Copula and tail dependence in risk deposit management

Norwegian Computing Center
APPLIED RESEARCH AND DEVELOPMENT

NOTE

SAMBA/29/03

Bård Storvik

December 2003

Norwegian Computing Center
APPLIED RESEARCH AND DEVELOPMENT

NR Note

Title: Copula and tail dependence in risk deposit management

Date: December 2003
Year: 2004
Note no.: SAMBA/29/03

Author: Bård Storvik <baard@nr.no>

Abstract: Total risk is a sum of different risk quantities. This report studies how the tail of the total risk variable behaves when the marginal distributions of the different risk quantities differ as well as the dependence between the risk quantities. The dependence is quantified through a so called *Copula*. Moreover, the tail of the total risk is compared with the sum of tails from each risk quantity.

Keywords: Total risk, marginal distribution, dependence, tail dependence, copula

Target group: All employees

Availability: Open

Project: GB-OPT

Project no.: 220191

Research field: Finance

No. of pages: 12

Norwegian Computing Center
Gaustadalléen 23, P.O. Box 114 Blindern, NO-0314 Oslo, Norway
Telephone: +47 2285 2500, telefax: +47 2269 7660, http://www.nr.no

Contents

1 Introduction **1**

2 Copulas **1**

3 Tail dependence and local dependence **3**

4 Credit Risk and Operational Risk **4**

4.1 The dependence structure in the DnB model 4

4.2 Tail dependence 5

5 Ratios of Value at risk **6**

6 SPLUS code **9**

References **12**

1 Introduction

In a former project done for DnB (an Norwegian bank), NR has built a model for the total risk of the financial institution. The total risk is defined as the sum of different risk variables. The different risk variables are uncertain quantities. In the project done for DnB they are modeled by different probability distributions. Because one risk variable may depend on other risk variables, it is important to specify a good dependence structure. This may mean that a high risk for one variable gives rise to a high risk for another risk variable. How the dependence structure is specified will determine how the distribution of the total risk variable will look like. Banks are especially interested in the tail of the total risk distribution (at a given quantile). Based on the quantile, they determine the amount of money they should put aside to meet potential problems. Previously, banks have based their analysis on adding the quantiles of the different risk distributions. This is not necessarily the same as taking the tail of a distribution base on the total risk (sum of the different risk variables). Taking the former way of producing a tail risk (as sum of tail risks) means that all tail risks are added up. However, this means that if we assume that the different risk are perfectly correlated, it could potentially mean that the risk is overestimated and to much money is put aside. This is a serious problem for the bank because the money can not be invested and more money could be made if a smaller amount is put aside.

The two most important risk variables are credit risk and operational risk. In the current version of the DnB's total risk system, the dependence structure of credit risk and operational risk are modeled in two steps. First, the marginal distributions of the credit and operational losses are assumed to be beta and lognormal distributed. Second, the dependence structure between the two variables are modeled on a normal scale and then transformed to the beta and lognormal scale. The correlation between the beta and the lognormal variables is fixed by the bank (qualified guess) and from this number it is possible to set the correlation on the normal scale (at least by simulation). This project will try to find out more about the features of the model. The model could be changed in two ways: First, the marginal distributions could change to other distribution than beta and lognormal. Second, the dependence structure could change from correlation on normal scale to other dependence structures. The other dependence structures could be modeled by using the copula theory. This copula theory is currently a very *popular* subject in the Finance literature. In Section 2 we will try to link copula theory to the method described above, where dependence is measured on normal scale by correlation. Section 3 studies tail dependence and local dependence. Furthermore, Section 4 and 5 define the model and Value at risk respectively. Finally, Section 6 gives the Splus code that produced the results in this report.

2 Copulas

A joint distribution can be estimated parametrically, semiparmetrically or nonparametrically. Deciding on a parametric multivariate distribution will limit ourself to quite strict dependence structure as well as distributional shape. In most textbooks on multivariate distributions, all the marginal distributions has the the same parametric representation. There are however, many situations where the marginal distributions are different in parametric representation. A standard trick is to transform the marginal variables to Gaussian ones, then measure the dependence on the Gaussian scale and finally transform back to get the multivariate distribution. This trick assumes however linear dependence on Gaussian scale (correlation). See Embrechts et al. (2002) for correlation and dependence in Risk Management. It is however possible that the dependence on the original scale is different than what is described through the dependence on the Gaussian scale. It is here the flexibility of copulas comes to its right, see Embrechts et al. (2003). Once the copula dependence function is set, as well as the marginals, the joint distribution can be found.

