Dependence Analysis via Copulas

Nikolai Kolev, Leandro Ferreira and Rafael Aguilera

IME-USP

Contact e-mails: kolev.ime@gmail.com
ferreira.laf@gmail.com and aguilera@ime.usp.br

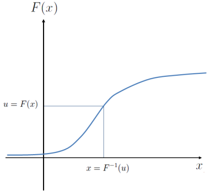
Outline

- Introduction
- Copula: Definition, Sklar's Theorem, Examples, Simulation
- Dependence Measures
- Copula Families
- Estimation and Model Fitting
- Real data analysis: Option pricing via Copulas

Continuous univariate distributions

Let X be a **continuous** random variable with distribution $F(x) = P(X \leq x)$.

1. $F(x)$ is **non-decreasing**;
2. **Inverse** function is given by $F^{-1}(u) = \inf_{x \in \Re} \{F(x) \geq u\}$, $u \in [0,1]$, which is **non-decreasing** as well;



Density function: $f(x) = \frac{d}{dx} F(x)$.

1. $f(x) \geq 0$ such that $\int_{-\infty}^{\infty} f(u)\, du = 1$;
2. $F(x) = \int_{-\infty}^{x} f(u)\, du$;
3. $P(a < X \leq b) = \int_{a}^{b} f(u)\, du = F(b) - F(a)$.

Continuous uniform distribution in $[0, 1]$

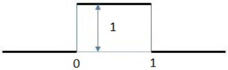
- The distribution function *of* $U(0, 1)$ is given by

$$F(x) = \begin{cases} 0, & \text{if} \quad x < 0, \\ x, & \text{if} \quad x \in [0, 1], \\ 1, & \text{elsewhere;} \end{cases}$$

- The density function of $U(0, 1)$ is

$$f(x) = \begin{cases} 1, & \text{if} \quad x \in [0, 1], \\ 0, & \text{elsewhere;} \end{cases}$$

- If $U \sim U(0, 1)$ then $E[U] = 1/2$ and $Var[U] = 1/12$.



Probability integral transform 1

- Let X be a continuous random variable with distribution function $F(x)$. The relation $U = F(X)$ is denominated **probability integral transform**.
- The distribution function of random variable $U = F(X)$ is

$$P(F(X) \leq u) = P(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u.$$

- Thus, $U \overset{d}{=} F(X) \sim U(0,1) \;\Leftrightarrow F^{-1}(U) \overset{d}{=} X$.
- Therefore, random variables $F^{-1}(U)$ and X **share the same distribution.**

Probability integral transform 2

Probability integral transform

Given a continuouis random variable X with distribution function $F(x)$, the random variable $U = F(X) \sim U(0,1)$. Moreover, $X \stackrel{d}{=} F^{-1}(U) \sim F(x)$.

This result is useful for **simulating** continuous random variables with **known distribution** function $F(x)$ using the **standard uniform** random numbers generator.

"Discrete" probability integral transform

- Let X be a **discrete** random variable defined by $p_i = P(X = x_i) \geq 0$, $\sum p_i = 1$, then $F(X)$ is not $U(0,1)$-distributed.

In fact, if X is given by

X	x_1	x_2	x_3	$\cdots$
Prob.	p_1	p_2	p_3	$\cdots$

with distribution function $F(x) = P(X \leq x) = \sum_{j \leq i} p_j$, for $x_i \leq x < x_{i+1}$. Then $Y = F(X)$ is discrete random variable

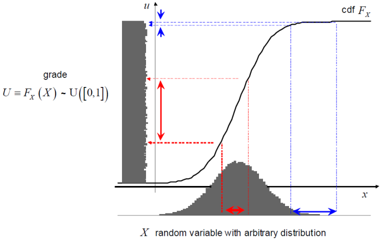
$Y = F(X)$	p_1	$p_1 + p_2$	$p_1 + p_2 + p_3$	$\cdots$
Prob.	p_1	p_2	p_3	$\cdots$

which is not $U(0,1)$, but

$$E[F(X)] = 1/2 + 1/2 \sum_{i=1}^{k} p_i^2 \xrightarrow{k \to \infty} 1/2 \text{ which is the mean of } U(0,1),$$

$$Var[F(X)] = 1/12 + f\left(\sum_{i=1}^{k} p_i^2\right) \xrightarrow{k \to \infty} 1/12 \text{ which is the variance of } U(0,1)$$

Graphical interpretation



- Note that the graph of $F(x)$ is **steeper** in the interval (in red) where there are more potential outcomes of the random variable X. This interval spreads out over a wider interval within $[0, 1]$;
- On the other hand, we observe an inverse effect in the interval where $F(x)$ is **flatter** (in blue), i.e., where there are less potential outcomes of the random variable X;
- Nevertheless, the probability integral transform **defects** the density (heavy tails and kurthosis disappear).

Simulating a logistic distribution

- If X follows a **logistic distribution** with parameters $\beta \geq 0$ and $\mu \in \Re = (-\infty, \infty)$, its distribution function is given by

$$F(x) = \frac{1}{1 + \exp(-\frac{x-\mu}{\beta})}, \quad x \in \Re = (-\infty, \infty);$$

- The inverse $F^{-1}(\cdot)$ of F can be found as a solution of

$$u = F(x) \Rightarrow x = F^{-1}(u) = \mu - \beta \ln(u^{-1} - 1);$$

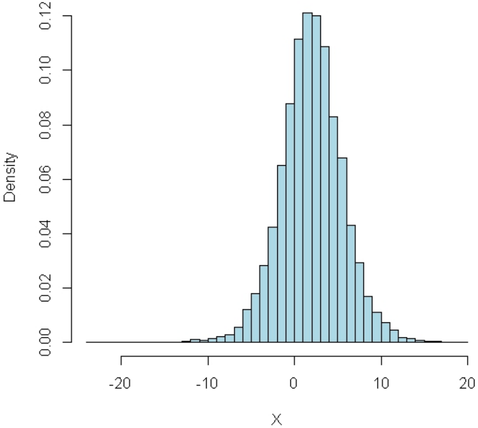
- **Simulating** logistic distribution with parameters β and μ:
  1. Using a standard uniform random numbers generator, generate $u \in [0, 1]$;
  2. Calculate $x = F^{-1}(u) = \mu - \beta \ln(u^{-1} - 1)$, which is the required observation (since $X \stackrel{d}{=} F^{-1}(U)$).

Commands in *R*

- We will simulate a sample with size $N = 10000$ of a logistic distribution with parameters $\mu = \beta = 2$.
- The commands in *R* are:

```
# Simulation of a logistic distribution

N = 10000
U = runif(N)
mi = 2
beta = 2
X = mi-beta*log((1/U)-1)
hist(X,30,freq=FALSE, main="")
```

Histogram of logistic distribution ($\mu = \beta = 2$)



Continuous bivariate distributions

- Joint distribution of (X, Y): $H(x, y) = P(X \leq x, Y \leq y)$;
- Marginal distributions:

$$F(x) = \lim_{y \to \infty} H(x, y) \text{ and } G(y) = \lim_{x \to \infty} H(x, y);$$

- Density function: $h(x, y) = \frac{\partial^2}{\partial x \partial y} H(x, y) \geq 0$, satisfying $\int_{u=-\infty}^{\infty} \int_{v=-\infty}^{\infty} h(u, v) du dv = 1$;
- Marginal density functions: $f(x) = \int_{-\infty}^{\infty} h(x, u) du = \frac{d}{dx} F(x)$ and $g(y) = \int_{-\infty}^{\infty} h(u, y) du = \frac{d}{dy} G(y)$.

Copula: definition and Sklar's Theorem

Definition (Bivariate copula)

A bivariate copula is a bivariate distribution function $C : [0,1]^2 \to [0,1]$, with standard uniform marginal distributions, i.e., $C(u,v) = P(U \leq u, V \leq v)$, where U and $V \sim U(0,1)$.Therefore, $C(u,v)$ is non-decreasing in its arguments; $C(0,0) = 0$; $C(1,1) = 1$.

Sklar's Theorem (Sklar, 1959)

Let $H(x,y)$ be a bivariate distribution function with marginal distributions $F(x)$ and $G(y)$. Then there exists a copula $C : [0,1]^2 \to [0,1]$ such that

$$H(x,y) = C(F(x), G(y)),$$

for all $(x,y) \in [-\infty, \infty]^2$. If $F(x)$ and $G(y)$ are **continuous**, then C is **unique**; otherwise, C is uniquely determined on $RanX \times RanY$.

Obtaining the copula $C(u, v)$ from $H(x, y)$

- From relations $F(x) = u$ and $G(y) = v$, where $u, v \in [0, 1]$, we obtain $x = F^{-1}(u)$ and $y = G^{-1}(v)$.
- Substituting $x = F^{-1}(u)$ and $y = G^{-1}(v)$ in $H(x, y) = C(F(x), G(y))$, we get the copula $C(u, v)$.

Obtaining copula $C(u, v)$ from joint distribution $H(x, y)$

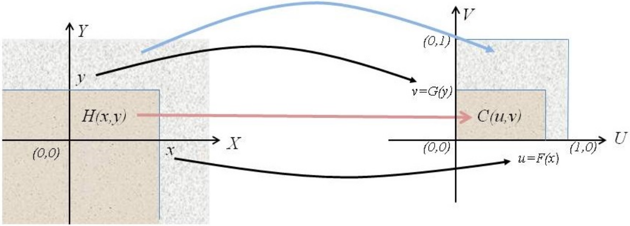
Given a bivariate distribution function $H(x, y)$, the corresponding copula is

$$C(u, v) = H\left(F^{-1}(u), G^{-1}(v)\right),$$

for all $(u, v) \in [0, 1]^2$.

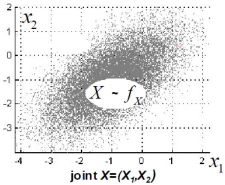
Graphical interpretation 1

$$H(x,y) = P(X \le x, Y \le y) \qquad H(x,y) = C(F(x), G(y))$$
$$F(x) = P(X \le x),\ G(y) = P(Y \le y) \quad C(u,v) = H(F^{-1}(u), G^{-1}(v))$$
$$U \stackrel{d}{=} F(X) \sim U(0,1), \qquad V \stackrel{d}{=} G(Y) \sim U(0,1)$$



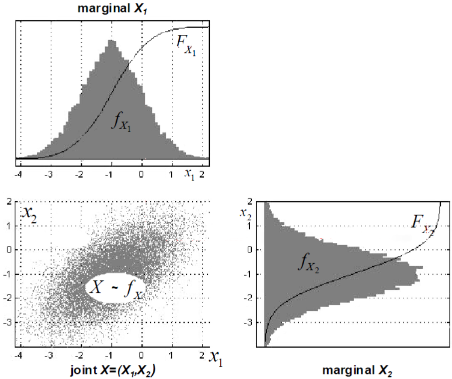
Graphical interpretation 1

- Starting from the joint density $h(x, y)$ we can obtain the marginal densities by $f(x) = \int_{-\infty}^{\infty} h(x, u)du$ and $g(y) = \int_{-\infty}^{\infty} h(u, y)du$.



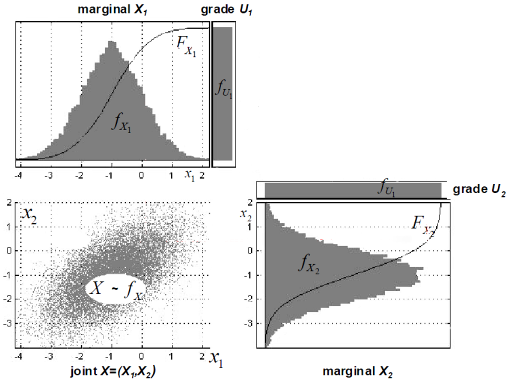
Graphical interpretation 1

- The marginal distribution functions are given by $F(x) = \int_{-\infty}^{x} f(u)\, du$ and $G(y) = \int_{-\infty}^{y} g(v)\, dv$.



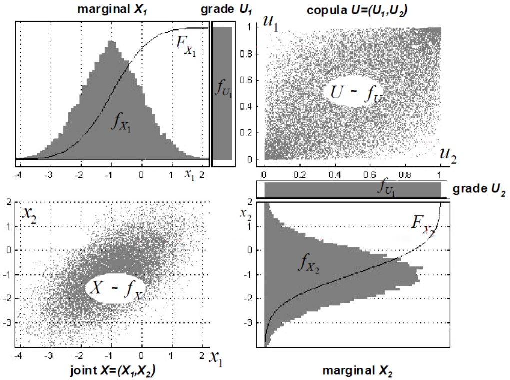
Graphical interpretation 1

- From the marginal distributions we have $U = F(X)$ and $V = G(Y)$, which are uniformly distributed in $[0, 1]$.

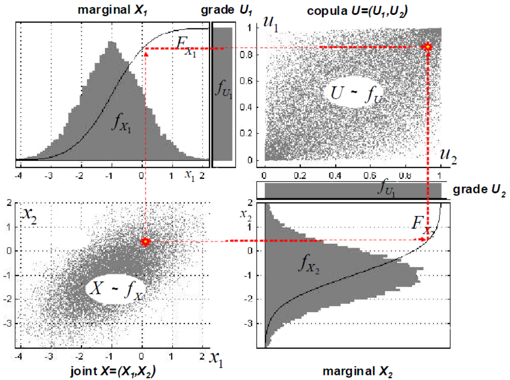


Graphical interpretation 1

- The joint distribution of (U, V) is copula $C(u, v) = P(U \leq u, V \leq v)$.



Graphical interpretation 1



Several reliable R Packages for Copulas

Package	**Title**
copula	Multivariate dependence with copula
copulaedas	Estimation of distribution, Algorithms based on copulas
CDVine	Statistical inference of C- and D-vine copulas
qcmr	Gaussian copula marginal regression
nacopula	Nested Archimedean copulas
qrm	Quantitative risk managment

Graphical interpretation 2: Commands in R

Generating a sample of size 10000 of:

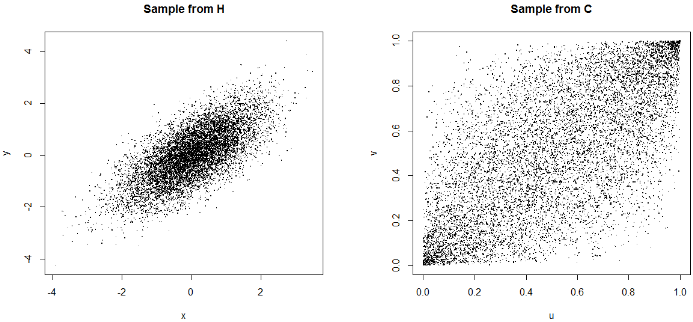
- a bivariate normal distribution $H(x, y)$ with standard normal $N(0, 1)$ marginals and correlation coefficient ρ has density

$$\phi_{2,\rho}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}[x^2 + y^2 - 2\rho xy]\right);$$

- corresponding copula $C(u, v)$ via Sklar's theorem with $\rho = 0.7$.

```
library(mvtnorm)
#Step 1: Generating a sample from H
c1<-c(1,.7);c2<-c(.7,1);R=cbind(c1,c2)
sample <- rmvnorm(n=10000, mean=c(0,0), sigma=R)
plot(sample[,1],sample[,2],xlab="x",ylab="y",pch=".",cex=1.5,main="Sample from H
    ")

#Step 2: Generating a sample from C via Sklar's Theorem
sample.copula=pnorm(sample)
plot(sample.copula,xlab="u",ylab="v",pch=".",cex=1.5,main="Sample from C")
```

Standard bivariate Normal distribution and its copula



Scatterplot of a sample of size 10000

Comments

- Copula contains **all the information about the dependence structure independent of marginal influence** since

$$C(F(x), G(y)) = H(x, y);$$

- Copulas enable us to model marginal distributions and the dependence structure **separately**;
- Copulas provide modelling **flexibility**: given a copula we can obtain many multivariate distributions by selecting different marginal distributions;
- **Any** bivariate distribution can be used to construct a copula: $C(u, v) = H\left(F^{-1}(u), G^{-1}(v)\right)$.

