# An “Introduction” to Vine Copulas

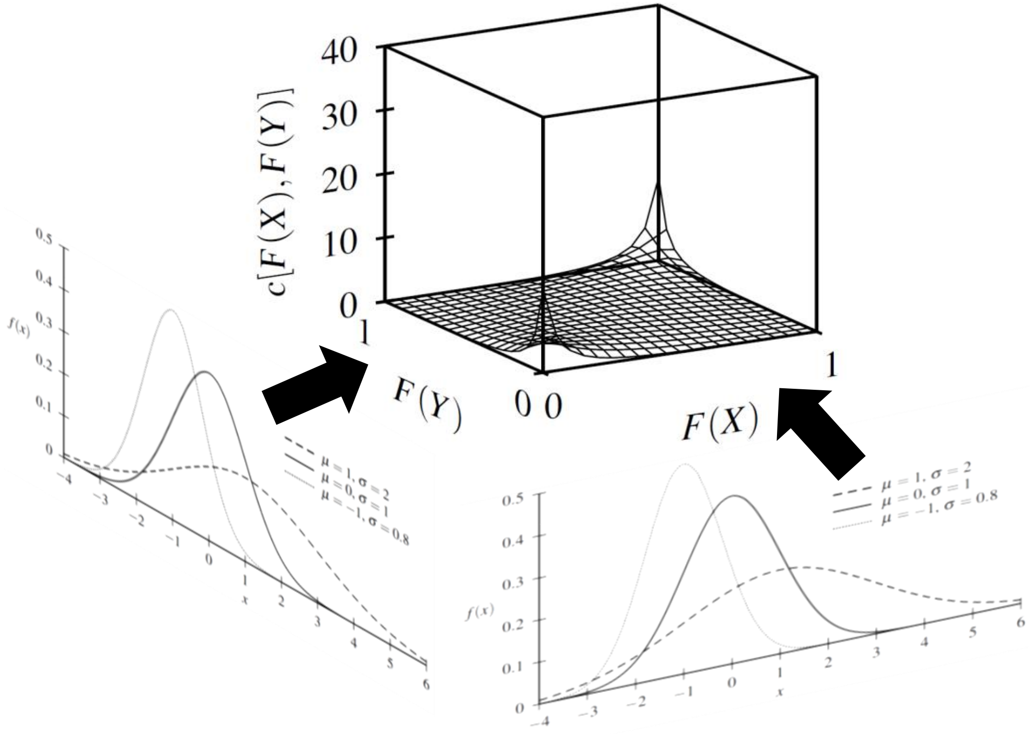
Paul Sweeting

# Recap: what is a copula?

- $C\big(F(x), F(y)\big)$ is the copula distribution function
- The inputs are the cumulative distribution functions for $x$ and $y$
- The output is the probability that the cumulative distribution functions of two random variables, $F(X)$ and $F(Y)$, are less than $F(x)$ and $F(y)$ respectively
- In this way it is independent of the shape of the distributions of $x$ or $y$

# Recap: bivariate normal copula

- Look at the two marginal distributions here…
- …and imagine stretching and reshaping them…
- …so they become (e.g.) gamma, lognormal, etc…
- …but the *order* of the observations stays the same




- A copula joins marginal distributions based on *order* not *value*

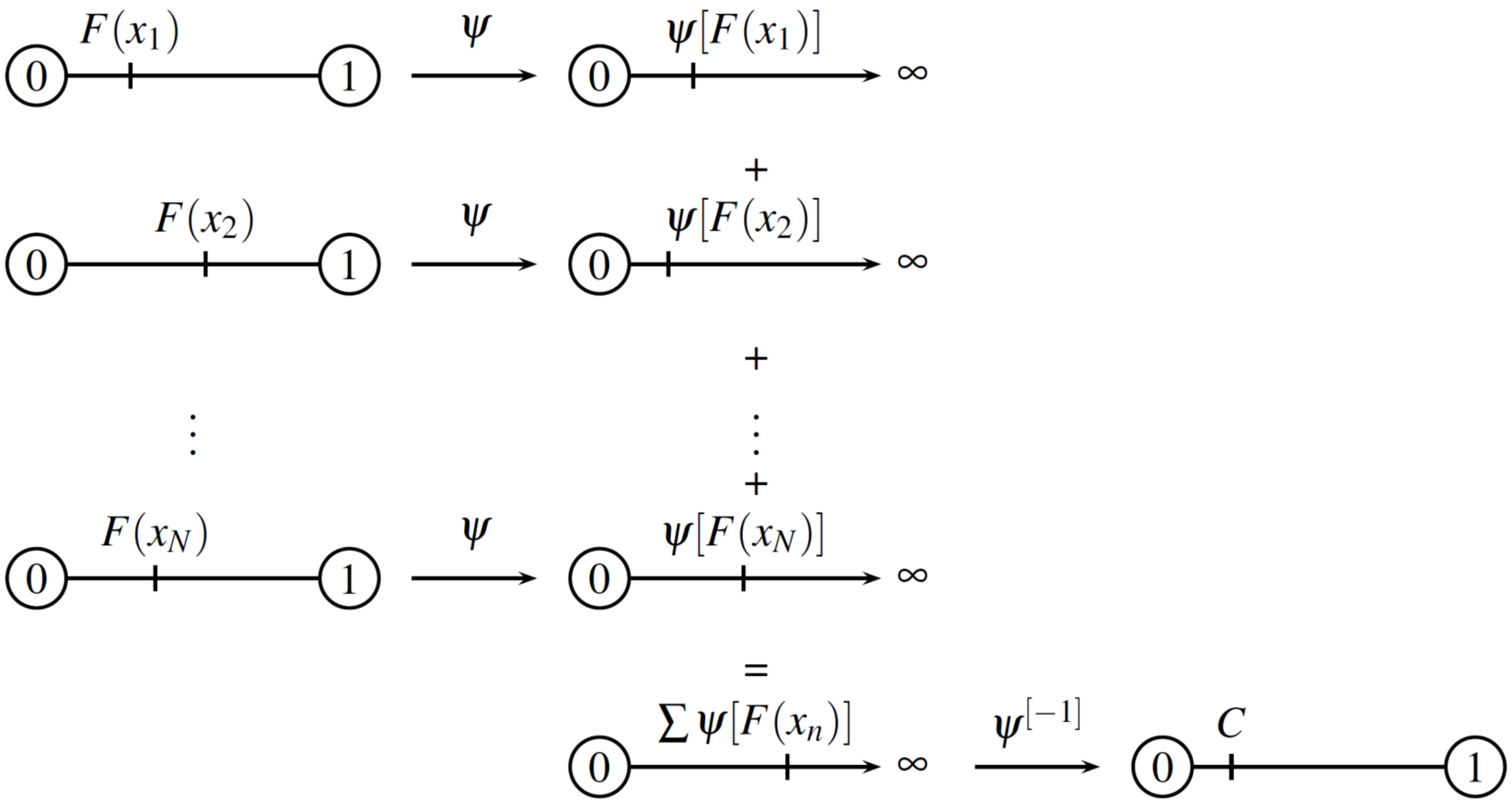
# From bivariate to multivariate

- Most examples use just 2 dimensions…
- …as 3+ dimensional copulas are hard to draw!
- But moving to higher dimensions is easy
- Consider elliptical copulas (normal, t)
- Start with multivariate elliptical distribution…
- …with corresponding correlation matrix…
- …and specified degrees of freedom for t
- For Archimedean copulas, just as simple

