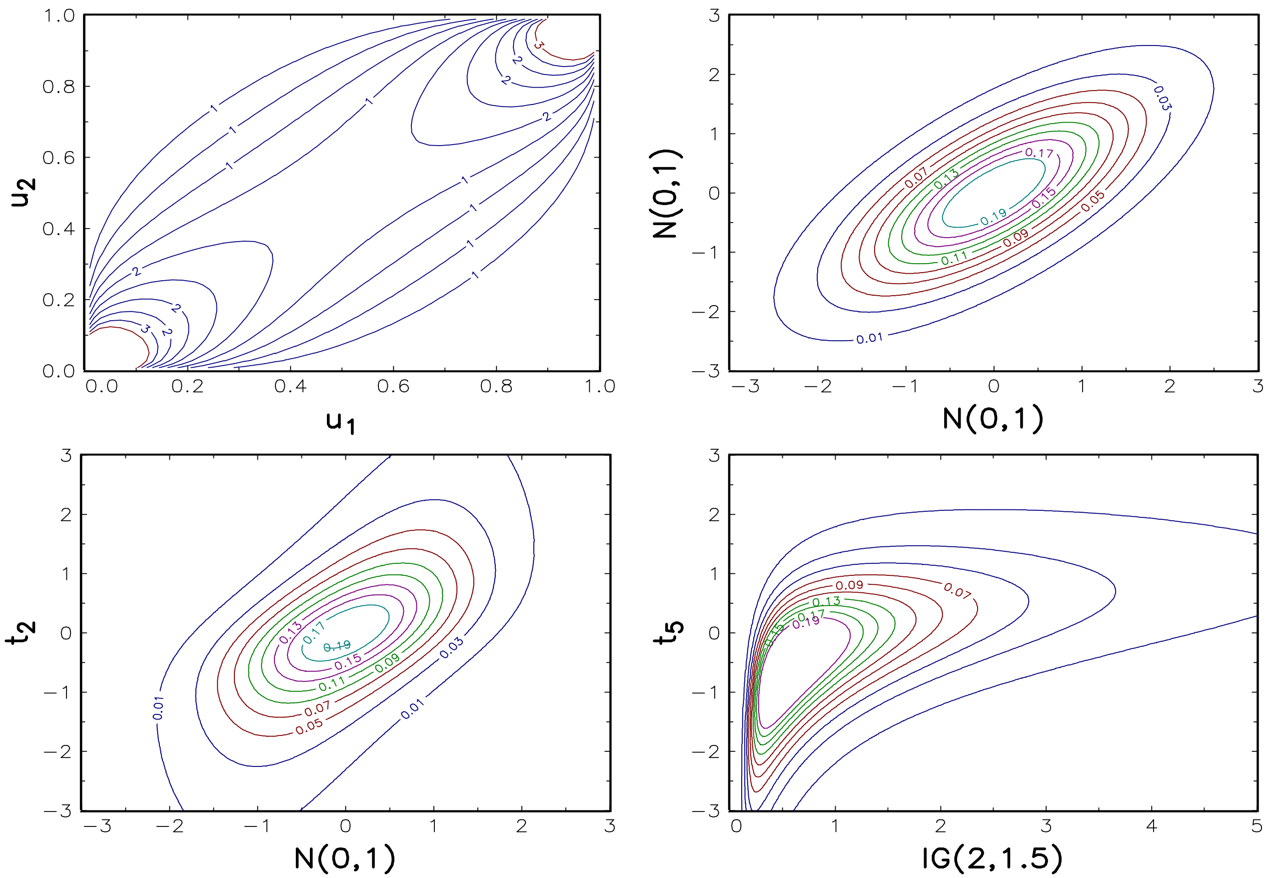


Figure: Contour lines of bivariate densities (Normal copula with $\tau = 50\%$)

Linear correlation

The linear correlation (or Pearson's correlation) is defined as follows:

$$\rho \langle X_1, X_2 \rangle = \frac{\mathbb{E}[X_1 \cdot X_2] - \mathbb{E}[X_1] \cdot \mathbb{E}[X_2]}{\sigma(X_1) \cdot \sigma(X_2)}$$

It satisfies the following properties:

- if $\mathbf{C}\langle X_1, X_2 \rangle = \mathbf{C}^{\perp}$, then $\rho \langle X_1, X_2 \rangle = 0$
- ρ is an increasing function with respect to the concordance measure:

$$\mathbf{C}_1 \succ \mathbf{C}_2 \Rightarrow \rho_1 \langle X_1, X_2 \rangle \geq \rho_2 \langle X_1, X_2 \rangle$$

- $\rho \langle X_1, X_2 \rangle$ is bounded:

$$\rho^- \langle X_1, X_2 \rangle \leq \rho \langle X_1, X_2 \rangle \leq \rho^+ \langle X_1, X_2 \rangle$$

  and the bounds are reached for the Fréchet copulas $\mathbf{C}^-$ and $\mathbf{C}^+$

Linear correlation

1. However, we don't have $\rho \langle \mathbf{C}^- \rangle = -1$ and $\rho \langle \mathbf{C}^+ \rangle = +1$. If we use the stochastic representation of Fréchet bounds, we have:

$$\rho^- \langle X_1, X_2 \rangle = \rho^+ \langle X_1, X_2 \rangle = \frac{\mathbb{E}\left[f_1(X) \cdot f_2(X)\right] - \mathbb{E}\left[f_1(X)\right] \cdot \mathbb{E}\left[f_2(X)\right]}{\sigma\left(f_1(X)\right) \cdot \sigma\left(f_2(X)\right)}$$

The solution of the equation $\rho^- \langle X_1, X_2 \rangle = -1$ is $f_1(x) = a_1 x + b_1$ and $f_2(x) = a_2 x + b_2$ where $a_1 a_2 < 0$. For the equation $\rho^+ \langle X_1, X_2 \rangle = +1$, the condition becomes $a_1 a_2 > 0$

2. Moreover, we have:

$$\rho \langle X_1, X_2 \rangle = \rho \langle f_1(X_1), f_2(X_2) \rangle \Leftrightarrow \left\{ \begin{array}{l} f_1(x) = a_1 x + b_1 \\ f_2(x) = a_2 x + b_2 \\ a_1 a_2 > 0 \end{array} \right.$$

Remark

The linear correlation is only valid for a linear (or Gaussian) world. **The copula function generalizes the concept of linear correlation in a non-Gaussian non-linear world**

Linear correlation

Example

We consider the bivariate log-normal random vector (X_1, X_2) where $X_1 \sim \mathcal{LN}\left(\mu_1, \sigma_1^2\right)$, $X_2 \sim \mathcal{LN}\left(\mu_2, \sigma_2^2\right)$ and $\rho = \rho\left\langle \ln X_1, \ln X_2 \right\rangle$.

We can show that:

$$\mathbb{E}\left[X_1^{p_1} \cdot X_2^{p_2}\right] = \exp\left(p_1\mu_1 + p_2\mu_2 + \frac{p_1^2\sigma_1^2 + p_2^2\sigma_2^2}{2} + p_1 p_2 \rho \sigma_1 \sigma_2\right)$$

and:

$$\rho\left\langle X_1, X_2 \right\rangle = \frac{\exp\left(\rho\sigma_1\sigma_2\right) - 1}{\sqrt{\exp\left(\sigma_1^2\right) - 1} \cdot \sqrt{\exp\left(\sigma_2^2\right) - 1}}$$

Linear correlation

If $\sigma_1 = 1$ and $\sigma_2 = 3$, we obtain the following results:

Copula	$\rho \langle X_1, X_2 \rangle$	$\tau \langle X_1, X_2 \rangle$	$\varrho \langle X_1, X_2 \rangle$
$\mathbf{C}^-$	-0.008	-1.000	-1.000
$\rho = -0.7$	-0.007	-0.494	-0.683
$\mathbf{C}^\perp$	0.000	0.000	0.000
$\rho = 0.7$	0.061	0.494	0.683
$\mathbf{C}^+$	0.162	1.000	1.000

Tail dependence

Definition

We consider the following statistic:

$$\lambda^{+} = \lim_{u \to 1^{-}} \frac{1 - 2u + \mathbf{C}(u, u)}{1 - u}$$

We say that $\mathbf{C}$ has an upper tail dependence when $\lambda^{+} \in (0, 1]$ and $\mathbf{C}$ has no upper tail dependence when $\lambda^{+} = 0$

