

FITTING BIVARIATE CUMULATIVE RETURNS WITH COPULAS

Werner Hürlimann
Value and Risk Management
Winterthur Life and Pensions
Postfach 300
CH-8401 Winterthur, Switzerland
Tel. +41-52-2615861
E-mail werner.huerlimann@winterthur.ch

Abstract.

We propose a copula based statistical method of fitting joint cumulative returns between a market index and a stock from the index family to daily data. Modifying the method of inference functions for margins (IFM method), we perform two separate maximum likelihood estimations of the univariate marginal distributions, assumed to be normal inverse gamma mixtures with kurtosis parameter equal to 6, followed by a minimization of the bivariate chi-square statistic associated to an adequate bivariate version of the usual Pearson goodness-of-fit test. Our copula fitting results for daily cumulative returns between the Swiss Market Index and a stock in the index family for an approximate one-year period are quite satisfactory. The best overall fits are obtained for the new linear Spearman copula, as well as for the Frank and Gumbel-Hougaard copulas. Finally, a significant application to covariance estimation for the linear Spearman copula is discussed.

Keywords : copula, normal inverse gamma mixture, IFM method, bivariate chi-square statistic, daily cumulative return, covariance estimation

AJUSTEMENTS COPULES DE RENDEMENTS CUMULATIFS

Résumé.

Nous proposons une méthode statistique de copule pour ajuster aux données quotidiennes des rendements cumulatifs bivariés associés à un indice de marché et une action particulière de cet indice. Par modification de la méthode des fonctions d'inférence pour marges (méthode IFM), nous estimons d'abord séparément par la méthode du maximum de vraisemblance les marges univariées, supposées de loi mixte normale inverse gamma de paramètre d'élongation égal à 6, et minimisons ensuite la statistique bivariée chi-carrée associée à une version bivariée appropriée du test habituel d'ajustement de Pearson. Nos résultats d'ajustement copule pour les rendements cumulatifs quotidiens de l'indice Suisse du marché et une action de cet indice pour une période approximative d'une année sont assez satisfaisants. Les meilleurs ajustements sont obtenus pour la nouvelle copule linéaire Spearman, ainsi que les copules de Frank et de Gumbel-Hougaard. Finalement, nous discutons une application significative à l'estimation de la covariance pour la copule linéaire Spearman.

Mots clés : copule, loi mixte normale inverse gamma, méthode IFM, statistique bivariée chi-carrée, rendement cumulatif quotidien, estimateur de covariance

1. Introduction.

The present paper is a sequel and synthesis of work done in Hürlimann(2000a/b). We link empirical results on fitting univariate daily cumulative returns with the representation of bivariate distributions by copulas to study the statistical fitting of joint cumulative returns between a market index and a stock from the index family to daily data.

In Section 2, the following empirical fact on fitting univariate distributions to daily cumulative returns is recalled. The normal inverse gamma mixture distribution with kurtosis parameter equal to 6, obtained by mixing the inverse variance of a normal distribution with a gamma prior, usually beats the lognormal and the logLaplace under a chi-square goodness-of-fit test with regrouped data. Therefore, our marginal distributions in bivariate fitting of cumulative returns are restricted to this analytically tractable two-parameter family of symmetric location-scale distributions.

Section 3 contains a short review of the representation of bivariate distributions by copulas. For our purposes, a number of attractive one-parameter families of copulas are retained. They include the copulas by Cuadras and Augé(1981), Gumbel(1960) and Hougaard(1986), Galambos(1975), Frank(1979), Hüsler and Reiss(1989), and Clayton(1978). The parameter of these copulas measures the degree of dependence between the margins. In case the widest possible range of dependence should be covered, one is especially interested in one-parameter families of copulas, which are able to model continuously a whole range of dependence between the lower Fréchet bound copula, the independent copula, and the upper Fréchet bound copula. Such families are called inclusive or comprehensive, and include the copulas by Frank and Clayton. Another simple copula with this property, first considered in Hürlimann(2000b), is the linear Spearman copula described in details in Section 4.

