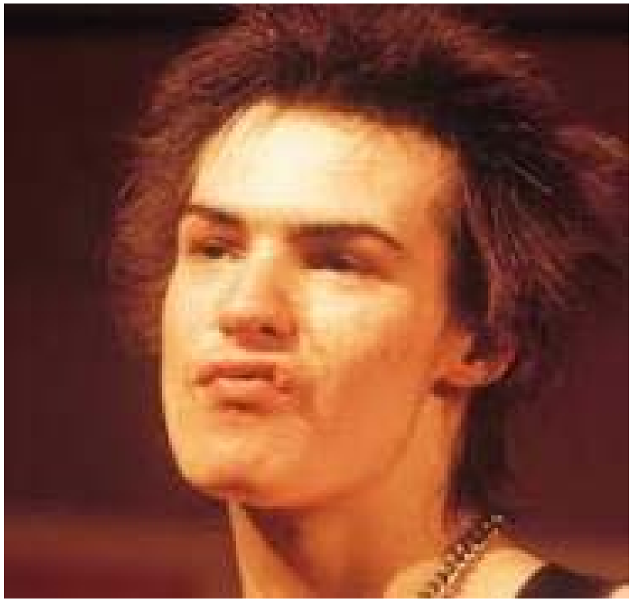
Banc Ceannais na hÉireann
Central Bank of Ireland
Eurosystem

Aggregation and the use of copulas

Tony Jeffery

4th March 2013

Disclaimer 1



Disclaimer 2

Any views or opinions expressed herein are my own and do not represent those of my employer the Central Bank of Ireland

Assumptions

- Everybody knows what a PDF and CDF is
- Everybody knows what "var" means
- Everybody knows that Solvency II uses a one year one in 200 var measure
- I use the term "varpoint" to mean the value calculated under this measure

Aggregation is

- Very important

Aggregation is

- Very important
- Very difficult due to lack of data

Aggregation is

- Very important
- Very difficult due to lack of data
- Conceptually difficult

Non-life

Household & Motor

Total Losses =

Household Losses

+ Motor Losses

Life

Annuity Book Interest rate & Mortality

Total Losses = F(Int, Mort)