# Generator functions

- Start with generator functions, denoted $\psi$ (“psi”)…
- …which turn distribution functions into summable functions
- Takes each of $F(x_1), F(x_2), \ldots, F(x_N)$…
- …and generates a number that lies between 0 and ∞
- When added together, result also between 0 and ∞
- …so using *inverse* generator gives number between 0 and 1…
- …which is the joint probability, $C[F(x_1), F(x_2), \ldots, F(x_N)]$

# Using an Archimedean copula



# Problems with multivariate copulas

- All of these suffer from a similar issue
- For Archimedean copulas, can have only parameter α
- For elliptical copulas, can have a correlation matrix…
- …but only one degree of freedom parameter
- In other words, there no flexibility over the shape…
- …between different pairs of variable

# What are vine copulas?

- Vien copulas seek to overcome this
- They are networks of copulas…
- …constructed of pair-wise relationships…
- …with each relationship defined by a bivariate copula
- This gives much greater freedom…
- …but also brings with it challenges

# Background: graph theory

- Vine copulas can have many different constructions
- Constructions are described using graph theory

| Term | Mathematical Representation | Plain English Explanation |
|---|---|---|
| Node | $N = \{x_1, x_2, \dots, x_n\}$ | Something connected to something else, the somethings being $x_1, x_2, \dots, x_n$, which make up $N$ |
| Edge | $E = \{\{x_1, x_2\}, \{x_1, x_3\}, \dots, \{x_{n-1} x_n\}\}$ | The thing that joins the nodes together, which together make up $E$ |
| Adjacent | Nodes connected by an edge | Nodes connected by an edge (!) |
| Neighbours | Degree of $x_a$, $d(x_a)$ where $x_a \in N$ | The number of edges joined to node $x_a$; the number of neighbours for $x_a$ |
| Graph | A pair $G = (N, E)$ of sets such that $E \subseteq \{\{x_a, x_b\}: x_a, x_b \in N\}$ | A group of nodes and edges |

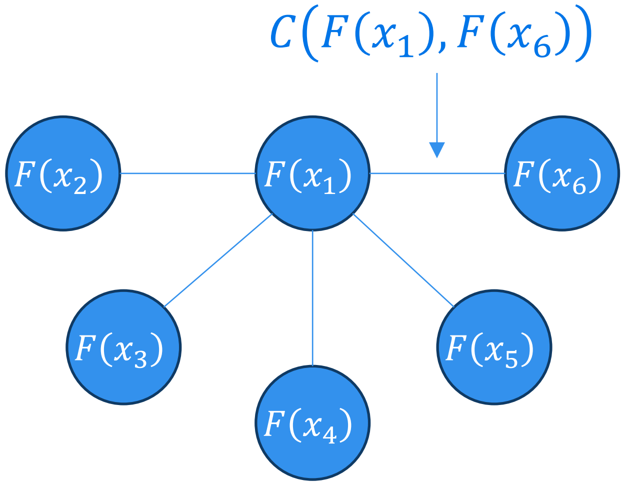
# Graph theory and vine copulas

- In vine copulas, graph theory has following meanings
- Node $\Rightarrow$ distribution function for a particular variable
- Edge $\Rightarrow$ bivariate copula joining two distribution functions
- Neighbours $\Rightarrow$ number of bivariate copulas attached to a particular variable (or its distribution function)
- In other words
  - Nodes could be described $F(x_1), F(x_2)$, etc
  - Edges could be described $C\big(F(x_1), F(x_2)\big)$, etc

# Vine structures

- Two basic vine constructions
- Canonical vine (C-vine)
- Useful if one variable "drives" others…
- …e.g. large stock market and lots of smaller stock markets

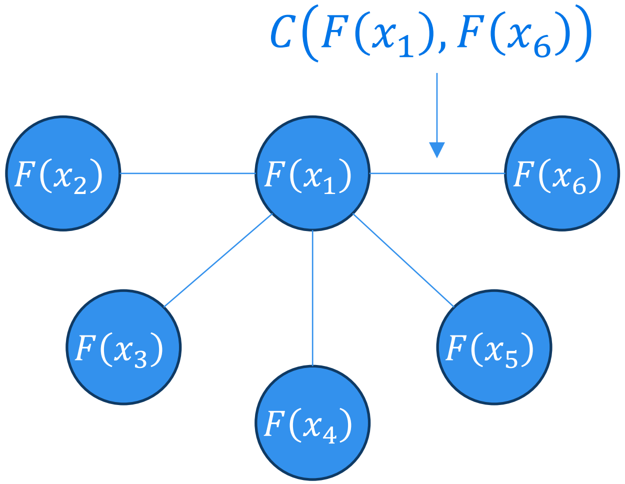
C-vine



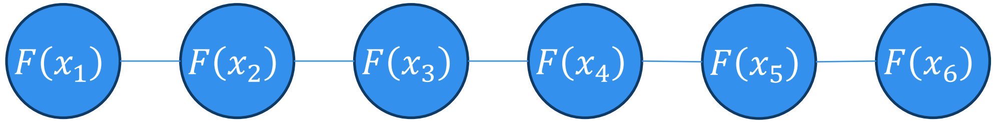
# Vine structures

- Two basic vine constructions
- Canonical vine (C-vine)
- Useful if one variable “drives” others…
- …e.g. large stock market and lots of smaller stock markets
- Drawable vine (D-vine)
- Useful if path of causation is linear

C-vine




D-vine

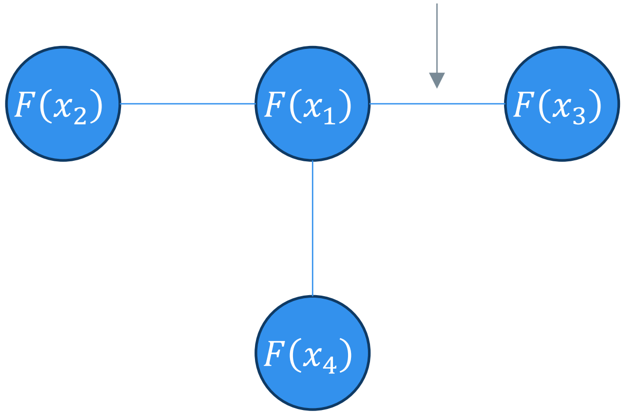


# Vine structures

- It can also be helpful to contrast C-vines and  D-vines with a “traditional” copula

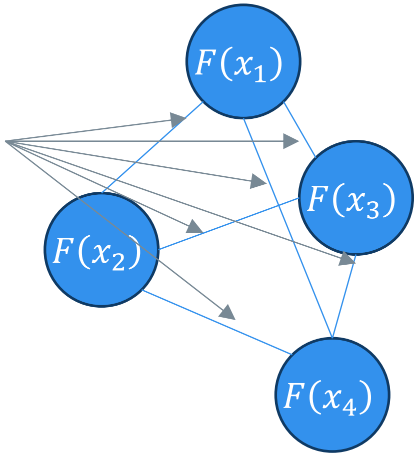
C-vine

$C(F(x_1), F(x_6))$

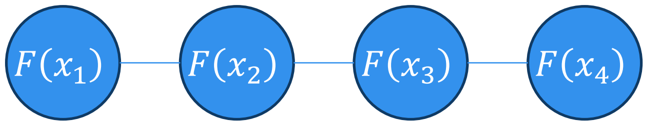



“Traditional” copula

$C(F(x_1), F(x_2), F(x_3), F(x_4))$



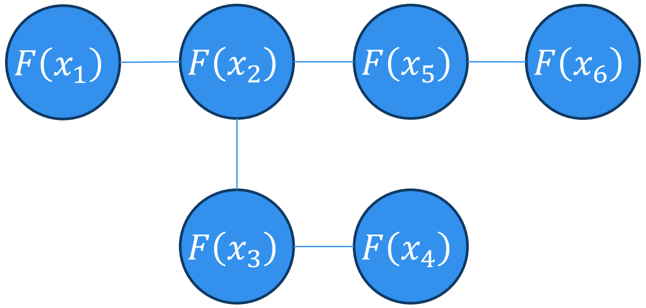

D-vine



# Vine structures

- C-vine and D-vines can be combined…
- ..to give regular vine (R-vine)
- All R-vines are made from C-vines and D-vines…
- …or C-vines and D-vines are special cases of R-vines

