The `nCopula` package

David Beauchemin
Simon-Pierre Gadoury

August 3, 2017

Abstract

The *nCopula* package aims to simplify the construction process and usage of hierarchical Archimedean copulas through compound distributions in R. Is is possible to build structures with clear representations, obtain expressions for Archimedean copulas, wether they are hierarchical or not, as well as other important functions (i.e. Laplace Stieltjes Transform, pgf, etc.), given a certain path and structure. Furthermore, the generation of random vectors is possible from any given structure.

Contents

1 Introduction

(...)

2 Archimedean copulas

2.1 Definition

Archimedean copulas are a very interesting family of copulas. Bla bla bla (...).

Let the copula's *generator* be defined as a decreasing function $\psi : [0,\infty) \to [0,1]$, where $\psi(0) = 1$ and $\lim_{t\to\infty} \psi(t) = 0$. In the same manner, $\psi^{-1} : [0,1] \to [0,\infty)$, for which $\psi^{-1}(0) = \inf\{t : \psi(t) = 0\}$, where

ψ^{-1} is the inverse of the generator. The set of all such functions is denoted by ψ_∞. Then, we can defined an Archimedean copulas C as

$$C(u_1,\ldots,u_d) = \psi\left(\psi^{-1}(u_1)+\ldots+\psi^{-1}(u_d)\right). \tag{1}$$

2.2 Sampling

(...)

2.3 Symbolical derivatives

To obtain symbolical expressions of Archimedean copulas, one only needs to known the derivatives of the corresponding distribution's LST and inverse LST. Indeed, an Archimedean copula can be written as follows:

$$C(u_1,\ldots,u_d) = \psi\left(\sum_{i=1}^{d}\psi^{-1}(u_i)\right), \tag{2}$$

where ψ is the generator of the copula. To compute its corresponding density, one should take the derivatives with respect to u_1 up to u_d of (2). Doing so, we obtain that

$$c(u_1,\ldots,u_d) = \prod_{i=1}^{d}\left(\psi^{-1}\right)'(u_i)\,\psi^{(d)}(v), \tag{3}$$

where $v = \sum_{i=1}^{d}\psi^{-1}(u_i)$. This expression is often easily computable, since it only requires the d^{th} derivative of a Laplace Stieltjes Transform (LST) and the first derivative of its inverse.

Remark 1. *Here, $\psi^{(d)}(v)$ implies that we take the d^{th} derivative of the* ***univariate*** *function ψ, and then evaluate it at v.*

In the case where the corresponding distribution is discrete, it is often easier to find the derivatives of the pgf. Thus, the following relation can be useful:

$$\begin{aligned}\mathscr{L}^{(d)}(t) &= \frac{\mathrm{d}^d}{\mathrm{d}t^d}\mathscr{P}\left(e^{-t}\right)\\ &= \sum_{r=1}^{d}\mathscr{P}^{(r)}\left(e^{-t}\right)e^{-td}\sum_{s=1}^{t}\frac{(-1)^{r-s}}{s!(r-s)!}(-s)^d.\end{aligned} \tag{4}$$

In the package, (4) is used for `GEO` and `LOG`.

2.4 Comparison with the `copula` package

```
library(nCopula)

## Loading required package:  copula
## Warning:  package 'copula' was built under R version 3.3.2
```

```r
## Copula package ##
cop <- claytonCopula(5, 10)
res1 <- rCopula(10000, cop)

logv <- function(x, data)
{
  cop2 <- claytonCopula(x, 10)
  -sum(log(dCopula(data, cop2)))
}

suppressWarnings(optimize(logv, c(1e-5, 10), data = res1))

## $minimum
## [1] 5.005595
##
## $objective
## [1] -107012.7

## nCopula package ##

cop <- Clayton(5, 10, density = TRUE)
res1 <- rCop(10000, cop)

dens <- cop@dens

logv <- function(x, data)
{
  alpha <- x
  for (i in 1:10)
    eval(parse(text = paste("u", i, " <- data[,", i, "]", sep = "")))
  -sum(log(eval(parse(text = dens))))
}

suppressWarnings(optimize(logv, c(1e-5, 10), data = res1))

## $minimum
## [1] 0.199368
##
## $objective
## [1] -107308.9
```

A Discrete distributions

A.1 Geometric

Let $M \sim \text{Geo}(p)$, such that

$$\mathscr{P}_M(t) = \frac{pt}{1-(1-p)t}.$$

Then,

$$\mathscr{P}_M^{(k)}(t) = \frac{k!}{t^{k-1}}\frac{1}{p}\left(\frac{\mathscr{P}_M(t)}{t}\right)^2\left(\frac{\mathscr{P}_M(t)}{tp}-1\right)^{k-1}.$$

Furthermore,

$$\left(\mathscr{P}_M^{-1}\right)'(t) = \frac{p}{(p+t(1-p))^2}.$$

A.2 Logarithmic

Let $M \sim \text{Log}(p)$, such that

$$\mathscr{P}_M(t) = \frac{\ln(1-pt)}{\ln(1-p)}.$$

Then,

$$\mathscr{P}_M^{(k)}(t) = -\frac{(k-1)!}{\ln(1-p)}\left(\frac{p}{1-pt}\right)^k.$$

Furthermore,

$$\left(\mathscr{P}_M^{-1}\right)'(t) = -\frac{\ln(1-p)}{p}(1-p)^t.$$

B Continuous distributions

B.1 Gamma

Let $B \sim \text{Gamma}\left(\frac{1}{\alpha}, 1\right)$, such that

$$\mathscr{L}_B(t) = (1+t)^{-\frac{1}{\alpha}}. \tag{5}$$

Then,

$$\mathscr{L}_B^{(k)}(t) = (-1)^k \prod_{j=0}^{k-1}\left(j+\frac{1}{\alpha}\right)(1+t)^{-\frac{1}{\alpha}-k}.$$

Furthermore,

$$\left(\mathscr{L}_B^{-1}\right)'(t) = -\alpha t^{-\alpha-1}.$$

B.2 Stable

Let $B \sim \text{Stable}(\ldots,\ldots)$, such that

$$\mathscr{L}_B(t) = e^{-t^\alpha}. \tag{6}$$

Then,

$$\begin{aligned}
\mathscr{L}_B^{(k)}(t) &= \sum_{r=1}^{k}\sum_{s=1}^{r} \mathscr{L}_B(t)(-1)^s \frac{t^{(r-s)\alpha}}{s!(r-s)!}\frac{\Gamma(\alpha s+1)}{\Gamma(\alpha s+1-k)}\, t^{\alpha s-k}\,\mathbb{1}_{\{k-\alpha s\notin\mathbb{N}^+\}} \\
&= \sum_{s=1}^{k}(-1)^s\left(\sum_{r=0}^{k-s}\frac{(t^\alpha)^r}{r!}\mathscr{L}_B(t)\right)\frac{\Gamma(\alpha s+1)}{s!\Gamma(\alpha s+1-k)}\, t^{\alpha s-k}\,\mathbb{1}_{\{k-\alpha s\notin\mathbb{N}^+\}} \\
&= \sum_{s=1}^{k}(-1)^s\frac{\Gamma(\alpha s+1)}{s!\Gamma(\alpha s+1-k)}\, t^{\alpha s-k}\,\mathbb{P}(N\le k-s)\,\mathbb{1}_{\{k-\alpha s\notin\mathbb{N}^+\}},\ N\sim\text{Poisson}\left(t^\alpha\right) \\
&= \mathbb{E}\left[\sum_{s=1}^{k-N}(-1)^s\frac{\Gamma(\alpha s+1)}{s!\Gamma(\alpha s+1-k)}\, t^{\alpha s-k}\times\mathbb{1}_{\{N\le k-1\,\cap\, k-\alpha s\notin\mathbb{N}^+\}}\right] \\
&= \mathbb{E}\left[g(N)\right],
\end{aligned}$$

where $g(t) = \sum_{s=1}^{k-t}(-1)^s\frac{\Gamma(\alpha s+1)}{s!\Gamma(\alpha s+1-k)}\, t^{\alpha s-k}\times\mathbb{1}_{\{t\le k-1\,\cap\, k-\alpha s\notin\mathbb{N}^+\}}$. Furthermore,

$$\left(\mathscr{L}_B^{-1}\right)'(t) = -\frac{1}{\alpha}\frac{(-\ln(t))^{\frac{1}{\alpha}-1}}{t}.$$

References

[Bedford and Cooke, 2002] Bedford, T. and Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*, pages 1031–1068.

[Feller, 1971] Feller, W. (1971). An introduction to probability and its applications, vol. ii. *Wiley, New York*.

[Górecki et al., 2016] Górecki, J., Hofert, M., and Holeňa, M. (2016). An approach to structure determination and estimation of hierarchical archimedean copulas and its application to bayesian classification. *Journal of Intelligent Information Systems*, 46(1):21–59.

[Hofert, 2010] Hofert, M. (2010). *Sampling nested Archimedean copulas with applications to CDO pricing*. PhD thesis, Universität Ulm.

[Hofert, 2011] Hofert, M. (2011). Efficiently sampling nested archimedean copulas. *Computational Statistics & Data Analysis*, 55(1):57–70.

[Hofert, 2012] Hofert, M. (2012). A stochastic representation and sampling algorithm for nested archimedean copulas. *Journal of Statistical Computation and Simulation*, 82(9):1239–1255.

[Hofert et al.,] Hofert, M., Mächler, M., et al. Nested archimedean copulas meet r: The nacopula package.

[Joe, 1997] Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.

[Kimberling, 1974] Kimberling, C. H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes mathematicae*, 10(2):152–164.

[McNeil, 2008] McNeil, A. J. (2008). Sampling nested archimedean copulas. *Journal of Statistical Computation and Simulation*, 78(6):567–581.