Multivariate Dependence Modeling using Pair–Copulas

Ernesto Schirmacher

Liberty Mutual Group

Agenda

1. A brief introduction to copulas
2. Sklar's Theorem
3. Dependence and Increasing Transformations
4. χ–Plots to help us visualize dependence
5. Archimedean Constructions
6. The Pair–Copula Construction
7. Example on currency rate changes
8. Conclusions

How should we think about copulas?

Multivariate Normal Distribution? We need two items:

1. a vector of means and
2. a variance–covariance matrix.

How should we think about copulas?

Multivariate Normal Distribution? We need two items:

1. a vector of means and
2. a variance–covariance matrix.

General Multivariate Distribution? We need two items:

1. one-dimensional marginal distributions $F_i(x_i)$ and
2. a copula function $C\colon [0,1]^n \to [0,1]$.

What is a copula?

A copula is a multivariate distribution function with all univariate margins being uniform on $[0,1]$.

Examples

1. $C(u_1,\ldots,u_n) = u_1 \cdot u_2 \cdots u_n$
2. $C(u_1,\ldots,u_n) = \min\{u_1, u_2, \ldots, u_n\}$
3. $C(u,v) = \max\{0, u+v-1\}$
4. $C(u,v) = uv + \theta uv(1-u)(1-v)$
5. $C(u,v) = uv \exp\left\{\left[(-\log u)^{-\delta} + (-\log v)^{-\delta}\right]^{-1/\delta}\right\}$

Fréchet bounds

$$\max\{0, u_1 + \cdots + u_n - (n-1)\} \le C(u_1,\ldots,u_n) \le \min\{u_1,\ldots,u_n\}$$

Sklar's Theorem

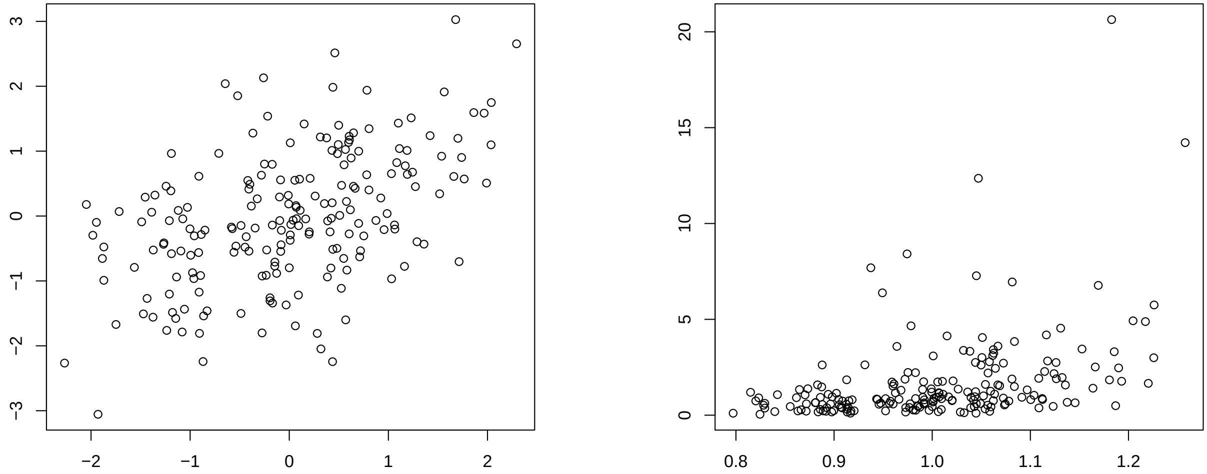
Let $F(x_1, \ldots, x_n)$ be an n–dimensional distribution function with continuous marginals $F_1, F_2, \ldots, F_n$. Then there exists a unique copula function $C\colon [0,1]^n \to [0,1]$ such that

$$F(x_1, x_2, \ldots, x_n) = C\left(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)\right).$$

One can also move in the other direction:

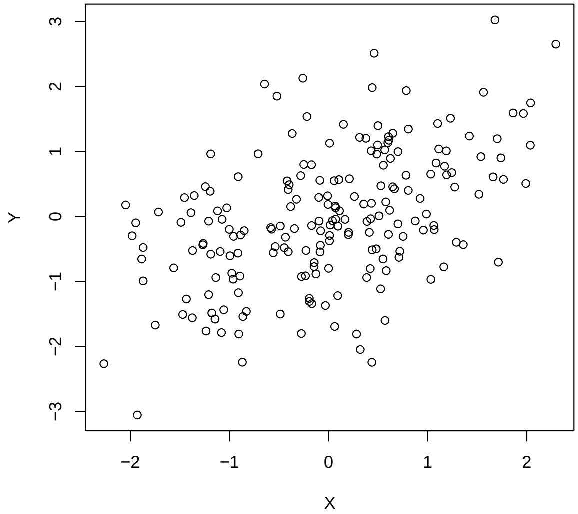
Copula + Marginals → Joint Distribution.

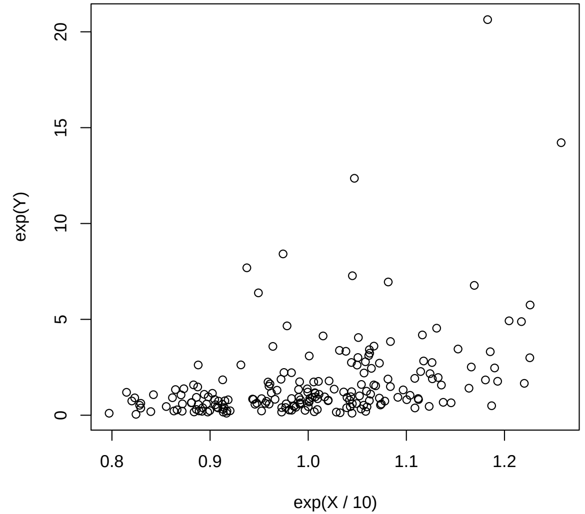
Copulas and increasing transformations



Do both panels shows the same dependence structure?