Methods of Aggregation: Covariance Matrix

$$VARP\ (X,Y)^2 =$$

$$VARP(X)^2 + 2rVARP(X)VARP(Y) + VARP(Y)^2$$

Or

$$VARP = V^TRV$$

Methods of Aggregation: Covariance Matrix

- Single VARPoint
- But big disadvantage is

Methods of Aggregation: Covariance Matrix

- Single VARPoint
- But big disadvantage is

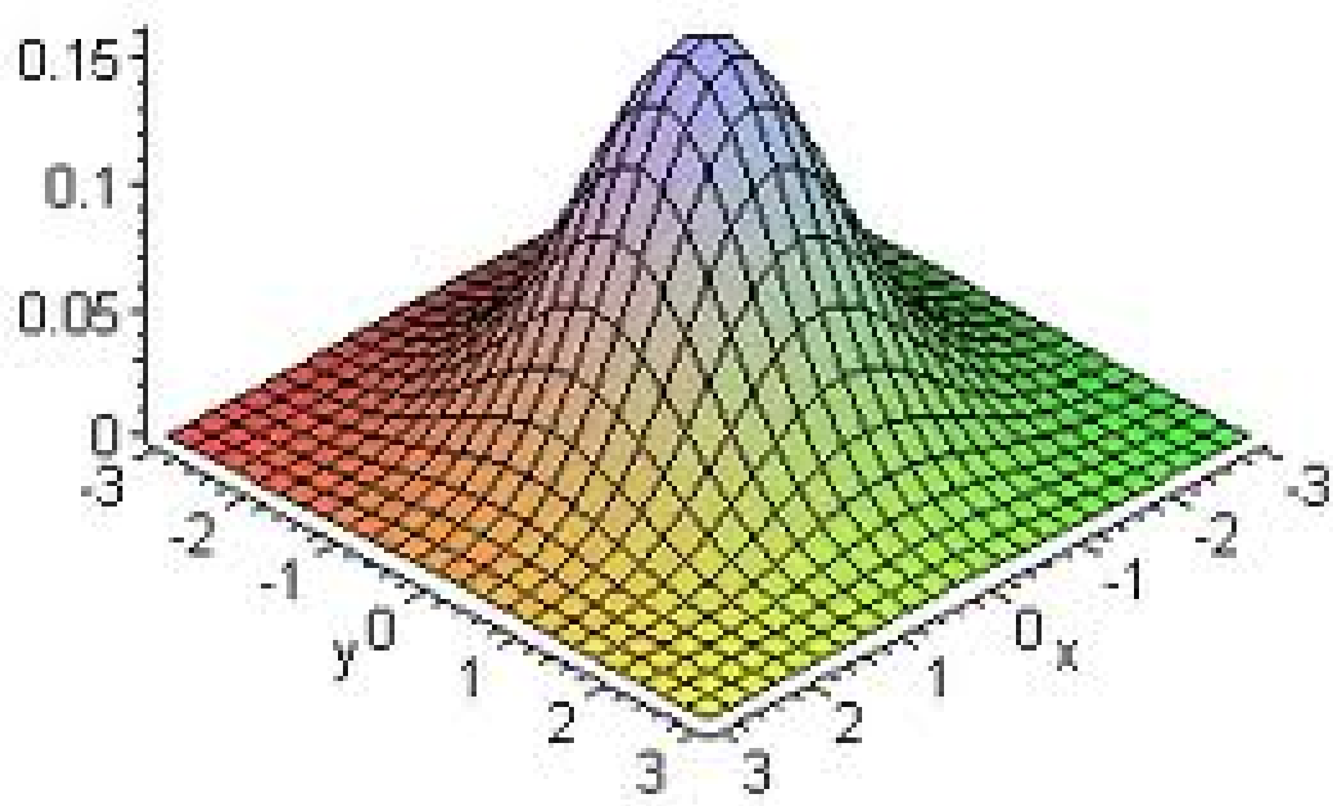
it's wrong!

(unless the JDF is elliptical)

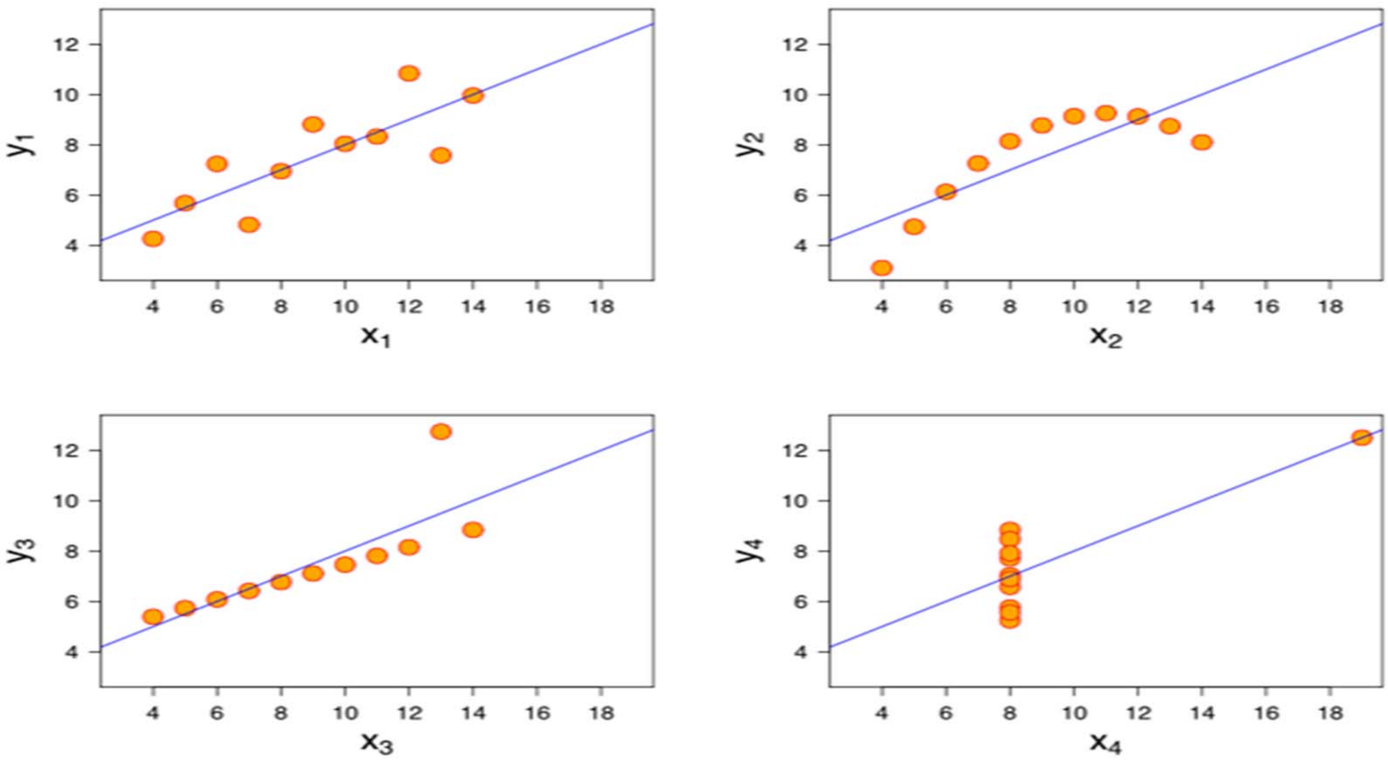
Methods of Aggregation: Risk Driver Approach