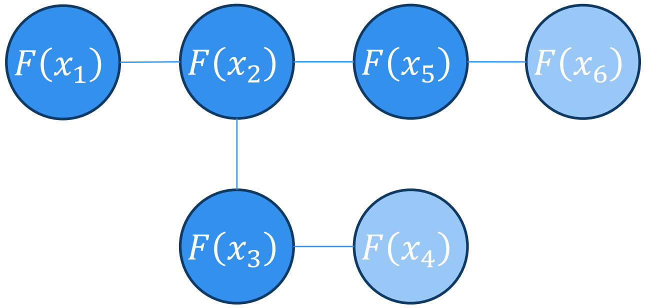
R-vine



# Vine structures

- C-vine and D-vines can be combined…
- ..to give regular vine (R-vine)
- All R-vines are made from C-vines and D-vines…
- …or C-vines and D-vines are special cases of R-vines

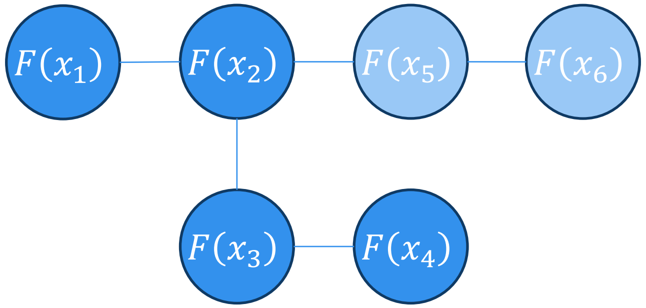
C-vine



# Vine structures

- C-vine and D-vines can be combined…
- ..to give regular vine (R-vine)
- All R-vines are made from C-vines and D-vines…
- …or C-vines and D-vines are special cases of R-vines

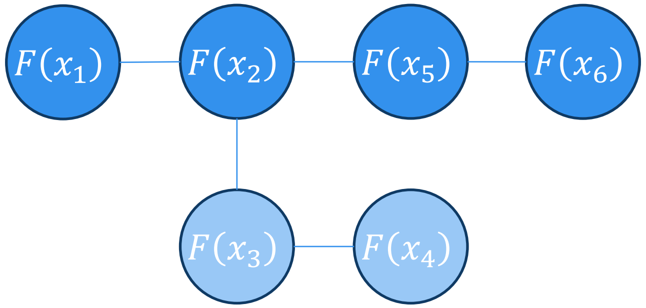
D-vine



# Vine structures

- C-vine and D-vines can be combined…
- ..to give regular vine (R-vine)
- All R-vines are made from C-vines and D-vines…
- …or C-vines and D-vines are special cases of R-vines

D-vine

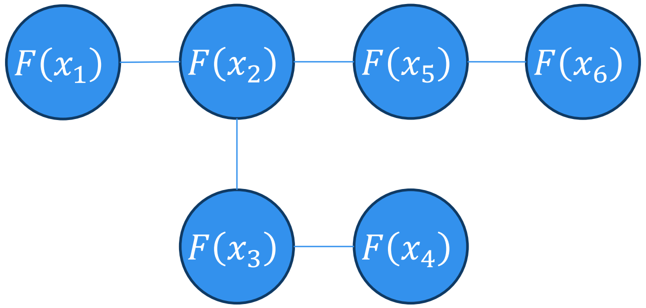



Paul Sweeting

# Vine structures

- An R-vine with $N$ nodes has $N - 1$ edges
- Each edge is part of probability density calculation for the next layer

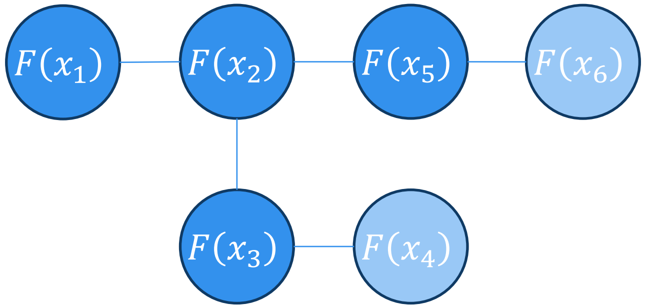
R-vine



# Vine structures

- An R-vine with $N$ nodes has $N - 1$ edges
- Each edge is part of probability density calculation for the next layer
- A C-vine must also have a "centre" component…

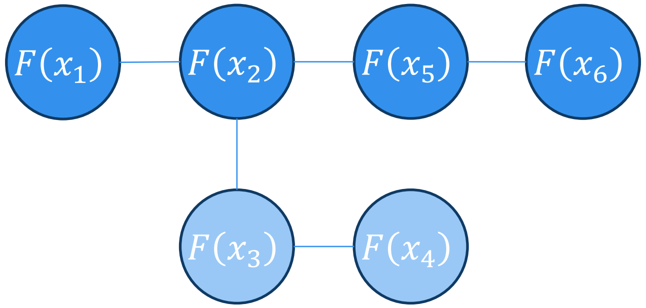
C-vine



# Vine structures

- An R-vine with $N$ nodes has $N - 1$ edges
- Each edge is part of probability density calculation for the next layer
- A C-vine must also have a "centre" component…
- …and a D-vine must form a path

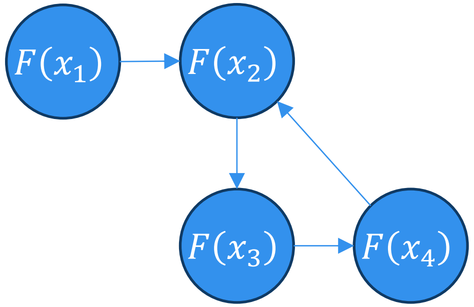
D-vine



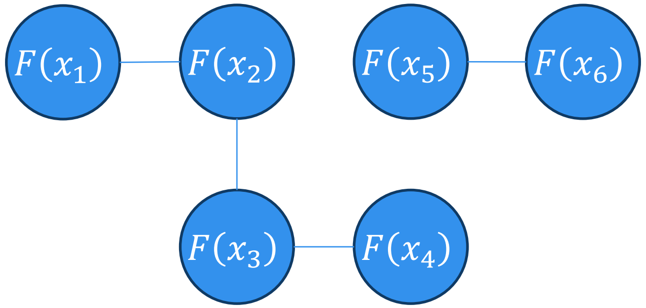
# Vine structures

- R-vines are *undirected* – $F(x_1)$ does not lead to $F(x_2)$ or vice versa; in a *directed* relationship, the edges would be arrows
- Each R-vine is also a *tree* ($T$) so all nodes are *connected* by a unique *path* (there is only one route via other nodes and edges between any two nodes), and there are no *cycles* (ways of following a path to a previously-visited node)