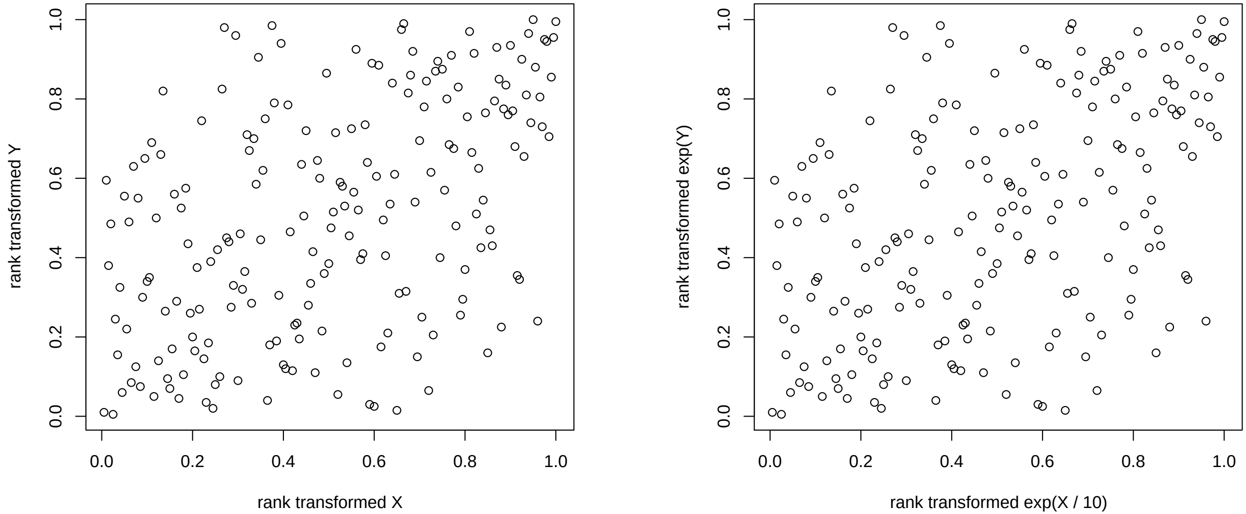
Copulas and increasing transformations





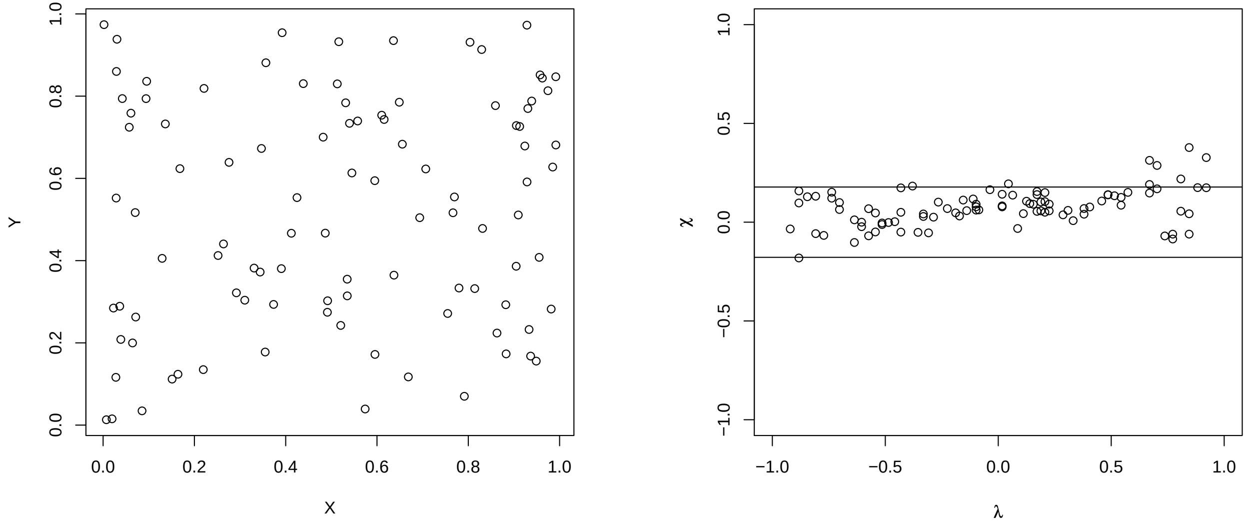
Do both panels shows the same dependence structure?

Remove marginals to study dependence

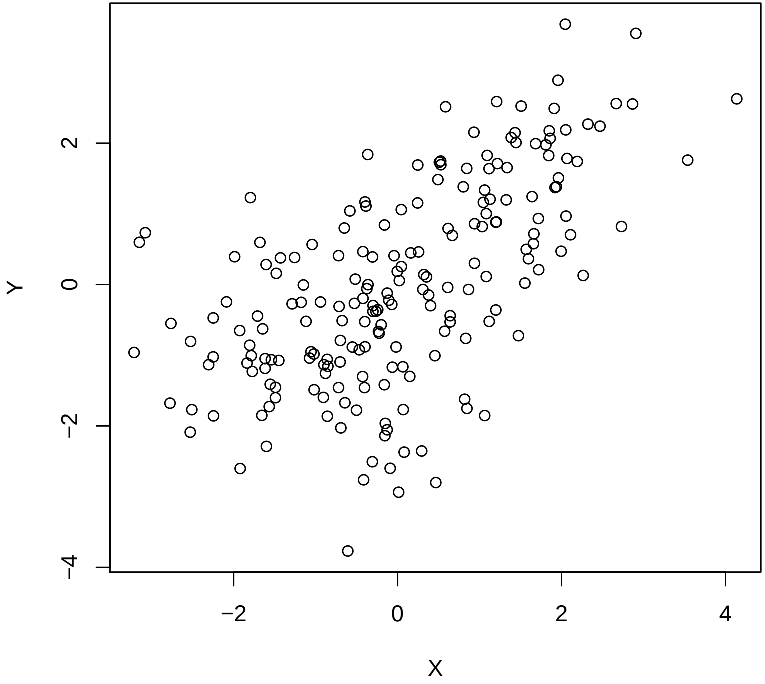


To understand dependence, rank transform your data to eliminate the marginals as **copulas are invariant under strictly increasing transformations**.