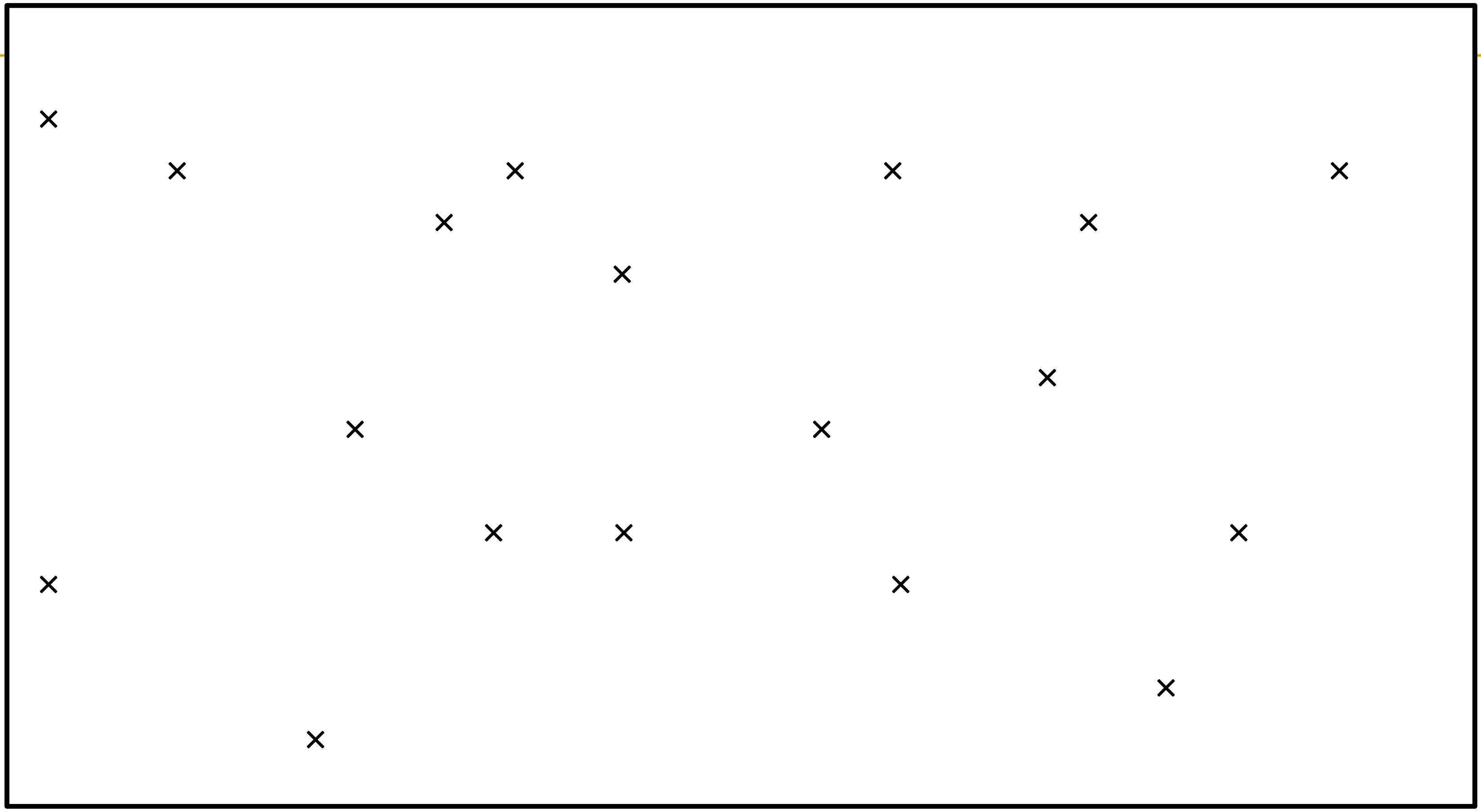
- Hard to do
- Not used in practice much