- For the lower tail dependence λ^{-}, the limit becomes:

$$\lambda^{-} = \lim_{u \to 0^{+}} \frac{\mathbf{C}(u, u)}{u}$$

- We notice that λ^{+} and λ^{-} can also be defined as follows:

$$\lambda^{+} = \lim_{u \to 1^{-}} \Pr\{U_2 > u \mid U_1 > u\}$$

and:

$$\lambda^{-} = \lim_{u \to 0^{+}} \Pr\{U_2 < u \mid U_1 < u\}$$

Tail dependence

- For the copula functions $\mathbf{C}^-$ and $\mathbf{C}^\perp$, we have $\lambda^- = \lambda^+ = 0$
- For the copula $\mathbf{C}^+$, we obtain $\lambda^- = \lambda^+ = 1$
- In the case of the Gumbel copula:

$$\mathbf{C}(u_1, u_2; \theta) = \exp\left(-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{1/\theta}\right)$$

  we obtain $\lambda^- = 0$ and $\lambda^+ = 2 - 2^{1/\theta}$
- In the case of the Clayton copula:

$$\mathbf{C}(u_1, u_2; \theta) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$$

  we obtain $\lambda^- = 2^{-1/\theta}$ and $\lambda^+ = 0$

Tail dependence

The quantile-quantile dependence function is equal to:

$$
\begin{aligned}
\lambda^{+}(\alpha) &= \Pr\left\{X_2 > \mathbf{F}_2^{-1}(\alpha) \mid X_1 > \mathbf{F}_1^{-1}(\alpha)\right\} \\
&= \frac{\Pr\left\{X_2 > \mathbf{F}_2^{-1}(\alpha), X_1 > \mathbf{F}_1^{-1}(\alpha)\right\}}{\Pr\left\{X_1 > \mathbf{F}_1^{-1}(\alpha)\right\}} \\
&= \frac{1 - \Pr\left\{X_1 \leq \mathbf{F}_1^{-1}(\alpha)\right\} - \Pr\left\{X_2 \leq \mathbf{F}_2^{-1}(\alpha)\right\}}{1 - \Pr\left\{X_1 \leq \mathbf{F}_1^{-1}(\alpha)\right\}} + \\
&\quad \frac{\Pr\left\{X_2 \leq \mathbf{F}_2^{-1}(\alpha), X_1 \leq \mathbf{F}_1^{-1}(\alpha)\right\}}{1 - \Pr\left\{\mathbf{F}_1(X_1) \leq \alpha\right\}} \\
&= \frac{1 - 2\alpha + \mathbf{C}(\alpha, \alpha)}{1 - \alpha}
\end{aligned}
$$

Tail dependence

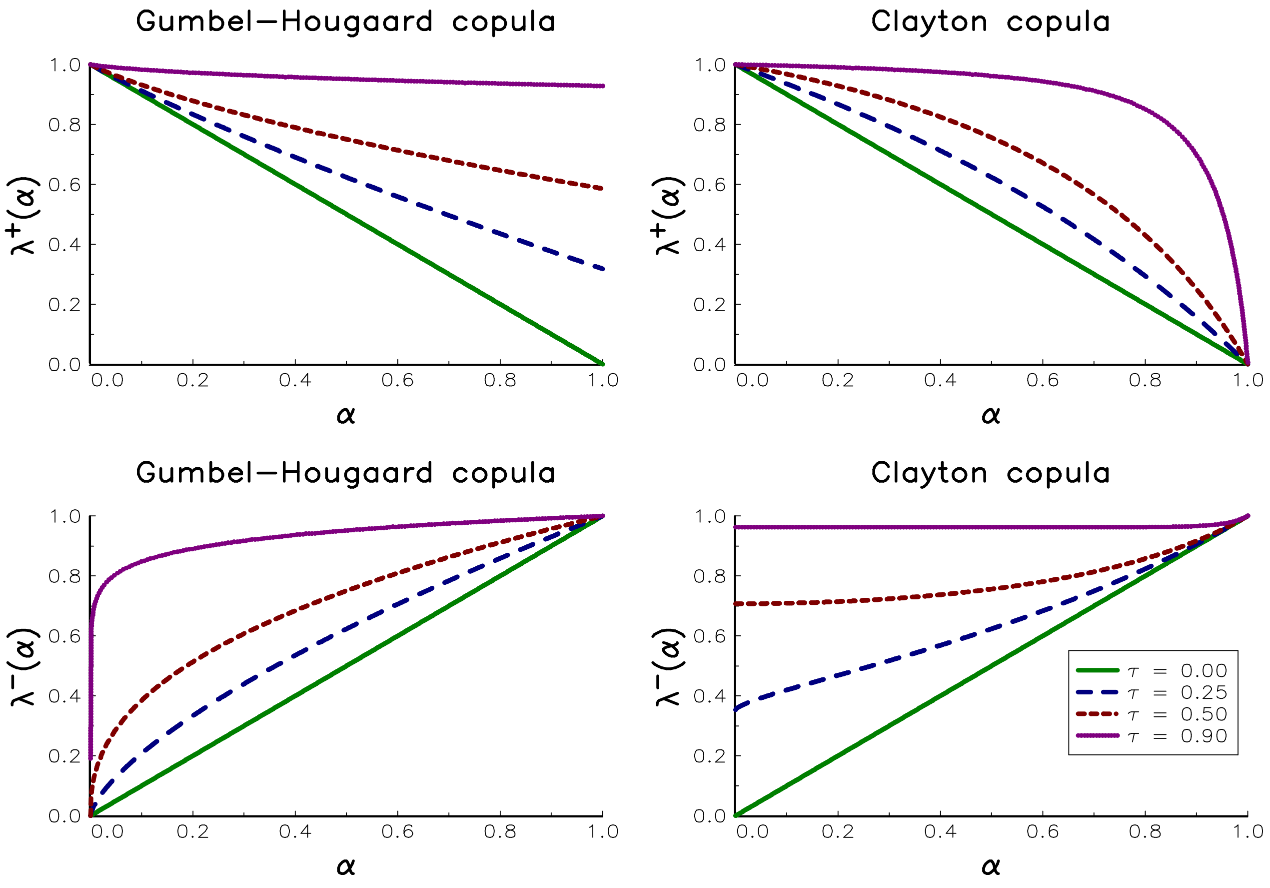


Figure: Quantile-quantile dependence measures $\lambda^{+}(\alpha)$ and $\lambda^{-}(\alpha)$

Risk interpretation of the tail dependence

We consider two portfolios, whose losses correspond to the random variables L_1 and L_2 with probability distributions $\mathbf{F}_1$ and $\mathbf{F}_2$. We have:

$$\begin{aligned} \lambda^+(\alpha) &= \Pr\left\{L_2 > \mathbf{F}_2^{-1}(\alpha) \mid L_1 > \mathbf{F}_1^{-1}(\alpha)\right\} \\ &= \Pr\left\{L_2 > \mathrm{VaR}_\alpha(L_2) \mid L_1 > \mathrm{VaR}_\alpha(L_1)\right\} \end{aligned}$$

Archimedean copulas

Definition

An Archimedean copula is defined by:

$$\mathbf{C}(u_1, u_2) = \begin{cases} \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) & \text{if } \varphi(u_1) + \varphi(u_2) \leq \varphi(0) \\ 0 & \text{otherwise} \end{cases}$$

where φ a C^2 is a function which satisfies $\varphi(1) = 0$, $\varphi'(u) < 0$ and $\varphi''(u) > 0$ for all $u \in [0, 1]$

$\Rightarrow \varphi(u)$ is called the generator of the copula function

Archimedean copulas

Example

If $\varphi(u) = u^{-1} - 1$, we have $\varphi^{-1}(u) = (1+u)^{-1}$ and:

$$\mathbf{C}(u_1, u_2) = \left(1 + \left(u_1^{-1} - 1 + u_2^{-1} - 1\right)\right)^{-1} = \frac{u_1 u_2}{u_1 + u_2 - u_1 u_2}$$

The Gumbel logistic copula is then an Archimedean copula

Remark

- *The product copula $\mathbf{C}^{\perp}$ is Archimedean and the associated generator is $\varphi(u) = -\ln u$*
- *Concerning Fréchet copulas, only $\mathbf{C}^{-}$ is Archimedean with $\varphi(u) = 1 - u$*

Archimedean copulas

Table: Archimedean copula functions

Copula	$\varphi(u)$	$\mathbf{C}(u_1, u_2)$
$\mathbf{C}^{\perp}$	$-\ln u$	$u_1 u_2$
Clayton	$u^{-\theta} - 1$	$\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$
Frank	$-\ln \dfrac{e^{-\theta u} - 1}{e^{-\theta} - 1}$	$-\dfrac{1}{\theta} \ln \left(1 + \dfrac{\left(e^{-\theta u_1} - 1\right)\left(e^{-\theta u_2} - 1\right)}{e^{-\theta} - 1}\right)$
Gumbel	$(-\ln u)^{\theta}$	$\exp\left(-\left(\tilde{u}_1^{\theta} + \tilde{u}_2^{\theta}\right)^{1/\theta}\right)$
Joe	$-\ln\left(1 - (1-u)^{\theta}\right)$	$1 - \left(\bar{u}_1^{\theta} + \bar{u}_2^{\theta} - \bar{u}_1^{\theta}\bar{u}_2^{\theta}\right)^{1/\theta}$

We use the notations $\bar{u} = 1 - u$ and $\tilde{u} = -\ln u$

Multivariate Normal copula

The Normal copula is the dependence function of the multivariate normal distribution with a correlation matrix ρ:

$$\mathbf{C}\left(u_1, \ldots, u_n; \rho\right) = \Phi_n\left(\Phi^{-1}\left(u_1\right), \ldots, \Phi^{-1}\left(u_n\right); \rho\right)$$