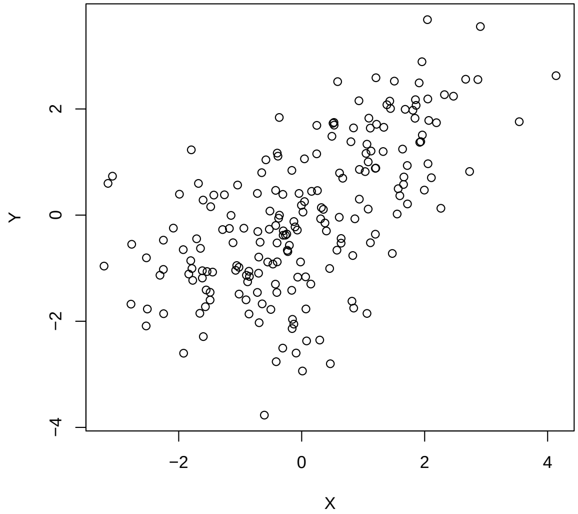
The χ-Plot helps visualize dependence

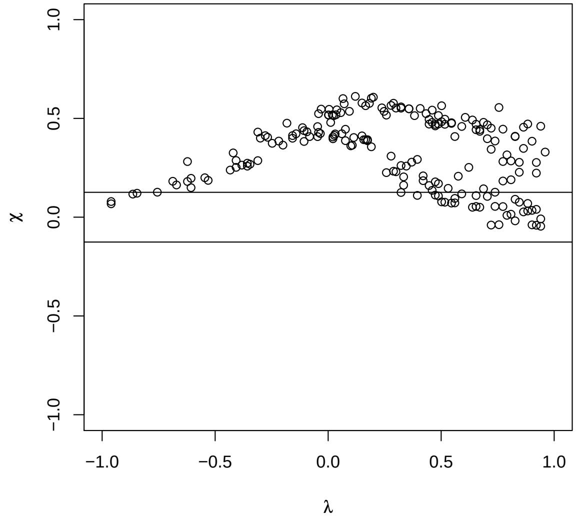


χ-Plot Examples

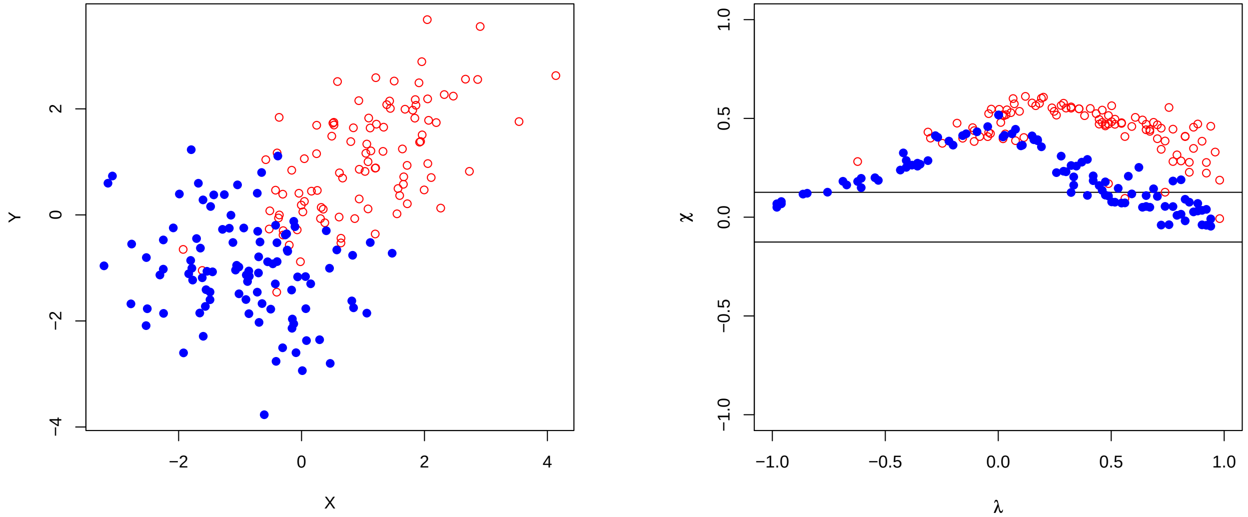