Cyclical, directed graph




Disconnected graph

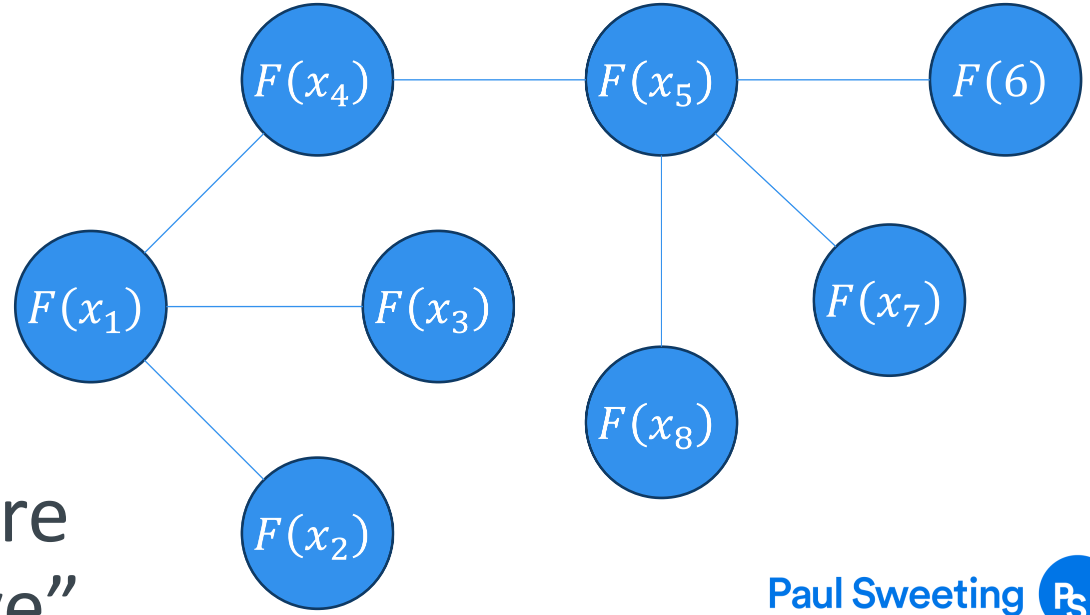


# Complication 1: structure

- As number of variables increases, so do possible structures
- In fact, for $n$ variables, the number of possible structures is

$$f(n) = \frac{n(n-1)(n-2)!2^{\frac{(n-2)(n-3)}{2}}}{2}$$

| n | F(n) |
|---|---|
| 3 | 3 |
| 4 | 24 |
| 5 | 480 |
| 6 | 23040 |




- So finding the "best" structure is "computationally expensive"

Paul Sweeting PS

# Finding the structure

- Finding the structure is the first aspect of fitting a vine copula
- For a D-vine, deciding the structure means working out the order of the nodes in the chain…
- …whilst for a C-vine, it's finding the central node
- For an R-vine, it all gets more complicated

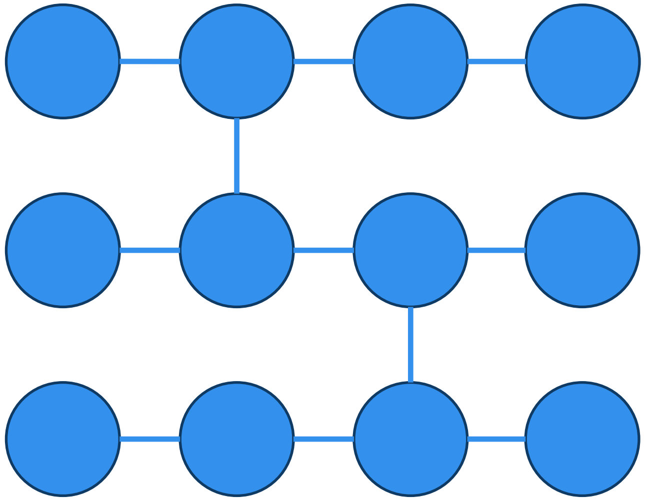
# Suggested approach

- Dißmann et al (2013)* propose a sequential approach
- Uses graph theory…
- …and is similar to decision trees
- Important concepts
  - Minimum spanning trees
  - Greedy algorithms

*Dißmann, J., Brechmann, E.C., Czado, C. and Kurowicka, D., 2013. Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis, 59, pp.52-69.

# Minimum spanning trees

- Remember, an R-vine is a tree, with all nodes connected by a unique path, with no cycles
- A spanning tree is a tree containing all nodes of a graph
- A minimum spanning tree is a spanning tree for which the sum of the edge weights is as small as possible…
- …or for a vine copula, where the sum of the absolute correlations is as high as possible



# Finding a minimum spanning tree

- Several algorithms exist
- Most are “greedy”, i.e. they find locally optimal solutions…
- …which might not be globally optimal
- For vine copulas, all start by calculating all pair correlations
- I use Kendall’s tau – rank correlation is better for copulas

# Prim’s algorithm

- Choose any node
- Find node with highest correlation to first node…
- …and connect it
- Find node with highest correlation to one of first two nodes…
- ..and connect it to that node
- Do not make any connections that create a cycle
- Continue until all nodes connected

# Kruskal’s algorithm

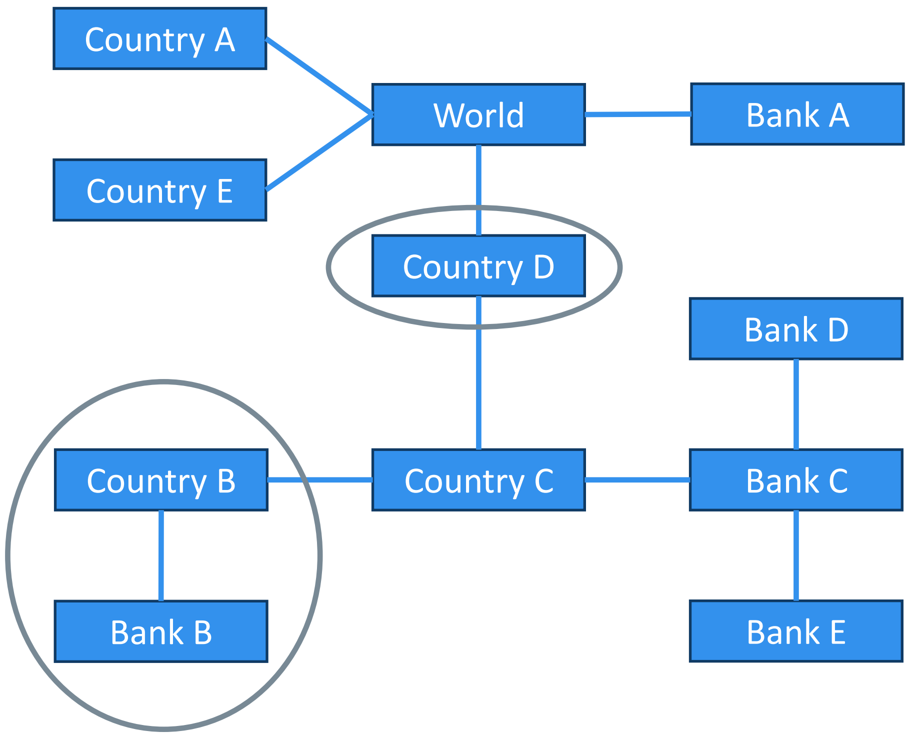
- Order all potential edges (node pairs) by correlation
- Work through connecting pairs of nodes…
- …starting with highest correlation and working down…
- …ignoring any connections that would create a cycle
- Continue until all nodes connected as a single tree
- (Reverse-delete algorithm is similar…
- …except it starts with a fully connected graph…
- …and deletes edges that create cycles)