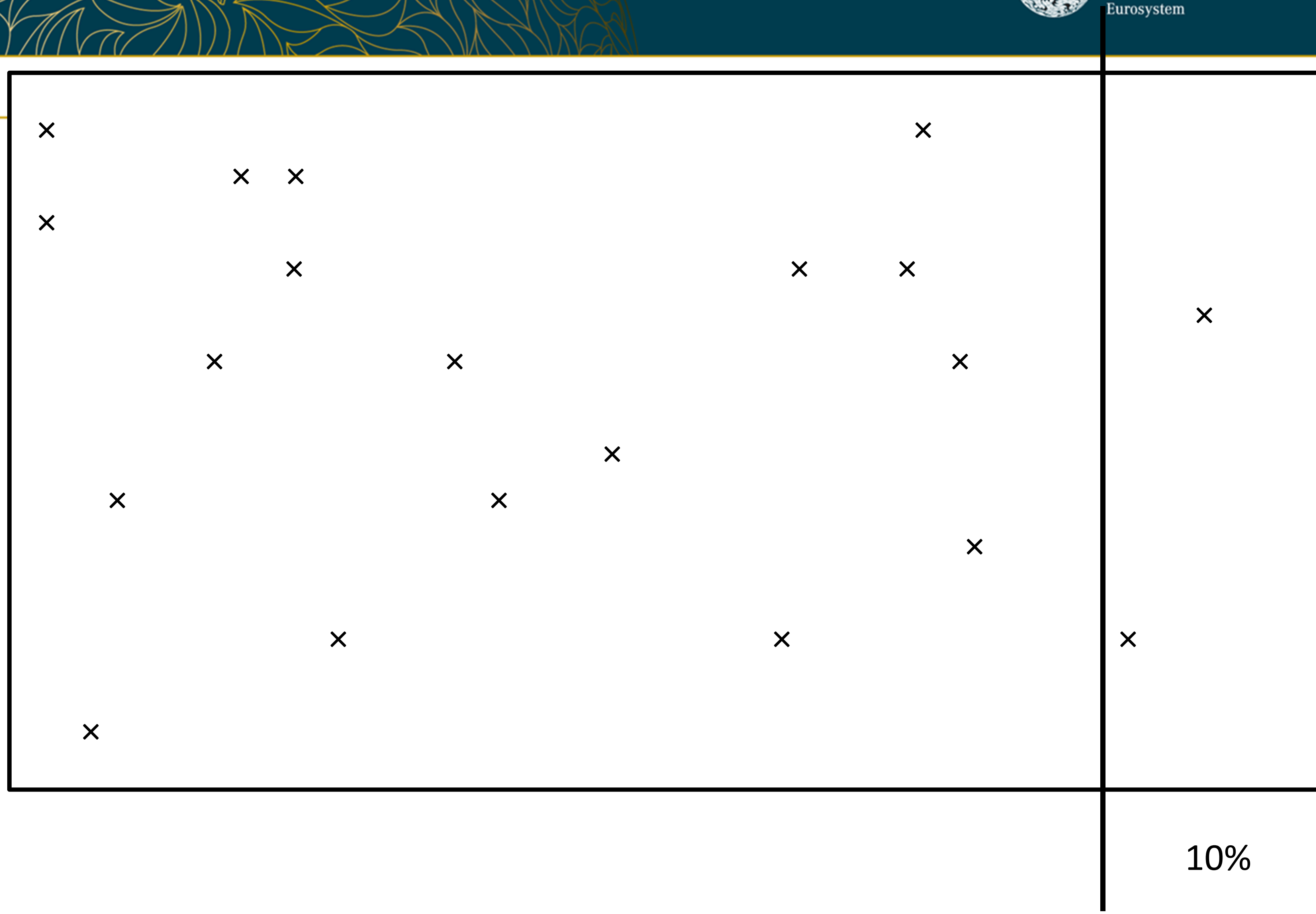
Methods of Aggregation: Multivariate Distributions