Example 1 - symmetric bivariate Gumbel distribution

- **Symmetric** $(H(x, y) = H(y, x))$ bivariate Gumbel distribution

$$H(x, y) = [1 + \exp(-x) + \exp(-y)]^{-1},$$

for all $x, y \in [-\infty, +\infty]$;

- The marginal distribution of X is $F(x) = \lim_{y \to \infty} H(x, y) = [1 + \exp(-x)]^{-1}$;
- The inverse $F^{-1}(u)$ is the solution of $u = F(x)$, i.e.

$$F(x) = [1 + \exp(-x)]^{-1} = u \;\Rightarrow\; \exp(-x) = \frac{1}{u} - 1 = \frac{1-u}{u}.$$

and therefore,

$$x = -\ln\left(\frac{1-u}{u}\right) = F^{-1}(u).$$

Example 1 - copula of bivariate Gumbel distribution

- By analogy, $y = -\ln\left(\frac{1-v}{v}\right) = G^{-1}(v)$;
- In $H(x,y) = [1+\exp(-x)+\exp(-y)]^{-1}$ we substitute $x = F^{-1}(u)$ and $y = G^{-1}(v)$ to get

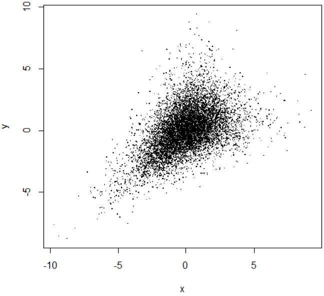
$$\begin{aligned} C(u,v) &= H\left(F^{-1}(u), G^{-1}(v)\right) \\ &= \left\{1+\exp\left[\ln\left(\tfrac{1-u}{u}\right)\right]+\exp\left[\ln\left(\tfrac{1-v}{v}\right)\right]\right)\right\}^{-1}; \end{aligned}$$

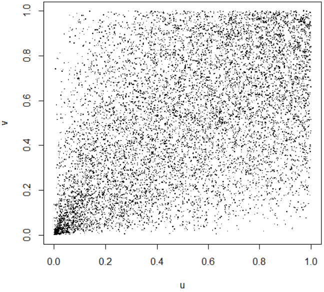
- Thus, the copula of bivariate symmetric Gumbel distribution is

$$C(u,v) = \left\{1+\frac{1-u}{u}+\frac{1-v}{v}\right\}^{-1} = \frac{uv}{u+v-uv}.$$

Example 1 - Graphs of bivariate Gumbel distribution

Bivariate Gumbel distribution and its copula





Scatterplot of a sample of size 10000

Example 2 - copula of asymmetric distribution

- Consider the **asymmetric** distribution $(H(x,y) \neq H(y,x))$

$$H(x,y) = \begin{cases} \frac{(x+1)[\exp(y)-1]}{x+2\exp(y)-1}, & \text{if} \quad (x,y) \in [-1,1] \times [0,\infty], \\ 1-\exp(-y), & \text{if} \quad (x,y) \in (1,\infty) \times [0,\infty], \\ 0, & \text{elsewhere.} \end{cases} \tag{1}$$

- *The distribution function of the marginal variable X is*

$$F(x) = \begin{cases} 0, & \text{if} \quad x < -1, \\ \frac{x+1}{2}, & \text{if} \quad x \in [-1,1], \\ 1, & \text{elsewhere,} \end{cases}$$

i.e., $X \sim U(-1,1)$. Therefore, $x = F^{-1}(u) = 2u - 1$.

Example 2 - copula of asymmetric distribution

- The distribution function of the random variable Y is

$$G(y) = \begin{cases} 0, & \text{if} \quad y < 0, \\ 1 - \exp(-y), & \text{elsewhere,} \end{cases}$$

i.e., $Y \sim Exp(1)$ and $y = G^{-1}(v) = -\ln(1 - v)$;
- Substituting solutions $x = F^{-1}(u) = 2u - 1$ and $y = G^{-1}(v) = -\ln(1 - v)$ in $H(x, y)$, given by (1), we obtain

$$C(u, v) = \left\{1 + \frac{1-u}{u} + \frac{1-v}{v}\right\}^{-1} = \frac{uv}{u + v - uv}.$$

Conclusion: symmetric and asymmetric distributions with the same copula, i.e. having the same dependence structure ???

- The **symmetric** bivariate Gumbel logistic distribution

$$H(x,y) = [1+\exp(-x)+\exp(-y)]^{-1}, \quad \text{for all } x,y \in [-\infty,+\infty]$$

and the **asymmetric** distribution

$$H(x,y) = \begin{cases} \frac{(x+1)[\exp(y)-1]}{x+2\exp(y)-1}, & \text{if} \quad (x,y) \in [-1,1]\times[0,\infty], \\ 1-\exp(-y), & \text{if} \quad (x,y) \in (1,\infty)\times[0,\infty], \\ 0, & \text{elsewhere} \end{cases}$$

share the **same** copula $C(u,v) = H(F^{-1}(u), G^{-1}(v)) = \frac{uv}{u+v-uv}$;

- Mathematically correct, but **confusing**! Believe in to the same dependence structure.

Fréchet-Hoeffding bounds for the joint distribution $H(x, y)$

- The bounds for the distribution function $H(x, y)$ are given by

$$\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y));$$

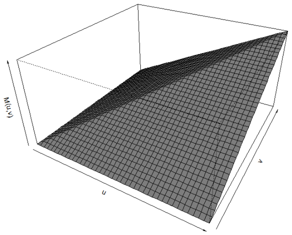
- In **absence of information about genuine dependence**, the joint distribution can be **bounded** by functions of marginals;
- These bounds can also be written in terms of copulas as

$$\max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v),$$

  (use the relations $F(x) = u$, $G(y) = v$ and $C(u, v) = H(F^{-1}(u), G^{-1}(v)))$;
- These bounds can **be sharper under additional** information (about the value of correlation coefficient, for example).

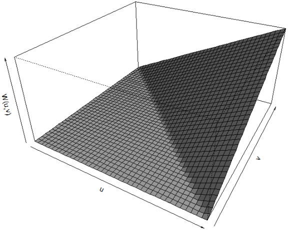
Comonotonic copula $M(u, v)$

- The **upper** Fréchet-Hoeffding bound is the copula $M(u, v) = P(U \leq u, V \leq v) = \min(u, v)$, i.e. $U = V$ almost surely and U and V are called **comonotonic** (meaning that they possess the **highest possible positive dependence**);
- The graph of the copula $M(u, v) = \min(u, v)$ is given below.



Countermonotonic copula $W(u, v)$

- The **lower** Fréchet-Hoeffding bound is the copula $W(u, v) = P(U \leq u, V \leq v) = \max(u + v - 1, 0)$;
- In this case, $U = 1 - V$ almost surely, and U and V are named **countermonotonic**, meaning that U and V exhibit **the extreme possible negative dependence**;
- The graph of the copula $W(u, v) = \max(u + v - 1, 0)$ is given below.

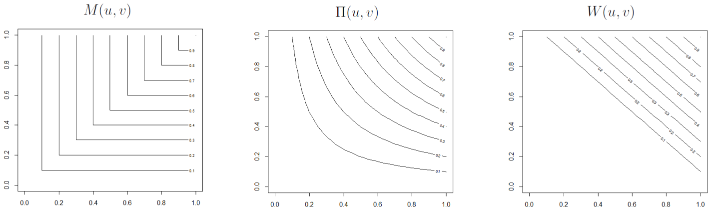


Independent copula and level curves

- The copula representing the **independence** structure between U and V is given by
  $\Pi(u, v) = P(U \leq u, V \leq v) = P(U \leq u)P(V \leq v) = uv$;
- The independent copula $\Pi(u, v)$ **characterizes** the independence between U and V;
- Let $H_1(x, y)$ and $H_2(x, y)$ have the **same** marginal distributions $F(x)$ and $G(y)$
  - If X and Y are independent, then $H_i(x, y) = F(x)G(y)$, i=1,2 ;
  - But the product $F(x)G(y)$ does not characterize **independence** uniquelly.

Level curves

- Some **level curves** are presented for the copulas M, Π and W, i.e., curves such that $C(u, v) = a = constant$ in $[0, 1]$.



Copula invariance under increasing transformation

Let X and Y be continuous random variables and let $C_{XY}(u,v)$ be its respective copula.

Copula invariance

If $\alpha(x)$ and $\beta(y)$ are strictly increasing functions in $DomX$ and $DomY$, then

$$C_{\alpha(X),\beta(Y)}(u,v) = C_{XY}(u,v),$$

i.e., $C_{XY}(u,v)$ **is invariant under strictly increasing transformation of** X **and** Y.

Invariance proof

Proof.

Denote by F_1, G_1, F_2 and G_2 the distribution functions of X, Y, $\alpha(X)$ and $\beta(Y)$, respectively.
Since $\alpha(x)$ and $\beta(y)$ are strictly increasing functions,
$F_2(x) = P[\alpha(X) \leq x] = P[X \leq \alpha^{-1}(x)] = F_1(\alpha^{-1}(x))$.
Analogously, $G_2(y) = G_1(\beta^{-1}(y))$.
Therefore, for all (x, y) in $\Re^2$ we have

$$
\begin{aligned}
C_{\alpha(X),\beta(Y)}(F_2(x), G_2(y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\
&= P[X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)] \\
&= C_{XY}(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))) = C_{XY}(F_2(x), G_2(y))
\end{aligned}
$$

Since X and Y are continuous, $DomF_2 = DomG_2 = [0, 1]$.
Therefore, $C_{\alpha(X),\beta(Y)}(u, v) = C_{X,Y}(u, v)$ in $[0, 1]^2$. $\square$

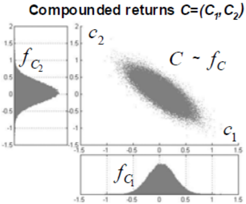
Example 3a: linear and compounded returns

- Consider the **linear** (R_n) and **compound** (C_n) returns of prices ($P_{n,t}$) between times t and $t+1$ for the **same stocks**

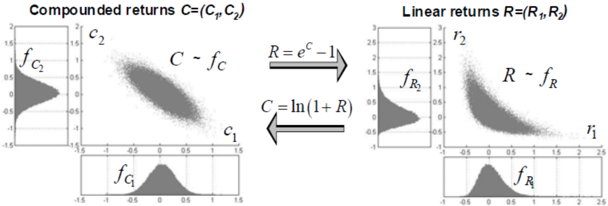
$$R_n \equiv \frac{P_{n,t+1}}{P_{n,t}} - 1, \quad C_n \equiv ln(\frac{P_{n,t+1}}{P_{n,t}});$$

- These two types of returns, although calculated on the same stock prices, are different: $R_n = e^{C_n} - 1$ and $C_n = ln(1 + R_n)$;
- For example, if **stock prices distribution** H_P is a **multivariate log-normal** distribution, **linear returns distribution** H_R follow a **multivariate shifted lognormal** distribution and **compounded returns** H_C follow a **multivariate normal** distribution.

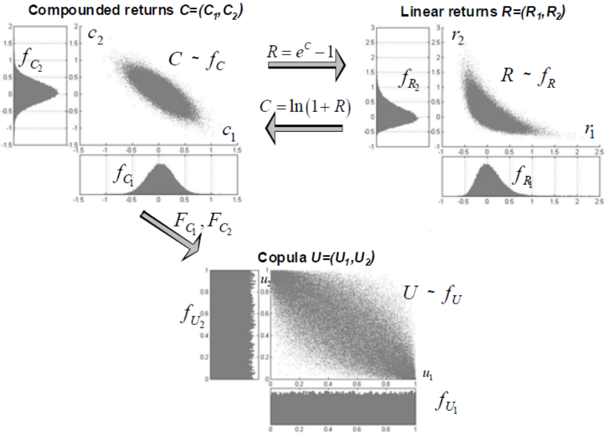
Example 3a: graphical interpretation



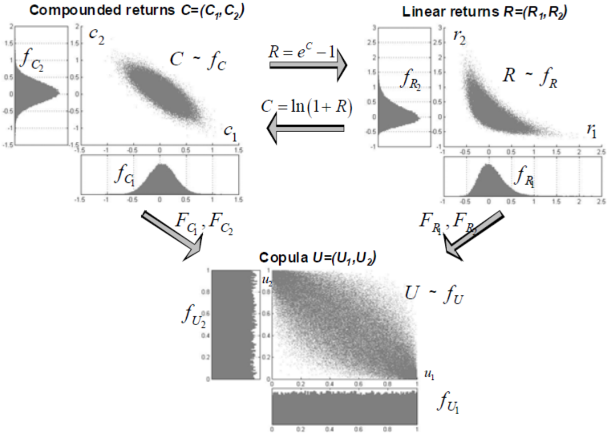
Example 3a: graphical interpretation



Example 3a: graphical interpretation



Example 3a: graphical interpretation



Example 3a: conclusions

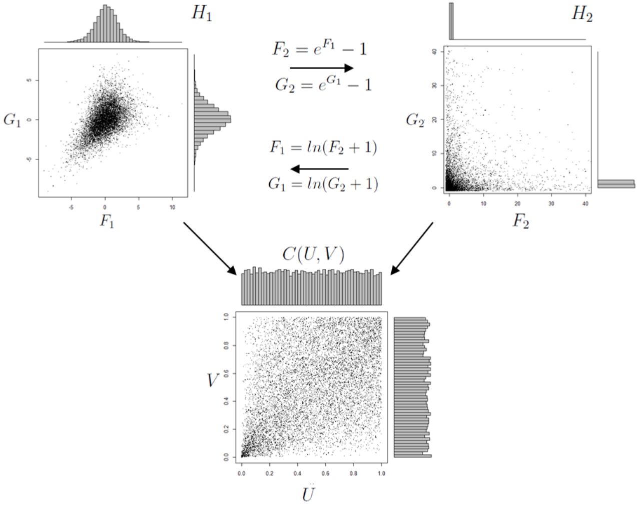
- Since $R_n = e^{C_n} - 1$ is an **increasing** transformation $\alpha(C_n)$ and $C_n = ln(1 + R_n)$ is as **increasing** transformation $\beta(R_n)$ then $C_{R_n,C_n}(u, v) = C_{\beta(R_n),\alpha(C_n)}(u, v)$;
- Thus, the copula that joins the linear returns and the copula that joins the compounded returns is the **same**.

Example 3a: Commands in R for bivariate Gumbel

Let $H_1(x, y)$ be a bivariate Gumbel distribution with copula $C(u, v) = \frac{uv}{u+v-uv}$, applying the increasing transformations $e^x - 1$ for marginals obtaining a new distribution $H_2(x, y)$ with the same copula $C(u, v)$.

```
scatterhist = function(x, y, xlab="", ylab="",xl=NULL,yl=NULL,m=NULL){
plot(x, y,xlab=xl, ylab=yl,pch=".",cex=1.5,main=m)
hist(x, freq=FALSE,xlab=xl,main=xl)
hist(y, freq=FALSE,xlab=yl,main=yl)
}

#Getting a sample from H1
H1=sample.gumbel(10000)
scatterhist(H1[,1], H1[,2],xlab="F1", ylab="G1")

#Getting a sample from Copula using marginals from H1
H1.copula = 1/(1+exp(-H1))
scatterhist(H1.copula[,1], H1.copula[,2],xlab="U", ylab="V")

#Getting a sample from H2 via transformation
H2 = exp(H1) - 1
scatterhist(H2[,1], H2[,2],xl=c(0,50),yl=c(0,50),xlab="F2", ylab="G2")

#Getting a sample from Copula using marginals from H2
H2.copula=(H2+1)/(H2+2)
scatterhist(H2.copula[,1], H2.copula[,2], xlab="U", ylab="V")
```

Example 3a: Graphs for bivariate Gumbel



Copula of increasing/decreasing transformations

- Analogously, we have the following relations:
  1. If $\alpha(x)$ is **strictly increasing** in *DomX* and $\beta(y)$ is **strictly decreasing** in *DomY*, then $C_{\alpha(X),\beta(Y)}(u,v) = u - C_{XY}(u,1-v)$;
  2. If $\alpha(x)$ is **strictly decreasing** in *DomX* and $\beta(y)$ is **strictly increasing** in *DomY*, then $C_{\alpha(X),\beta(Y)}(u,v) = v - C_{XY}(1-u,v)$;
  3. If $\alpha(x)$ and $\beta(y)$ are **strictly decreasing** in *DomX* and *DomY*, then $C_{\alpha(X),\beta(Y)}(u,v) = u + v - 1 + C_{XY}(1-u,1-v)$.