Copulas is a quite general tool for constructing models with dependence other than the standard ones. Copulas is a cumulative distribution function where the random variables have specific marginals. More specific, let $X_1, \dots, X_n$ be random variables with continuous distribution functions $F_1, \dots, F_n$, respectively, and the joint distribution function is F. The copula is defined as

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)). \tag{1}$$

The relationship between the joint distribution and the copula can then be written as

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \tag{2}$$

From (1) and (2) it is seen that the joint distribution is defined if both the copula and the marginals are given. To be able to understand copulas better, it might help to look at (1) more closely. Before we do that we need the following result: If X have distribution function F_X, then $F_X(X)$ is uniformly distributed on $[0,1]$. This is seen from

$$\begin{aligned} Pr\{F_X(X) \le y\} &= Pr\{X \le F_X^{-1}(y)\} \\ &= F_X(F_X^{-1}(y)) \\ &= y. \end{aligned}$$

Since the distribution function of a uniform variable is identical to the distribution function of the variable $F_X(X)$, the result is established. Returning to the copula definition

$$\begin{aligned} C(u_1, \dots, u_n) &= F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \\ &= Pr\{X_1 \le F_1^{-1}(u_1), \dots, X_n \le F_n^{-1}(u_n)\} \\ &= Pr\{F_1(X_1) \le u_1, \dots, F_n(X_n) \le u_n\} \\ &= Pr\{U_1 \le u_1, \dots, U_n \le u_n\}, \end{aligned} \tag{3}$$

where $U_i = F_i(X_i)$, $i = 1, \dots, n$ is uniformly distributed in the interval $[0,1]$, but not necessarily independent. There is a dependence between the U_i's when the X_i's are dependent. The point made here is that the copula does not have the marginal aspects built in, only the dependence structure.

3 Tail dependence and local dependence

The so called *tail dependence* characterizes an important property of the extreme dependence between two variables X and Y with marginal distribution functions F_X and F_Y, respectively. The tail dependence allows us to quantify the dependence in the tail. It is defined by

$$\lambda \quad = \quad \lim_{u \to 1} Pr\{X > F_X^{-1}(u) | Y > F_Y^{-1}(u)\}, \tag{4}$$

provided that the limit $\lambda \in [0, 1]$ exists. If $\lambda \in (0, 1]$, X and Y are said to be asymptotically dependent in the upper tail. If $\lambda = 0$, X and Y are said to be asymptotically independent in the upper tail. That is, λ quantifies the probability to observe a large X, assuming Y is large itself. It is also possible to link the copula to the tail dependence defined in (4) by the following alternative definition of the tail dependence

$$\lambda \quad = \quad \lim_{u \to 1} \frac{1 - 2u + C(u,u)}{1 - u} \;=\; \lim_{u \to 1} 2 - \frac{\log(C(u,u))}{\log(u)}, \tag{5}$$

where C is the bivariate copula of X and Y. Notice that independence between X and Y will give $\lambda = 0$. It is often the case that we want to look at tail dependence in general, that is the dependence far out in the tail but not completely in the limit. Therefore, we define

$$\lambda_u \quad = \quad Pr\{X > F_X^{-1}(u) | Y > F_Y^{-1}(u)\} \tag{6}$$

where $\lambda = \lim_{u \to 1} \lambda_u$. λ_u will tell us the tail dependence for a certain quantile u of the marginal distributions. For the Gaussian copula there is independence in the limit. Nevertheless, we are not interested in what happens in the limit or when the quantile goes to one. Our real interest is what happens for a fixed quantile close to one, e.g. $u = 0.9997$. Let X, Y have the density $f(x, y)$. The local dependence function is defined by

$$\gamma_f(x,y) \quad = \quad \frac{\partial^2 \log(f(x,y))}{\partial x \partial y}$$

The local dependence function has the following properties:

- X and Y are independent iff $\gamma_f(x, y) = 0$
- $\gamma_f(x, y)$ is margin free in the sense that $\gamma_f(x, y) = \gamma_h(x, y)$ if $h(x, y) = f(x, y)\phi_1(x)\phi_2(y)$
- If $f_{1|2}$ and $f_{2|1}$ are the conditional density functions, then $\gamma_f = \gamma_{f_{1|2}} = \gamma_{f_{2|1}}$.

Furthermore, for any integrable local dependence function defined over a compact area and any given continuous marginal density functions, there exists a unique bivariate density function.

For the standard bivariate normal density function, the local dependence function is $\gamma_f(x, y) = \frac{\rho}{1-\rho^2}$, where ρ is the correlation in the bivariate distribution. This means that the local dependence function is constant. This is not the case for the T-distribution. Part of this project studies the dependence structure in the tail for for different copulas. The local dependence function might help us to understand dependence in the tail better, because of the ability to study dependence locally. The tail dependence can be thought of as a more global (integrated) measure of dependence.

4 Credit Risk and Operational Risk

As we mentioned in the introduction, the model for Credit risk and Operational Risk built in the DnB tool for total risk management in the following way. First, the credit risk and operational risk is modeled marginally, that is a separate model for the two variables. Second, the dependence structure is modeled. The dependence structure is not modeled through a copula, but we will see in this section this is an option. Let us first look at the marginal model of the credit risk and operational risk:

$$
\begin{aligned}
C &= eB^{-1}(\Phi(X)) && (7)\\
O &= \exp\{\xi + \tau Y\} && (8)
\end{aligned}
$$

where X and Y are standard normally distributed.

4.1 The dependence structure in the DnB model

In order to get a dependence structure between the variables C and O, the variables X and Y has to be dependent. Let

$$
\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right). \qquad (9)
$$

We can write

$$
\begin{aligned}
X &= \Phi^{-1}(B(C/e)) && (10)\\
Y &= \frac{\log(O) - \xi}{\tau} && (11)
\end{aligned}
$$

The marginal distribution functions are

$$
\begin{aligned}
F_C(c) &= B(\frac{c}{e}) && (12)\\
F_O(o) &= \Phi(\frac{\log(o) - \xi}{\tau}) && (13)
\end{aligned}
$$

The inverse of the marginal distribution functions are

$$
\begin{aligned}
F_C^{-1}(u) &= eB^{-1}(u) && (14)\\
F_O^{-1}(v) &= \exp(\tau\Phi^{-1}(v) + \xi). && (15)
\end{aligned}
$$

The joint distribution function is

$$F(c,o) \quad = \quad \Phi_\rho(\Phi^{-1}(B(c/e)), \frac{\log(o)-\xi}{\tau}), \tag{16}$$

where Φ_ρ is the cumulative distribution function of X, Y. The copula of C, O is then

$$\begin{aligned} C(u,v) &= F(F_C^{-1}(u), F_O^{-1}(v)) \\ &= \Phi_\rho(\Phi^{-1}(B(F_C^{-1}(u)/e)), \frac{\log(F_O^{-1}(v))-\xi}{\tau}) \\ &= \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)). \end{aligned} \tag{17}$$

The copula defined in (17) is exactly the normal copula, meaning that the joint credit and operational loss model has a dependence structure which is the same as normal copula. One of the objections (in the literature) against using the normal copula is t hat extreme events happen independent and the 99.97% quantile is an extreme event. Therefore, it is interesting to see what happens with other dependence structures or other copulas. The copulas we will study is the t (with different degree of freedom, so called heavy tailed), Gumbel, Joe, Galambos and Kimeldorf/Samson. Let $T_{f,\rho}$ be the bivariate cumulative distribution function of the t-distribution with f degrees of freedom and correlation ρ. Furthermore, let T_f be the marginal cumulative distribution functions. The copula of the t-distribution is:

$$C(u,v) \quad = \quad T_{f,\rho}(T_f^{-1}(u), T_f^{-1}(v)). \tag{18}$$

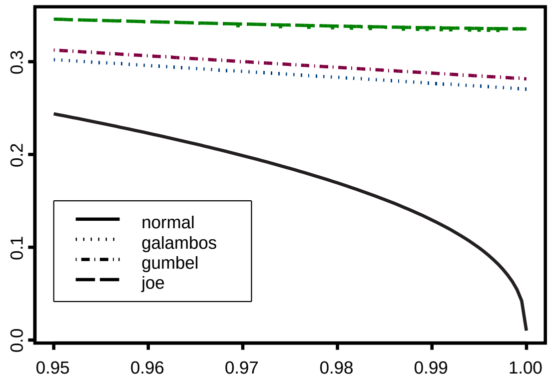
4.2 Tail dependence

We will now study the tail dependence of some copulas. The tail dependence for a certain quantile u is

$$\lambda_u \quad = \quad \frac{1-2u+C(u,u)}{1-u}. \tag{19}$$

The normal, galambos, gumbel and joe copula have zero dependence in the limit. However, Figure 1 shows that far out in the tail there is quite a distinct difference in the quantity of dependence. The galambos, gumbel and joe tail dependence function is higher than the normal and tends to go very slowly towards zero. For $q = 0.997$, the tail dependence is 0.08, 0.27, 0.28 and 0.34, when the dependence copula is normal, galambos, gumbel and joe respectively.

Figure 1: Displaying $\lambda_u(q)$ for high values of q and for four different copulas.



5 Ratios of Value at risk

Let the total risk be defined as a sum of the two risk variables operational and credit risk, that is $T = O + C$. Let

$$
\begin{aligned}
z_T(q) &= F_T^{-1}(q) && (20)\\
z_C(q) &= F_C^{-1}(q) && (21)\\
z_O(q) &= F_O^{-1}(q) && (22)
\end{aligned}
$$

be the quantiles (or Value-at-risk (VaR) in the financial literature) of the different distributions T, C and O, respectively. DnB is interested in the quantity $r(q)$ which is given by

$$r(q) = \frac{z_T(q)}{z_C(q) + z_O(q)}. \tag{23}$$

Now, in most cases (dependent on the dependence between C and O and the marginal distributions) we have

$$z_T(q) \leq z_C(q) + z_O(q), \tag{24}$$

but we can construct situations where the opposite yields. In this section we study different values of $r(q)$ for different quantiles q and different copulas for the dependence structure between credit and operational losses. In all cases we have specified the correlation between C and O to be approximated by 0.43.

In Table 1 we show the results. The t-distribution is not in the EVANESCE implementation, so an independent S-PLUS code was made for this case. All other dependence

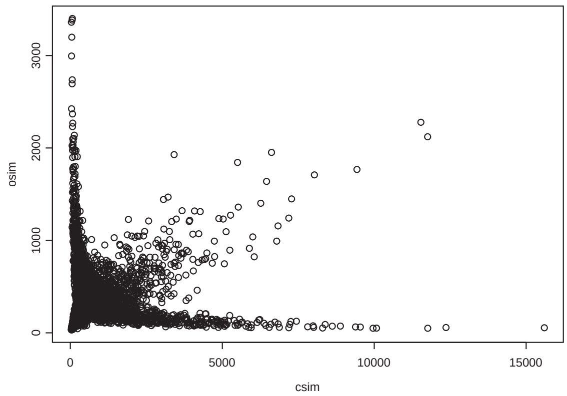
Table 1: Displaying $r(q)$ for $q = 0.95, 0.99, 0.997$ for $500,000$ simulations. The dependence is modeled by different copulas. C is beta distributed and O is lognormally distributed.

method	$\widehat{r}(0.95)$	$\widehat{r}(0.99)$	$\widehat{r}(0.9997)$	$\widehat{\rho}_{C,O}$	Parameter	$\widehat{\lambda}_{0.9997}$ $(\widehat{sd})$
Normal	0.9358	0.9046	0.8422	0.426	ρ=0.500	0.033 (0.000)
t_2	0.9207	0.9159	0.9143	0.428	ρ=0.401	0.324 (0.014)
t_4	0.9248	0.9017	0.8812	0.430	ρ=0.378	0.221 (0.014)
t_6	0.9380	0.8928	0.8778	0.422	ρ=0.387	0.147 (0.011)
Gumbel	0.9230	0.9162	0.9177	0.425	δ=1.280	0.276 (0.016)
Kimeldorf.Sampson	0.9379	0.8796	0.7887	0.424	δ=2.500	0.001 (0.001)
Galambos	0.9216	0.9163	0.9145	0.424	δ=0.530	0.250 (0.011)
Joe	0.9233	0.9210	0.9367	0.429	δ=1.360	0.343 (0.014)

structures was implemented in the EVANESCE system. Most of these copulas have two parameters. We focused on copulas with one parameter since copulas with two parameters involves finding two parameter values that gives correlation between C and O to be approximately $\rho_{C,O} = 0.43$ which is time consuming. It is also possible that several values of the two parameters gives $\rho_{C,O} = 0.43$.

Using correlation as a measure of dependence does not necessarily tell us how strong the relationship between two variables are. The correlation measures the linear dependence, but other types of dependencies may not be captured. In the case of beta and lognormal risk variables, correlation seems to be a fair measure of dependence. We will now show a case where the dependence is far from linear. Let C and O be inverse normally distributed with the same means and variances as in the beta and lognormal case. If the correlation on normal scale is 0.5, the correlation on inverse normal scale is about zero. Now, if the correlation on normal scale is 0.9, the correlation on inverse normal scale is still about zero. Even though the correlation is about zero, it does not mean that there is no dependence. The dependence structure is seen in Figure 2 where it is seen that it is far from linear. This tells us that correlation is a bad measure for dependence in this particular case.

Figure 2: Displaying 5000 simulated credit and operational risk variables, which is inverse normal distributed variables. The correlation is 0.9 on the normal scale before transformation to inverse normal variables where the correlation is approximately zero.



In Table 2, it is difficult to set the correlation between credit and operational (inverse normal scale) to be 0.43 as in Table 1 because of other dependence structure than linear. Therefore we set the operational correlation to be 0.5 for normal copula. For the other copulas the same parameter value as in Table 1 were set.

We learn from Table 1 that the dependence structure (different copulas) does not matter very much when it comes to $r(q)$ for $q = 0.95, 0.99$. For the quantile $q = 0.9997$, there is a distinct difference. The tail dependence at 0.9997 differs depending on the copula. The Kimeldorf Sampson copula has the smallest tail dependence (0.001) while the Joe copula has highest tail dependence (0.343). The connection between tail dependence and $r(0.9997)$ is clear, high tail dependence gives high $r(0.9997)$. The opposite also yields, low tail dependence gives low $r(0.9997)$. From Table 2, $r(q)$ does not vary very much for $q = 0.95, 0.99$ as well as $q = 0.9997$ for different copulas. The value of $r(0.9997)$ is however smaller in most cases compared to Table 1. If we compare Table 1 and Table 2, it is not only how we model the dependence that matters, but also how we model the marginal distributions. Therefore, both dependence and marginal distributions should be addressed when modeling total risk.

Acknowledgments. Thanks to Kjersti Aas for helpful comments.

Table 2: Displaying $r(q)$ for $q = 0.95, 0.99, 0.997$ for $500,000$ simulations. The dependence is modeled by different copulas. C and O are inverse normally distributed.

method	$\widehat{r}(0.95)$	$\widehat{r}(0.99)$	$\widehat{r}(0.9997)$	$\widehat{\rho}_{C,O}$	Parameter
Normal	0.857	0.838	0.843	0.001	ρ=0.500
Gumbel	0.857	0.839	0.835	0.001	δ=1.280
Kimeldorf.Sampson	0.852	0.836	0.841	0.001	δ=2.500
Joe	0.856	0.839	0.843	0.000	δ=1.360
t_6	0.864	0.821	0.822	0.004	ρ=0.500