By using the canonical decomposition of the multivariate density function:

$$f\left(x_1, \ldots, x_n\right) = c\left(\mathbf{F}_1\left(x_1\right), \ldots, \mathbf{F}_n\left(x_n\right)\right) \prod_{i=1}^{n} f_i\left(x_i\right)$$

we deduce that the probability density function of the Normal copula is:

$$c\left(u_1, \ldots, u_n, ; \rho\right) = \frac{1}{\left|\rho\right|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} x^\top \left(\rho^{-1} - I_n\right) x\right)$$

where $x_i = \Phi^{-1}\left(u_i\right)$

Bivariate Normal copula

In the bivariate case, we obtain:

$$c(u_1, u_2; \rho) = \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1-\rho^2)} + \frac{x_1^2 + x_2^2}{2}\right)$$

It follows that the expression of the bivariate Normal copula function is also equal to:

$$\mathbf{C}(u_1, u_2; \rho) = \int_{-\infty}^{\mathbf{\Phi}^{-1}(u_1)} \int_{-\infty}^{\mathbf{\Phi}^{-1}(u_2)} \phi_2(x_1, x_2; \rho)\, \mathrm{d}x_1\, \mathrm{d}x_2$$

where $\phi_2(x_1, x_2; \rho)$ is the bivariate normal density:

$$\phi_2(x_1, x_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1-\rho^2)}\right)$$

Bivariate Normal copula

Remark

Let (X_1, X_2) be a standardized Gaussian random vector, whose cross-correlation is ρ. Using the Cholesky decomposition, we write X_2 as follows: $X_2 = \rho X_1 + \sqrt{1-\rho^2} X_3$ where $X_3 \sim \mathcal{N}(0,1)$ is independent from X_1 and X_2. We have:

$$\begin{aligned}
\Phi_2(x_1, x_2; \rho) &= \Pr\{X_1 \leq x_1, X_2 \leq x_2\} \\
&= \mathbb{E}\left[\Pr\left\{X_1 \leq x_1, \rho X_1 + \sqrt{1-\rho^2} X_3 \leq x_2 \mid X_1\right\}\right] \\
&= \int_{-\infty}^{x_1} \Phi\left(\frac{x_2 - \rho x}{\sqrt{1-\rho^2}}\right) \phi(x) \, \mathrm{d}x
\end{aligned}$$

It follows that:

$$\mathbf{C}(u_1, u_2; \rho) = \int_{-\infty}^{\Phi^{-1}(u_1)} \Phi\left(\frac{\Phi^{-1}(u_2) - \rho x}{\sqrt{1-\rho^2}}\right) \phi(x) \, \mathrm{d}x$$

Bivariate Normal copula

- We deduce that:

$$\mathbf{C}\left(u_1, u_2; \rho\right) = \int_0^{u_1} \Phi\left(\frac{\Phi^{-1}\left(u_2\right) - \rho\Phi^{-1}\left(u\right)}{\sqrt{1-\rho^2}}\right) \mathrm{d}u$$

- We have:

$$\tau = \frac{2}{\pi} \arcsin \rho$$

and:

$$\varrho = \frac{6}{\pi} \arcsin \frac{\rho}{2}$$

- We can show that:

$$\lambda^+ = \lambda^- = \begin{cases} 0 & \text{if } \rho < 1 \\ 1 & \text{if } \rho = 1 \end{cases}$$

Bivariate Normal copula

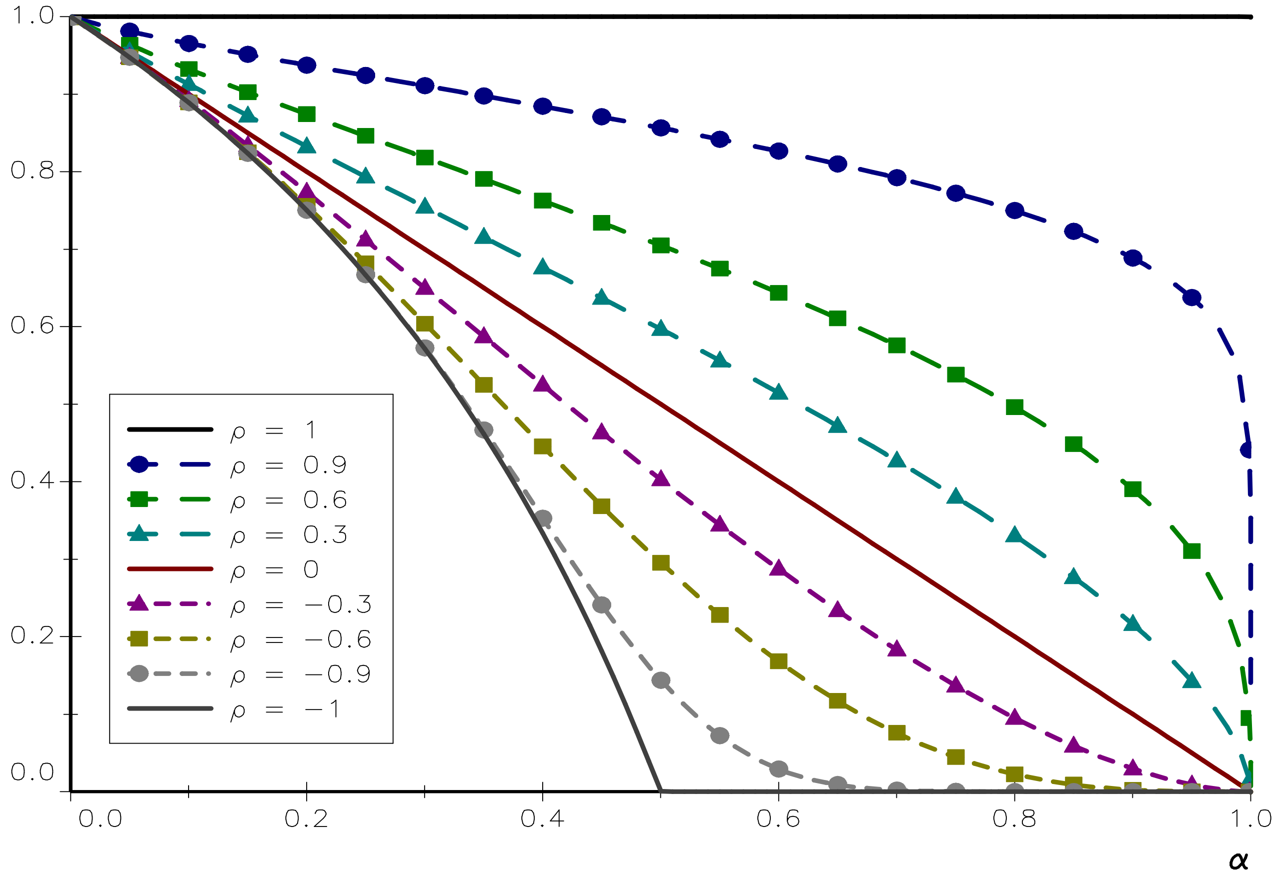


Figure: Tail dependence $\lambda^{+}(\alpha)$ for the Normal copula

Multivariate Student's t copula

We have:

$$\mathbf{C}\left(u_1,\ldots,u_n;\rho,\nu\right)=\mathbf{T}_n\left(\mathbf{T}_\nu^{-1}\left(u_1\right),\ldots,\mathbf{T}_\nu^{-1}\left(u_n\right);\rho,\nu\right)$$

By using the definition of the cumulative distribution function:

$$\mathbf{T}_n\left(x_1,\ldots,x_n;\rho,\nu\right)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\frac{\Gamma\left(\frac{\nu+n}{2}\right)\left|\rho\right|^{-\frac{1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)\left(\nu\pi\right)^{\frac{n}{2}}}\left(1+\frac{1}{\nu}x^\top\rho^{-1}x\right)^{-\frac{\nu+n}{2}}\,\mathrm{d}x$$

we can show that the copula density function is then:

$$c\left(u_1,\ldots,u_n,;\rho,\nu\right)=\left|\rho\right|^{-\frac{1}{2}}\frac{\Gamma\left(\frac{\nu+n}{2}\right)\left[\Gamma\left(\frac{\nu}{2}\right)\right]^n}{\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^n\Gamma\left(\frac{\nu}{2}\right)}\frac{\left(1+\frac{1}{\nu}x^\top\rho^{-1}x\right)^{-\frac{\nu+n}{2}}}{\prod_{i=1}^n\left(1+\frac{x_i^2}{\nu}\right)^{-\frac{\nu+1}{2}}}$$

where $x_i=\mathbf{T}_\nu^{-1}\left(u_i\right)$

Bivariate Student's t copula

- We have:

$$\mathbf{C}\left(u_1, u_2; \rho, \nu\right) = \int_0^{u_1} \mathbf{C}_{2|1}\left(u, u_2; \rho, \nu\right) \,\mathrm{d}u$$

where:

$$\mathbf{C}_{2|1}\left(u_1, u_2; \rho, \nu\right) = \mathbf{T}_{\nu+1}\left(\left(\frac{\nu+1}{\nu+\left[\mathbf{T}_\nu^{-1}\left(u_1\right)\right]^2}\right)^{1/2} \frac{\mathbf{T}_\nu^{-1}\left(u_2\right) - \rho\mathbf{T}_\nu^{-1}\left(u_1\right)}{\sqrt{1-\rho^2}}\right)$$

- We have:

$$\lambda^+ = 2 - 2 \cdot \mathbf{T}_{\nu+1}\left(\left(\frac{\left(\nu+1\right)\left(1-\rho\right)}{\left(1+\rho\right)}\right)^{1/2}\right) = \left\{\begin{array}{ll} 0 & \text{if } \rho = -1 \\ > 0 & \text{if } \rho > -1 \end{array}\right.$$

Bivariate Student's t copula

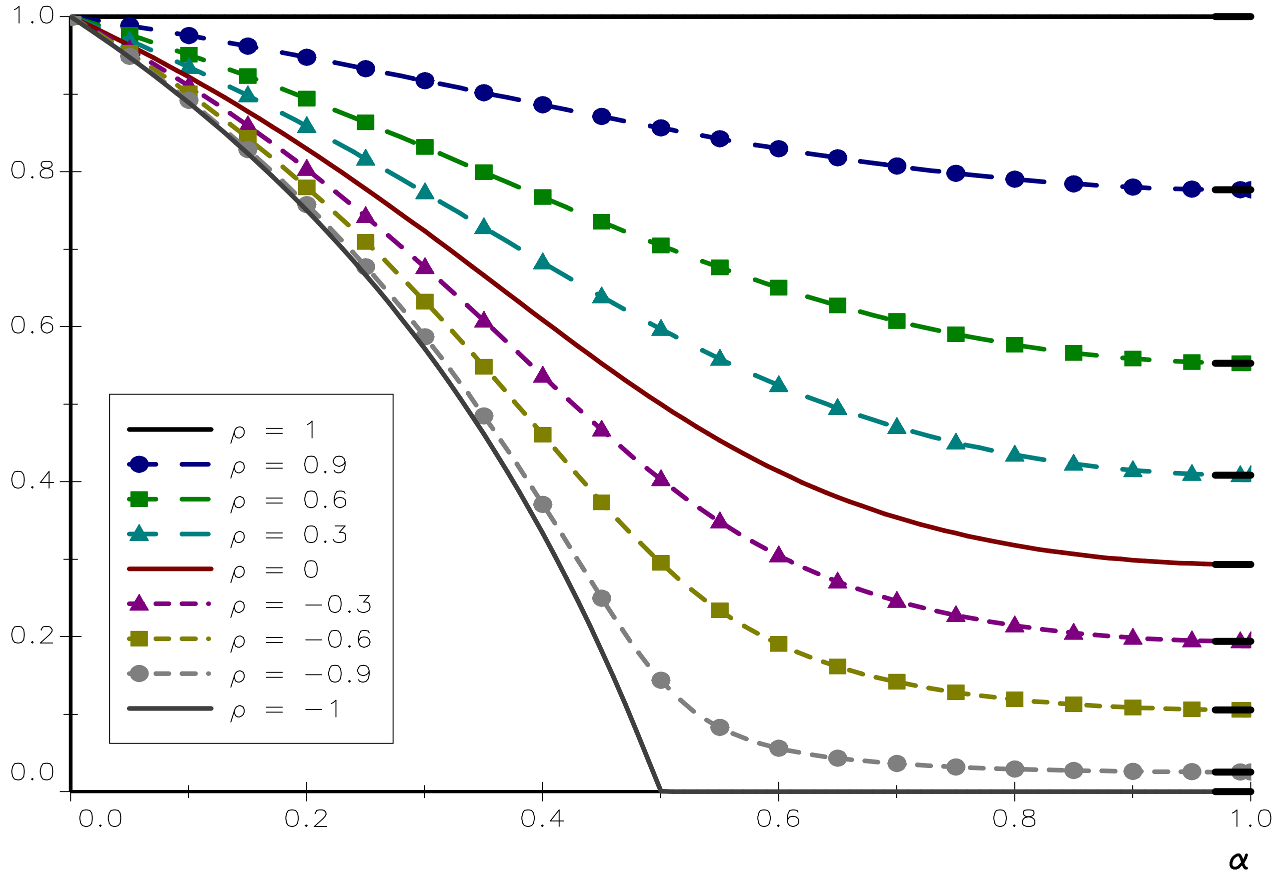


Figure: Tail dependence $\lambda^{+}(\alpha)$ for the Student's t copula ($\nu = 1$)

Bivariate Student's t copula

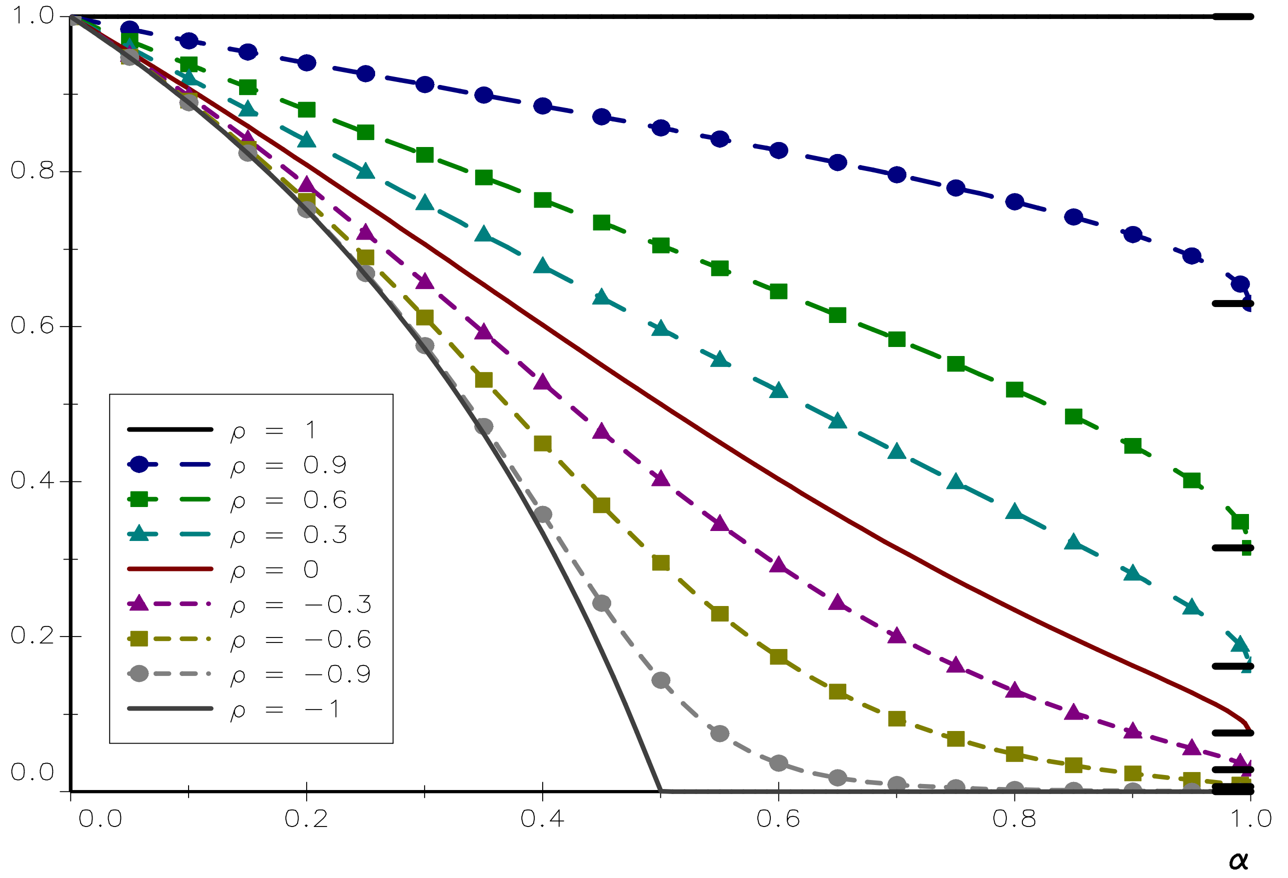


Figure: Tail dependence $\lambda^{+}(\alpha)$ for the Student's t copula ($\nu = 4$)

Dependogram

The dependogram is the scatter plot between $u_{t,1}$ and $u_{t,2}$ where:

$$u_{t,i} = \frac{1}{T+1}\mathfrak{R}_{t,i}$$

and $\mathfrak{R}_{t,i}$ is the rank statistic (T is the sample size)

Example

$x_{t,1}$	-3	4	1	8
$x_{t,2}$	105	65	17	9
$\mathfrak{R}_{t,1}$	1	3	2	4
$\mathfrak{R}_{t,2}$	4	3	2	1
$u_{t,1}$	0.20	0.60	0.40	0.80
$u_{t,2}$	0.80	0.60	0.40	0.20

Dependogram

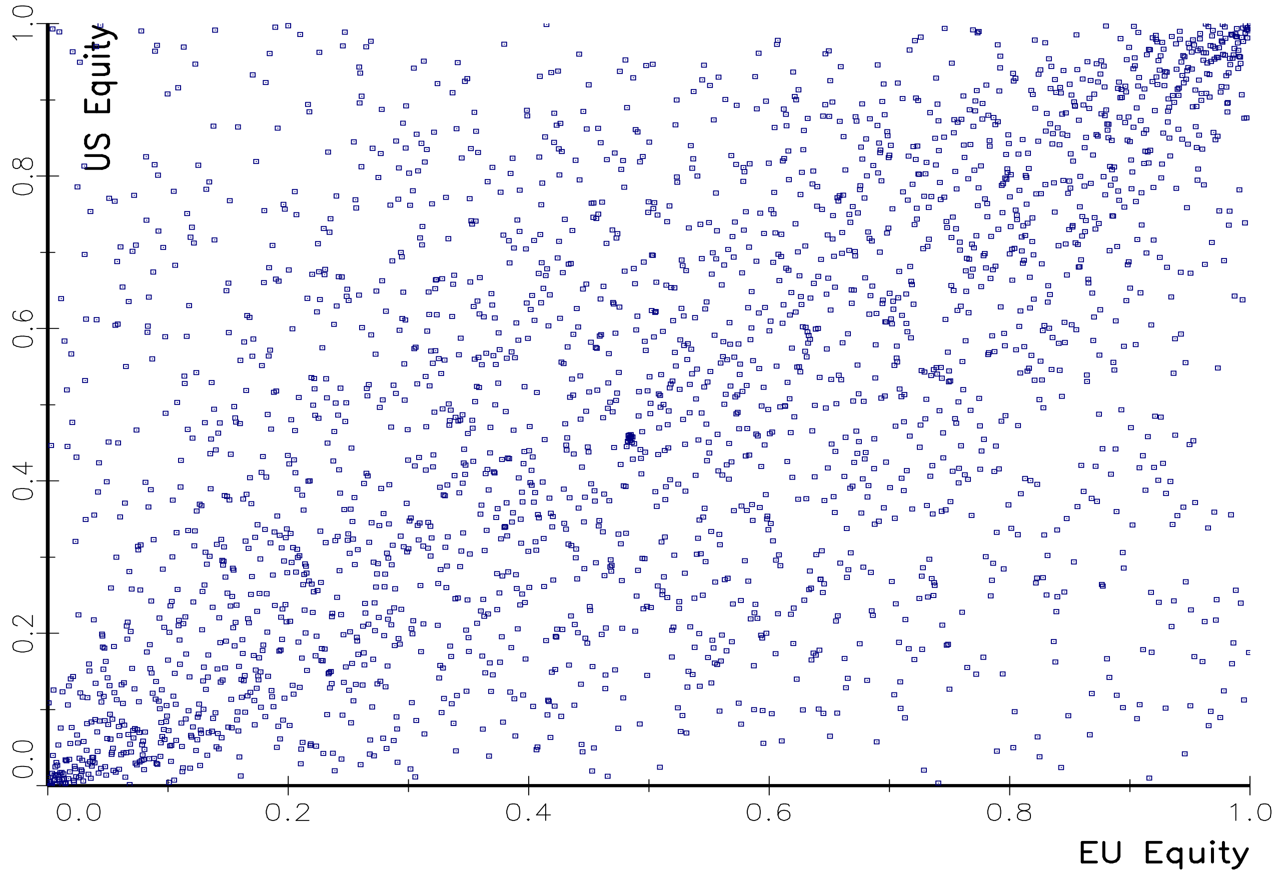


Figure: Dependogram of EU and US equity returns ($\rho = 57.8\%$)

Dependogram

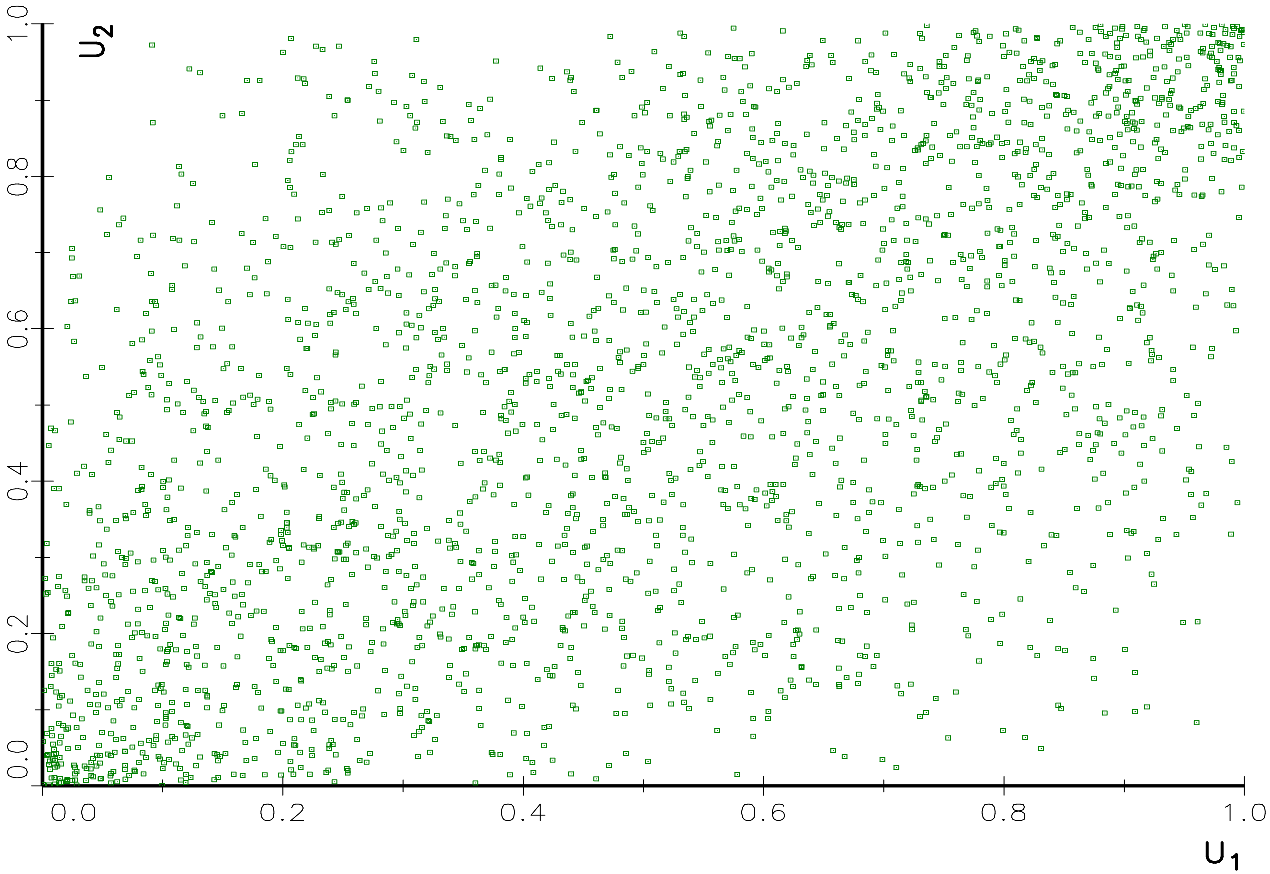


Figure: Dependogram of simulated Gaussian returns ($\rho = 57.8\%$)

The method of moments

If $\tau = f_\tau(\theta)$ is the relationship between θ and Kendall's tau, the MM estimator is simply the inverse of this relationship:

$$\hat{\theta} = f_\tau^{-1}(\hat{\tau})$$

where $\hat{\tau}$ is the estimate of Kendall's tau based on the sample

Remark

We have:

$$\hat{\tau} = \frac{c-d}{c+d}$$

where c and d are the number of concordant and discordant pairs

For instance, in the case of the Gumbel copula, we have:

$$\tau = \frac{\theta - 1}{\theta}$$

and:

$$\hat{\theta} = \frac{1}{1-\hat{\tau}}$$

The method of maximum likelihood

We have:

$$\mathbf{F}(x_1,\ldots,x_n)=\mathbf{C}\left(\mathbf{F}_1(x_1;\theta_1),\ldots,\mathbf{F}_n(x_n;\theta_n);\theta_c\right)$$

with two types of parameters:

- the parameters $(\theta_1,\ldots,\theta_n)$ of univariate distribution functions
- the parameters θ_c of the copula function

The expression of the log-likelihood function is:

$$\begin{aligned}\ell(\theta_1,\ldots,\theta_n,\theta_c) &= \sum_{t=1}^{T}\ln c\left(\mathbf{F}_1(x_{t,1};\theta_1),\ldots,\mathbf{F}_n(x_{t,n};\theta_n);\theta_c\right)+ \\ &\quad \sum_{t=1}^{T}\sum_{i=1}^{n}\ln f_i(x_{t,i};\theta_i)\end{aligned}$$

The ML estimator is then defined as follows:

$$\left(\hat{\theta}_1,\ldots,\hat{\theta}_n,\hat{\theta}_c\right)=\arg\max\ell(\theta_1,\ldots,\theta_n,\theta_c)$$

The method of inference functions for marginals

The IFM method is a two-stage parametric method:

1. the first stage involves maximum likelihood from univariate marginals
2. the second stage involves maximum likelihood of the copula parameters θ_c with the univariate parameters $\hat{\theta}_1, \ldots, \hat{\theta}_n$ held fixed from the first stage:

$$\hat{\theta}_c = \arg\max \sum_{t=1}^{T} \ln c \left(\mathbf{F}_1 \left(x_{t,1}; \hat{\theta}_1 \right), \ldots, \mathbf{F}_n \left(x_{t,n}; \hat{\theta}_n \right); \theta_c \right)$$

The omnibus method

The omnibus method replaces the marginals $\mathbf{F}_1, \ldots, \mathbf{F}_n$ by their non-parametric estimates:

$$\hat{\theta}_c = \arg\max \sum_{t=1}^{T} \ln c\left(\hat{\mathbf{F}}_1\left(x_{t,1}\right), \ldots, \hat{\mathbf{F}}_n\left(x_{t,n}\right); \theta_c\right)$$

where:

$$\hat{\mathbf{F}}_i\left(x_{t,i}\right) = u_{t,i} = \frac{1}{T+1}\mathfrak{R}_{t,i}$$

Estimation of the Normal copula

In the case of the Normal copula, the matrix ρ of the parameters is estimated with the following algorithm:

1. we first transform the uniform variates $u_{t,i}$ into Gaussian variates:

$$n_{t,i} = \Phi^{-1}\left(u_{t,i}\right)$$

2. we then calculate the correlation matrix $\hat{\rho}$ of the Gaussian variates $n_{t,i}$.