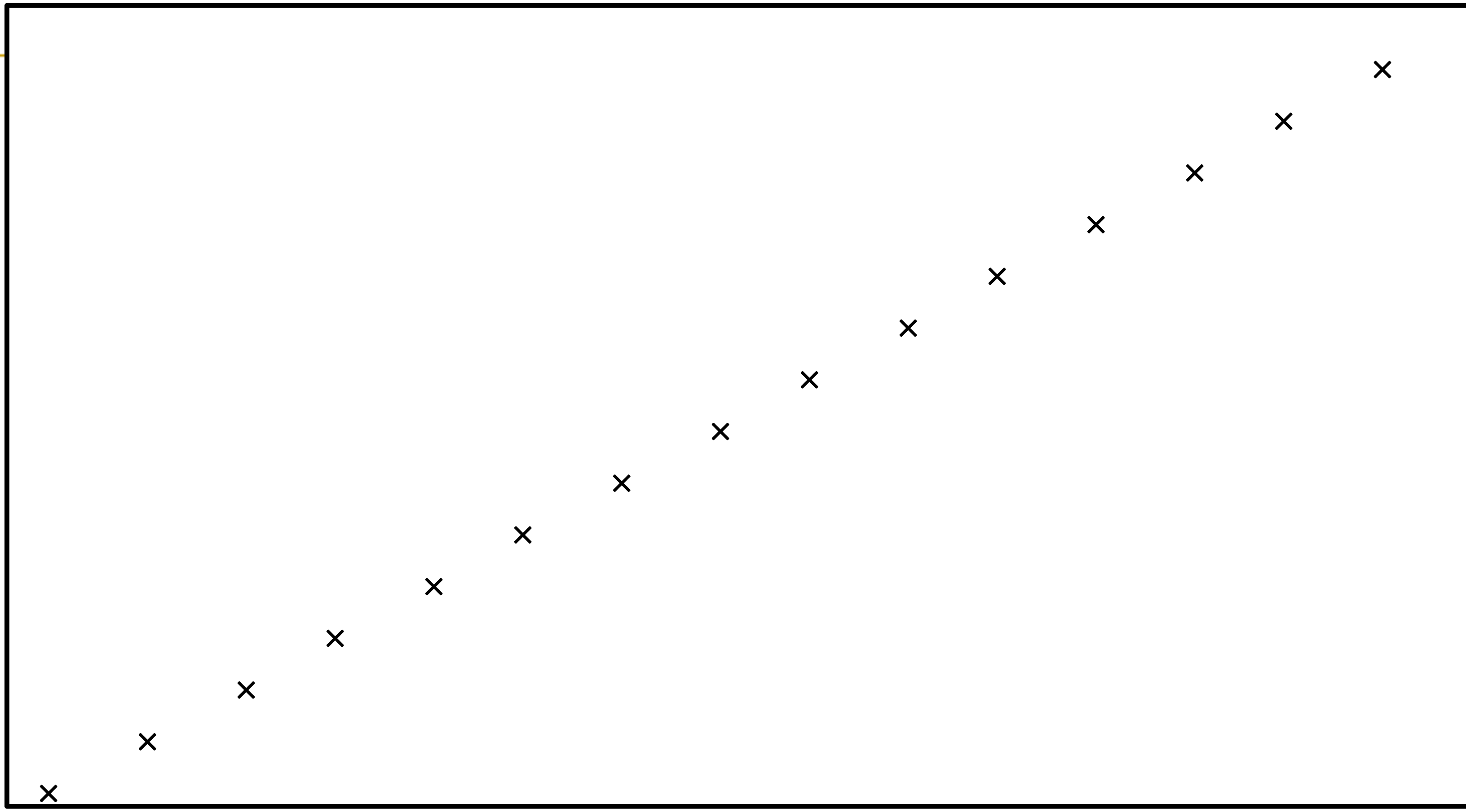


Anscombe's Quartet






10%



In probabilistic terms, $C : [0,1]^d \to [0,1]$ is a d-dimensional **copula** if C is a joint cumulative distribution function of a d-dimensional random vector on the unit cube $[0,1]^d$ with uniform marginals.

