
Copulas

In statistics, a copula can be thought of as defining the correlation structure for a family of multivariate distributions. Each distribution in the family is constructed by "gluing" together two or more univariate distributions, with the copula supplying the "glue". Copulas can be applied to a wide range of simulation problems - for example, in financial modelling.

At Mark 9 of the [NAG C Library](#), we have available the copulas:

- [Student's t](#) (g05rcc);
- [Gaussian](#) (g05rdc);
- [Clayton/Cook-Johnson](#) (g05rec);
- [Frank](#) (g05rfc) bivariate;
- [Plackett](#) (g05rgc);
- [Clayton](#) (g05rhc);
- [Frank](#) (g05rjc) multivariate;
- [Gumbel-Hougaard](#) (g05rkc).

Bivariate copulas use the conditional sampling approach as different from the less computationally efficient mixture of powers method. The Student's t and Gaussian copulas in the NAG C Library, Mark 9, use the new pseudorandom base generators, and supersede their Mark 8 counterparts.

To illustrate, suppose we wish to simulate the joint distribution of several random variables. The kind of copula we use - Gaussian, Student's t, etc - describes the correlation structure between the variables, while a collection of univariate distributions (usually called the marginal distributions) define the distribution within each of the variables. Compared to more conventional multivariate distributions (such as, for example, the multivariate normal), the use of a copula allows each variable to have a different marginal distribution. Thus, for example, one variable may have a Gaussian marginal distribution, while another has a Beta distribution, a third is uniformly distributed, and so on.

In the below graph, we construct a bivariate distribution using a Gaussian copula with Beta marginal distributions. The input for the calculation of the multivariate distribution includes its covariance matrix, which contains details of the correlations between the different variables. The copula is created using NAG function g05rdc. The shape of the Beta distribution is defined by two parameters (usually called alpha and beta), and can range from the symmetric bell curve of the normal distribution (when alpha and beta are both large) to being highly skewed or, at the other extreme when alpha and beta are both 1, a uniform distribution. Here, we set alpha = 12.0 and beta

= 5.0 for the first variable - which gives a skewed distribution - and alpha = beta = 5.0 for the second variable, which is distributed symmetrically:

