Advanced Non-Life Insurance: Home Assignment about Copulas

1 Instructions

Deadline for assignment: October, 22 2018.
Send your assignment to tim.verdonck@kuleuven.be.

Individual assignment or in group of two students.
Hand in a R-script with clear R-code (that can be run without errors). Answers to the questions may be in the same R-script (in comment) or may be given in a separate file in latex, word, ...).
The assignment counts for 4 bonuspoints (i.e. if you obtain $x/30$ on exam and $y/4$ on assignment, then final score is $\max(x, (x+y)\frac{30}{34})$) on the exam of Advanced Non-Life Insurance.

Install and load the R packages `copula`, `fCopulae`, `Ecdat`, `fGarch` and `MASS`.

2 Assignment

We will fit copulas to a bivariate data set of returns on IBM and CRSP index. First, we will fit a model with univariate t-distributions and a t-copula. Run and study the following R code (ignore the warnings).

```
data(CRSPday,package="Ecdat")
ibm = CRSPday[,5]
crsp = CRSPday[,7]

est.ibm = as.numeric(fitdistr(ibm,"t")$estimate)
est.crsp = as.numeric(fitdistr(crsp,"t")$estimate)
est.ibm[2] = est.ibm[2]*sqrt(est.ibm[3]/(est.ibm[3]-2))
est.crsp[2] = est.crsp[2]*sqrt(est.crsp[3]/(est.crsp[3]-2))
```

1. We need an estimate `omega` of the correlation coefficient in the t-copula. Obtain this value using Kendall's tau.

2. Define `cop_t_dim2` as the t-copula using the obtained value `omega` as correlation parameter and 4 degrees-of-freedom.

Run the following R code:

```
n = length(ibm)
data1 = cbind(pstd(ibm,mean=est.ibm[1],sd=est.ibm[2],nu=est.ibm[3]),
pstd(crsp,mean=est.crsp[1],sd=est.crsp[2],nu=est.crsp[3]))
data2 = cbind(rank(ibm)/(n+1), rank(crsp)/(n+1))
ft1 = fitCopula(cop_t_dim2, data1, optim.method="L-BFGS-B", method="ml",
start=c(omega,5),lower=c(0,2.5),upper=c(.5,15) )
ft2 = fitCopula(cop_t_dim2, data2, optim.method="L-BFGS-B", method="ml",
start=c(omega,5),lower=c(0,2.5),upper=c(.5,15) )
```

3. Study the code and explain what the code does. Explain also the difference between methods used to obtain the two estimates `ft1` and `ft2`.

4. Do the two estimates seem significantly different (in a practical sense)?

Run the following code (this takes some time).

```
mvdc_t_t = mvdc( cop_t_dim2, c("std","std"),
list(list(mean=est.ibm[1],sd=est.ibm[2],nu=est.ibm[3]),
list(mean=est.crsp[1],sd=est.crsp[2],nu=est.crsp[3]) ) )

start=c(est.ibm,est.crsp,ft1@estimate)
objFn = function(param)
{
-loglikMvdc(param, cbind(ibm,crsp), mvdc_t_t)
}
t1 = proc.time()
fit_cop = optim(start,objFn,method="L-BFGS-B",
lower = c(-0.1,0.001,2.5,-0.1,0.001,2.5,0.2,2.5),
upper = c(0.1,0.03,15,0.1,0.03,15,0.8,15)
)
t2 = proc.time()
total_time = t2-t1
total_time[3]/60
```

5. Explain what the code does. What are the estimates of the copula parameters in `fit_cop`?

6. What are the estimates of the parameters in the univariate marginal distributions?

7. Was the estimation method maximum likelihood, parametric pseudo-maximum likelihood or semi-parametric pseudo-maximum likelihood?

8. Estimate the coefficient of lower tail dependence for this copula.

9. Fit normal, Gumbel, Frank and Clayton copulas to the data and compare the estimated copulas (CDF's) with the empirical copula.

10. Do you see any difference between the parametric estimates of the copula? If so, which seem closest to the empirical copula?

11. Find AIC for the t, normal, Gumbel, Frank and Clayton copulas. Which copula model fits best by AIC?