χ-Plot Examples

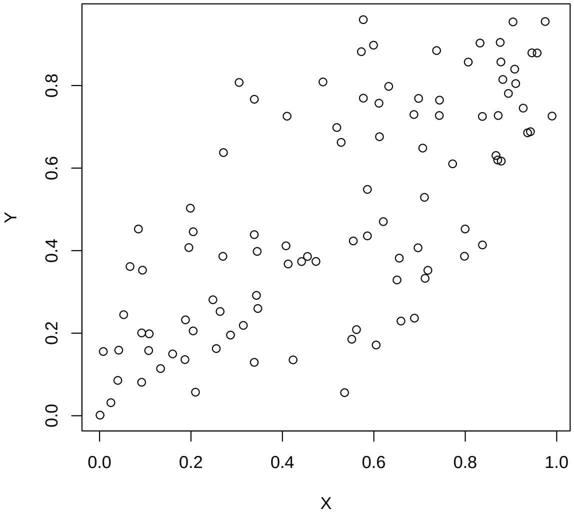


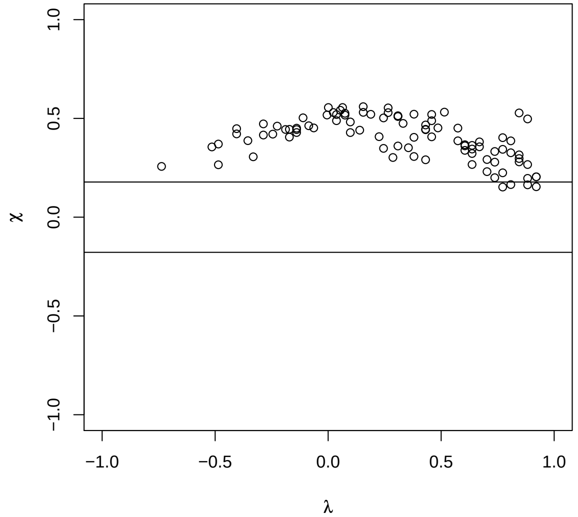


χ-Plot Examples

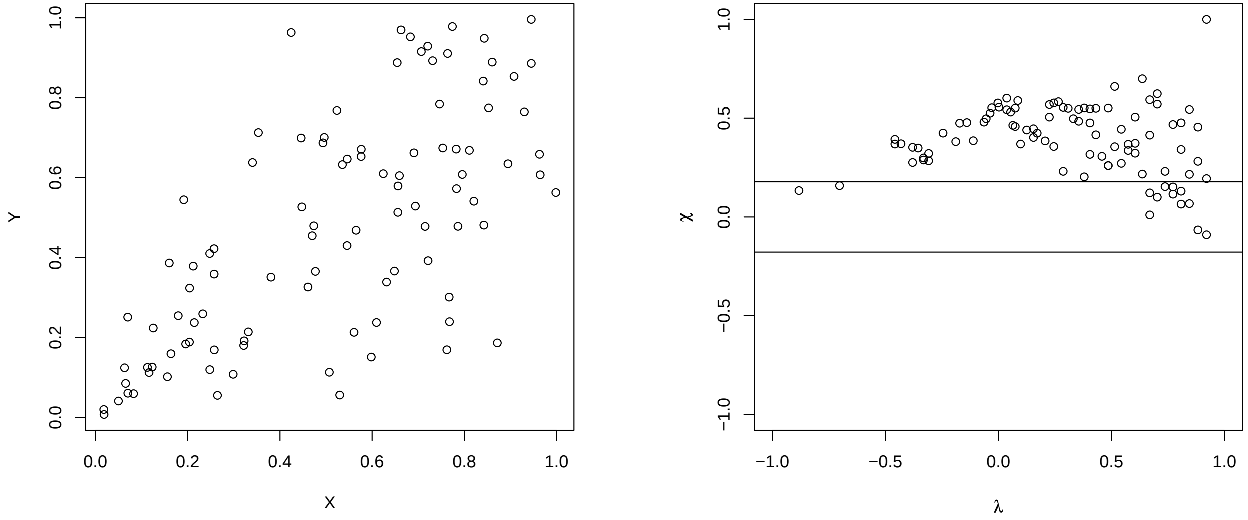


χ-Plot Examples (Normal copula)