Conditional copula

$$
\begin{aligned}
c_u(v) &= P(V \le v | U = u) \\
&= \lim_{\Delta u \to 0} P(V \le v | u \le U \le u + \Delta u) \\
&= \lim_{\Delta u \to 0} \frac{P(V \le v, U \le u + \Delta u) - P(V \le v, U \le u)}{P(u \le U \le u + \Delta u)} \\
&= \lim_{\Delta u \to 0} \frac{C(u + \Delta u, v) - C(u, v)}{\Delta u} \\
&= \frac{\partial}{\partial u} C(u, v) = C(v|u)
\end{aligned}
$$

Definition (Conditional copula)

Conditional copula in u: $c_u(v) = \frac{\partial C(u,v)}{\partial u} = P(V \le v | U = u)$ and conditional copula in v: $c_v(u) = \frac{\partial C(u,v)}{\partial v} = P(U \le u | V = v)$.

- Let us calculate $c_u(v)$ of copula $C(u, v) = \frac{uv}{u+v-uv}$.

$$
c_u(v) = \frac{\partial}{\partial u}\left(\frac{uv}{u+v-uv}\right) = \frac{v(u+v-uv) - uv(1-v)}{(u+v-uv)^2} = \frac{v^2}{(u+v-uv)^2}.
$$

Generating random variables using conditional copulas

$$
\begin{aligned}
P(U \le u, V \le v) &= \underbrace{P(U \le u)}_{u} \cdot \underbrace{P(V \le v | U = u)}_{c_u(v) = t} \\
&\Rightarrow \quad v = c_u^{-1}(t)
\end{aligned}
$$

Generating random variables using conditional copulas

To generate one observation of a given copula $C(u, v)$:

1. generate two standard uniform observations u and t;
2. fix $v = c_u^{-1}(t)$, where $c_u^{-1}(t)$ is inverse of conditional copula $c_u(t)$;
3. (u, v) is the required observation.

Example 3b

- To generate observation of copula $C(u,v) = \frac{uv}{u+v-uv}$: we proceed as follows:
  1. calculate $c_u(v) = \frac{\partial C(u,v)}{\partial u} = (\frac{v}{u+v-uv})^2$;
  2. generate two independent standard uniform random variables u and t ;
  3. from $t = c_u(v)$ obtain $v = c_u^{-1}(t) = \frac{u\sqrt{t}}{1-(1-u)\sqrt{t}}$;
  4. set $v = \frac{u\sqrt{t}}{1-(1-u)\sqrt{t}}$;
  5. (u,v) is the required observation of (U,V).

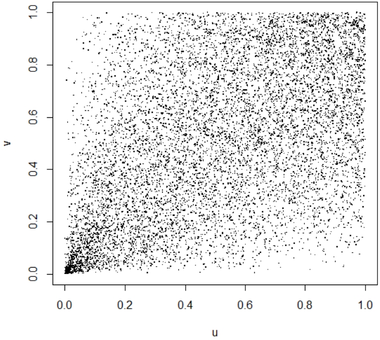
Example 3b: Commands in R

The R code of previously algorithm is given below.

```
#set the size of sample
n <- 10000
#set uniforms
u <- runif(n)
t <- runif(n)

#fix v using the conditional copula inverse
v <- u
for(i in 1:n){
  v[i]=u[i]*sqrt(t[i])/(1-(1-u[i])*sqrt(t[i]))
}
#scatterplot of a sample (U,V) required
plot(u,v, xlab="U", ylab="V", pch = ".", cex = 1.5)
```

Example 3b: Graphs

Scatterplot of a copula $C(u,v) = \frac{uv}{u+v-uv}$ obtained by previously algorithm.



Sample of size 10000

Application: Monte Carlo integration using copulas

- **Aim**: to obtain the **expected value** of a continuous function $q(x, y)$ of a bivariate random vector (X, Y) having joint distribution $H(x, y)$, i.e.

$$E(q(X, Y)) = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} q(x, y) dH(x, y);$$

- Given the copula $C(u, v) = H(F^{-1}(v), G^{-1}(v))$ and marginal distributions $F(x) = \lim_{y \to \infty} H(x, y)$ and $G(y) = \lim_{x \to \infty} H(x, y)$, we can use the following algorithm to approximate the value of $E(q(X, Y))$:
  1. generate n observations of the bivariate random vector (X, Y);
  2. for each observation i, calculate $q_i = q(x_i, y_i)$, $i = 1, 2, ..., n$;
  3. $E(q(X, Y)) \approx \frac{1}{n} \sum_{i=1}^{n} q_i$

Example 4

1. Let (X, Y) is the Gumbel bivariate distributed, i.e.
$$H(x, y) = [1 + \exp(-x) + \exp(-y)]^{-1}.$$
2. Marginal inverses are $F^{-1}(u) = -\ln\left(\frac{1-u}{u}\right)$ and $G^{-1}(v) = -\ln\left(\frac{1-v}{v}\right)$;
3. We intend to estimate $E(q(X, Y))$, where $q(x, y) = \sqrt{x^2 + y^2}$;
4. The algorithm is presented in the sequel.

Simulation for Example 4

1. For $i = 1$ to n do:
    1. generate two standard uniform random variables u_i and t_i;
    2. fix $v_i = \frac{u_i\sqrt{t_i}}{1-(1-u_i)\sqrt{t_i}}$;
    3. fix $x_i = -\ln\left(\frac{1-u_i}{u_i}\right)$ and $y_i = -\ln\left(\frac{1-v_i}{v_i}\right)$;
    4. calculate $q_i = \sqrt{x_i^2 + y_i^2}$
2. obtain $E(\sqrt{X^2 + Y^2}) \approx \frac{1}{n}\sum_{i=1}^{n} q_i$.

Commands in R for Example 4

```
#Example 4

n = 1000
q = 0
u = runif(n)
t = runif(n)
x = y = u

for(i in 1:n){
  v[i]=u[i]*sqrt(t[i])/(1-(1-u[i])*sqrt(t[i]))
  x[i]=-log((1-u[i])/u[i])
  y[i]=-log((1-v[i])/v[i])
  q=q+sqrt(x[i]^2+y[i]^2)
}
E=q/n #Estimation of expected value
E
```

The Result is $E\left(\sqrt{X^2+Y^2}\right) \approx 2.127499.$

Dependence measures

- We will present **four dependence measures** between two random variables X and Y: the Pearson linear correlation coefficient and its **local version**, Kendall's tau $\tau(X, Y)$, Spearman's rho $\rho(X, Y)$ and Blest's measure of rank correlation $\nu(X, Y)$;
- The measures $\tau(X, Y)$, $\rho(X, Y)$ and $\nu(X, Y)$ depend only on the copula $C(u, v)$ corresponding to (X, Y). Therefore, their values **do not change under strictly increasing transformations** of X and Y (since copula is **time invariant**).

Pearson linear correlation coefficient

- Correlation coefficient should be used with **caution** when working outside the class of elliptical distributions. It is defined by

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}};$$

- $Corr(X, Y)$ is defined only when we have **finite variances** and it measures **linear dependence** and assumes values in the interval $[-1, 1]$;
- If two random variables X and Y are **independent**, then $Corr(X, Y) = 0$. The **inverse statement is not always true**.

Pearson linear correlation coefficient - some pitfalls

- The Pearson linear correlation coefficient is **invariant under strictly increasing linear transformations**, i.e.,
$$Corr(X, Y) = Corr(a_1X + b_1, a_2Y + b_2);$$
- $Corr(X, Y) \neq Corr(\alpha(X), \beta(Y))$, for **monotone increasing non-linear** functions $\alpha(x)$ and $\beta(y)$.

Pearson linear correlation coefficient - some pitfalls

- **It is not true** that given two marginal distributions $F(x)$ and $G(y)$ and a value for the Pearson linear correlation coefficient it is **always possible** to obtain a bivariate distribution with these characteristics (the statement is valid for elliptical world);
- We know that

$$\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y)),$$

$$W(x, y) \leq H(x, y) \leq M(x, y);$$

- The following relations are valid

$$r_{min} = r_W \leq r_H = Corr(X, Y) \leq r_M = r_{max}.$$

Example - McNeil et al. (2005)

- Consider two **log-normally** distributed random variables X and Y, i.e., $\ln X \sim \mathcal{N}(0,1)$ and $\ln Y \sim \mathcal{N}(0,\sigma^2)$;
- It is important to note that if $\sigma^2 \neq 1$ then X and Y **are not of the same type**, i.e., do not exist real constants a and b such that $X =^d a + bY$, i.e., X and Y are neither comonotonic nor countermonotonic when $\sigma^2 \neq 1$.

Example - McNeil et al. (2005)

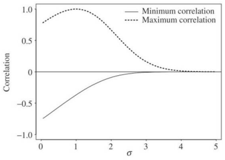
- The maximum ($r_{max}(X, Y)$) and minimum ($r_{min}(X, Y)$) values that $Corr(X, Y)$ may assume in this case are given by the following expressions

$$r_{max}(X, Y) = \frac{\exp(\sigma) - 1}{\sqrt{(e-1)(\exp(\sigma^2) - 1)}} \underset{\sigma\to\infty}{\to} 0$$

  and

$$r_{min}(X, Y) = \frac{\exp(-\sigma) - 1}{\sqrt{(e-1)(\exp(\sigma^2) - 1)}} \underset{\sigma\to\infty}{\to} 0;$$

- The graph below illustrates the maximum and minimum values that $Corr(X, Y)$ may assume as a function of the parameter σ.



Comments

- Note how both limits tend fast to 0 as σ increases.
- The graph shows that we may have **comonotonic** random variables (maximally positive dependent) exhibiting values of linear correlation coefficient close to 0;
- Since comonotonicity is the **strongest** form of positive dependence, this **example provides a correction** to the usual view that **small correlation imply weak dependence**;
- Therefore, the concept of Pearson linear correlation coefficient is **meaningless** unless applied in the context of elliptical world.

Local Pearson linear correlation coefficient

While the Pearson linear correlation coefficient

$$\rho(X,Y) = \frac{E[(X - E[X])(Y - E[Y])]}{\sqrt{E[(X - E[X])^2]E[(Y - E[Y])^2]}}$$

is a number in $[-1,1]$, the **local Pearson linear correlation coefficient**

$$\rho_{local}(x,y) = \frac{E[(X - E(X|Y=y))(Y - E(Y|X=x))]}{\sqrt{E(X - E(X|Y=y))^2 E(Y - E(Y|X=x))^2}}$$

is a surface depending of (x,y) and $\rho_{local}(x,y) \in [-1,1]$.

Kendall's tau

Definition: Kendall's tau

The **population version of Kendall's tau**, $\tau(X, Y)$, for the bivariate random vector (X, Y) is defined as the difference between the probabilities of **concordance** and **discordance**, i.e.,

$$\tau(X, Y) = P[(X - X^{'})(Y - Y^{'}) > 0] - P[(X - X^{'})(Y - Y^{'}) < 0],$$

where $(X^{'}, Y^{'})$ is an **independent** copy of (X, Y).

Definition: sample version of Kendall's tau

Let (X_1, Y_1), ..., (X_n, Y_n) be a sample of (X, Y). Denote by R_i and S_i the ranks in the sets $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, respectively, $1 \leq i \leq n$. The **sample version** of Kendall's tau, τ_n, is given by

$$\tau_n = \frac{2}{n^2 - n} \sum_{1 \leq i < j \leq n} sign(R_i - R_j) sign(S_i - S_j).$$

Kendall's tau

Theorem

Let (X, Y) be a vector of continuous random variables with copula $C(u, v)$. Then the **population Kendall's tau** is given by

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) \ - 1,$$

Note that if $U, V \sim U(0, 1)$, then

$$\tau(X, Y) = 4E[C(U, V)] - 1.$$

Spearman's rho

Definition: Spearman's rho, $\rho(X, Y)$

The **population Spearman's rho**, $\rho(X, Y)$, for the vector (X, Y) is

$$\rho(X, Y) = 3P[(X - X^{'})(Y - Y^{'}) > 0] - P[(X - X^{'})(Y - Y^{''}) < 0],$$

where (X, Y), $(X^{'}, Y^{'})$ and $(X^{''}, Y^{''})$ are **independent** copies of (X, Y) and $X^{'}$ and $Y^{''}$ are **independent**.

Definition: sample version Spearman's rho, ρ_n

The **sample version** ρ_n of Spearman's rho is

$$\rho_n = \frac{12}{n^3 - n} \sum_{1=1}^{n} R_i S_i - \frac{3(n+1)}{n-1}.$$

Spearman's rho

Theorem

Let (X, Y) be a vector of continuous random variables with copula $C(u, v)$. Then **population Spearman's rho** for (X, Y) is given by

$$\rho(X, Y) = 12 \int_0^1 \int_0^1 uvdC(u, v) \, - 3 = 12 \int_0^1 \int_0^1 C(u, v) dudv \, - 3.$$

Theorem

Let (X, Y) be a vector of **continuous** random variables, $X \sim F(x)$, $Y \sim G(y)$, $U =^d F(X) \sim U(0, 1)$ and $V =^d G(Y) \sim U(0, 1)$. Then

$$\rho(X, Y) = \frac{Cov(U, V)}{\sqrt{Var(U)Var(V)}} = Corr(F(X), G(Y)).$$

Spearman's rho

Theorem

Let (X, Y) be a vector of continuous random variables with copula $C(u, v)$. Then the measure Spearman's rho for (X, Y) is given by

$$\rho(X, Y) = 12 \int_0^1 \int_0^1 [C(u, v) - uv] dudv.$$

- This result provides a **geometric interpretation** for the coefficient $\rho(X, Y)$: it is proportional to the **volume between the surfaces of copula** $C(u, v)$ and **independence copula** $\Pi(u, v) = uv$.

The measures of Kendall, Spearman and Pearson

- The classical non-parametric Kendall's tau and Spearman's rho are **preferable** dependence measures than $Corr(X, Y)$, since they are **invariant under increasing variable** transformations (since the corresponding copula is invariant);
- If X and Y are continuous random variables with copula $C(u, v)$, then

$$C(u, v) = M(u, v) = \min(u, v) \Leftrightarrow \tau(X, Y) = \rho(X, Y) = 1;$$

$$C(u, v) = W(u, v) = \max(u+v-1, 0) \Leftrightarrow \tau(X, Y) = \rho(X, Y) = -1.$$

Blest's measure of rank correlation, ν

- The **sample versions** of Kendalls tau and Spearman's rho,

$$\tau_n = \frac{2}{n^2 - n} \sum_{1 \leq i < j \leq n} sign(R_i - R_j) sign(S_i - S_j)$$

and

$$\rho_n = \frac{12}{n^3 - n} \sum_{1=1}^{n} R_i S_i - \frac{3(n+1)}{n-1},$$

attribute the **same importance to the difference between** the ranks $R_i - S_i$, $i = 1, \ldots, n$;

Idea: The correlation in the pairs (R_i, S_i) provides an idea of consistency between two ranks and the **difference between two extreme ranks should be emphasized**. Thus, Blest (2000) proposed an alternative non-parametric correlation measure of ranks

Blest's measure of rank correlation, ν

Definition: sample version of Blest's measure of rank correlation, ν_n

The **sample version** of Blest's measure of rank correlation, ν_n, is

$$\nu_n = \frac{2n+1}{n-1} - \frac{12}{n^2-n}\sum_{1=1}^{n} S_i \left(1 - \frac{R_i}{n+1}\right)^2.$$

Definition: Population version of Blest's measure of rank correlation, ν

The **population** Blest's measure of rank correlationfor the vector (X, Y) is

$$\begin{aligned}\nu(X, Y) &= 2 - 12\int_{\Re^2}\left[1 - F^2(x)\right] G(y) dH(x, y)\\ &= 2 - 12\int_{[0,1]^2}(1-u)^2 v dC(u, v),\end{aligned}$$

where $\nu(X, Y) \in [0, 1]$.

