Dependencies: The briefest guide!

Pragmatically, what is a copula?

It is possible that thing *x* and thing *y* tend to be large at the same time, as shown with the blue dots in Fig. 1. Various words are used to describe this state of non-independence, i.e. *correlation*, *dependency*, *co-occurrence*, *interaction*, *compounding*, although which you prefer is not important. A copula is the next step from correlation coefficients (i.e. Pearson's and Spearman's *r*) taught in school. It is a statistically well-established means of associating two random variables, such as two types of loss (*x*, *y*). But, most importantly for risk, a copula is vital to quantify with any reliability how very large and rare losses might co-occur in a way that is an existential problem for an insurer.

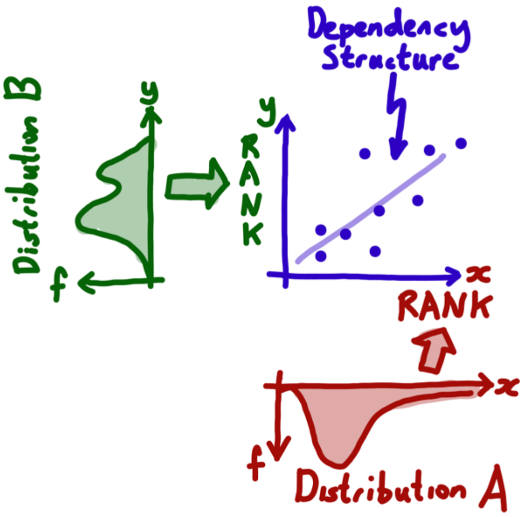

Fig 1: Illustration of what a copula is. It's the blue bit. It is a method for formally describing (approximating) how ranks of two (or more) observed quantities (e.g. A and B) relate to each other, called their '*dependency structure*'. The shape of the distributions that are joined is irrelevant to the copula! A and B are shown as frequency distributions, and their data can be ranked.

Some key copulas

The strength with which common copulas force a correlation in the tail (i.e. rare extremes happening together) differs. They are listed below from the weakest (i.e. closer to independence) to the strongest (i.e. happen together).

- *Gaussian* – weakest, non-zero at pragmatic levels (e.g. 1-in-200) although zero in the limit.
- *t-copula*: tends to Gumbel at df = 1, and Gaussian as df increases (i.e. little different df > 25).
- *Gumbel*
- *Flipped Clayton* – strongest.

The European Banking Authority proposes that Gaussian Copulas are not to be used for operational risk modelling, suggesting instead that a T-Student copula with 3 or 4 degrees of freedom is typically more appropriate to capture the dependencies between the operational risk events (EBA, 2015).

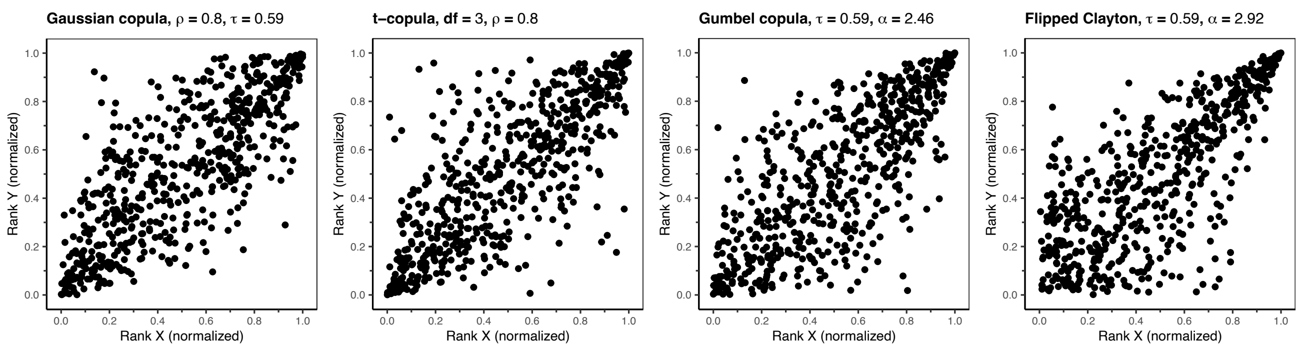

Fig 2: Illustration of some important copulas, displayed as scatter plots. Tight, linear, concentrations of dots, particularly in the top right corner (largest events) will most affect tail-end risk. Without correlation, the dots would be evenly and randomly scattered.

How might dependency arise? Namely, thing causing us to need copulas in our modelling.

Dependency prevents diversification, creating additional accumulation risk with implications (e.g. for required capital to be held for solvency). The following are ways in which dependency might arise. A copula is a tool that can be applied to describe and explore (e.g. by simulation modelling) any of these dependencies.

- Exposure of two classes of asset to the same driver such as a hazard (e.g. life and property in a wildfire, or motor and residential property in a hail-storm).
- Spatial correlation of assets at risk (e.g. proximal houses in a flood). Think of the idea that an event might have a 'footprint' and area on the ground that it affects, which likely contain more than one asset at risk.
- Exposure of a single class of assets (e.g. residential property) to hazards that have a tendency to co-occur (e.g. flooding and wind).
- Supply chain mechanisms whereby supplier or customers, or internal equivalents for larger distributed entities, cannot fulfil obligations or deviate from expected behaviours.
- Financial mechanisms, such as stock market movement, interest rates or other shared influences.
- It is likely that you can think of some more examples.

Impact on diversification benefit - Some key points

Diversification benefit (i.e. reduction of 1-in-200 year risk from a completely correlated and un-diversified case) is greater:

- for heavier tailed loss distributions (e.g. Pareto/power law distributions instead of normal/Gaussian distributions)
- where losses in the two classes/lines of business that are dependent are of roughly equal magnitude
- for data or copulas with stronger upper-tail dependency (e.g. Gumbel *vs* Gaussian) although, somewhat counterintuitively correlation metrics for the whole distribution (i.e. Pearson/Spearman) *might* actually decrease in this scenario. This simply highlights the need to use appropriate measures of dependency.

And, more years in physical or statistical simulations of potential realisations of risk (e.g. of next year) are needed to achieve a stable result for

- Heavier tails (i.e. the more relatively extreme rare events are).
- Higher thresholds of what extreme is (e.g. 1-in-200 years *vs* 1-in-10 years), because there are fewer relevant events.
- For correlation rather than an individual peril, yet again as there are fewer relevant events in the same simulation length (e.g. 10,000 years).

Questions to ask of an insurer (or broker / reinsurer / regulator)

- Have you considered dependencies when quantifying your tail risks including diversification benefit?
- What metric? (Pearson's r, Spearman's r, Kendall's τ, co-occurrence over threshold, copula parameters) And, why?
- Why did you use a Gaussian Copula?
- Have you tried any copulas other than Gaussian?
- Why are you comfortable using a Gaussian copula? i.e. do you know, even indicatively how sensitive your analysis is to this assumption as compared to another choice?
- How many degrees of freedom did you use for the t-Copula, and why?
- Have you done a sensitivity test on the method & assumptions you used to assess dependency?
- Precisely what quantity is dependent? E.g. hazard metric, loss, and on what time-frame? (e.g. day, week, year).
- Do you understand the origin of key dependencies, such that you can justify their presence or absence in your analysis?

EBA (2015) EBA: Final report: Final Draft Regulatory Technical Standards on the specification of the assessment methodology under which competent authorities permit institutions to use Advanced Measurement Approaches (AMA) for operational risk in accordance with Article 312 of Regulation (EU) No 575/2013 (EBA/RTS/2015/02)