- a) A copula is an increasing function in each of its inputs
- b) if all the marginal distribution functions bar one are equal to one then the copula must be equal to the value of the remaining marginal
- c) the copula must always return a non-negative probability
- .

Sklar's theorem states that a multivariate cumulative distribution function

$$H(x_1,\ldots,x_d) = \mathbb{P}[X_1 \leq x_1,\ldots,X_d \leq x_d]$$

of a random vector $(X_1, X_2, \ldots, X_d)$ with marginals $F_i(x) = \mathbb{P}[X_i \leq x]$ can be written as

$$H(x_1,\ldots,x_d) = C\left(F_1(x_1),\ldots,F_d(x_d)\right),$$

where C is a copula.

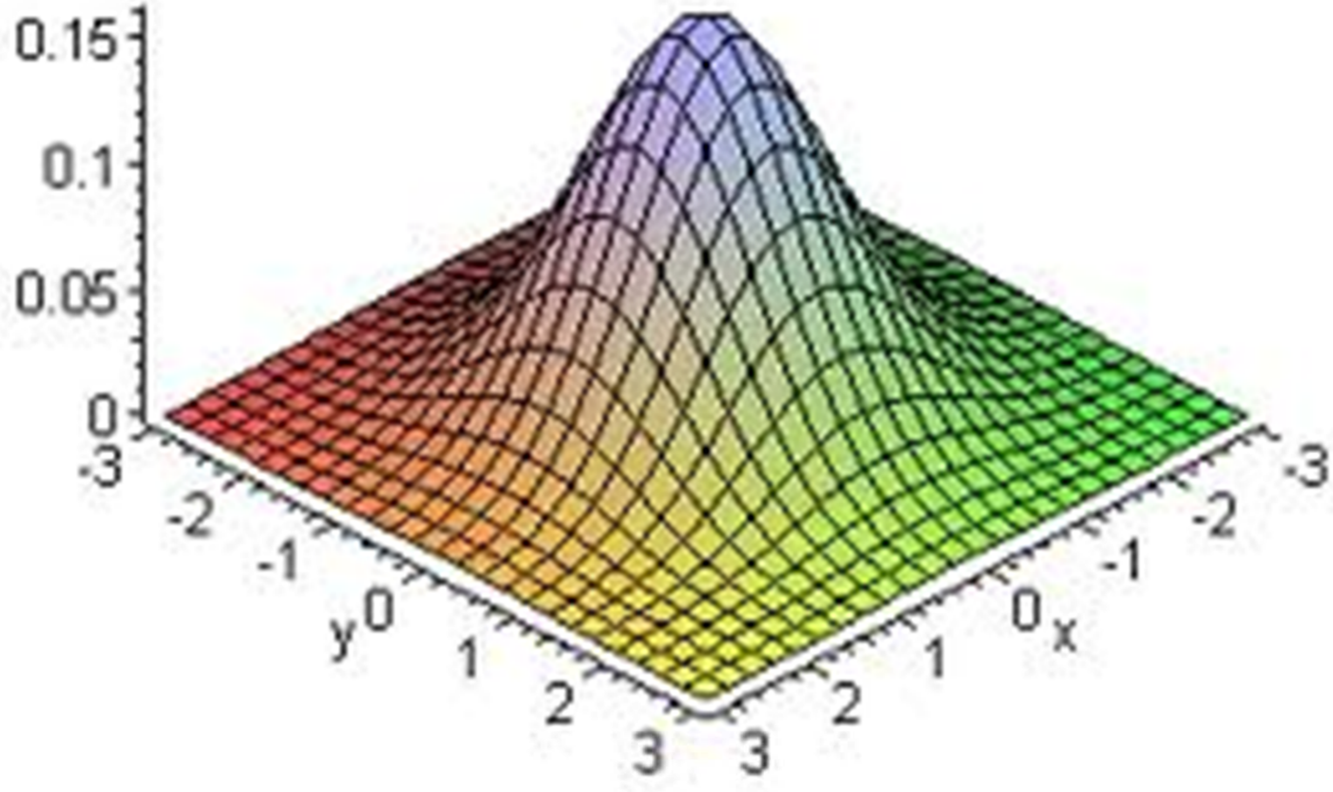
The theorem also states that, given H, the copula is unique on $\operatorname{Ran}(F_1) \times \cdots \times \operatorname{Ran}(F_d)$, which is the cartesian product of the ranges of the marginal cdf's. This implies that the copula is unique if the marginals F_i are continuous.