Blest's measure of rank correlation, ν

- The **extreme values** of ν_n occur when we have $(R_i = S_i)$ or $(R_i = n + 1 - S_i)$;
- Meanwhile is valid the property $\nu(X, -Y) = -\nu(X, Y)$;
- Explicit expressions for $\nu(X, Y)$ can be obtained for several bivariate distributions only.

Example. Suppose (X, Y) follows a bivariate standard normal distribution with correlation coefficient r. Then $\nu(X, Y) = \rho(X, Y) = \frac{6}{\pi} \arcsin(\frac{r}{2})$, while $\tau(X, Y) = \frac{2}{\pi} \arcsin(r)$.

Gini measure

Population Kendall's tau: $\tau(X, Y) = 4E[C(U, V)] - 1 = 4E[K(Z)] - 1.$
Population Spearman's rho: $\rho(X, Y) = 12 \int_0^1 \int_0^1 [C(u, v) - uv] dudv,$

i.e. both measures serve for **ordinal variables** X and Y (since copula is invariant on increasing transformation).

- But, when one variable is **ordinal** (say X) and the other is **nominal** (say Y), then $\rho(X, Y)$ and $\tau(X, Y)$ are **not** appopriate.
- In this case, one can use **Gini measure** $\Gamma(X, Y)$. If can be shown that $Cov[Y, F(X)] = E[Y \cdot F(X)] - \frac{E[Y]}{2}$,

$$\Gamma(X, Y) \overset{def}{=} \frac{Cov[Y, F(X)]}{Cov[Y, G(Y)]},$$

where

$$Cov[Y, G(Y)] = \frac{1}{4} E[(Y_1 - Y_2) sign(Y_1 - Y_2)] = \frac{1}{4} E[Y_1 - Y_2],$$

with Y_1 and Y_2 being **independent** copies of Y.

Coefficients of tail dependence

Definition: Upper tail dependence coefficient, λ_U

Let X and Y be two continuous random variables with copula $C(u, v)$. The **upper tail dependence coefficient** λ_U between X and Y is a property of the copula $C(u, v)$. It is defined by

$$\lambda_U = \lim_{u \to 1^-} \frac{\overline{C}(u, u)}{1 - u} = P(U \geq u | V \geq u) = \frac{1 - 2u + C(u, v)}{1 - u},$$

provided the limit exists and belongs to the interval $[0, 1]$.
$\overline{C}(u, v)$ is the survival copula, given by

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

- If $\lambda_U \in (0, 1]$, then X and Y display **upper tail dependence**, or **extreme dependence in the upper tail**;
- If $\lambda_U = 0$, X and Y are **asymptotically independent** in the upper tail.

Coefficients of tail dependence

Definition: Lower tail dependence, λ_L

Let X and Y be two continuous random variables with copula $C(u, v)$. The **lower tail dependence coefficient** λ_L between X and Y is defined by

$$\lambda_L = P(U \leq u | V \leq u) = \lim_{u \to 0^+} \frac{C(u, u)}{u},$$

provided the limit exists and belong to the interval $[0, 1]$.

- If $\lambda_L \in (0, 1]$, then X and Y display **lower tail dependence**, or **extreme dependence in the lower tail**;
- If $\lambda_L = 0$, X and Y are **asymptotically independent** in the lower tail.

Archimedean copulas

Definition (Archimedean copulas)

A copula $C(u, v)$ belongs to the Archimedean family if

$$C(u, v) = \varphi(\varphi^{-1}(u) + \varphi^{-1}(v)) \text{ with } (u, v) \in [0, 1]^2,$$

for continuous, positive non-increasing and convex functions $\varphi : [0, \infty) \to [0, 1]$ such that $\varphi(0) = 1$. The function $\varphi(.)$ is denominated **generator function** of the copula $C(u, v)$.

- Typical examples of Archimedean family are the Clayton, Frank and Gumbel copulas presented in the sequel.

Clayton copula

- The **Clayton copula** is given by

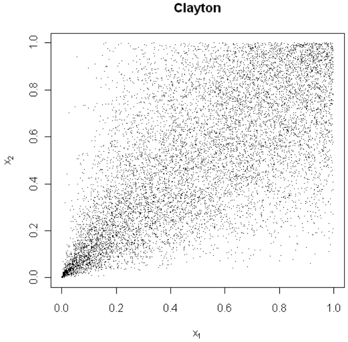
$$C_\alpha(u,v) = \max[(u^{-\alpha} + v^{-\alpha} - 1)^{-\frac{1}{\alpha}}, 0];$$

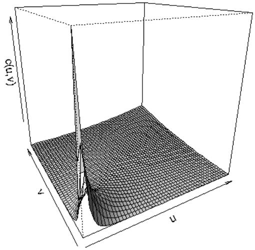
- The generator function is $\varphi(t) = \frac{1}{\alpha}(t^{-\alpha} - 1)$, where $\alpha \in [-1, \infty) \setminus \{0\}$;
- The relation between its parameter α and corresponding Kendall's tau $\tau = 4\int_0^1 \int_0^1 C(u,v)dC(u,v) - 1$ is given by $\alpha = \frac{2\tau}{1-\tau}$;
- In this case, $\lambda_L = 2^{-\frac{1}{\alpha}}$, for $\alpha > 0$. So, **Clayton copula displays lower tail dependence**, which tends to 1 as $\alpha \to \infty$.

Clayton copula - commands in R

```
library(copula)

# Clayton copula
cc <- claytonCopula(2)
sample <- rCopula(10000,cc)
#Scatterplot
plot(sample, xlab="U", ylab="V", pch = ".", cex = 1.5)

#Density
persp(cc,dCopula, xlab="u", ylab="v", zlab="c(u,v)")
```

Clayton copula - scatterplot and copula density





Clayton copula with parameter $\alpha = 2$

Frank copula

- The **Frank copula** is given by

$$C_\alpha(u,v) = -\frac{1}{\alpha}\ln\left\{1 + \frac{(\exp(-\alpha u)-1)(\exp(-\alpha v)-1)}{\exp(-\alpha)-1}\right\};$$

- The generator function is $\varphi(t) = -ln\left\{\frac{(\exp(-\alpha t)-1)}{\exp(-\alpha)-1}\right\}$, where $\alpha \in (-\infty,\infty)\backslash\{0\}$;
- The relation between its parameter α and Kendall's tau (τ) is

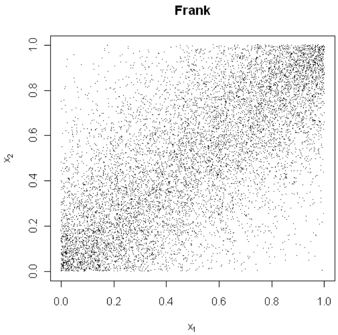
$$\frac{D_1(\alpha)-1}{\alpha} = \frac{1-\tau}{4}, \text{ where } D_1(\alpha) = \frac{1}{\alpha}\int_0^\alpha \frac{t}{\exp(t)-1}dt$$

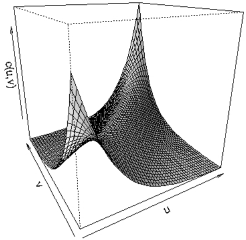
($D_1(\alpha)$ is the Debye function of the first kind).

Frank copula - commands in R

```
library(copula)

#Frank Copula
fr <- frankCopula(10)
sample <- rCopula(10000,fr)
#Scatterplot
plot(sample, xlab="U", ylab="V", pch = ".", cex = 1.5)

#Density
persp(fr,dCopula, xlab="u", ylab="v", zlab="C(u,v)", shade=.0001)
```

Frank copula - scatterplot and copula density





Frank copula with parameter $\alpha = 10$

Gumbel-Hougard copula

- The **Gumbel-Hougard** copula is given by

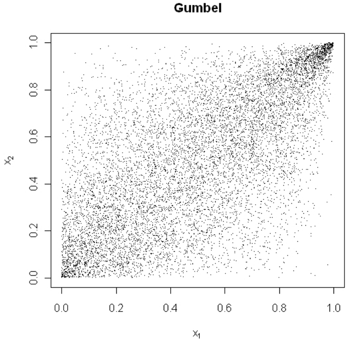
$$C_\alpha(u, v) = \exp\left\{ -[(-\ln u)^\alpha + (-\ln v)^\alpha]^{\alpha^{-1}} \right\};$$

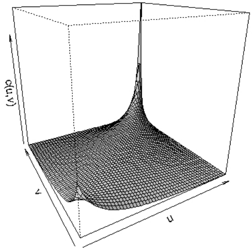
- The generator function is given by $\varphi(t) = (-\ln t)^\alpha$, where $\alpha \in [1, \infty)$;
- The relation between its parameter α and Kendall's τ is $\alpha = \frac{1}{1-\tau}$;
- $\lambda_U = 2 - 2^{\frac{1}{\alpha}}$. If $\alpha > 1$, the copula displays **upper tail dependence**. This dependence tends to 1 as $\alpha \to \infty$, (what is to be expected, since in this situation, the Gumbel-Hougard copula tends to the comonotonic copula).

Gumbel copula - commands in R

```
library(copula)

# Gumbel copula
gu <- gumbelCopula(4)
sample <- rCopula(10000,gu)
#Scatterplot
plot(sample, xlab="U", ylab="V", pch = ".", cex = 1.5)

#Density
persp(gu,dCopula, xlab="u", ylab="v", zlab="C(u,v)")
```

Gumbel copula - scatterplot and copula density





Gumbel-Hougard copula with parameter $\alpha = 4$

Standard univariate normal distribution

- The distribution function of a random variable X that follows a standard normal $N(0,1)$ distribution is given by

$$\Phi(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-t^2/2)dt$$

Bivariate normal distribution

- The density function of a random vector (X, Y) that follows a **standard bivariate normal** distribution with correlation coefficient ρ is given by

$$\phi_{2,\rho}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[x^2 + y^2 - 2\rho xy\right]\right),$$

  where $-\infty < x < \infty$, $-1 \leq \rho \leq 1$ and $X, Y \sim N(0, 1)$;
- The corresponding joint distribution function is given by

$$\Phi_{2,\rho}(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} \phi_{2,\rho}(u, v) du dv.$$

Univariate $t-$Student distribution

- A random variable η follows the Student$-t$ distribution with ν degrees of freedom whenever it can be written as $\eta \stackrel{d}{=} \frac{X}{\sqrt{\frac{\xi}{\nu}}}$, $X \sim \mathcal{N}(0,1)$ and is independent of $\xi \sim \chi^2_\nu$;
- The density function is given by

$$t_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$

where $x \in \Re$ and $\nu > 0$.

Bivariate t-Student distribution

- A random vector $\mathbf{T} = (T_1, T_2)$ follows a **bivariate t-Student** distribution with ν **degrees** of freedom whenever it can be written as

$$(T_1, T_2) = \left(\frac{X}{\sqrt{\frac{\xi}{\nu}}}, \frac{Y}{\sqrt{\frac{\xi}{\nu}}} \right).$$

  The bivariate random vector (X, Y) has standard bivariate normal distribution with correlation coefficient ρ being independent of $\xi \sim \chi^2_\nu$;

- The joint density function is given by

$$t_{\nu,\rho}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \left\{ 1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)^2} \right\}^{-\frac{\nu+2}{2}},$$

  where $x, y \in (-\infty, \infty)$, $\rho \in [-1, 1]$.

Elliptical distributions

- The name "elliptical" comes from the **elliptical** (sum of squares) **form of the level curves** of the joint density function, i.e., $f(x, y) = a = constant$;
- The bivariate random vector $\mathbf{Z} = (Z_1, Z_2)$ follows a **spherical** distribution if and only if its characteristic function can be represented by $E[\exp(it^T Z)] = \psi(z_1^2 + z_2^2)$, $t \in \Re^2$, for some function $\psi : \Re \to \Re$;
- The bivariate random vector $\mathbf{W} = (W_1, W_2)$ follows a **elliptical** distribution if $\mathbf{W} = \mu + A\mathbf{Z}$, where $\mu \in \Re^2$, $A \in \Re^2 \times \Re^2$ and $\mathbf{Z}$ follows a spherical distribution;
- **Particular cases** of elliptical distributions are the bivariate normal as well as the bivariate t-Student distributions.

Elliptical copulas

- **Elliptical copulas** became very popular in finance and risk management because they are easily implemented. The ease in obtaining the marginal distribution functions is another advantage when one uses this elliptical copulas for forecast, see Frees e Wang (2005);
- "Elliptical" because they are associated with **a quadratic form of correlation coefficient between the marginals**. It means, the elliptical family of copulas is **symmetric**. The dependence structure is determined by the correlation matrix;
- The **Gaussian copula** and the t-**copula are particular cases of elliptical copulas**, with dispersion matrix inherited from the elliptical distributions (correlation coefficient ρ). The t-copula possesses an additional degrees of freedom parameter ν, which **modify the shape of the copula for given level of dependence** governed by ρ.

Bivariate Gaussian copula

- The bivariate Gaussian copula with parameter $\rho \in [-1, 1]$ is given by

$$C_\rho(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \phi_{2,\rho}(x, y) dxdy,$$

  where $\phi_{2,\rho}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[x^2 + y^2 - 2\rho xy\right]\right)$ and $\Phi^{-1}(\cdot)$ is the inverse of standard normal distribution;
- The relation between Kendall's tau and the correlation coefficient is $\rho = \sin(\frac{\pi}{2}\tau)$, i.e. $\tau = \frac{2}{\pi} \arcsin \rho$.

Tail dependence of Gaussian copula

Due to **symmetry** of Gaussian copula we have **equal** upper and lower tail dependence coefficients, i.e. $\lambda_U = \lambda_L = \lambda$ with

$$\lambda = 2 \lim_{x \to -\infty} \Phi\left(\frac{x\sqrt{1-\rho}}{\sqrt{1+\rho}}\right) = 0.$$

Interpretation: Independently of the value of the correlation coefficient, **asymptotically the Gaussian copula displays independence in both tails**, meaning that **regardless of how high a correlation** coefficient we choose, if we go far enough into the tail, **extreme events appear to occur independently in each margin**.