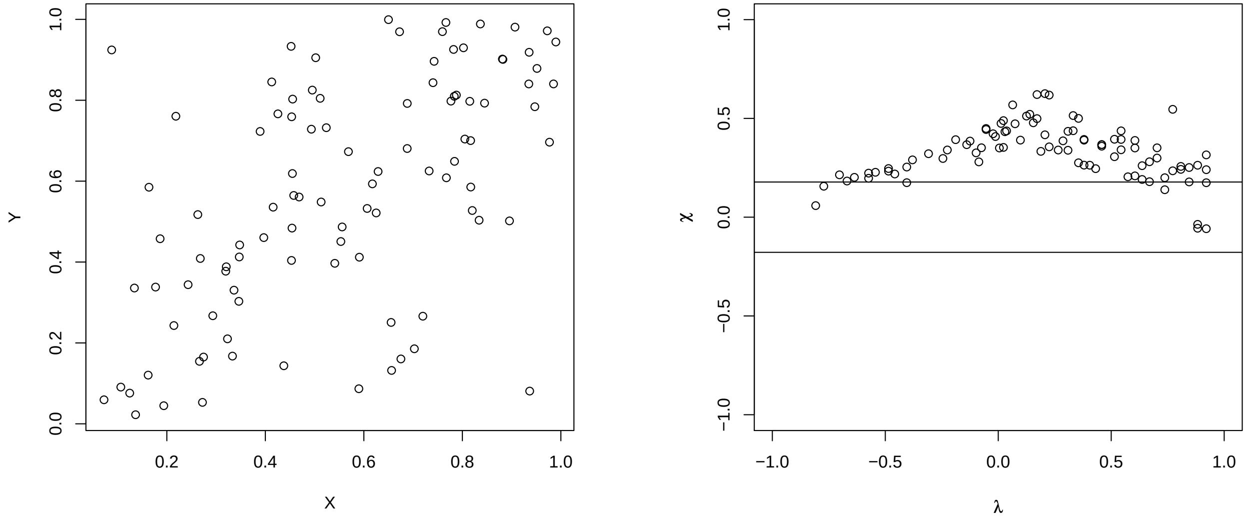




χ-Plot Examples (Clayton copula)



χ-Plot Examples (Frank copula)



Exchangeable Archimedean Copula

A multivariate Archimedean copula can be defined as

$$C(u_1, \ldots, u_n) = \varphi^{-1}\{\varphi(u_1) + \cdots + \varphi(u_n)\}$$

where φ is a decreasing function known as the *generator* of the copula.

1. Independent: $\varphi(t) = -\log(t)$
2. Gumbel: $\varphi(t) = (-\log(t))^{\theta}$
3. Clayton: $\varphi(t) = (t^{-\theta} - 1)/\theta$
4. Frank: $\varphi(t) = \log \frac{1-\theta}{1-\theta^t}$

Other constructions are also possible (partially–, fully–, hierarchically–nested) but their properties are very restrictive.

Liberty Mutual.

The Pair–Copula Construction

Given an n–dimensional joint density function $f(x_1, \ldots, x_n)$ do the following:

1. 'Factorize' it into a product of conditional densities
2. Rewrite each conditional density from the previous step into a product of bivariate copulas and marginal densities
3. Model each bivariate copula via one of the many choices: normal, t, Frank, Gumbel, Galambos, Clayton, etc. . .