Order statistics

Definition

- Let $X_1, \ldots, X_n$ be *iid* random variables, whose probability distribution is denoted by $\mathbf{F}$
- We rank these random variables by increasing order:

$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n-1:n} \leq X_{n:n}$$

- $X_{i:n}$ is called the i^{th} order statistic in the sample of size n
- We note $x_{i:n}$ the corresponding random variate or the value taken by $X_{i:n}$

Order statistics

We have:

$$\begin{aligned}
\mathbf{F}_{i:n}(x) &= \Pr\{X_{i:n} \leq x\} \\
&= \Pr\{\text{at least } i \text{ variables among } X_1, \ldots, X_n \text{ are less or equal to } x\} \\
&= \sum_{k=i}^{n} \Pr\{k \text{ variables among } X_1, \ldots, X_n \text{ are less or equal to } x\} \\
&= \sum_{k=i}^{n} \binom{n}{k} \mathbf{F}(x)^k (1 - \mathbf{F}(x))^{n-k}
\end{aligned}$$

and:

$$f_{i:n}(x) = \frac{\partial\, \mathbf{F}_{i:n}(x)}{\partial\, x}$$

Order statistics

Example

If $X_1, \ldots, X_n$ follow a uniform distribution $\mathcal{U}_{[0,1]}$, we obtain:

$$\mathbf{F}_{i:n}(x) = \sum_{k=i}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \mathcal{IB}(x; i, n-i+1)$$

where $\mathcal{IB}(x; \alpha, \beta)$ is the regularized incomplete beta function:

$$\mathcal{IB}(x; \alpha, \beta) = \frac{1}{\mathfrak{B}(\alpha, \beta)} \int_0^x t^{\alpha-1} (1-t)^{\beta-1} \, \mathrm{d}t$$

We deduce that $X_{i:n} \sim \mathcal{B}(i, n-i+1)$ and[a]:

$$\mathbb{E}[X_{i:n}] = \mathbb{E}[\mathcal{B}(i, n-i+1)] = \frac{i}{n+1}$$

[a]We recall that $\mathbb{E}[\mathcal{B}(\alpha, \beta)] = \alpha / (\alpha + \beta)$

Extreme order statistics

The extreme order statistics are:

$$X_{1:n} = \min\left(X_1, \ldots, X_n\right)$$

and:

$$X_{n:n} = \max\left(X_1, \ldots, X_n\right)$$

We have:

$$\begin{aligned}
\mathbf{F}_{1:n}\left(x\right) &= \sum_{k=1}^{n}\binom{n}{k}\mathbf{F}\left(x\right)^{k}\left(1-\mathbf{F}\left(x\right)\right)^{n-k} = 1-\binom{n}{0}\mathbf{F}\left(x\right)^{0}\left(1-\mathbf{F}\left(x\right)\right)^{n} \\
&= 1-\left(1-\mathbf{F}\left(x\right)\right)^{n}
\end{aligned}$$

and:

$$\begin{aligned}
\mathbf{F}_{i:n}\left(x\right) &= \sum_{k=n}^{n}\binom{n}{k}\mathbf{F}\left(x\right)^{k}\left(1-\mathbf{F}\left(x\right)\right)^{n-k} = \binom{n}{n}\mathbf{F}\left(x\right)^{n}\left(1-\mathbf{F}\left(x\right)\right)^{n-n} \\
&= \mathbf{F}\left(x\right)^{n}
\end{aligned}$$

Alternative proof

We have:

$$\begin{aligned} \mathbf{F}_{1:n}(x) = \Pr\{\min(X_1,\ldots,X_n) \le x\} &= 1 - \Pr\{\min(X_1,\ldots,X_n) \ge x\} \\ &= 1 - \Pr\{X_1 \ge x, X_2 \ge x, \ldots, X_n \ge x\} \\ &= 1 - \prod_{i=1}^{n} \Pr\{X_i \ge x\} \\ &= 1 - \prod_{i=1}^{n} (1 - \Pr\{X_i \le x\}) \\ &= 1 - (1 - \mathbf{F}(x))^n \end{aligned}$$

and:

$$\begin{aligned} \mathbf{F}_{n:n}(x) = \Pr\{\max(X_1,\ldots,X_n) \le x\} &= \Pr\{X_1 \le x, X_2 \le x, \ldots, X_n \le x\} \\ &= \prod_{i=1}^{n} \Pr\{X_i \le x\} \\ &= \mathbf{F}(x)^n \end{aligned}$$

Extreme order statistics

We deduce that the density functions are equal to:

$$f_{1:n}(x) = n(1 - \mathbf{F}(x))^{n-1} f(x)$$

and

$$f_{n:n}(x) = n\mathbf{F}(x)^{n-1} f(x)$$

Extreme order statistics

We consider the daily returns of the MSCI USA index from 1995 to 2015

$\mathcal{H}_1$ Daily returns are Gaussian, meaning that:

$$R_t = \hat{\mu} + \hat{\sigma} X_t$$

where $X_t \sim \mathcal{N}(0,1)$, $\hat{\mu}$ is the empirical mean of daily returns and $\hat{\sigma}$ is the daily standard deviation

$\mathcal{H}_2$ Daily returns follow a Student's t distribution[2]:

$$R_t = \hat{\mu} + \hat{\sigma}\sqrt{\frac{\nu - 2}{\nu}} X_t$$

where $X_t \sim \mathbf{t}_\nu$. We consider two alternative assumptions: $\mathcal{H}_{2a} : \nu = 3$ and $\mathcal{H}_{2b} : \nu = 6$

[2] We add the factor $\sqrt{(\nu - 2)/\nu}$ in order to verify that $\operatorname{var}(R_t) = \hat{\sigma}^2$

Extreme order statistics

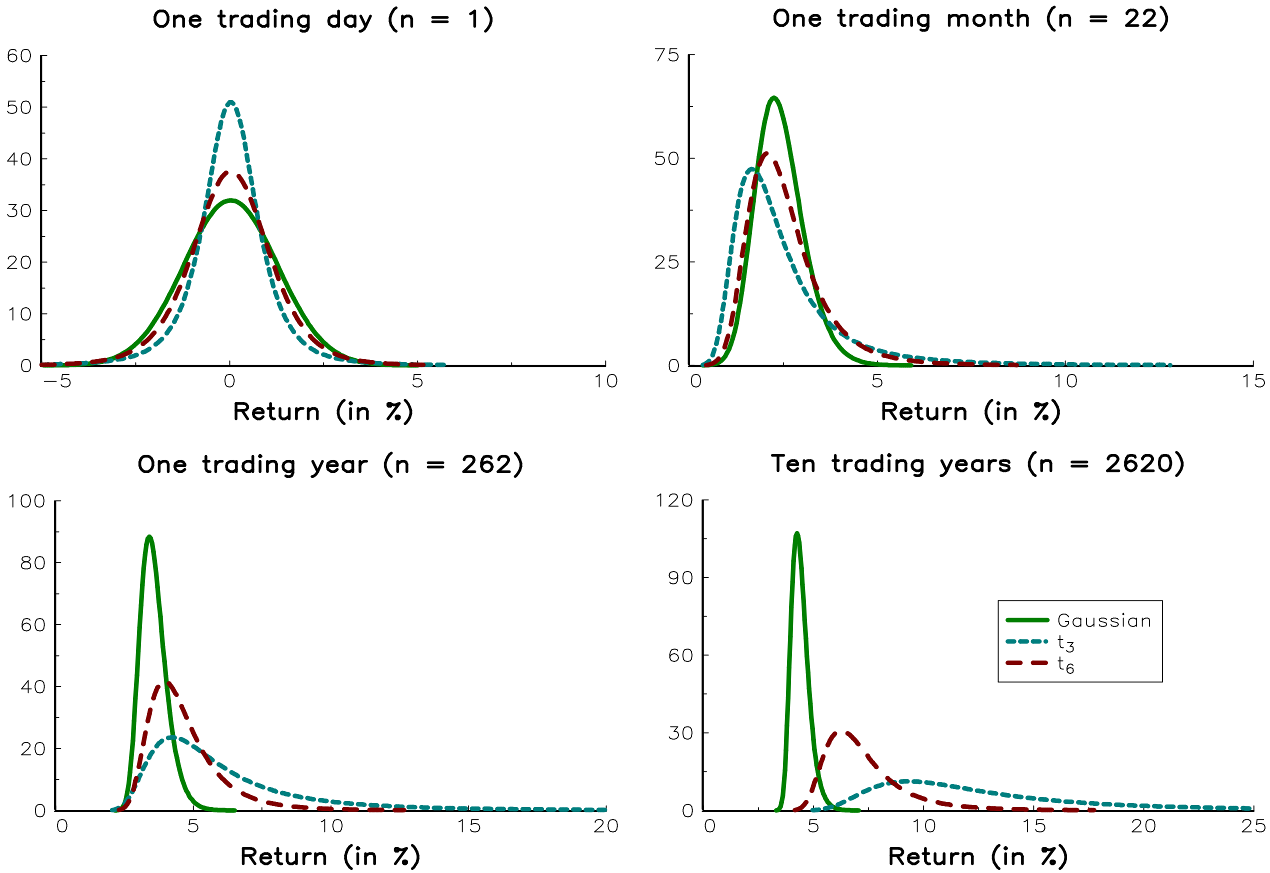


Figure: Density function of the maximum order statistic (daily return of the MSCI USA index, 1995-2015)