The linear Spearman copula represents a mixture of perfect dependence and independence. If X and Y are uniform(0,1), $Y = X$ with probability $q \geq 0$ and Y is independent of X with probability $1 - q$, then (X, Y) has the linear Spearman copula with the positive dependence structure. In the statistical literature it has been considered by Konijn(1959) and motivated in Cohen(1960). The chosen nomenclature for this copula suggests that it has piecewise *linear* sections and that the parameter of dependence q coincides with *Spearman's* grade correlation coefficient. Besides the useful fact that this copula is suitable for analytical calculation, it has many important properties. It satisfies two extremal properties, one of which is related to a conjecture by Hutchinson and Lai(1990). Of great significance for financial modelling is its simple tail dependence structure. The coefficient of (upper) tail dependence coincides with the dependence parameter q , which implies asymptotic positive dependence in case $q > 0$. This is a desirable property in insurance and financial modelling because data often tend to be dependent in their extreme values. In contrast to this, the ubiquitous Gaussian copula yields always asymptotic independence, unless perfect correlation holds.

The main issue of copula fitting is discussed in Section 5. We apply a method close in spirit to the method of inference function for margins or IFM method studied in McLeish and Small(1988), Xu(1996), and Joe(1997), Section 10.1. This estimation method proceeds by doing two separate maximum likelihood estimations of the univariate marginal distributions, followed by an optimization of the bivariate likelihood as a function of the dependence parameter. In our proposal, we do not maximize the bivariate likelihood. Instead, we determine the dependence parameter, which maximizes the p-value (respectively minimizes the bivariate chi-square statistic) of a bivariate version of the usual Pearson goodness-of-fit test. The reason for considering a modified method lies in the observation that the IFM method reduces the p-value in some cases rather drastically, leading eventually to a rejection

of the model. Our copula fitting results for daily cumulative returns between the Swiss Market Index and a stock in the index family for an approximate one-year period are quite satisfactory. In particular, the analytically tractable linear Spearman copula does very well. The Frank and Gumbel-Hougaard copulas provide competitive best overall fits.

Finally, Section 6 illustrates our estimation results at a significant application. First, based on a general covariance formula for the linear Spearman copula, derived in Theorem 6.1, we show that for the linear Spearman copula model with the chosen normal inverse gamma mixture margins, the Spearman grade correlation coefficient coincides with the Pearson linear correlation coefficient. This allows one to compare the standard product-moment correlation estimator with the estimated dependence parameter from the linear Spearman copula fitting. We observe a considerable discrepancy between the absolute values of both estimators, but on a relative scale both estimators rank the strength of dependence quite similarly.

2. Fitting univariate cumulative returns.

There exist many distributions, which are able to fit the daily cumulative returns on individual stocks and corresponding market indices. As the application of the penalized likelihood scoring method by Schwartz(1978) suggests (so-called Schwartz Bayesian Criterion), it is reasonable to restrict the attention to two-parameter distributions, as shown in Hürlimann(2000a). The only analytically tractable distributions we retain for comparative purposes are the lognormal, the logLaplace, and a normal inverse gamma mixture with fixed kurtosis equal to 6.

Let X represent the daily cumulative return of a market index or a stock in the index family. If $X \sim \ln N(\mathbf{m}, \mathbf{S})$ follows a lognormal distribution, then its distribution is clearly $F_X(x) = \Phi\left(\frac{\ln x - \mathbf{m}}{\mathbf{s}}\right)$, with $\Phi(x)$ the standard normal distribution. The random variable $X \sim \ln L(\mathbf{m}, \mathbf{S})$ follows a logLaplace distribution provided one has

$$F_X(x) = \begin{cases} \frac{1}{2}(e^{-\mathbf{m}}x)^{\frac{\sqrt{2}}{\mathbf{s}}}, & x \leq e^{\mathbf{m}}, \\ 1 - \frac{1}{2}(e^{-\mathbf{m}}x)^{-\frac{\sqrt{2}}{\mathbf{s}}}, & x \geq e^{\mathbf{m}}. \end{cases} \quad (2.1)$$