# Example in Excel

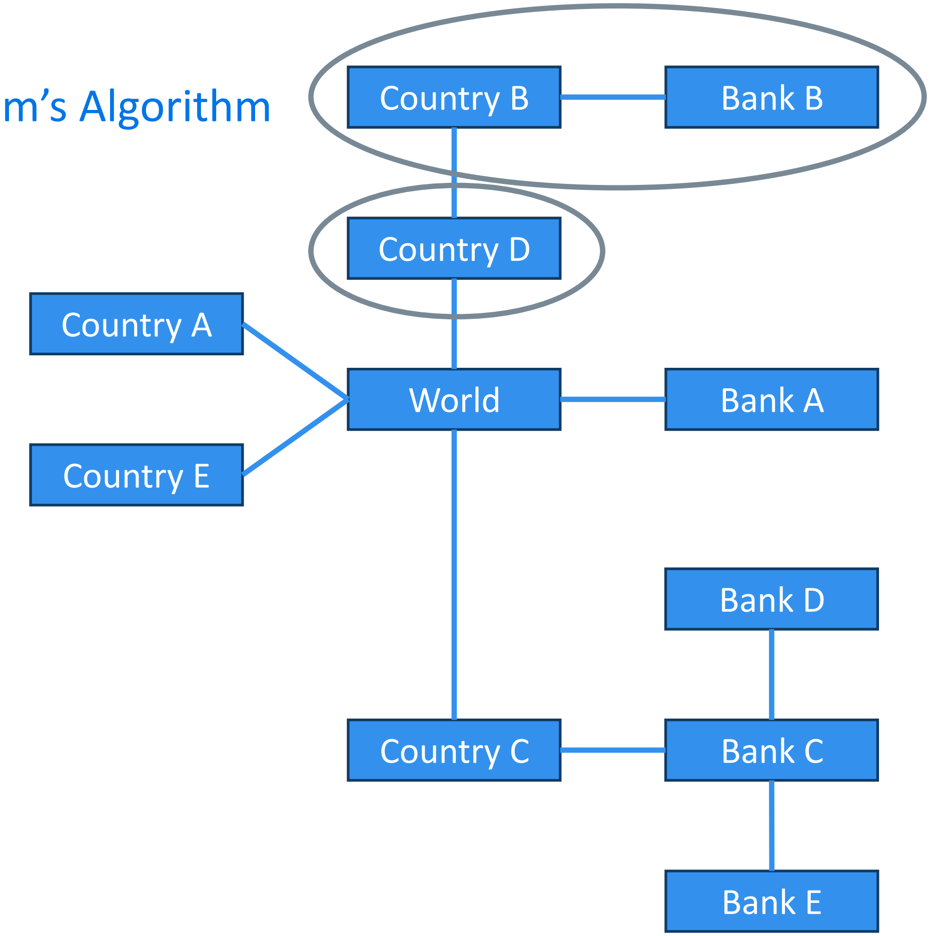
- 120 months of equity return data covering
  - 1 world equity index
  - 5 country equity indices (labelled A-E)
  - 5 banking stocks, one from each country (labelled A-E)
- Demonstrates international nature of some banks…
- …and national links of others
- Also show that different algorithms give different results

# Comparison of Algorithms

## Kruskal's Algorithm



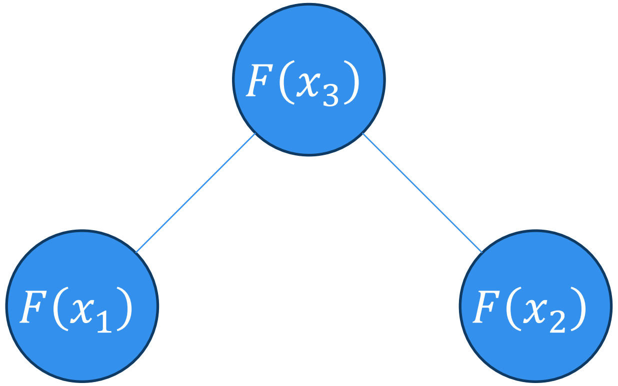

## Prim's Algorithm




Paul Sweeting

# Complication 2: conditionality

- Consider this simple vine copula…
- …built using $C\left(F(x_1), F(x_3)\right)$ and $C\left(F(x_2), F(x_3)\right)$
- Choice and parameterisation of these two copulas implies relationship between $F(x_1)$ and $F(x_3)$…
- …even if it is not explicitly defined
- This needs to be allowed for…
- …as we want to find the structure and parameters that maximise $f(x_1, x_2, x_3)$



# Maximum likelihood estimation

- Checking best fit and finding the parameters can be linked using maximum likelihood estimation (MLE)
- MLE not only gives parameters, also gives you the value of the likelihood function…
- …which can be used in goodness of fit measures
  - likelihood ratio test
  - Akaike Information Criterion (AIC)
  - Bayesian Information Criterion (BIC)
- Even if MLE not used to find parameters, likelihood function is useful in checking goodness of fit

# Maximum likelihood estimation

- Start with probability density function, $f(x_{1,t}, x_{2,t}, \dots, x_{N,t})$…
- …for $N$ variables at time $t$, where $t = 1,2, \dots, T$
- Find parameters that maximise likelihood given observations
- Multiply likelihoods (density functions) for all observations…
- …so likelihood function $L = \prod_{t=1}^{T} f(x_{1,t}, x_{2,t}, \dots, x_{N,t})$…
- …and find parameters that maximise $L$
- In practice, use $\log L$ – turns the product into a sum

# Relevance to conditional density

- For for three dimensions, $L = \prod_{t=1}^{T} f(x_{1,t}, x_{2,t}, x_{3,t})$
- This means we are interested in $f(x_1, x_2, x_3)$
- To evaluate this, we need to break this down
  - to univariate marginal density functions…
  - …and to tractable copula density functions
- Can then find the likelihood function, and optimize…
- …though this can be tricky if large number of parameters
- Two compromises make estimation more manageable

# Simplifying MLE approaches

- One approach is a two-stage process: inference-functions-(for)-margins (IFM) method
- Marginal distributions are fitted first…
- …then resulting parameters are taken as fixed when MLE is used to fit whole distribution, including the copula.
- Second approach works if it is just the copula being fitted: canonical maximum likelihood (CML) estimation
- This is fitting the copula with the distribution functions as inputs, removing the direct link to the underlying data

# Conditional density for Two variables

- Initially focus on two variable approach
- We have distribution functions $F(x_1)$ and $F(x_2)$
- Need to use density functions $f(x_1)$ and $f(x_2)$…
- …found either by differentiation, as $f(x_1) = \frac{\partial F(x_1)}{\partial x_1}$…
- …or by definition, if distribution is integral of density (e.g. Gaussian, Student's t)
- Then define $f(x_1, x_2) = f(x_1|x_2)f(x_2) = f(x_2|x_1)f(x_1)$
- This means $f(x_1|x_2) = \frac{f(x_1,x_2)}{f(x_2)}$

# Conditional density for Two variables

- Similar approach needed for copulas
- We have distribution functions, e.g. $C\big(F(x_1), F(x_2)\big)$
- Need to use density functions, e.g. $c\big(F(x_1), F(x_2)\big)$
- First, define $F(x_1)$, $F(x_2)$ etc as $u_1$, $u_2$ etc
- We can then define $c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}$