Extreme order statistics

Remark

The limit distributions of minima and maxima are degenerate:

$$\lim_{n\to\infty} \mathbf{F}_{1:n}(x) = \lim_{n\to\infty} 1 - (1 - \mathbf{F}(x))^n = \begin{cases} 0 & \text{if } \mathbf{F}(x) = 0 \\ 1 & \text{if } \mathbf{F}(x) > 0 \end{cases}$$

and:

$$\lim_{n\to\infty} \mathbf{F}_{n:n}(x) = \lim_{n\to\infty} \mathbf{F}(x)^n = \begin{cases} 0 & \text{if } \mathbf{F}(x) < 1 \\ 1 & \text{if } \mathbf{F}(x) = 1 \end{cases}$$

Remark

We only consider the largest order statistic $X_{n:n}$ because the minimum order statistic $X_{1:n}$ is equal to $Y_{n:n}$ by setting $Y_i = -X_i$

Univariate extreme value theory

Fisher-Tippet theorem

Let $X_1, \ldots, X_n$ be a sequence of *iid* random variables, whose distribution function is $\mathbf{F}$. If there exist two constants a_n and b_n and a non-degenerate distribution function $\mathbf{G}$ such that:

$$\lim_{n \to \infty} \Pr \left\{ \frac{X_{n:n} - b_n}{a_n} \leq x \right\} = \mathbf{G}(x)$$

then $\mathbf{G}$ can be classified as one of the following three types:

Type I	(Gumbel)	$\mathbf{\Lambda}(x) = \exp\left(-e^{-x}\right)$
Type II	(Fréchet)	$\mathbf{\Phi}_\alpha(x) = \mathbb{1}(x \geq 0) \cdot \exp\left(-x^{-\alpha}\right)$
Type III	(Weibull	$\mathbf{\Psi}_\alpha(x) = \mathbb{1}(x \leq 0) \cdot \exp\left(-(-x)^{\alpha}\right)$

$\mathbf{\Lambda}$, $\mathbf{\Phi}_\alpha$ and $\mathbf{\Psi}_\alpha$ are called extreme value distributions

Fisher-Tippet theorem $\approx$ an extreme value analog of the central limit theorem

Univariate extreme value theory

We recall that:

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots = \exp\left(x\right)$$

Univariate extreme value theory

- We consider the exponential distribution: $\mathbf{F}(x) = 1 - \exp(-\lambda x)$. We have:

$$\begin{aligned}\lim_{n\to\infty} \mathbf{F}_{n:n}(x) &= \lim_{n\to\infty} \left(1 - e^{-\lambda x}\right)^n = \lim_{n\to\infty} \left(1 - \frac{ne^{-\lambda x}}{n}\right)^n \\ &= \lim_{n\to\infty} \exp\left(-ne^{-\lambda x}\right) = 0\end{aligned}$$

  We verify that the limit distribution is degenerate
- If we consider the affine transformation with $a_n = 1/\lambda$ et $b_n = (\ln n)/\lambda$, we obtain:

$$\begin{aligned}\Pr\left\{\frac{X_{n:n} - b_n}{a_n} \leq x\right\} &= \Pr\left\{X_{n:n} \leq a_n x + b_n\right\} = \left(1 - e^{-\lambda(a_n x + b_n)}\right)^n \\ &= \left(1 - e^{-x-\ln n}\right)^n = \left(1 - \frac{e^{-x}}{n}\right)^n\end{aligned}$$

  and:

$$\mathbf{G}(x) = \lim_{n\to\infty} \left(1 - \frac{e^{-x}}{n}\right)^n = \exp\left(-e^{-x}\right) = \mathbf{\Lambda}(x)$$

Generalized extreme value distribution

- We combine the three distributions $\mathbf{\Lambda}$, $\mathbf{\Phi}_\alpha$ et $\mathbf{\Psi}_\alpha$ into a single distribution function $\mathcal{GEV}(\mu, \sigma, \xi)$:

$$\mathbf{G}(x) = \exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right)$$

  defined on the support $\Delta = \left\{x : 1 + \xi\sigma^{-1}(x-\mu) > 0\right\}$
- the limit case $\xi \to 0$ corresponds to the Gumbel distribution $\mathbf{\Lambda}$
- $\xi = -\alpha^{-1} > 0$ defines the Fréchet distribution $\mathbf{\Phi}_\alpha$
- the Weibull distribution $\mathbf{\Psi}_\alpha$ is obtained by considering $\xi = -\alpha^{-1} < 0$

Generalized extreme value distribution

The density function is equal to:

$$g(x) = \frac{1}{\sigma}\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-(1+\xi)/\xi} \exp\left(-\left(1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right)$$

Block maxima approach

The log-likelihood function is equal to:

$$\ell_t = -\ln\sigma - \left(\frac{1+\xi}{\xi}\right)\ln\left(1 + \xi\left(\frac{x_t-\mu}{\sigma}\right)\right) - \left(1 + \xi\left(\frac{x_t-\mu}{\sigma}\right)\right)^{-1/\xi}$$

where x_t is the observed maximum for the t^{th} period (or block maximum)

Generalized extreme value distribution

- We consider the example of the MSCI USA index
- Using daily returns, we calculate the block maximum for each period of 22 trading days and estimate the GEV distribution using the method of maximum likelihood
- We compare the estimated GEV distribution with the distribution function $\mathbf{F}_{22:22}(x)$ when we assume that daily returns are Gaussian:

α	90%	95%	96%	97%	98%	99%
Gaussian	3.26%	3.56%	3.65%	3.76%	3.92%	4.17%
GEV	3.66%	4.84%	5.28%	5.91%	6.92%	9.03%