Density of a bivariate copula

To calculate the density $c(u,v)$ of a bivariate copula proceed as follows

$$c(u,v) = \frac{\partial^2}{\partial u\,\partial v} C(u,v) = \frac{\partial^2}{\partial u\,\partial v} H(F^{-1}(u), G^{-1}(v)) = \frac{h(F^{-1}(u), G^{-1}(v))}{f(F^{-1}(u))g(G^{-1}(v))}.$$

Therefore,

- If we **know the bivariate density** $h(x,y)$, then we can obtain $f(x)$, $F^{-1}(x)$, $g(y)$ and $G^{-1}(y)$, to calculate

$$c(u,v) = \frac{h(F^{-1}(u), G^{-1}(v))}{f(F^{-1}(u))g(G^{-1}(v))};$$

- If we **know copula density** $c(u,v)$ and the **marginal densities** $f(x)$ and $g(y)$, we can calculate $F(x)$ and $G(y)$. Thus,

$$h(x,y) = c(F(x), G(y))f(x)g(y).$$

Density of the bivariate Gumbel logistic copula

- The copula corresponding to the bivariate Gumbel logistic distribution is given by

$$C(u,v) = H(F^{-1}(u), G^{-1}(v)) = \frac{uv}{u+v-uv};$$

- Its density function is

$$c(u,v) = \frac{\partial^2}{\partial u \, \partial v}\left(\frac{uv}{u+v-uv}\right) = \frac{2uv}{(u+v-uv)^3}.$$

Density of bivariate Gaussian copula

- The density function $c_\rho(u, v)$ of a bivariate Gaussian copula is obtained by calculating $\frac{\partial^2}{\partial u \partial v} C_\rho(u, v)$;
- Therefore

$$c_\rho(u, v) = \frac{\partial^2}{\partial u \partial v} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \phi_{2,\rho}(x, y) dx dy;$$

- Thus, if bivariate Gaussian density is $\phi_{2,\rho}$, the corresponding copula density is given by

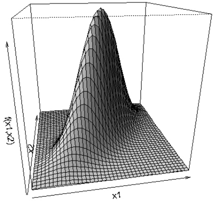
$$c_\rho(u, v) = \frac{\phi_{2,\rho}[\Phi^{-1}(u), \Phi^{-1}(v)]}{\phi[\Phi^{-1}(u)]\phi[\Phi^{-1}(v)]};$$

- If we know the copula density and the marginal densities then $h(x, y) = c(F(x), G(y))f(x)g(y)$ and we are able to reconstruct the bivariate Gaussian density function from the relation $\phi_{2,\rho}(x, y) = c_\rho(\Phi(x), \Phi(y))\phi(x)\phi(y)$.

R code to visualize the density functions of a standard bivariate normal distribution and its copula

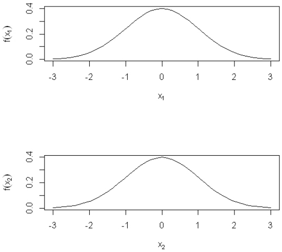
```
library(mvtnorm)
d=2;x = seq(-3,3,6*.025);x.cop = seq(0,1,.025)
#Covariance matrix
c1<-c(1,.7);c2<-c(.7,1);R=cbind(c1,c2)
dens = dens.cop = dens2 = dens3 = matrix(0,nrow=length(x),ncol=length(x))

#Calculating densities of distributions
for(i in 1:dim(dens)[1])
{for(j in 1:dim(dens)[2])
{dens[i,j] = dmvnorm(x=c(x[i],x[j]),mean=rep(0,d),sigma=R)
 dens.cop[i,j] = dmvnorm(x=c(qnorm(x.cop[i]),qnorm(x.cop[j])),mean=rep(0,d),
      sigma=R)/(dnorm(qnorm(x.cop[i]))*dnorm(qnorm(x.cop[j])))}}

#Density of a standard bivariate normal distribution
persp(x,x,dens,xlab="x",ylab="y",zlab="f(x,y)",shade = 0.75)
#Density of normal copula
persp(x.cop,x.cop, dens.cop,xlab="u",ylab="v",zlab="c(u,v)",shade = 0.75)
```

Graph of the standard bivariate normal density function



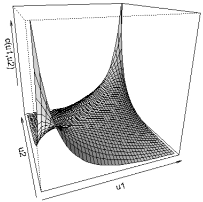
The density $\phi_{2,\rho}(x, y)$ of a standard bivariate normal distribution, with $\rho = 0.7$

Graphs of two standard univariate normal densities



divided by the product of the corresponding marginal standard normal densities $\phi(x)$ and $\phi(y)$ ...

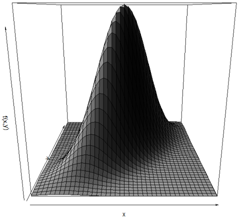
Graph of the Gaussian copula density

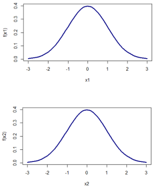


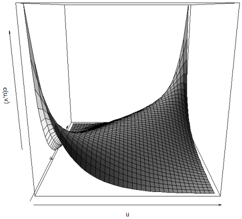
provides the Gaussian copula density

$$c_\rho(u,v) = \frac{\phi_{2,\rho}(\Phi^{-1}(u), \Phi^{-1}(v))}{\phi(\Phi^{-1}(u))\phi(\Phi^{-1}(v))}$$

Summarizing

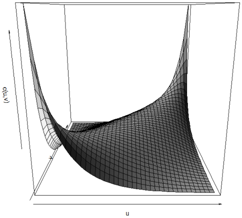


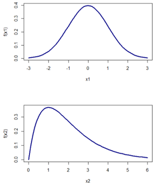


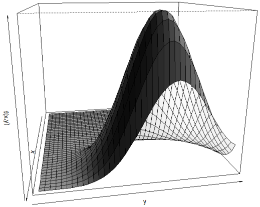


The standard bivariate normal density $\phi_{2,\rho}$ divided by the product of the corresponding marginal standard normal density functions results in the bivariate Gaussian copula density $c_\rho(u, v)$

Generating a density of new bivariate distribution







If we multiply the bivariate Gaussian copula density $c_\rho(u, v)$ by two arbitrarily density functions we will obtain a **new bivariate density** function: $h(x, y) = c_\rho(F(x), G(y))f(x)g(y)$. It keeps the dependence structure of the standard bivariate normal distribution but the marginal distributions are just $F(x)$ and $G(y)$

Generating a density of new bivariate distribution

Commands in R to generating a density of a new bivariate distribution with gamma(5,1) and $N(0,1)$ marginals

```
#Calculating densities of distributions
for(i in 1:dim(dens)[1])
  {for(j in 1:dim(dens)[2])
    {#using normal and normal as marginals
    dens2[i,j]=dens.cop[i,j]*(dnorm(x[i])*dnorm(x[j]))
    #using gamma and  normal as marginals
    dens3[i,j]=dens.cop[i,j]*(dgamma(x[i],shape=5,scale=1)*dnorm(x[j]))}}

#Density of the density using normal marginals
persp(x,x,dens2,xlab="x",ylab="y",zlab="f(x,y)")
#Density of the density using gamma and normal marginals
persp(x,x,dens3,xlab="x",ylab="y",zlab="f(x,y)",xlim=c(-3,3),ylim=c(-3,3),theta
     =80,shade = 0.75,expand=0.8)
```

Bivariate t-copula

- The bivariate t-copula is given by

$$C_{\nu,\rho}(u,v) = T_{\nu,\rho}(T_{\nu}^{-1}(u), T_{\nu}^{-1}(v)),$$

where $\rho \in [-1, 1]$, $T_{\nu}(.)$ is the univariate distribution function of a random variable that follow the t-Student distribution with ν degrees of freedom and $T_{\nu,\rho}(.,.)$ is the joint distribution function of a random bivariate vector $\mathbf{T} = (T_1, T_2)$ that follows a bivariate t-Student distribution with ν degrees of freedom;

- The corresponding bivariate t-copula density is

$$c_{\nu,\rho}(u,v) = \frac{t_{\nu,\rho}(T_{\nu}^{-1}(u), T_{\nu}^{-1}(v))}{t_{\nu}(T_{\nu}^{-1}(u))t_{\nu}(T_{\nu}^{-1}(v))}.$$

Tail dependence of t-copula

Due to the **symmetry** of t-copula for tail dependence coefficient we have

$$\lambda_U = \lambda_L = \lambda = 2T_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right),$$

when $T_{\nu+1}$ means the distribution function of a random variable that follows t-Student distribution with $\nu+1$ degrees of freedom.

Provided $\rho > -1$, **t-copula is asymptotically dependent** in **both tails**.

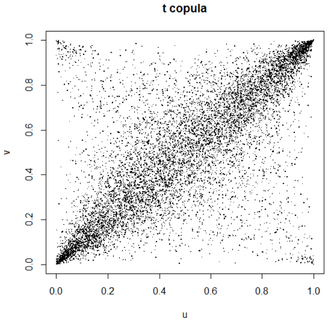
Bivariate t-copula

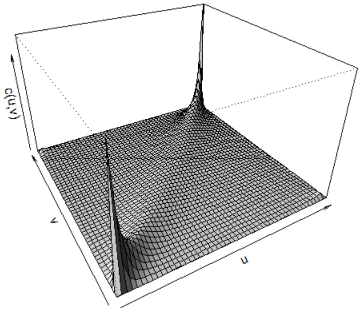
- The Kendall's tau $\tau = 4\int_0^1\int_0^1 C_{\nu,\rho}(u,v)dC_{\nu,\rho}(u,v) - 1$ is $\frac{2}{\pi}\arcsin(\rho)$, i.e. **the same expression** as the Gaussian copula with correlation coefficient ρ;
- Just like the univariate t-Student distribution, the **degrees of freedom ν control the weight in the tails, i.e., the smaller ν the heavier the tails** (modify the copula shape);
- A bivariate Gaussian copula with correlation coefficient ρ can be considered as the limiting case of a bivariate t-copula with the same parameter ρ, when $\nu \to \infty$.

Bivariate *t*-copula: Commands in R

```
library(copula)

#The bivariate t copula
tc<- tCopula(0.8, dim=2, dispstr = "un", df = 1)
sample <- rCopula(10000,tc)

#Scatterplot
plot(sample, xlab="u", ylab="v", pch = ".", cex = 1.5)

#Density
persp(tc,dCopula, xlab="u", ylab="v", zlab="c(u,v)",shade=0.35)
```

Bivariate t-copula - scatterplot and density





Bivariate t-copula with parameters $\rho = 0.8$ and $\nu = 1$

Other examples of copulas

Besides the families of copulas we have already seen there are **others that exhibit tail dependence** (frequently used in practice). We will consider

- Rotated Gumbel Copula (lower tail);
- Symmetrized Joe-Clayton Copula (upper and lower tail).

Rotated Gumbel copula

- The **Rotated Gumbel copula** is given by

$$C_{RG}(u, v|\alpha) = u + v - 1 + C_\alpha(1 - u, 1 - v|\alpha),$$

where $\alpha \in [1, +\infty)$ and C_α is the Gumbel-Hougard copula, which is given by

$$C_\alpha(u, v) = \exp\left\{ -[(-\ln u)^\alpha + (-\ln v)^\alpha]^{\alpha^{-1}} \right\};$$

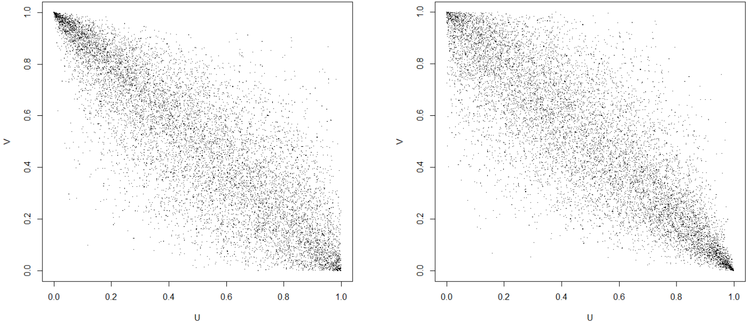
- This copula exhibits only **lower tail dependence**.

Rotated Gumbel copula: Commands in R

```
library("CDVine")
# simulate from a bivariate Rotated-Gumbel (90 degrees) copula
rg90 = BiCopSim(10000,24,-3)
plot(rg90,xlab="U", ylab="V",pch = ".",cex=1.5)

# simulate from a bivariate Rotated-Gumbel (270 degrees) copula
rg270 = BiCopSim(10000,34,-3)
plot(rg270,xlab="U", ylab="V",pch = ".",cex=1.5)
```

Rotated Gumbel copula - scatterplot

Rotated Gumbel copula with $\alpha = 90$ and $\alpha = 270$



Scatterplot of a sample of size 10000

Symmetrized Joe-Clayton copula

- The **Symmetrized Joe-Clayton copula** C_{SJC} is given by

$$C_{SJC}(u, v|\tau^U, \tau^L) = \frac{1}{2} \cdot [C_{JC}(u, v|\tau^U, \tau^L) + C_{JC}(1-u, 1-v|\tau^U, \tau^L) + u + v - 1],$$

where C_{JC} is the **Joe-Clayton copula** represented by

$$C_{JC}(u, v|\tau^U, \tau^L) = 1 - (1 - \{[1-(1-u)^{\kappa}]^{-\gamma} + [1-(1-v)^{\kappa}]^{-\gamma} - 1\}^{-1/\gamma})^{-1/\kappa},$$

with $\kappa = 1/\log_2(2-\tau^U)$, $\gamma = -1/\log_2(\tau^L)$ and $\tau^U, \tau^L \in (0, 1)$.

- The *SJC* has **both upper and lower tail dependence parameters**. Its own dependence parameters, τ^U and τ^L, are the measures of dependence of the upper and lower tail, respectively. Furthermore, τ^U and τ^L range freely and are not dependent on each other.

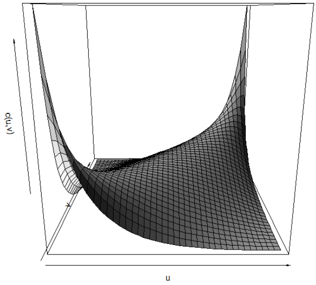
Power copula

- Let C be an arbitrary copula and define the **power copula** P_C as following

$$P_C(u, v) = u^{\theta_1} v^{\theta_2} C(u^{1-\theta_1}, v^{1-\theta_2}),$$

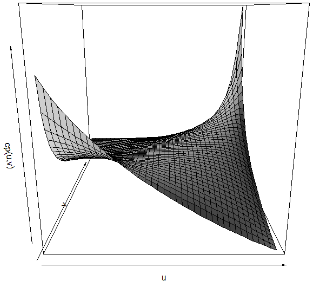
  where the parameters $\theta_1, \theta_2 \in [0, 1]$.
- When we choose $\theta_1 = \theta_2 = 0$ then $P_C(u, v) = C(u, v)$.
- In financial derivatives, for example, the parameters θ_1 and θ_2 control the **slope and the curvature of the implied volatility smile**.

Comparison of a Gaussian copula and Power Gaussian copula densities



Gaussian copula,

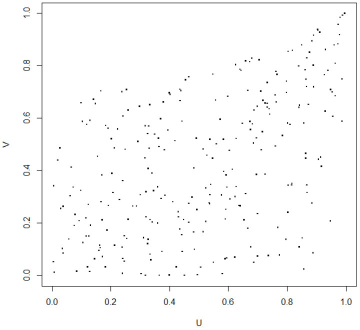
$\rho = 0.7$.



Power Gaussian copula,

$\rho = 0.7$, $\theta_1 = 0.7$ and $\theta_2 = 0.3$.

Scatterplot of Power Gaussian copula



Sampling of Power Gaussian copula,

$\rho = 0.7$, $\theta_1 = 0.7$ and $\theta_2 = 0.3$.

Example: Marshall-Olkin copula

If our base copula is

$$M(u, v) = \min(u, v),$$

then the resulting power copula is

$$P_M(u, v) = u^{\theta_1} v^{\theta_2} \min\left(u^{1-\theta_1}, v^{1-\theta_2}\right),$$

being the **Marshall-Olkin copula**.

Kendall distribution

Recall that, according to Probability Integral Transform

$$U = F(X) \sim U(0,1), \text{ i.e. } P(U \leq w) = w$$

- "Bivariate Prob. Int. Transform" is the **Kendall distribution**
  $$K(w) = P[H(X,Y) \leq w) = P[C(X,Y) \leq w), \quad w \in [0,1];$$
- $K(w)$ is **univariate summary** of dependence embodied in C;
- $K(w)$ depends only on the copula C associated with H, and hence not on the marginals F and G;
- $w \leq K(w) \leq 1, \quad w \in [0,1];$
- If U and V are independent the
  $$K(w) = P(UV \leq w) = w - w\log(w);$$
- **In general**, $K(w) \neq w$, i.e. $H(X,Y)$ is not $U(0,1)$.