Three dimensional example

Given $f(x_1, x_2, x_3)$ we can apply steps (1) and (2) to get:

$$
\begin{aligned}
f(x_1, x_2, x_3) &= f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|12}(x_3|x_1, x_2) \\
&= f_1(x_1) \cdot \\
&\quad c_{12}\Big(F_1(x_1), F_2(x_2)\Big) \cdot f_2(x_2) \cdot \\
&\quad c_{13|2}\Big(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\Big) \cdot \\
&\quad c_{23}\Big(F_2(x_2), F_3(x_3)\Big) \cdot f_3(x_3).
\end{aligned}
$$

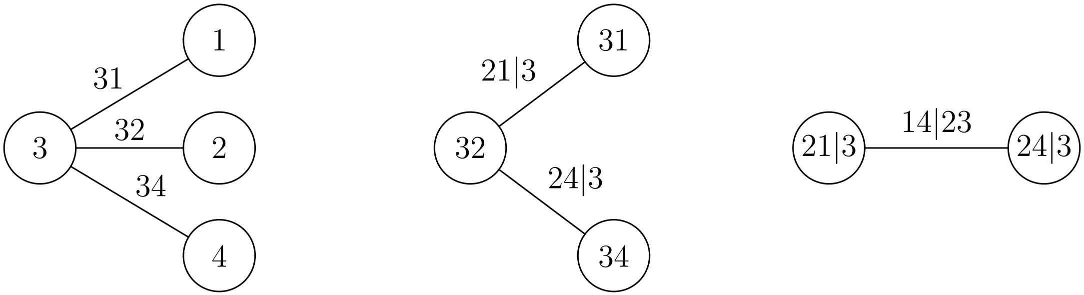
Vines to organize decompositions

The decomposition of $f(x_1, \ldots, x_n)$ in the previous slide into pair–copulas and marginal densities is not unique.

D-vines and canonical vines are two graphical models that help us organize a subset of all possible decompositions.

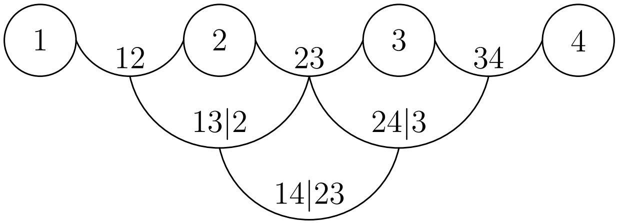
Both consists of sequences of trees that show us how to write a joint density function into pair–copulas and marginal densities.

Four dimensional canonical vine



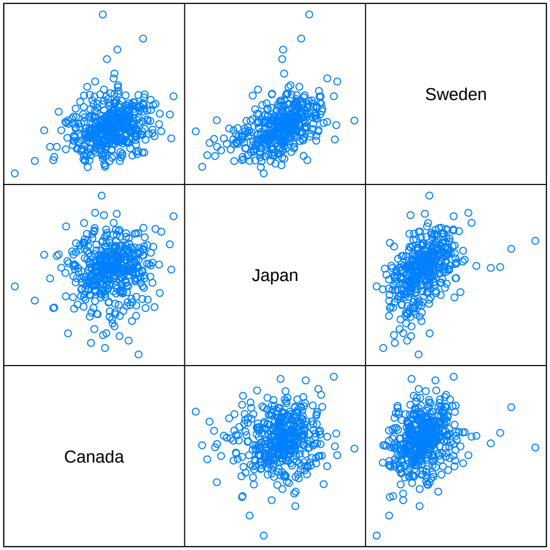
$$f(x_1, x_2, x_3, x_4) = f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4)$$
$$c_{31}\Big(F_3(x_3), F_1(x_1)\Big) c_{32}\Big(F_3(x_3), F_2(x_2)\Big) c_{34}\Big(F_3(x_3), F_4(x_4)\Big)$$
$$c_{21|3}\Big(F_{2|3}(x_2|x_3), F_{1|3}(x_1|x_3)\Big) c_{24|3}\Big(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)\Big)$$
$$c_{14|23}\Big(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)\Big)$$