# Some copula density functions (Archimedean)

| Copula Name | Distribution Function C(u,v) | Density Function c(u,v) |
|---|---|---|
| Gumbel | $\exp\left[-\{(-\ln u)^{\alpha}+(-\ln v)^{\alpha}\}^{1/\alpha}\right]$ | $C(u,v)\frac{1}{uv}\{(-\ln u)^{\alpha} + (-\ln v)^{\alpha}\}^{\frac{2}{\alpha}-2}(\ln u \ln v)^{\alpha-1} \times \left\{1+(\alpha-1)[(-\ln u)^{\alpha}+(-\ln v)^{\alpha}]^{-\frac{1}{\alpha}}\right\}$ |
| Frank | $-\frac{1}{\alpha}\ln\left(\frac{1}{1-e^{-\alpha}}[(1-e^{-\alpha}) - (1-e^{-\alpha u})(1-e^{-\alpha v})]\right)$ | $\frac{\alpha(1-e^{-\alpha})e^{-\alpha(u+v)}}{[(1-e^{-\alpha})-(1-e^{-\alpha u})(1-e^{-\alpha v})]^2}$ |
| Clayton | $(u^{-\alpha}+v^{-\alpha}-1)^{-1/\alpha}$ | $(1+\alpha)(uv)^{-1-\alpha}(u^{-\alpha}+v^{-\alpha}-1)^{-\frac{1}{\alpha}-2}$ |
| Independence | $uv$ | $1$ |

Paul Sweeting

# Some copula density functions (Elliptical, bivariate)

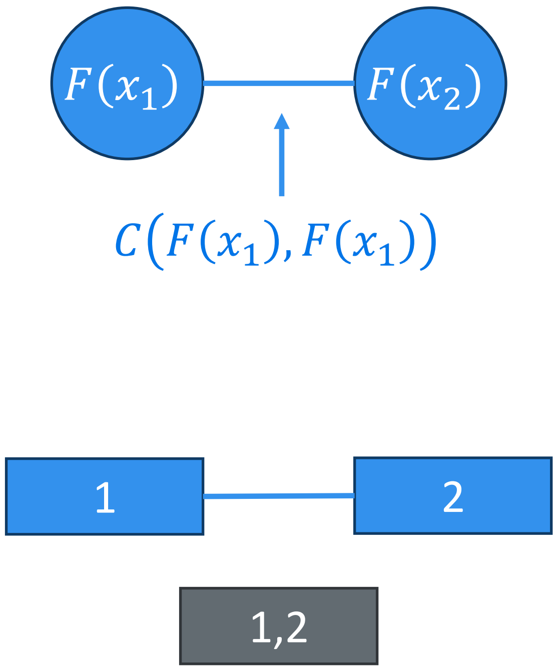
| Copula Name | Distribution Function C(u,v) | Density Function c(u,v) |
|---|---|---|
| Gaussian (normal) | $\Phi_\rho[\Phi^{-1}(u), \Phi^{-1}(v)]$ <br> (where $\Phi_\rho$ is the integral of the bivariate probability density function, and $\Phi$ the univariate) | $\frac{1}{\sqrt{1-\rho^2}} \times \exp\left\{-\frac{\rho^2[\Phi^{-1}(u)^2+\Phi^{-1}(v)^2]-2\rho\Phi^{-1}(v)\Phi^{-1}(v)}{2(1-\rho^2)}\right\}$ |
| Student's t | $t_{\rho,\nu}[t_\nu^{-1}(u), t_\nu^{-1}(v)]$ <br> (where $t_{\rho,\nu}$ is the integral of the bivariate probability density function, and $t_\nu$ the univariate) | $\frac{\Gamma\left(\frac{\nu+2}{2}\right)/\Gamma\left(\frac{\nu}{2}\right)}{\nu\pi dt_\nu[t_\nu^{-1}(u)]dt_\nu[t_\nu^{-1}(v)]\sqrt{1-\rho^2}} \times \left(1+\frac{t_\nu^{-1}(u)^2+t_\nu^{-1}(v)^2-2\rho t_\nu^{-1}(u)t_\nu^{-1}(v)}{\nu(1-\rho^2)}\right)^{\frac{\nu+1}{2}}$ <br> (where $dt_\nu$ is univariate probability density function) |

# Back to conditional density

- We know $C(u_1, u_2) = C\big(F(x_1), F(x_2)\big) = F(x_1, x_2)$
- We also know $f(x_1, x_2) = \frac{\partial^2 F(x_1,x_2)}{\partial x_1 \partial x_2}$…
- …so $f(x_1, x_2) = \frac{\partial^2 C(u_1,u_2)}{\partial x_1 \partial x_2} = \frac{\partial^2 C(u_1,u_2)}{\partial u_1 \partial u_2} \times \frac{\partial F(x_1)}{\partial x_1} \times \frac{\partial F(x_2)}{\partial x_2}$
- Substituting gives $f(x_1, x_2) = c(u_1, u_2) f(x_1) f(x_2)$

# Back to conditional density

- $f(x_1, x_2) = c(u_1, u_2)f(x_1)f(x_2)$ reflects structure of relationship between two variables
- This is helpful when looking at more complex vine structures



# Back to conditional density

- So we have $c(u_1, u_2) = \frac{f(x_1,x_2)}{f(x_1)f(x_2)}$
- But we also have $f(x_1|x_2) = \frac{f(x_1,x_2)}{f(x_2)}$
- Substituting gives $f(x_1|x_2) = c(u_1, u_2)\, f(x_1)$
- This is key in evaluating vine copulas

# MLE and conditional density

- We need to maximise the (log) likelihood function, $L$
- (Or find it to test fits found in other ways)
- For for three dimensions, $L = \prod_{t=1}^{T} f\left(x_{1,t}, x_{2,t}, x_{3,t}\right)$…
- … where $t = 1,2, \ldots, T$
- This means we are interested in $f(x_1, x_2, x_3)$
- To evaluate this, we need to break this down
  - to univariate marginal density functions…
  - …and to tractable copula density functions

# Densities for 3 dimensions

- We know that $f(x_1, x_2) = f(x_1|x_2)f(x_2)$
- Moving to 3 dimensions, we can say that
$$f(x_1, x_2, x_3) = f(x_1|x_2, x_3)f(x_2, x_3)$$
- Using the same relation, we can expand the last term to give
$$f(x_1, x_2, x_3) = f(x_1|x_2, x_3)f(x_2|x_3)f(x_3)$$
- This can be decomposed to include copula densities

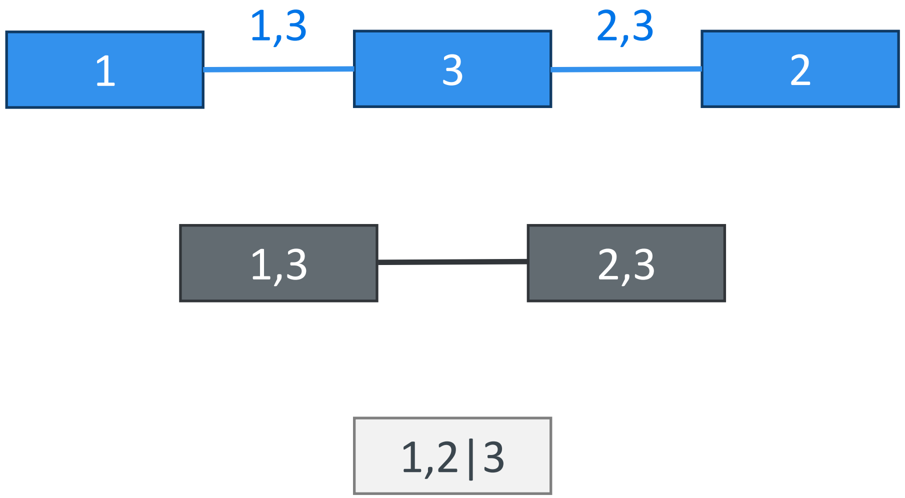
# Densities for 3 dimensions

- Start with $f(x_1, x_2, x_3) = f(x_1|x_2, x_3)f(x_2|x_3)f(x_3)$…
- …and use $f(x_1|x_2) = c(u_1, u_2)\, f(x_1)$ on individual terms
- $$\begin{aligned} f(x_1|x_2, x_3) &= c\left(u_{1|3}, u_{2|3}\right) f(x_1|x_3) \\ &= c\left(u_{1|3}, u_{2|3}\right) c(u_1, u_3) f(x_1) \end{aligned}$$
- $f(x_2|x_3) = c(u_2, u_3)\, f(x_2)$
- Combine these with final term, $f(x_3)$, and rearrange:

$$\begin{aligned} f(x_1, x_2, x_3) = {} & f(x_1)f(x_2)f(x_3) \\ & \times c(u_1, u_3)c(u_2, u_3) \\ & \times c\left(u_{1|3}, u_{2|3}\right) \end{aligned}$$

# Densities for 3 dimensions

- There is a graphical analogue for this…
- …as with the first probability and copula density relation…
- $$f(x_1, x_2, x_3) = f(x_1)f(x_2)f(x_3) \times c(u_1, u_3)c(u_2, u_3) \times c\left(u_{1|3}, u_{2|3}\right)$$




- …although this decomposition is not unique…
- …unlike the vine copula shown

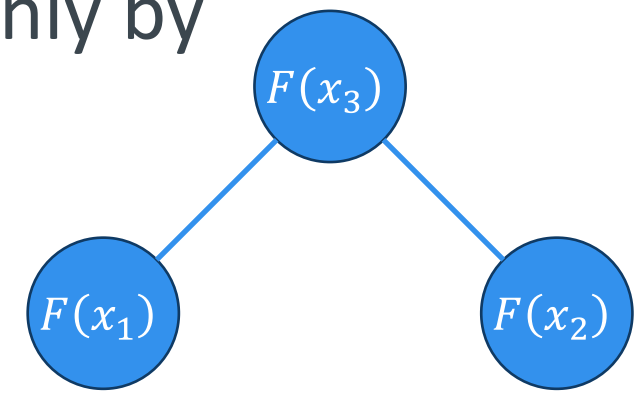
Paul Sweeting

# Densities for 3 dimensions

- In fact, there are 6 possible decompositions…
- $$\begin{aligned} f(x_1, x_2, x_3) &= f(x_1|x_2, x_3)f(x_2|x_3)f(x_3) \\ &= f(x_1|x_2, x_3)f(x_3|x_2)f(x_2) \\ &= f(x_2|x_1, x_3)f(x_1|x_3)f(x_3) \\ &= f(x_2|x_1, x_3)f(x_3|x_1)f(x_1) \\ &= f(x_3|x_1, x_2)f(x_1|x_2)f(x_2) \\ &= f(x_3|x_1, x_2)f(x_2|x_1)f(x_1) \end{aligned}$$
- …so six versions of the previous equations!

# Densities for 3 dimensions

- However, the structure shown is given only by
$$f(x_1,x_2,x_3) = f(x_1)f(x_2)f(x_3) \times c(u_1,u_3)c(u_2,u_3) \times c\left(u_{1|3},u_{2|3}\right)$$




- The challenge (for MLE) is to find $c\left(u_{1|3},u_{2|3}\right)$…
- …since this is the only term not initially defined explicitly

# Conditional independence

- One solution: assume that $x_1$ and $x_2$ independent given $x_3$…
- …implying an independence copula
- This implies $c\left(u_{1|3}, u_{2|3}\right) = 1$…
- …so joint density simplifies to

$$f(x_1, x_2, x_3) = f(x_1)f(x_2)f(x_3) \\ \times c(u_1, u_3)c(u_2, u_3)$$

- If this assumption made for all conditional copulas, life much easier – only need to worry about "specified" copulas…
- …but need to be confident it is realistic!

# Densities for 3 dimensions
# (The hard way)

- If $x_1$ and $x_2$ *not* independent given $x_3$…
- …need to select a copula…
- …whose fit can be assessed as part of MLE process
- However, still need to find $u_{1|3}$ and $u_{2|3}$
- Do this by partially differentiating selected copula
- $u_{1|3} = F(x_1|x_3) = \dfrac{\partial C\big(F(x_1),F(x_3)\big)}{\partial F(x_3)} = \dfrac{\partial C(u_1,u_3)}{\partial u_3}$
- $u_{2|3} = F(x_2|x_3) = \dfrac{\partial C\big(F(x_2),F(x_3)\big)}{\partial F(x_3)} = \dfrac{\partial C(u_2,u_3)}{\partial u_3}$

Paul Sweeting

# h function

- $\partial C(u_1, u_3)/\partial u_3$ also known as the h function, $h(u_1, u_3)$
- First term in expression is conditioned on second
- This is used in likelihood functions, and in simulation

# Some h functions (Archimedean)

| Copula Name | h Function h(u,v) |
|---|---|
| Gumbel* | $C(u,v)\frac{1}{v}(-\ln u)^{\alpha} \times [(-\ln u)^{\alpha} + (-\ln v)^{\alpha}]^{\frac{1}{\alpha}-1}$ |
| Frank** | $\frac{e^{-\alpha v}(e^{-\alpha u}-1)}{(e^{-\alpha}-1)+(e^{-\alpha u}-1)(e^{-\alpha v}-1)}$ |
| Clayton* | $\left[(uv^{\alpha+1})^{-\frac{\alpha}{\alpha+1}} + 1 - v^{-\alpha}\right]^{-\frac{1}{\alpha}}$ |

*Aas, K., Czado, C., Frigessi, H. and Bakken H., 2006. Pair-copula constructions of multiple dependence, Sonderforschungsbereich 386, Paper 487;
**Ahn, K.-H., 2021. Streamflow estimation at partially gaged sites using multiple-dependence conditions via vine copulas, Hydrol. Earth Syst. Sci., 25, 4319–4333

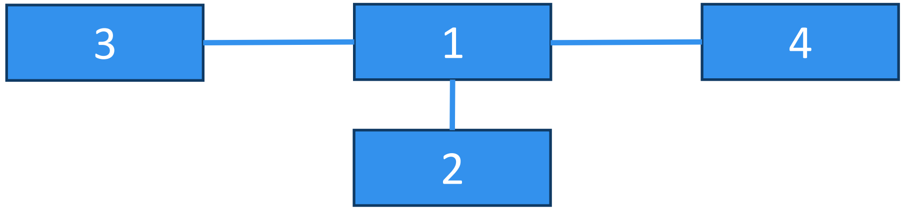
# Some h functions (Elliptical, bivariate)

| Copula Name | h Function h(u,v) |
|---|---|
| Gaussian (normal)* | $\Phi\left(\frac{\Phi^{-1}(u)-\rho\Phi^{-1}(v)}{\sqrt{1-\rho^2}}\right)$ |
| Student's t* | $t_{\nu+1}\left(\frac{t_\nu^{-1}(u)-\rho t_\nu^{-1}(v)}{\sqrt{\frac{\left(\nu+\left(t_\nu^{-1}(u)\right)^2\right)(1-\rho^2)}{\nu+1}}}\right)$ |