The use of a logLaplace distribution in Finance has been somewhat motivated in Hürlimann(1995). The normal inverse gamma mixture model is constructed as follows. Let $(X|\mathbf{q})$ be conditional on \mathbf{q} normally distributed with mean \mathbf{m} and variance $1/\mathbf{q}$, and suppose \mathbf{q} follows a conjugate gamma prior $\Gamma(\frac{1}{2}c^2, \mathbf{a})$, $\mathbf{a} > 0$. Then $X \sim NI\Gamma(\mathbf{m}, c, \mathbf{a})$ has the normal inverse gamma mixture density

$$f_X(x) = \frac{1}{B(\frac{1}{2}, \mathbf{a})} \cdot \frac{1}{c} \left[\frac{c^2}{c^2 + (x - \mathbf{m})^2} \right]^{a + \frac{1}{2}}, \quad (2.2)$$

where $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is a beta coefficient. Recall that the location-scale transform

$Z = \frac{X - \mathbf{m}}{c}$ has a Pearson type VII density

$$f_Z(z) = \frac{1}{B(\frac{1}{2}, \mathbf{a}) \cdot (1+z^2)^{\mathbf{a}+\frac{1}{2}}}. \quad (2.3)$$

If $\mathbf{a} = \frac{u}{2}$, $u = 1, 2, 3, \dots$ is an integer, the random variable $\sqrt{u} \cdot Z$ has a Student t-distribution with u degrees of freedom. The substitution $t = \frac{z^2}{1+z^2}$ shows the integral identity

$$\int_0^x \frac{dz}{(1+z^2)^{\mathbf{a}+\frac{1}{2}}} = \frac{1}{2} \int_0^{\frac{x^2}{1+x^2}} t^{-\frac{1}{2}} (1-t)^{\mathbf{a}-1} dt, \quad (2.4)$$

from which one derives the distribution function

$$F_Z(z) = \begin{cases} \frac{1}{2} \left[1 - \mathbf{b}\left(\frac{1}{2}, \mathbf{a}; \frac{z^2}{1+z^2}\right) \right] & z \leq 0, \\ \frac{1}{2} \left[1 + \mathbf{b}\left(\frac{1}{2}, \mathbf{a}; \frac{z^2}{1+z^2}\right) \right] & z \geq 0, \end{cases} \quad (2.5)$$

where $\mathbf{b}(a, b; x) = \frac{1}{B(a, b)} \int_0^x t^{a-1} (1-t)^{b-1} dt$ is a beta density. The variance (if $\mathbf{a} > 1$) and kurtosis (if $\mathbf{a} > 2$) are equal to

$$\mathbf{s}_Z^2 = \text{Var}[Z] = \frac{1}{2(\mathbf{a}-1)}, \quad \mathbf{g}_{2,Z} = \frac{E[Z^4]}{\text{Var}[Z]^2} = 3 \cdot \left(\frac{\mathbf{a}-1}{\mathbf{a}-2} \right). \quad (3.6)$$

The kurtosis takes values in $[3, \infty)$, and is therefore capable to model leptokurtic data, a typical feature of observed financial market returns, which tend to have heavier tails than those predicted by a normal distribution. This model, first suggested by Praetz(1972), has been fitted to stock returns by Blattberg and Gonedes(1974) and Kon(1984) (see Taylor(1992), Section 2.8). These authors find maximum likelihood estimates for \mathbf{a} between $\frac{3}{2}$ and 3. With the suggested restriction to two-parameter distributions our proposal is $\mathbf{a} = 3$, which yields a kurtosis parameter $\mathbf{g}_{2,Z} = 6$. This choice is motivated by the fact that it yields the less dangerous model among those models for which $\mathbf{a} \in [\frac{3}{2}, 3]$. For comparison, the kurtosis parameter of the logLaplace is also equal to 6, an additional argument for this choice.

As an application we have fitted the daily cumulative returns of the SMI (Swiss Market Index) as well as some typical stocks from this index family using the maximum likelihood method. In our calculations we use the so-called scoring method, where the observed data is grouped into a finite number of classes (Hogg and Klugman(1984), chap. 3.7 and 4.3, Klugman et al.(1998)). We find that the model $NIT(\hat{\mathbf{m}}, \hat{c}, 3)$ beats the lognormal and the logLaplace under a chi-square goodness-of-fit test with regrouped data as in Hürlimann(2000a), Table 6.5. In view of this satisfactory empirical result, our marginal distributions in bivariate fitting of cumulative returns are assumed to be throughout $NIT(\mathbf{m}, c, 3)$ distributions.