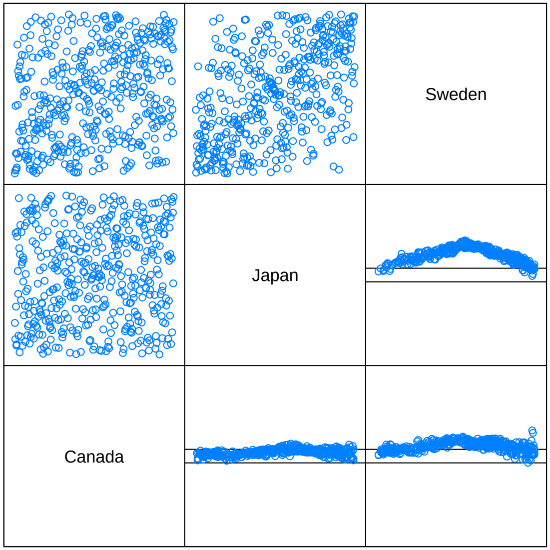
Four dimensional D-vine



$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4) \\
& c_{12}\Big(F_1(x_1), F_2(x_2)\Big) c_{23}\Big(F_2(x_2), F_3(x_3)\Big) c_{34}\Big(F_3(x_3), F_4(x_4)\Big) \\
& c_{13|2}\Big(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\Big) c_{24|3}\Big(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)\Big) \\
& c_{14|23}\Big(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)\Big)
\end{aligned}
$$

Example: Currency Rate Changes

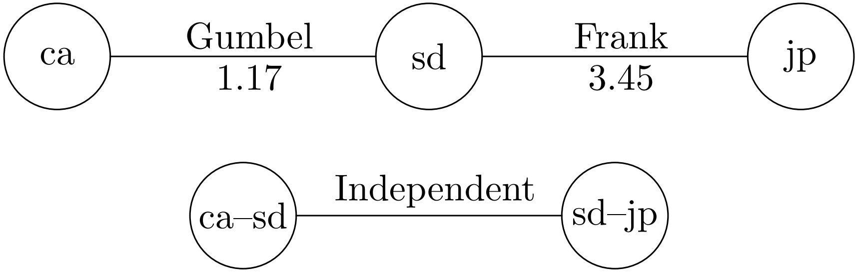




Monthly changes in foreign currency rates to US dollar.

Data Source: FRED database from the Federal Reserve Bank of St. Louis.

Initial ML–estimates for canonical vine



1. Bivariate ML–estimates are easy to calculate
2. These are just initial estimates used to start a global ML–estimation

Maximum likelihood parameter estimates

Pair-copula	Family	ML estimate
Canada–Sweden	Gumbel	1.11
Japan–Sweden	Frank	1.62
Canada–Japan *given* Sweden	Independent	

Liberty Mutual.

Conclusions

1. Copulas encapsulate dependence
2. Remove marginals when studying dependence
3. χ-plots help visualize dependence
4. Archimedean copulas are too restrictive
5. Pair–copula construction is flexible
6. Canonical– and D–vines are easy to work with

References

[1] K. Aas, C. Czado, A. Frigessi, and H. Bakken, *Pair-copula constructions of multiple dependence*, Tech. Rep. SAMBA/24/06, Norwegian Computing Center, Postboks 114, Blindern, NO-0314 Oslo, Norway (2006).
http://www.nr.no/files/samba/bff/SAMBA2406.pdf

[2] D. Berg and K. Aas, *Models for construction of multivariate dependence*, Tech. Rep. SAMBA/23/07, Norwegian Computing Center, Postboks 114, Blindern, NO-0314 Oslo, Norway (2007).
http://www.nr.no/files/samba/bff/SAMBA2307.pdf

[3] N. I. Fisher and P. Switzer, *Chi-plots for assessing dependence*, Biometrika **72** (1985), no. 2, 253–65.

[4] N. I. Fisher and P. Switzer, *Graphical assessment of dependence: Is a picture worth 100 tests?*, The American Statistician **55** (2001), no. 3, 233–239.

[5] C. Genest and A.-C. Favre, *Everything you always wanted to know about copula modeling but were afraid to ask*, Journal of Hydrologic Engineering **12** (2007), no. 4, 347–368.
http://archimede.mat.ulaval.ca/pages/genest/publi/JHE-2007.pdf

[6] H. Joe, *Multivariate Models and Dependence Concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC, Boca Raton, Florida, 1997b.