

Sujet de mémoire statistique

Eté 2014

Johan Segers

johan.segers@uclouvain.be

Nathan Uyttendaele

nathan.uyttendaele@uclouvain.be

One of the big challenges among statisticians today is the modelization of the dependencies between more than two random variables. To model these dependencies, the usual approach is to decompose the multivariate distribution of the random variables under consideration into a set of univariate distributions (the margins) and a copula, the later taking into account the dependencies.

An example of copula is the nested Archimedean copula. This copula is a rather flexible copula while still being easy to handle. It is made of a tree structure and of a set of univariate functions. However, based on a sample of observations from the vector of random variables under consideration, how to pick the nested Archimedean copula that fits best the observed data - best tree and best set of univariate functions for the data at hand - is still a hot topic of research.

Segers and Uyttendaele (2014) developed a method allowing to pick the best tree for the observed data. Further researches by Uyttendaele (2014) led to even more new ways to estimate the best tree for the observed data.

One problem remains: all the developed methods rely on a tuning parameter. In case the true tree for the observed data is known, the optimal value of the tuning parameter can be found by simulation. But in the real world the true tree is of course unknown and, therefore, so is the optimal tuning parameter.

The aim of this mémoire is to build an estimator for this optimal tuning parameter based on the observed data.

To apply for this the mémoire or for any question, please contact Nathan Uyttendaele.

The ideal candidate for this mémoire should have

- *good knowledge of English;
- *excellent programming skills, and some experience with the R language;
- *followed with distinction at least one course dealing with multivariate random variables.

The candidate can expect a mémoire

- *with well defined boundaries and goals while still allowing for creativity;
- *that might end up in a publication in a statistical journal;
- *where he will not be left on its own and will always get help when he requests it.

Some reading materials to get started:

<http://cran.r-project.org/web/packages/copula/vignettes/nacopula-pkg.pdf>

<http://arxiv.org/pdf/1304.1384v2.pdf>