6 SPLUS code

The EVANESCE implementation in SPlus FinMetrics Module works on a PC platform.
Remember to open the EVANESCE system by the SPLUS code:
> module(finmetrics)
Some of the functions used below is found in the files:
/nr/project/stat/GB-optimering/Copula/creditOper.ssc
and */nr/project/stat/GB-optimering/baard/functions.ssc*

```
delta <- 0.53
#found by trying many different values
b1c = galambos.copula(delta)
galambos <- rep(0,10)
u = rcopula(b1c,500000)
vv  <- simCreditandOperationalLossunif(u$x,u$y)
cat("delta=",delta,"quantilratio=",quantileratio(vv$osim,vv$csim),
"cor=",cor(vv$osim,vv$csim),"\n")

delta <- 1.28
#found by trying many different values
b1c = gumbel.copula(delta)
u = rcopula(b1c,500000)
vv <- simCreditandOperationalLossunif(u$x,u$y)
cat("delta=",delta,"quantilratio=",quantileratio(vv$osim,vv$csim),
"cor=",cor(vv$osim,vv$csim),"\n")

delta <- 1.36
#found by trying many different values
b1c = joe.copula(delta)
u = rcopula(b1c,500000)
vv <- simCreditandOperationalLossunif(u$x,u$y)
cat("delta=",delta,"quantilratio=",quantileratio(vv$osim,vv$csim),
"cor=",cor(vv$osim,vv$csim),"\n")

delta <- 2.5
#found by trying many different values
```

```
b1c = kimeldorf.sampson.copula(delta)
u = rcopula(b1c,500000)
vv <- simCreditandOperationalLossunif(u$x,u$y)
cat("delta=",delta,"quantilratio=",quantileratio(vv$osim,vv$csim),
"cor=",cor(vv$osim,vv$csim),"\n")

rho <- 0.5
b1c = normal.copula(rho)
u = rcopula(b1c,500000)
vv <- simCreditandOperationalLossunif(u$x,u$y)
cat("rho=",rho,"quantilratio=",quantileratio(vv$osim,vv$csim),
"cor=",cor(vv$osim,vv$csim),"\n")

#source("/nr/project/stat/GB-optimering/baard/Ssource/kjor.s")
#source("/nr/project/stat/GB-optimering/baard/Ssource/functions.s")
source("/nr/project/stat/GB-optimering/Copula/creditOper.ssc")
# Here we optimize such that C and O have correlation close to 0.43
# for the t-distribution copula.
df1_2
vv <- optimize(minfunc, lower = 0.35, upper = 0.5,max=F)
zz_simCreditandOperationalLossMethod(operasjonellKorr=vv$minimum,
numSim=500000,method="Tdist",df=df1)
cat("kor=",vv$minimum,"df=",df1,"quantilratio=",quantileratio(zz$osim,zz$csim),
"cor=",cor(zz$osim,zz$csim),"\n")
df1_4
#vv <- nlminb(0.39, minfunc, lower = 0.35, upper = 0.5)
vv <- optimize(minfunc, lower = 0.35, upper = 0.5,max=F)
zz_simCreditandOperationalLossMethod(operasjonellKorr=vv$minimum,
numSim=500000,method="Tdist",df=df1)
cat("kor=",vv$minimum,"df=",df1,"quantilratio=",quantileratio(zz$osim,zz$csim),
"cor=",cor(zz$osim,zz$csim),"\n")
df1_6
vv <- optimize(minfunc, lower = 0.35, upper = 0.5,max=F)
zz_simCreditandOperationalLossMethod(operasjonellKorr=vv$minimum,
numSim=500000,method="Tdist",df=df1)
cat("kor=",vv$minimum,"df=",df1,"quantilratio=",quantileratio(zz$osim,zz$csim),
"cor=",cor(zz$osim,zz$csim),"\n")



lambda <- 1.28
b1c = gumbel.copula(lambda)
u = rcopula(b1c,500000)
vv <- simCreditandOperationalLossIN(u$x,u$y)
cat("lambda=",lambda,"quantilratio=",quantileratio(vv$osim,vv$csim),
"cor=",cor(vv$osim,vv$csim),"\n")


lambda <- 0.5
norm = normal.copula(lambda)
u = rcopula(norm,500000)
vv <- simCreditandOperationalLossIN(u$x,u$y)
```

```
cat("lambda=",lambda,"quantilratio=",quantileratio(vv$osim,vv$csim),
"cor=",cor(vv$osim,vv$csim),"\n")

zz_simCreditandOperationalLossMethod(operasjonellKorr=0.5, numSim=500000,
method="TdistIN",df=6)
cat("kor=",0.5,"df=",6,"quantilratio=",quantileratio(zz$osim,zz$csim),
"cor=",cor(zz$osim,zz$csim),"\n")

b1c = normal.copula(0.5)
q_(0:100)*0.04999/100+0.95
lambda.u <- function(u,b1c) {
(1-2*u+pcopula(b1c,u,u))/(1-u)
}
delta <- 0.53
b2c = galambos.copula(delta)
delta <- 1.28
b3c = gumbel.copula(delta)
delta <- 1.36
b4c = joe.copula(delta)
matplot(cbind(q,q,q,q),cbind(lambda.u(q,b1c),lambda.u(q,b2c),lambda.u(q,b3c),
lambda.u(q,b4c)),cex=1.5,lwd=6,type="l",xlab="",ylab="")
legend(0.95,0.15,c("normal","galambos","gumbel","joe"),lty=1:4,cex=1.5,lwd=6,type="l")
```

References

P. Embrechts, F. Lindskog, and J. McNeil, A. *Modelling dependence with copulas and applications to risk management.* Elsevier/North-Holland, Amserdam, 2003. 2

P. Embrechts, J. McNeil, A, and Straumann D. *Correlation and dependence in risk management: properties and pitfalls.* Cambridge University Press, Cambridge, 2002. 2