Empirical copula and empirical Kendall distribution

- Let $(X_1, Y_1), \ldots, (X_n, Y_n)$, $n \geq 2$, be a random sample of a continuous distribution and let $X_{(i)}$ and $Y_{(j)}$ be the order statistics of the sample.
- The **empirical copula** C_n is defined as

$$C_n = \frac{1}{n} \left(\text{number of points } (X_m, Y_m) \text{ such that } X_m \leq X_{(i)} \text{ and } Y_m \leq Y_{(j)}\right).$$

An equivalent form of empirical copula is given by

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(u_i \leq u, v_i \leq v),$$

where $\mathbf{1}(\cdot)$ is the indicator function,

$$u_i = \frac{rank(X_i)}{n+1} = \frac{R_i}{n+1} \quad \text{and} \quad v_i = \frac{rank(Y_i)}{n+1} = \frac{S_i}{n+1};$$

- The **empirical Kendall distribution function** $K_n(t)$ for all t is given by

$$K_n(t) = \frac{1}{n} \left(\text{number of points } (X_m, Y_m) \text{ such that } C_n(u, v) \leq t\right).$$

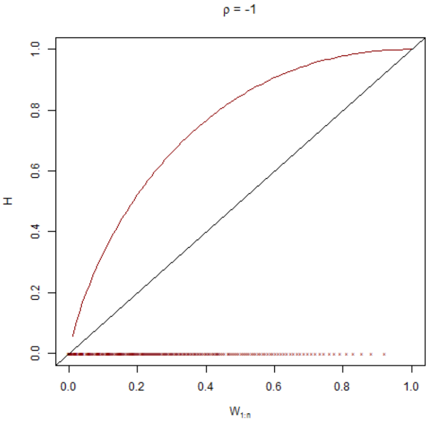
Construction of Kendall plot

- The Kendall plot (K-plot) is based on Kendall distribution function $K(w) = P(C(U,V) \leq w) = P(H(X,Y) \leq w)$ and it is analog of the classic Q-Q plot. It is introduced by Genest and Boies (2003) to evaluate the dependence structure in a bivariate sample. The construction follows the steps:
  1. Calculate $H_i = \frac{\#\{j \neq i : X_j \leq X_i, Y_j \leq Y_i\}}{n-1}$ for each $i = 1, \ldots, n$;
  2. Order the $H_i's$, such that $H_{(1)} \leq \ldots \leq H_{(n)}$;
  3. Plot the pairs $(W_{i:n}, H_{(i)})$, $1 \leq i \leq n$, where $W_{i:n}$ represents the expected value of the ith order statistic in a sample with size n of the distribution $K_0(w) = w - w\log(w)$, $0 \leq w \leq 1$ (Kendall distribution in independence case).

Kendall plot - commands in R

```r
library(CDVine)
library(MASS)

n<-250
nu<-c(0,0)
sigma1<-1
sigma2<-1
rho<-seq(-1,1,0.2)
for (i in 1:length(rho))
{
  S<-cbind(c(sigma1^2,rho[i]*sigma1*sigma2),c(rho[i]*sigma1*sigma2,sigma2^2))
  X<-mvrnorm(n=n, mu=nu, Sigma=S)
  vipt<-apply(X,2,rank)/n
  #Kendall Plot
  BiCopKPlot(vipt[,1],vipt[,2], family='B')
}
```

Animation of Kendall plot



Tail copula

Remind that

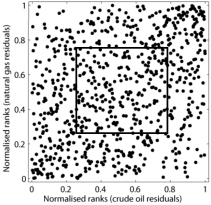
$$\lambda_L = \lim_{u \to 0^+} \frac{C(u,u)}{u} \text{ and } \lambda_U = \lim_{u \to 0^+} \frac{\overline{C}(u,u)}{u}$$

can be regarded as **directional derivatives** of $C(u,v)$. Assuming different directions (uq, vq), we get **lower tail copulas.** The lower tail copula $\Lambda_L(u,v)$ and upper tail copula $\Lambda_U(u,v)$ associated with X and Y is a function of their copula $C(u,v)$ or survival copula $\overline{C}(u,v)$, respectively and are defined by

$$\Lambda_L(u,v) = \lim_{q \to 0^+} \frac{C(uq,vq)}{q} \quad \text{and} \quad \Lambda_U(u,v) = \lim_{q \to 0^+} \frac{\overline{C}(uq,vq)}{q}.$$

for all $(u,v) \in [0,\infty)^2 = R_+^2$.

- A general goodness-of-fit test for copulas does not necessarily provide a good model for tail dependence;
- The reason is that the estimation techniques are usually based on **entire** available data set in the unit square $[0, 1]^2$;



- However, due the fact **that the center of distribution** say $[0.1, 0.9] \times [0.1, 0.9]$ **does not contain any information** about tail performance, the conventional estimation techniques yield **biased estimates for the tail dependence**.

- In order to overcome this problem, we apply **the concept of tail copulas** being function of the underlying copula (which describes the dependence structure in the tail of multivariate distributions)

$$\Lambda_L(u,v) = \lim_{q \to 0^+} \frac{C(uq, vq)}{q} \quad \text{and} \quad \Lambda_U(u,v) = \lim_{q \to 0^+} \frac{\overline{C}(uq, vq)}{q};$$

- The tail copulas can be considered as a local version of tail dependence coefficients λ_L and λ_U because

$$\Lambda_L(1,1) = \lambda_L \text{ and } \Lambda_U(1,1) = \lambda_U;$$

- To obtain estimators that are **more robust with respect** to the center of distribution, there are 2 approaches:
- **to use tail copula** (which enable to model tail **dependence of arbitrary form**);
- to rely a **extreme value copula** (Pareto one, say).

Tail empirical copula

- If C_n is the empirical copula, the **empirical lower tail copula** is defined by

$$
\begin{aligned}
\Lambda_{L,n}(u,v) &= \frac{n}{k} C_n\left(\frac{ku}{n}, \frac{kv}{n}\right) \\
&\approx \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}\left(u_i \le \frac{ku}{n+1}, v_i \le \frac{kv}{n+1}\right);
\end{aligned}
$$

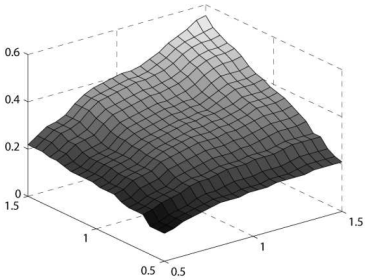
- If $\overline{C_n}$ is the empirical survival copula, the **empirical upper tail copula** is

$$
\begin{aligned}
\Lambda_{U,n}(u,v) &= \frac{n}{k} \overline{C}_n\left(\frac{ku}{n}, \frac{kv}{n}\right) \\
&\approx \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}\left(u_i > \frac{n-ku}{n+1}, v_i > \frac{n-kv}{n+1}\right)
\end{aligned}
$$

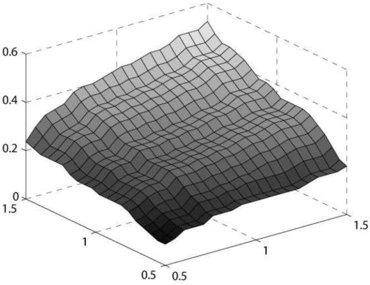
for some parameter $k \in \{1, \cdots, n\}$ (see Schmidt and Stadtmller (2006)).

Tail empirical copula for the data

Implementing the procedure described by Schmidt and Stadtmuller (2006), one can calculate the empirical tail copulas for the log-returns of the crude oil and natural gas futures.



Empirical lower tail copula.



Empirical upper tail copula.

Example: Copula for discrete bivariate distribution

Consider the joint discrete distribution

Y\X	0	1	$G(y)$
0	1/8	2/8	3/8
1	2/8	3/8	5/8
$F(x)$	3/8	5/8	1

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < 0 \\ 3/8, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases} \quad \text{and } G(y) = \begin{cases} 0, & y < 0 \\ 3/8, & 0 \leq y < 1 \\ 1, & y \geq 1 \end{cases}$$

i.e. marginal distributions are the same. Furthermore, from Sklar's theorem

$$H(x, y) = P(X \leq x, Y \leq y) = C(F(x), G(y)), \quad \text{for all } x, y \text{ and some copula } C.$$

Since $RanX = RanY = \{0, 3/8, 1\}$, the only constraint for copula C is $C(3/8, 3/8) = 1/8$, i.e. there are infinitely many such copulas, because $C(0, 0) = 0$ and $C(1, 1) = 1$.

Conclusion: There is more than one copula in discrete bivariate case.

As conclusions for copula pitfalls

- Any bivariate distribution with $H(x, y) = P(X \leq x, Y \leq y)$ can be trasformed to copula via

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)), \quad u, v \in [0, 1]. \tag{2}$$

Thus, copula C represents the class of all bivariate distributions

$$\mathcal{H} = [H(x, y)| \text{ continuous marginals are exactly } F(x) \text{ and } G(y)] \,.$$

- In general, any result valid for bivariate distributions can be transported to copula theory via (2).
- But any result obtained for a copula

$$C(u, v) = P(U \leq u, V \leq v), \quad U \sim U(0, 1), V \sim U(0, 1),$$

can be translated **only** for bivariate distributions $E(x, y)$ with **same** marginals (i.e. for exchangeable variables)

$$E(x, y) = P(X \leq x, Y \leq y), \text{ with } E(x, \infty) = F(x) \text{ and } E(\infty, y) = G(y).$$

Parameter estimation

- We will present three different methodologies (two parametric and one semi-parametric) to estimate unknown parameters:
- The classical maximum likelihood estimation;
- The inference for the marginals;
- Semi-parametric approach, see Genest et *al.* (1995).

Observation: Exist Bayesian based approaches as well.

Maximum likelihood estimation

- To apply the **maximum likelihood (ML) methodology**, we need to find the joint density in terms of the copula density. We have already seen that

$$h(x, y; \Theta) = c(F(x; \theta_1), G(y; \theta_2); \theta) f(x; \theta_1) g(y : \theta_2),$$

  where θ_1 and θ_2 are parameters of the marginal distributions and θ is copula parameter.
- Therefore, $\Theta = (\theta_1, \theta_2, \theta)$ is the parameter vector.

Maximum likelihood estimation

- Given a random sample (X_j, Y_j), $j = 1, \ldots, m$, the **likelihood function** is

$$L(\mathbf{X}; \Theta) = \prod_{j=1}^{m} \left[c\left(F\left(x_j; \theta_1\right), G\left(y_j; \theta_2\right), \theta\right) f\left(x_j; \theta_1\right) g\left(y_j; \theta_2\right)\right]$$
$$= \prod_{j=1}^{m} c\left(F\left(x_j; \theta_1\right), G\left(y_j; \theta_2\right), \theta\right) \prod_{j=1}^{m} f\left(x_j; \theta_1\right) g\left(y_j; \theta_2\right)$$

- For the log-likelihood $\ln L(\mathbf{X}; \Theta)$ we have

$$\ln L(\mathbf{X}; \Theta) = \sum_{j=1}^{m} \ln c\left(F\left(x_j; \theta_1\right), G\left(y_j; \theta_2\right), \theta\right)$$
$$+ \sum_{j=1}^{m} \left[\ln f\left(x_j; \theta_1\right) + \ln g\left(y_j; \theta_2\right)\right].$$

Maximum likelihood estimation

Maximum likelihood estimators

The Maximum likelihood estimators are given by

$$\widehat{\theta} = \arg\max \sum_{j=1}^{m} \frac{\partial \ln c\left(F\left(x_j;\theta_1\right) G\left(y_j;\theta_n\right),\theta\right)}{\partial \theta},$$

$$\hat{\theta}_1 = \arg\max \sum_{j=1}^{m} \frac{\partial \ln c\left(F\left(x_j;\theta_1\right) G\left(y_j;\theta_n\right),\theta\right)}{\partial \theta_1} + \sum_{j=1}^{m} \frac{\partial \ln f\left(x_j;\theta_1\right)}{\partial \theta_1},$$

$$\hat{\theta}_2 = \arg\max \sum_{j=1}^{m} \frac{\partial \ln c\left(F\left(x_j;\theta_1\right) G\left(y_j;\theta_n\right),\theta\right)}{\partial \theta_2} + \sum_{j=1}^{m} \frac{\partial \ln g\left(y_j;\theta_2\right)}{\partial \theta_2}.$$

In general, we **do not have closed form expressions** for the parameter estimators, so numerical optimization is needed.

Simulation of data set

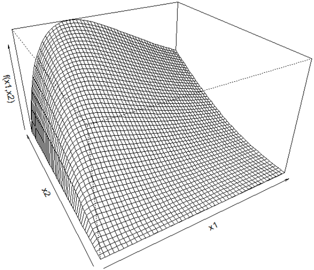
- Let us assume that marginals are Gamma distributed with parameters $(\alpha_1, \beta_1) = (2, 1)$ and $(\alpha_2, \beta_2) = (3, 2)$, i.e., the corresponding density function is

$$g(x; \alpha, \beta) = \frac{\beta^{\alpha} x^{\alpha - 1} e^{-x\beta}}{\Gamma(\alpha)} \text{ for } x \geq 0 \text{ and } \alpha, \beta > 0.$$

We will generate **a sample of size n=200 of a bivariate distribution with selected Gamma marginals and Gaussian copula with parameter** $\rho = 0.5$ using *R* package **copula**, see Yan (2007).

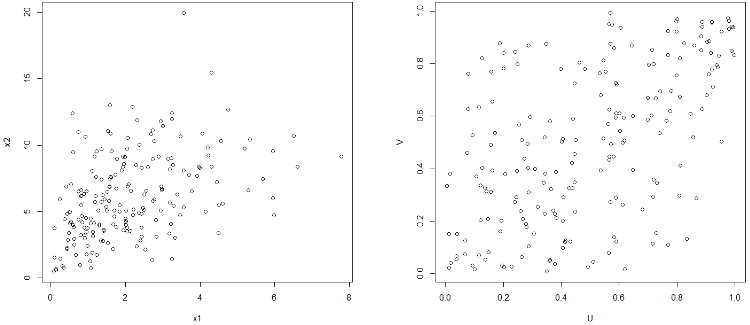
```
library(copula)

myMvd = mvdc(copula = ellipCopula(family = "normal", param = 0.5), margins
       = c("gamma", "gamma"), paramMargins = list(list(shape = 2,scale =
                            1),list(shape = 3, scale = 2)))
n = 200
dat = rmvdc(myMvd, n)
persp(myMvd,dMvdc,xlim=c(0,1), ylim=c(0,1),  xlab="x1", ylab="x2", zlab="f
      (x1,x2)")
plot(dat[,1], dat[,2],xlab="x1", ylab="x2")

u=pgamma(dat[,1],shape = 2,scale = 1)
v=pgamma(dat[,2],shape = 3,scale = 2)

plot(u,v,xlab="U", ylab="V")
```

Simulated data



Density of Simulated data

Simulated data



Scatterplots of Simulated data and its Copula

Model fitting - maximum likelihood method

- Let us use the **maximum likelihood method** to estimate the parameters of the model, i.e., $\theta_1 = (\alpha_1, \beta_1)$, $\theta_2 = (\alpha_2, \beta_2)$ and $\theta = \rho$.
- As **initial estimates** for the parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ and θ we will use their estimators obtained by **method of moments**.

```
mm <- apply(dat, 2, mean) #mean of data
vv <- apply(dat, 2, var) #variance of data
b1.0 <- c(mm[1]^2/vv[1], vv[1]/mm[1])
b2.0 <- c(mm[2]^2/vv[2], vv[2]/mm[2])
a.0 <- sin(cor(dat[, 1], dat[, 2], method = "kendall") * pi/2)
start <- c(b1.0, b2.0, a.0)
fit <- fitMvdc(dat, myMvd, start = start, optim.control = list(trace = TRUE,
    maxit = 2000))
fit
```

Results - maximum likelihood estimates

- True values are: $(\alpha_1, \beta_1) = (2, 1)$, $(\alpha_2, \beta_2) = (3, 2)$ and $\rho = 0,5$.

The Maximum Likelihood estimation is based on 200 observations.

	Marginal 1	
	Estimate	Std. Error
α_1	2,046	0,190
β_1	1,034	0,109
	Marginal 2	
	Estimate	Std. Error
α_2	3,048	0,290
β_2	2,003	0,207
	Copula	
	Estimate	Std. Error
ρ	0,522	0,051

The maximized loglikelihood shows value -774,236

Comments

- Usually, the **estimator of** the copula dependence parameter θ is **affected by** the parametric structure of the marginal distributions, i.e. by parameters θ_1 and θ_2. **One can observe the inverse effect as well**.
- When we are interested **mainly on the dependence structure**, it is profitable to have a tool that **relax this influence**.
- Two appropriate methodologies are
  - the **inference for the marginals**;
  - the **semi-parametric** method (when the marginal distributions are estimated using some non-parametric approach);
- Both methods are conceptually straightforward.

Inference for marginals

- The inference for marginals simplifies the estimation procedure of the parameters involved, since it divides the problem in two stages.
- In the **first stage**, the marginal parameters θ_1 and θ_2 are estimated by maximum likelihood method.

Estimators for the marginal density parameters.

$$\widehat{\theta}_1^* = \arg\max \sum_{j=1}^{m} \frac{\partial \ln f\left(x_j; \theta_1\right)}{\partial \theta_1},$$

$$\widehat{\theta}_2^* = \arg\max \sum_{j=1}^{m} \frac{\partial \ln g\left(y_j; \theta_2\right)}{\partial \theta_2}.$$

Inference for marginals

- In the **second stage**, the dependence parameter θ is estimated using a pseudo-likelihood.

Dependence parameter estimator.

$$\widehat{\theta}^* = \arg\max \sum_{j=1}^{m} \frac{\partial \ln c\left(\widehat{u}_j, \widehat{v}_j, \theta\right)}{\partial \theta},$$

where $\widehat{u}_j = F\left(x_j; \widehat{\theta}_1^*\right)$, $\widehat{v}_j = G\left(y_j; \widehat{\theta}_2^*\right)$ for all $j = 1, \ldots, m$.

- The name **pseudo-likelihood** comes from the fact that we use $\widehat{u}_j = F\left(x_j; \widehat{\theta}_1^*\right)$ and $\widehat{v}_j = G\left(y_j; \widehat{\theta}_2^*\right)$ instead of the observed values u_j and v_j, $j = 1, \ldots, m$.

Clayton, Frank and Gaussian copulas

- Let us apply the **inference for marginals method** to estimate the parameters of Clayton, Frank and Gaussian copula joining the marginal gamma distributions.

```
#loglikelihood for margins
loglik.marg <- function(b, x) sum(dgamma(x, shape = b[1], scale = b[2],log =
    TRUE))
ctrl <- list(fnscale = -1)

#First stage for estimation process
b1hat <- optim(b1.0, fn = loglik.marg, x = dat[, 1], control = ctrl)$par
b2hat <- optim(b2.0, fn = loglik.marg, x = dat[, 2], control = ctrl)$par
udat <- cbind(pgamma(dat[, 1], shape = b1hat[1], scale = b1hat[2]),pgamma(dat[,
    2], shape = b2hat[1], scale = b2hat[2]))
```

Inference for marginals

```
#second stage for estimation process
#Defining copula object
myCop.clayton <- claytonCopula(dim=2, param = 2)
myCop.frank <- frankCopula(dim=2, param = 2)
myCop.gaussian <- normalCopula(dim=2, param = 0.9)

#Model fitting
fit.if1 <- fitCopula(myCop.clayton, udat, start = a.0)
fit.if2 <- fitCopula(myCop.frank, udat, start = a.0)
fit.if3 <- fitCopula(myCop.gaussian,udat, start = a.0)
```

Results - inference for marginals

- The estimators of copula parameters **using the inference for marginals** method are:

		Estimate	Std. Error	z value	$Pr(>\|z\|)$
Clayton	param	0,644	0,127	5,061	4,17E-07
Frank	param	2,964	0,499	5,945	2,76E-09
Gaussian	ρ	0,452	0,060	7,541	4,66E-14

ML of ρ was 0.522 (the true value in Gaussian case is 0.5)

- How can **we select the model that best fits the data?**

Inference for marginals - comments

- The estimator of copula parameter θ obtained from the maximum likelihood and the inference for marginals methods **are different** (0,522 and 0,452 in Gaussian case).
- Xu (1996) performed Monte Carlo simulations to compare the results and verified that **in almost all simulations the relative efficiency** was very close to 1.
- In general, the **inference for marginals method is preferable** than the exact maximum likelihood approach.

Inference for marginals - comments

- The estimation of the parameters in two steps **leads to a loss in efficiency and standard errors cannot be obtained as the inverse of the Fisher Information**.
- Patton (2006a) shows in a simulation study that method does not perform very well applying one step of the Newton-Rhapson algorithm to the full likelihood function.
- Alternatively **when the marginal model is unknown** Genest et al. (1995) suggest modeling the marginal distribution with the empirical distribution and estimating the copula via ranks of the data.

Model selection

- A general criteria for model selection is the Akaike Information Criterion (AIC).
- It was developed by Hirotsugu Akaike (Akaike, 1974) and is based in the concept of information entropy, offering **a relative measure of the information lost when some model** is used to describe the reality.

Model selection

- In the general case, *AIC* is given by:

$$\mathcal{AIC} = 2k - 2\ln(L)$$

where k is the number of parameter in the statistical model and L is the maximized value of the likelihood function of the estimated model.
- For a set of candidate models for the data, the **"best"** one is with **minimum corresponding** *AIC* value.
- An alternative measure to *AIC* is the Bayesian Information Criterion (*BIC*), given by:

$$\mathcal{BIC} = k\ln(n) - 2\ln L,$$

where n is the sample size. Again, the **minimum** of *BIC* indicates better fit.

Inference for marginals

- Below we present the code in R to obtain results applying $\mathcal{AIC}$ and $\mathcal{BIC}$ criteria

```
#Aic criterion
aicClayton <- -2*fit.if1@loglik+2
aicFrank <- -2*fit.if2@loglik+2
aicGaussian <- -2*fit.if3@loglik+2

#Bic criterion
bicClayton <- -2*fit.if1@loglik+log(200)
bicFrank <- -2*fit.if2@loglik+log(200)
bicGaussian <- -2*fit.if3@loglik+log(200)
```

Model selection - inference for marginals method

- Comparison among the models obtained by the **inference for margins method**.

	Maximized LogL	AIC	BIC
Clayton	16,36	-30,73	-27,43
Frank	22,08	-42,16	-38,86
Gaussian	**22,74**	**-43,48**	**-40,18**

i.e. **the Gaussian copula gives better fit**.

Semi-parametric inference

- Many **semi-parametric** methods can be found in the literature, e.g., Clayton (1978), Clayton and Cuzick (1985), Genest et al. (1995) and Oakes (1994).
- In all these approaches, the **first step** consists in estimating the marginal distribution functions $\left(\widehat{F}(x), \widehat{G}(y)\right)$ applying some non-parametric methodology and, in the **second step**, performing a data transformation

$$\begin{pmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_m & y_m \end{pmatrix} \overset{\left(\ \widehat{F}(x_j) \quad \widehat{G}(y_j)\ \right)}{\longrightarrow} \begin{pmatrix} \widehat{u}_1 & \widehat{v}_1 \\ \vdots & \vdots \\ \widehat{u}_m & \widehat{v}_m \end{pmatrix}.$$

- In the **third step**, some parametric copula family that better fits the data is chosen.

Marginal distribution estimation

- **Several semi-parametric methodologies** to estimate a density function $f(x)$ exist. For example, one can use **kernel smoothing** to obtain $\widehat{f}(x)$ and to calculate

$$\widehat{F}(x) = \int \widehat{f}(x)\,dx.$$

- In Genest et al. (1995), **empirical distribution** functions were used,

$$\widehat{F}(x_j) = \frac{r_{X,j}}{m+1},\ \widehat{G}(y_j) = \frac{r_{Y,j}}{m+1},$$

where $r_{X,j}$ is the rank of jth sample observation of random variable X (the same for Y).

Dependence parameter estimation

- The dependence parameter can be estimated via

$$\widehat{\theta} = \arg\max \sum_{j=1}^{m} \frac{\partial \ln c\left(\widehat{u}_j, \widehat{v}_j, \theta\right)}{\partial \theta}.$$

- The estimator $\widehat{\theta}$ is **consistent and asymptotically normal** distributed.

Semi-parametric inference

- Below we present the code in *R* used to perform the semi-parametric inference:

```
#First stage: Getting empirical distribution
eu <- cbind((rank(dat[, 1]) - 0.5)/n, (rank(dat[, 2]) - 0.5)/n)

#second stage for estimation process
#Defining copula object
myCop.clayton <- claytonCopula(dim=2, param = 2)
myCop.frank <- frankCopula(dim=2, param = 2)
myCop.gaussian <- normalCopula(dim=2, param = 0.9)

#Model fitting
fit.cm1 <- fitCopula(myCop.clayton,eu, start = a.0)
fit.cm2 <- fitCopula(myCop.frank,eu, start = a.0)
fit.cm3 <- fitCopula(myCop.gaussian,eu, start = a.0)
```

Semi-parametric inference

- Below we present the code in *R* to obtain results applying *AIC* and *BIC* criteria

```
#Aic criterion
aicClayton <- -2*fit.cm1@loglik+2
aicFrank <- -2*fit.cm2@loglik+2
aicGaussian <- -2*fit.cm3@loglik+2

#Bic criterion
bicClayton <- -2*fit.cm1@loglik+log(200)
bicFrank <- -2*fit.cm2@loglik+log(200)
bicGaussian <- -2*fit.cm3@loglik+log(200)
```

Semi-parametric inference

- In the table below we present the results for the selected models using the **semi-parametric inference method**.

		Estimate	Std. Error	z value	$Pr(>\|z\|)$
Clayton	param	0,503	0,102	4,940	7,82E-07
Frank	param	2,894	0,508	5,699	1,21E-08
Gaussian	ρ	**0,447**	**0,060**	7,402	1,34E-13

- For comparison, see **inference for marginals** results

		Estimate	Std. Error	z value	$Pr(>\|z\|)$
Clayton	param	0,644	0,127	5,061	4,17E-07
Frank	param	2,964	0,499	5,945	2,76E-09
Gaussian	ρ	**0,452**	**0,060**	7,541	4,66E-14

Semi-parametric inference: model selection

- In the table below we present a comparison among the models using *AIC* and *BIC* criterion.

	Maximized LogL	AIC	BIC
Clayton	13,23	-24,46	-21,16
Frank	21,19	-40,39	-37,09
Gaussian	**33,07**	**-42,09**	**-38,80**

For comparison (inference for marginals method)

	Maximized LogL	AIC	BIC
Clayton	16,36	-30,73	-27,43
Frank	22,08	-42,16	-38,86
Gaussian	**22,74**	**-43,48**	**-40,18**

Real data analysis

Consider the **bivariate stochastic process** $\{X_t\}_{t=1}^{T}$ with $X_t = (X_{1t}, X_{2t})^{'}$. Let $H(x_{1t}, x_{2t})$ be the **joint distribution**. Whereas $F(x_{1t,\theta_1}), G(x_{2t,\theta_2})$ the **marginal** distribution functions and $f(x_{1t,\theta_1}), g(x_{2t,\theta_2})$ the correponding density functions.
By **Sklar's theorem** there exists a copula function $C(\cdot,\cdot|\theta) : [0,1]^2 \longrightarrow [0,1]$ mapping the marginal distributions of X_{1t} and X_{2t} to their joint distribution through

$$H(x_{1t}, x_{2t}) = C(F(x_{1t,\theta_1}), G(x_{2t,\theta_2})|\theta).$$

To satisfy all this (theorical) requirement implies **that we limit ourselves to the specific case that** each processes $\{X_{it}\}$ only depends on its own past, but **not on the past of** the **other process** $\{X_{jt},\ (i \neq j)\}$, and **that there is only instantaneous** causality between the variables (described by copula).

We assume that the **marginals** can be modeled **parametrically**. The **probability integral transform** of marginal distributions are given by $U_t = F(x_{1t}; \theta_1)$ and $V_t = G(x_{2t}; \theta_2)$ where θ_1 and θ_2 are the vector of parameters. In financial econometrics X_{it} the marginal processes $(i = 1, 2)$ are usually modeled by **ARMA-GARCH** type model, whose residuals are treated as **independent** and **identically distributed** (i.i.d.) random variables.

Fair rainbow options

Options consisting of two or more underlying stocks are called **rainbow options**. The price of these options are influenced by the dependence structure between the stocks since the fair option price still is the expected value.

Fair price of a bivariate rainbow option $V(t, S_1, S_2)$

Let S_1 and S_2 be two stocks traded on a complete and arbitrage free market. Let t be the present time and T the time of maturity, then the price $V(t, S_1, S_2)$ of an option with a given payoff function $g(S_1(T), S_2(T))$ is

$$V(t, S_1, S_2) = \exp\{-r(T-t)\} \int_0^\infty \int_0^\infty g(x,y) f^Q_{S_1,S_2}(x,y)\, dxdy.$$

Where $f^Q_{S_1,S_2}$ is the joint probability distribution of the two stocks under the risk-neutral probability measure Q and the sigma-algebra $\mathcal{F}_t$ is the filtration containing all information about the two stocks up to time t.

We will use in our analysis the so called **Exchange option.**, i.e, the payoff function is given by $g(S_1(T), S_2(T)) = \max(S_2(T) - S_1(T), 0)$. This means that the option will be exercised only if stock 2 is worth more than stock 1

Real data analysis: Option pricing via Copulas

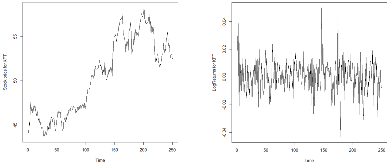
- We use the observed data for 2 stocks to obtain a **fair rainbow** option price:
    1. The food processing company Kraft Foods (KFT), representing $\{X_{1t}\}$;
    2. The technology company Hewlett-Packard (HPQ), representing $\{X_{2t}\}$.

  KFT and HPQ are traded on the New York Stock Exchange.
- We will analyse a one year time period from 1 August 2012 to 30 September 2013 (250 trading days). We use **daily close prices** (P_{it}) in USD that are adjusted for dividends and splits.
- The LogReturn of individual series is given by

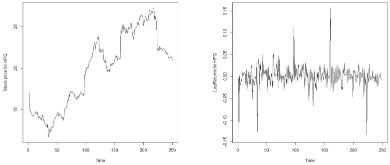
$$R_{it} = \log(P_{it}) - \log(P_{i(t-1)}), t = 1, 2, \ldots, n = 249.$$

  where P_{it} is the stock price at time t with $i = 1, 2$.

KFT and HPQ: Stock prices and LogReturns



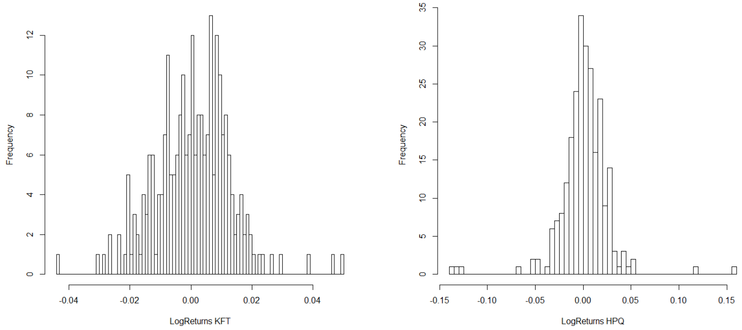
Daily stock prices and LogReturns for **KFT (above)** and **HPQ (below)** over the sample period.



First deductions:

1. The LogReturns range of KFT is more volatile that HPQ;
2. It is not evident if both series are independent or dependent.

KFT and HPQ: Histograms



Histograms of the LogReturns for KFT and HPQ.