The converse is also true: given a copula $C : [0,1]^d \rightarrow [0,1]$ and margins $F_i(x)$ then $C\left(F_1(x_1),\ldots,F_d(x_d)\right)$ defines a d-dimensional cumulative distribution function.

A better way of looking at it

Consider a pair of random variables and an outcome (X,Y). Look at the probability (CDF) of X and then Y and then of X and Y together. This gives a mapping from the probabilities as a pair of X and Y i.e. from the unit square to the unit line

This mapping is a copula


0.15
0.1
0.05
0
-3
-2
-1
y 0
1
2
3
3
2
1
0 x
-1
-2
-3

Measures of Dependency

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)}.$$

$$\tau = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{\frac{1}{2}n(n-1)}.$$

Cholesky Decomposition

$$C^T C = M$$

Cookbook

1. Establish the marginals from your experience
2. Workout the pairwise dependency to give your first try at a covariance matrix
3. Make it PSD
4. Apply Cholesky decomposition
5. Generate loads of sample vectors

Cookbook

6. Use the lower triangular matrix to get sample points with the right dependency. These are now your “runs” of the model but only on the order of the variables
7. Go backwards from these orders of each marginal to get the actual value (so this is where fat tail ness of the variables come in).
8. Add up the value of the marginals to get the total impact of that “run”
9. Rank all the runs’ impacts
10. Take the 1 in 200

Cookbook

11. Pour yourself a beer

Solvency II: Level 1 says:

"*As regards diversification effects, insurance and reinsurance undertakings may take account in their internal model of dependencies within and across risk categories, provided that supervisory authorities are satisfied that the system used for measuring those diversification effects is adequate.*"

Solvency II: Level 2 says

"*the system used for measuring diversification effects shall take into account: (a) any material non-linear dependence and any material lack of diversifications under extreme scenarios…*"

What the bard says

"when sorrows come, they come not single spies but in battalions"

Hamlet Act 4 Scene 5

Tail dependence

$$\text{Limit } \Pr\{x > F_q^{-1}(x) \mid F_q^{-1}(y)\}$$

Tail dependence

- You have thought about it and believe that you are not exposed to it (brave)
- You are using a tail dependent copula (so how much effect does that have)
- You are using a Gaussian Copula but are augmenting the effect in some way, such as adding some home made cat scenarios (so how much effect, are you double counting?)
- You are using a Gaussian copula but have adjusted the dependency factors to allow for the extra bit in the tail (so how much does that add up to)

Other Copulas

- t copula

Degrees of freedom	increase
10	8%
5	20%
2	31%

Other Copulas

- Frechet- Höffding

Other Copulas

- Archimedean
  - Gumbel
  - Clayton
  - Frank
  -

Archimedean

You are modelling the returns on a portfolio of ten high-yield corporate bonds. Your definition of default is that the return in any one year is less than minus 60%. The probability that a single bond will default is 10%. You believe that the returns on the bonds are linked by a Gumbel copula, with a single parameter α = 2.

The generator function for the Gumbel copula is (−ln F(x))α.

(i) Calculate the probability that all ten bonds will have defaulted in one year's time.

Archimedean

The generator function for this Gumbel copula is (−ln F(x))2 .

We are interested here in x = −0.6, for which F(x) = 0.1.

So the generated result for a single loss is (−ln(0.1))2 = 5.3019.

Multiply this by 10 to allow for the number of bonds in the portfolio, giving

53.019.

Apply the pseudo inverse of the generator function (square root, take thenegative, take the exponential) to obtain 0.000688 or 0.0688%.

Final thoughts

- Life example
- PSD problems