Generalized extreme value distribution

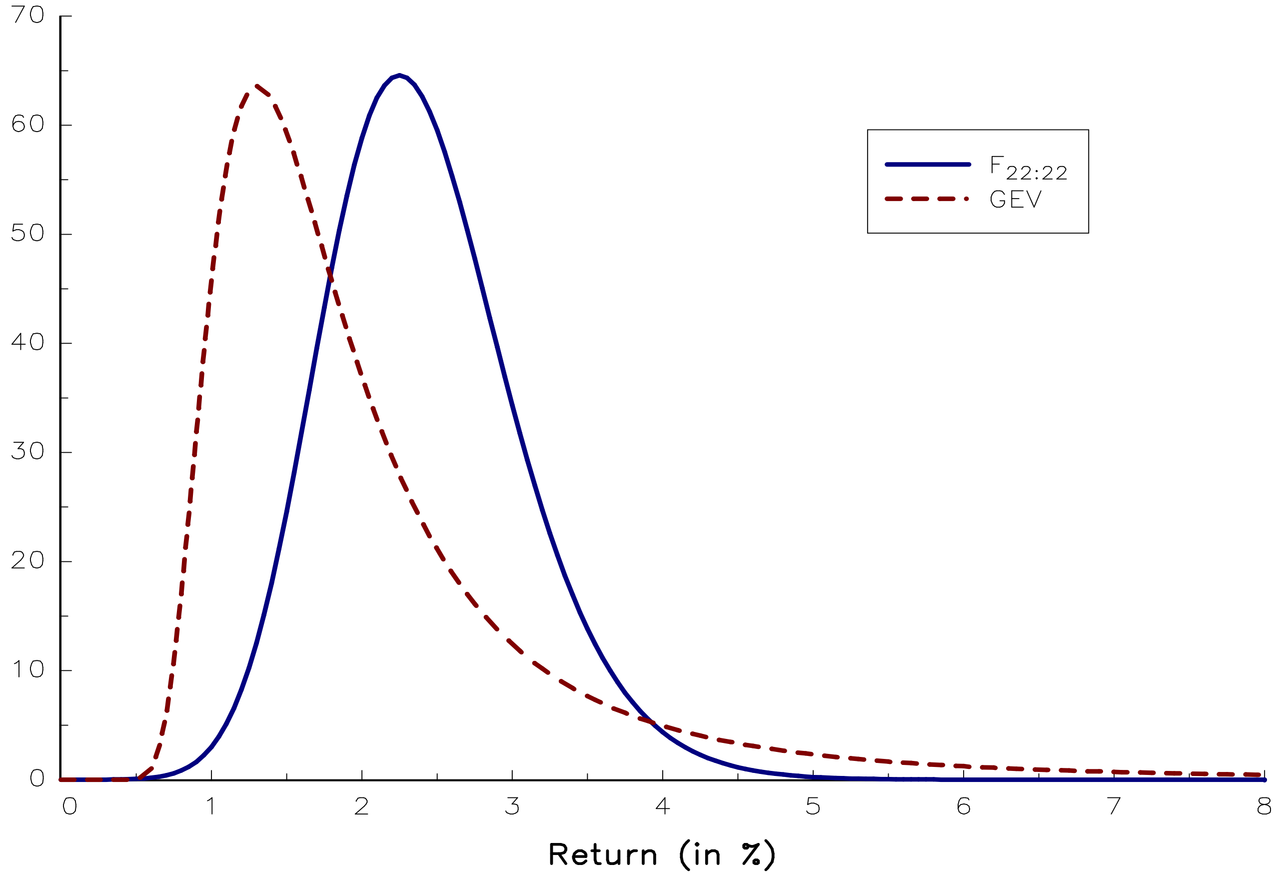


Figure: Probability density function of the maximum return $R_{22:22}$

Value-at-risk estimation

We recall that the P&L between t and $t+1$ is equal to:

$$\Pi(w) = P_{t+1}(w) - P_t(w) = P_t(w) \cdot R(w)$$

We have:

$$\mathrm{VaR}_\alpha(w) = -P_t(w) \cdot \hat{\mathbf{F}}^{-1}(1-\alpha)$$

We now estimate the GEV distribution $\hat{\mathbf{G}}$ of the maximum of $-R(w)$ for a period of n trading days. The confidence level must be adjusted in order to obtain the same return time:

$$\frac{1}{1-\alpha} \times 1 \text{ day} = \frac{1}{1-\alpha_{\mathrm{GEV}}} \times n \text{ days} \Leftrightarrow \alpha_{\mathrm{GEV}} = 1 - (1-\alpha) \cdot n$$

It follows that the value-at-risk is equal to:

$$\mathrm{VaR}_\alpha(w) = P(t) \cdot \hat{\mathbf{G}}^{-1}(\alpha_{\mathrm{GEV}}) = P(t) \cdot \left(\hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left(1 - \left(-\ln \alpha_{\mathrm{GEV}}\right)^{-\hat{\xi}} \right) \right)$$

because we have $\mathbf{G}^{-1}(\alpha) = \mu - \frac{\sigma}{\xi}\left(1 - (-\ln \alpha)^{-\xi}\right)$

Value-at-risk estimation

Table: Comparing Gaussian, historical and GEV value-at-risk measures

VaR	α	Long US	Long EM	Long US Short EM	Long EM Short US
Gaussian	99.0%	2.88%	2.83%	3.06%	3.03%
	99.5%	3.19%	3.14%	3.39%	3.36%
	99.9%	3.83%	3.77%	4.06%	4.03%
Historical	99.0%	3.46%	3.61%	3.37%	3.81%
	99.5%	4.66%	4.73%	3.99%	4.74%
	99.9%	7.74%	7.87%	6.45%	7.27%
GEV	99.0%	2.64%	2.61%	2.72%	2.93%
	99.5%	3.48%	3.46%	3.41%	3.82%
	99.9%	5.91%	6.05%	5.35%	6.60%

Expected shortfall estimation

We use the peak over threshold approach (HFRM, pages 773-777)

Extreme value copulas

Definition

An extreme value (EV) copula satisfies the following relationship:

$$\mathbf{C}\left(u_1^t, \ldots, u_n^t\right) = \mathbf{C}^t\left(u_1, \ldots, u_n\right)$$

for all $t > 0$

Extreme value copulas

The Gumbel copula is an EV copula:

$$
\begin{aligned}
\mathbf{C}\left(u_1^t, u_2^t\right) &= \exp\left(-\left(\left(-\ln u_1^t\right)^\theta + \left(-\ln u_2^t\right)^\theta\right)^{1/\theta}\right) \\
&= \exp\left(-\left(t^\theta\left(\left(-\ln u_1\right)^\theta + \left(-\ln u_2\right)^\theta\right)\right)^{1/\theta}\right) \\
&= \left(\exp\left(-\left(\left(-\ln u_1\right)^\theta + \left(-\ln u_2\right)^\theta\right)^{1/\theta}\right)\right)^t \\
&= \mathbf{C}^t\left(u_1, u_2\right)
\end{aligned}
$$

Extreme value copulas

The Farlie-Gumbel-Morgenstern copula is not an EV copula:

$$\begin{aligned}
\mathbf{C}\left(u_1^t, u_2^t\right) &= u_1^t u_2^t + \theta u_1^t u_2^t \left(1 - u_1^t\right)\left(1 - u_2^t\right) \\
&= u_1^t u_2^t \left(1 + \theta - \theta u_1^t - \theta u_2^t + \theta u_1^t u_2^t\right) \\
&\neq u_1^t u_2^t \left(1 + \theta - \theta u_1 - \theta u_2 + \theta u_1 u_2\right)^t \\
&\neq \mathbf{C}^t\left(u_1, u_2\right)
\end{aligned}$$

Extreme value copulas

Show that:

- $\mathbf{C}^{+}$ is an EV copula
- $\mathbf{C}^{\perp}$ is an EV copula
- $\mathbf{C}^{-}$ is not an EV copula

Multivariate extreme value theory

Let $X = (X_1, \ldots, X_n)$ be a random vector of dimension n. We note $X_{m:m}$ the random vector of maxima:

$$X_{m:m} = \begin{pmatrix} X_{m:m,1} \\ \vdots \\ X_{m:m,n} \end{pmatrix}$$

and $\mathbf{F}_{m:m}$ the corresponding distribution function:

$$\mathbf{F}_{m:m}(x_1, \ldots, x_n) = \Pr \{X_{m:m,1} \leq x_1, \ldots, X_{m:m,n} \leq x_n\}$$

The multivariate extreme value (MEV) theory considers the asymptotic behavior of the non-degenerate distribution function $\mathbf{G}$ such that:

$$\lim_{m \to \infty} \Pr \left(\frac{X_{m:m,1} - b_{m,1}}{a_{m,1}} \leq x_1, \ldots, \frac{X_{m:m,n} - b_{m,n}}{a_{m,n}} \leq x_n \right) = \mathbf{G}(x_1, \ldots, x_n)$$

Multivariate extreme value theory

Using Sklar's theorem, there exists a copula function $\mathbf{C}\langle\mathbf{G}\rangle$ such that:

$$\mathbf{G}(x_1,\ldots,x_n) = \mathbf{C}\langle\mathbf{G}\rangle(\mathbf{G}_1(x_1),\ldots,\mathbf{G}_n(x_n))$$

We have:

- The marginals $\mathbf{G}_1,\ldots,\mathbf{G}_n$ satisfy the Fisher-Tippet theorem
- $\mathbf{C}\langle\mathbf{G}\rangle$ is an extreme value copula

Remark

An extreme value copula satisfies the PQD property:

$$\mathbf{C}^{\perp} \prec \mathbf{C} \prec \mathbf{C}^{+}$$

Tail dependence of extreme values

We can show that the (upper) tail dependence of $\mathbf{C}\langle\mathbf{G}\rangle$ is equal to the (upper) tail dependence of $\mathbf{C}\langle\mathbf{F}\rangle$:

$$\lambda^{+}(\mathbf{C}\langle\mathbf{G}\rangle)=\lambda^{+}(\mathbf{C}\langle\mathbf{F}\rangle)$$

$\Rightarrow$ Extreme values are independent if the copula function $\mathbf{C}\langle\mathbf{F}\rangle$ has no (upper) tail dependence

Advanced topics

- Maximum domain of attraction
  - Univariate extreme value theory (HFRM, pages 765-770)
  - Multivariate extreme value theory (HFRM, pages 779 and 781-782)
- Deheuvels-Pickands representation (HFRM, pages 779-781)
- Generalized Pareto distribution $\mathcal{GPD}(\sigma, \xi)$ (HFRM, pages 773-777)

Exercises

- Copulas
  - Exercise 11.5.5 – Correlated loss given default rates
  - Exercise 11.5.6 – Calculation of correlation bounds
  - Exercise 11.5.7 – The bivariate Pareto copula

- Extreme value theory
  - Exercise 12.4.2 – Order statistics and return period
  - Exercise 12.4.4 – Extreme value theory in the bivariate case
  - Exercise 12.4.5 – Maximum domain of attraction in the bivariate case

References

NELSEN, R.B. (2006)
An Introduction to Copulas, Second edition, Springer.

RONCALLI, T. (2020)
Handbook of Financial Risk Management, Chapman and Hall/CRC Financial Mathematics Series, Chapters 11 and 12.

RONCALLI, T. (2020)
Handbook of Financial Risk Management – Companion Book, Chapters 11 and 12.