KFT and HPQ: Dependence Analysis

To test independence between LogReturns of KFT and HPQ we use the following statistics

- for correlation coefficient: $\frac{\rho_n}{\sqrt{n-1}} \sim N(0,1)$, where ρ_n is sample Pearson's;
- for Kendall's tau: $\sqrt{\frac{9}{2}\frac{n(n-1)}{2n+5}}\tau_n \sim N(0,1)$, where τ_n is sample Kendall's.

In table below several calculus values of measures of dependence between the LogReturns of KFT and HPQ are displayed along with corresponding **p-values for the null hypothesis of independence**.

	Spearman's rho	Kendall's tau	Correl. coef.
	0.206	0.139	0.153
p-value	0.0053	0.0052	0.008

Measures of dependence between LogReturns of KFT and LogReturns of HPQ.

Conclusion: The KFT and HPQ stocks **are statistically dependent** (of level 0.8%).

Time series GARCH process

We will model both series (KFT and HPQ) of LogReturns by GARCH(p, q) model, defined by

$$R_t = \sqrt{h_t}\varepsilon_t, \quad \varepsilon_t \sim N(0,1),$$

where the **errors** (residuals) ε_t are assumed independent and **(conditional) variance** h_t is specified by

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i R_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i}, \quad t = 1, 2, ..., n.$$

GARCH parameters estimates: STEP 1A

- For both vectors of stock LogReturns, we use the **R package fGarch** to estimate the parameters of GARCH(p,q) model;
- We obtain a good fit for both series adopting GARCH(1,1). The estimate of parameters α_0, α_1 and β_1 are given below:

Stock	α_0	α_1	β_1
KFT	6.665×10^{-5}	0.298	0.296
HPQ	6.988×10^{-10}	10^{-8}	0.9991

Verification validity of GARCH(1,1) model: STEP 1B

Use the estimated parameters in **Step 1A**

Stock	α_0	α_1	β_1
KFT	6.665×10^{-5}	0.298	0.296
HPQ	6.988×10^{-10}	10^{-8}	0.9991

to calculate conditional variance

$$h_t = \alpha_0 + \alpha_1 R_{t-1}^2 + \beta_1 h_{t-1}, \quad t = 1, 2, \ldots, n$$

and residuals ε_t for each stock;

- The **residuals** in the GARCH model, i.e. ε_t's, should follow a standard normal distribution $N(0, 1)$.
- Furthermore ε_t should also have **independent increments**.

We carry out two tests in order to validate the GARCH(1,1) process

1. The *Kolmogorov-Smirnov (KS1)* test (to test $\varepsilon_{it} \sim N(0, 1)$, $i = 1, 2$);
2. The *Ljung-Box (LB)* test (for independence of residuals).

The *Kolmogorov-Smirnov* (KS1) test

- The **Kolmogorov-Smirnov** (KS1) test is applied for each stocks residual $\varepsilon_{it}, \; i = 1, 2$ with a 5% significance level.
- The residuals empirical distribution $F_{emp}(x)$ and $G_{emp}(y)$ are used to check $N(0, 1)$ for residuals.
- The statistical tests and p-values are presented below:

$$KS1 = \max_x (|\Phi(x) - F_{emp}(x)|) \text{ and } KS1 = \max_y (|\Phi(y) - G_{emp}(y)|), \;\; \Phi(\cdot) \sim N(0, 1).$$

	KFT	HPQ
KS1	0.558	0.566
p-value	0.888	0.872

The KS1 distance and p-values testing $N(0, 1)$ of the residuals.

We **can not reject the null hypothesis that residual are indeed standard normally distributed**.

The *Ljung-Box* (LB) test

- The **increments independence of** $\varepsilon_{it},\ i = 1, 2$ is tested with the *Ljung-Box* (LB) test

$$LB = N(N+2)\sum_{k=1}^{M}\frac{\rho_k}{N-k} \ \sim\ \chi^2(M),$$

where N is the sample size, M is the number of **autocorrelation lags** and $\rho_k = \frac{\sum_{i=1}^{k}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{k}(x_i-\overline{x})^2}\sqrt{\sum_{i=1}^{k}(y_i-\overline{y})^2}}$ is the **autocorrelation** at lag k.

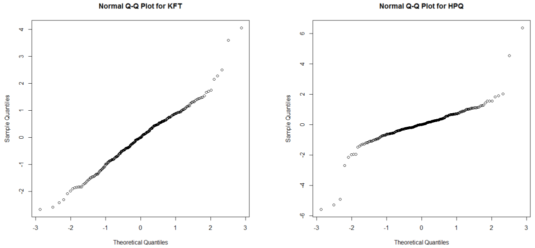
- The **null hypothesis is that there is no autocorrelation.** For $N(0,1)$ residuals, this **implies independent increments** of $\varepsilon_{it},\ i = 1, 2$. The results are given below:

	KFT	HPQ
LB	19.243	7.860
p-value	0.203	0.929

P-values for independence of the residuals using Ljung-Box test.

It can be seen, that **we can not reject the null hypothesis that residuals are independent**, (i.e., there is no autocorrelation), at 5% significance level.

QQ-Plot of Residuals



QQ-plot of Residuals for KFT and HPQ.

By now, we justified the choice of GARCH(1,1) model for KFT and HPQ.

Conclusions by now

- GARCH(1,1) serves to model individual time series (KFT and HPQ).

  This means that
  - residuals ($\varepsilon_{it}, i = 1, 2$) are $N(0, 1)$ distributed
  - ($\varepsilon_{it}, i = 1, 2$) are independent
- Nevertheless, both series seem weak dependent (Spearman's rho=0.206 and Kendall's tau=0.1390.)

Therefore the first steps are

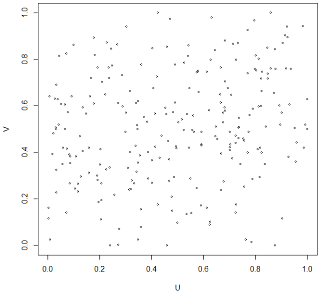
Step 1A Get the GARCH parameters estimates;
Step 1B Validate the GARCH(1,1) model;

So we will complement the analysis joining the error terms of both series with an appropriate copula. The algorithm can be summarized as follows

Step 2A Transform the error terms into $U(0, 1)$ distributed random variables;
Step 2B Fit the data set using some copula models via maximum likelihood method and select the best one using AIC and BIC criteria;
Step 2C Generate a sample from the copula selected;
Step 3A Transform the marginals from previously step into $N(0, 1)$ (error terms);
Step 3B Get the stock prices at time of maturity T;
Step 3C Get the option price at time of maturity T.

Transform the error terms into marginal uniforms: STEP 2A

In order to use copulas to join residuals in both series, we need to transform the **error terms** into $U(0,1)$ distributed random variables. We invert residuals ε_{it} obtained in step 1B for each stock into $U(0,1)$ by $u = \Phi(\varepsilon_{1i})$ and $v = \Phi(\varepsilon_{2i})$, where $\Phi(.) \sim N(0,1)$, $i = 1, 2, ..., 249$.



Copula fitting and model selection: STEP 2B

Apply the **maximum likelihood** to fit a copula to (u, v) from **Step 2A**. That is, for the transformed data $\{u_t, v_t\}_{t=1}^{n}$, estimate the copula parameters θ_C through

$$ArgMax_{\theta_C} \sum_{t=1}^{n} \ln[c((u_t, v_t); \theta_C)],$$

where $c(u, v)$ is **the density function** of the selected copula $C(u, v|\theta_C)$.

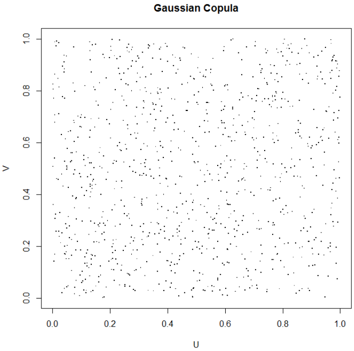
Copula	Gaussian	Student's t	Clayton	Gumbel
Parameter	0.135	(0.139, 308.48)	0.028	1.059
AIC	**-3.087**	-0.976	1.567	0.562
BIC	**0.430**	6.059	5.085	4.079

The estimated parameters for the different copulas and corresponding values of AIC and BIC criteria.

i.e. **the Gaussian copula gives better fit.**

Generate a sample from the copula: STEP 2C

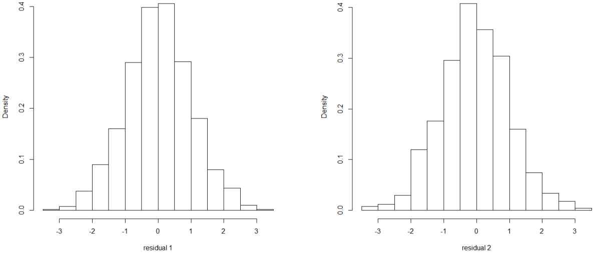
Generate a sample $\{u_t^*, v_t^*\}_{t=1}^T$ from a $U(0,1)$ marginal distribution using the chosen copula. Here T is the time to maturity for the option;



Scatterplot of a sample $\{u_t^*, v_t^*\}_{t=1}^T$ from a Gaussian copula.

Transform the generated marginals into N(0,1) margins: STEP 3A

For each time instant $t = 1, ..., T$, transform the generated marginals into N(0,1) margins (in the risk-neutral world), by $\varepsilon^*_{1t} = \Phi^{-1}(u^*_t)$ and $\varepsilon^*_{2t} = \Phi^{-1}(v^*_t)$;



Stock price at time of maturity T: STEP 3B

Use ε^*_{it} from **Step 3A** to calculate the conditional variances h_{it} give as considering **GARCH parameters** estimated in **Step 1A**. The **two future stock prices** at time T are

$$S_i(T) = S_i(0)\ \exp\Big\{\sum_{t=1}^{T}\sqrt{h_{it}}\varepsilon^*_{it}\Big\}, \quad i = 1, 2;$$

Stock	Price
KFT	44.95
HPQ	17.50

Stock prices at maturity time T = 260 using Gaussian Copula (i.e. 10 days after our last observations).

Application: Monte Carlo integration using copulas

- **Aim**: to obtain the **expected value** of a continuous function $q(x, y)$ of a bivariate random vector (X, Y) having joint distribution $H(x, y)$, i.e.

$$E(q(X, Y)) = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} q(x, y) dH(x, y);$$

- Given the copula $C(u, v) = H(F^{-1}(v), G^{-1}(v))$ and marginal distributions $F(x) = \lim_{y \to \infty} H(x, y)$ and $G(y) = \lim_{x \to \infty} H(x, y)$, we can use the following algorithm to approximate the value of $E(q(X, Y))$:
    1. generate n observations of the bivariate random vector (X, Y);
    2. for each observation i, calculate $q_i = q(x_i, y_i)$, $i = 1, 2, ..., n$;
    3. $E(q(X, Y)) \approx \frac{1}{n} \sum_{i=1}^{n} q_i$

Option price at time of maturity T: STEP 3C

Let $g(S_1(T), S_2(T)) = \max(S_2(T) - S_1(T), 0)$ be the **payoff function** and repeat **Steps 2C to 3B for N runs**. Finally, we obtain the **Monte Carlo** option price:

$$V(t, S_1, S_2) = \frac{\exp\{-r(T-t)\}}{N} \sum_{i=1}^{N} g(S_{1i}(T), S_{2i}(T)).$$

Exchange option	24.23

Option prices for Exchange option at time of maturity T=260.

Comments

- The marginal distributions that describe the individual behavior of each variables and the copula that fully captures the dependence between the variables.
- Furthermore, given a set of marginal distributions and a copula a multivariate distribution can be constructed by coupling the marginals with the copula. The flexibility of the way dependencies can be modeled independently of the marginal distributions.

Several conclusions

- Copula is useful tool in many applied areas where the interest is in analysis of multivariate dependence and when the multivariate normal distribution is controversial.
- In **actuarial science**, copulas are used in modeling dependence between mortality and losses, e.g., Frees, Carrière and Valdez (1996), Frees and Valdez (1998), Frees and Wang (2005).
- In **finance**, copulas are successfully applied in asset allocation, credit scoring, default risk modeling, derivative pricing and risk management, e.g., Bouyè, Durrleman, Bikeghbali, Riboulet and Roncalli (2000), Embrechts, Lindskog and McNeil (2003) and Cherubini, Luciano and Vecchiato (2004).
- In **biomedical studies**, copulas are used in modeling correlated events times and competing risks, e.g., Wang and Wells (2000), Escarela and Carrière (2003).

Several conclusions

- In many real situations, empirical evidence has proved the **inadequate** use of the normal multivariate distribution:
  - Empirical marginal distributions are skewed and heavy-tailed;
  - Possibilities of extreme co-movements, in contrast to the multivariate normal distribution.
- Copulas provide an alternative solution and often more useful representations of multivariate distribution functions compared to traditional approaches such as multivariate normal distribution.
- The use of linear correlation coefficient should be restricted to multivariate elliptical distributions. **Copula based dependence measures (Spearman's ρ_S and Kendall's τ_K) are free of such limitations.**

References

- Bouyè, E., Durrleman, V., Bikeghbali, A., Riboulet, G. and Roncalli, T. (2000). Copulas for Finance - A Reading Guide and Some Applications. *Working paper*, Groupe de Recherche Opérationnelle, Crédit Lyonnais.
- Blest, D. C.(2000). Rank correlation: an alternative measure. *Australian and New Zealand Journal of Statistics*, 42, 101-111.
- Cherubini, U., Luciano, E. and Vecchiato, W. (2004). *Copula Methods in Finance*. John Wiley & Sons.
- Embrechts, P., Lindskog, F. and McNeil, A. (2003). Modelling Dependence with Copulas and Applications to Risk Management. In S Rachev (ed.), *Handbook of Heavy Tailed Distribution in Finance.* pp. 329-384. Elsevier.
- Escarela, G. and Carrière, J.F. (2003). Fitting Competing Risks with an Assumed Copula. *Statistical Methods in Medical Research* **12**, 333-349.
- Frees, E.W, Carrière, J. and Valdez, E.A. (1996). Annuity Valuation with Dependent Mortality. *Journal of Risk and Insurance* **63**, 229-261.

References

- Frees, E.W. and Valdez, E.A. (1998). Understanding Relationships Using Copulas. *North American Actuarial Journal* **2**, 1-25.
- Frees, E.W. and Wang, P. (2005). Credibility Using Copulas. *North American Actuarial Journal* **9**, 31-48.
- Meucci, A. (2011). A short, comprehensive, practical guide to copulas. *Garp Risk Professional*, October, 22-27.
- Sklar, A. (1959). Fonctions de répartition n dimensions et leurs marges. *Publ Inst Statist Univ Paris* **8**, 229-231.
- Wang, W. and Wells, M.T. (2000). Model Selection and Semiparametric Inference for Bivariate Failure-time Data. *Journal of the American Statistical Association* **95**, 62-72.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19** 716-723.
- Clayton, D.G. (1978). A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141-151.
- Clayton, D.G. and Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. *Journal of Royal Statistics Society A* **148**, 82-117.
- Fisher, N. and Switzer, P. (1985). Chi-plots for assessing dependence. *Biometrika 72*, 253-265.
- Fisher, N. and Switzer, P. (2001). Graphical assessment of dependence: is a picture worth 100 tests? *The American Statistician* **55**, 233-239.

References

- Frees, E.W. and Wang, P. (2005). Credibility Using Copulas. *North American Actuarial Journal* **9**, 31-48.
- Genest, C. and Boies, J.C. (2003). Detecting dependence with Kendall plots. *The American Statistician* **57**, 275-284.
- Genest, C., Ghoudi, K. and Rivest, L.P. (1995). A semiparametric estimation procedure of dependence paramaters in multivariate families of distributions. *Biometrika* **82**, 543-552.
- McNeil, A., Frey, L. and Embrechts, P. (2005). *Quantitative Risk Management.* Princeton University Press.
- Oakes, D. (1994). Multivariate survival distributions. *Journal of Nonparametric Statistics* **3**, 343-354.
- Yan, J. (2007). Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software* **21**, 1-21.