*Aas, K., Czado, C., Frigessi, H. and Bakken H., 2006. Pair-copula constructions of multiple dependence, Sonderforschungsbereich 386, Paper 487

# Densities for 4 dimensions: C-vine

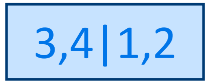
$$f(x_1, x_2, x_3, x_4) = f(x_1)f(x_2)f(x_3)f(x_4)$$




$$\times\, c(u_1, u_3)c(u_1, u_2)c(u_1, u_4)$$


$$\times\, c\left(u_{3|1}, u_{2|1}\right)c\left(u_{2|1}, u_{4|1}\right)$$


$$\times\, c\left(u_{3|1,2}, u_{4|1,2}\right)$$



# Densities for 4 dimensions: D-vine

$$f(x_1, x_2, x_3, x_4) = f(x_1)f(x_2)f(x_3)f(x_4)$$

1 — 2 — 3 — 4

$$\times\, c(u_1, u_2)c(u_2, u_3)c(u_3, u_4)$$

1,2 — 2,3 — 3,4

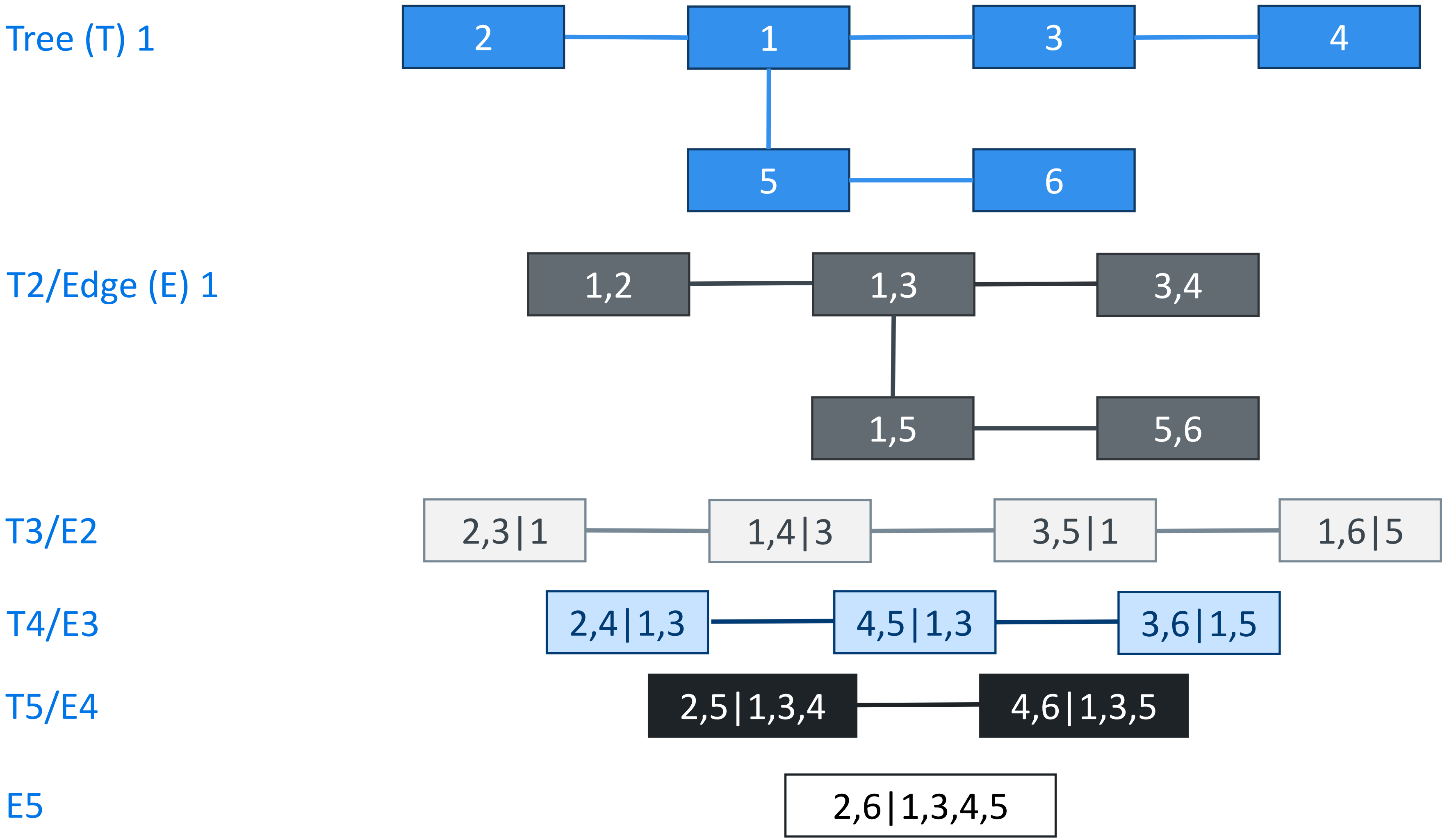
$$\times\, c\left(u_{1|2}, u_{3|2}\right)c\left(u_{2|3}, u_{4|3}\right)$$

1,3|2 — 2,4|3

$$\times\, c\left(u_{1|2,3}, u_{4|2,3}\right)$$

1,4|2,3

# Densities for an R-vine



# Densities for an R-vine

- This can also be used to derive the formula for the density of the function $f(x_1, x_2, x_3, x_4, x_5, x_6)$ as:

$$
\begin{gathered}
f(x_1)f(x_2)f(x_3)f(x_4)f(x_5)f(x_6) \\
\times c(u_1, u_2)c(u_1, u_3)c(u_1, u_5)c(u_3, u_4)c(u_5, u_6) \\
\times c\left(u_{2|1}, u_{3|1}\right)c\left(u_{1|3}, u_{4|3}\right)c\left(u_{3|1}, u_{5|1}\right)c\left(u_{1|5}, u_{6|5}\right) \\
\times c\left(u_{2|1,3}, u_{4|1,3}\right)c\left(u_{4|1,3}, u_{5|1,3}\right)c\left(u_{3|1,5}, u_{6|1,5}\right) \\
\times c\left(u_{2|1,3,4}, u_{5|1,3,4}\right)c\left(u_{4|1,3,5}, u_{6|1,3,5}\right) \\
\times c\left(u_{2|1,3,4,5}, u_{6|1,3,4,5}\right)
\end{gathered}
$$

# Densities for an R-vine

- Can also work this out from first principles recursively using $f(x_1, x_2) = f(x_1|x_2)f(x_2) = f(x_2|x_1)f(x_1)$…
- …and $f(x_1|x_2) = c(u_1, u_2)\, f(x_1)$
- $f(x_1, x_2, x_3, x_4, x_5, x_6) = f(x_1)f(x_2|x_1)f\left(x_3|x_{1,2}\right)$ $f\left(x_4|x_{1,2,3}\right)f\left(x_5|x_{1,2,3,4}\right)f\left(x_6|x_{1,2,3,4,5}\right)$
- $f(x_2|x_1) = c(u_1, u_2)f(x_2)$
- $f\left(x_3|x_{1,2}\right) = c\left(u_{2|1}, u_{3|1}\right)f\left(x_{3|1}\right)$

  $$= c\left(u_{2|1}, u_{3|1}\right)c(u_1, u_3)f(x_3)$$
- …and so on!

# Take-aways

- Vine copulas are hugely flexible…
- …and fairly easy to fit
- However, checking whether fit beats other options…
- …such as “ordinary” copulas…
- …or even multivariate distributions…
- …requires some complex maths!
- Of course, you could always do that bit in R…