3. Bivariate cumulative returns from copulas.

Though copulas have been introduced since Sklar(1959) their use in insurance and finance is more recent (e.g. Carriere(1994), Frees et al.(1996), Frees and Valdez(1998), Embrechts et al.(1999), Li(1999), Klugman and Prasa(1999), Bouyé et al.(2000), Hürlimann(2000b)). A lot of background on copulas can be found in Schweizer and Sklar(1983), Hutchinson and Lai(1990), Dall'Aglio et al.(1991), Rüschendorf et al.(1996), Benes and Stepen(1997), Joe(1997), Nelsen(1999).

Recall that the copula representation of a continuous multivariate distribution allows for a separate modelling of the univariate marginals and the multivariate or dependence structure. Denote by $R(F_1, \dots, F_n)$ the class of all continuous multivariate random variables (X_1, \dots, X_n) with given marginals F_i of X_i . If F denotes the multivariate distribution of (X_1, \dots, X_n) , then the *copula associated with* F is a distribution function $C : [0,1]^n \rightarrow [0,1]$ that satisfies

$$F(x) = C(F_1(x_1), \dots, F_n(x_n)), \quad x = (x_1, \dots, x_n) \in R^n. \quad (3.1)$$

Reciprocally, if $F \in R(F_1, \dots, F_n)$ and F_i^{-1} are quantile functions of the margins, then

$$C(u) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)), \quad u = (u_1, \dots, u_n) \in [0,1]^n, \quad (3.2)$$

is the unique copula satisfying (2.1) (theorem of Sklar(1959)).

In the present paper we are only interested in the joint cumulative returns between a market index and a stock from the index family, that is in the most important bivariate case $n = 2$. Joe(1997) and Nelsen(1999) provide extensive lists of one-parameter families of copulas $C_q(u, v)$, where q is a measure of dependence, from which we retain the following candidates for copula fitting.

Cuadras-Augé (1981)

$$C_q(u, v) = [\min(u, v)]^q [uv]^{1-q}, \quad 0 \leq q \leq 1. \quad (3.3)$$

Gumbel-Hougaard (Gumbel(1960), Hougaard(1986), Hutchinson and Lai(1990))

$$C_q(u, v) = \exp \left\{ - \left[(-\ln u)^q + (-\ln v)^q \right]^{\frac{1}{q}} \right\}, \quad q \geq 1. \quad (3.4)$$

Galambos (1975) (Joe(1997), p. 142)

$$C_q(u, v) = \exp \left\{ \left[(-\ln u)^{-q} + (-\ln v)^{-q} \right]^{\frac{1}{q}} \right\}, \quad q \geq 0. \quad (3.5)$$

Frank (1979)

$$C_q(u, v) = -\frac{1}{q} \ln \left\{ 1 + \frac{(e^{-qu} - 1)(e^{-qv} - 1)}{e^{-q} - 1} \right\}, \quad q \neq 0. \quad (3.6)$$

Hüsler-Reiss (1989) (Joe(1997), p. 143)

$$C_q(u, v) = \exp \left\{ \ln u \cdot \Phi \left[\frac{1}{q} + \frac{1}{2} q \ln \left(\frac{\ln u}{\ln v} \right) \right] + \ln v \cdot \Phi \left[\frac{1}{q} + \frac{1}{2} q \ln \left(\frac{\ln v}{\ln u} \right) \right] \right\}, \quad q \geq 0, \quad (3.7)$$

with $\Phi(x)$ the standard normal distribution.

Clayton (1978) (Kimeldorf and Sampson(1975a), Cook-Johnson(1981))

$$C_q(u, v) = \max \left(\left[u^{-q} + v^{-q} - 1 \right]^{\frac{1}{q}}, 0 \right), \quad q \geq 0. \quad (3.8)$$

In all these copula families, the parameter q of the joint cumulative distribution $F(x, y) = C_q[F_X(x), F_Y(y)]$ associated to a random couple (X, Y) measures the degree of dependence between X and Y . The larger q is in absolute value, the stronger the dependence. A positive value of q indicates a positive dependence. Sometimes one is especially interested in one-parameter families of copulas, which are able to model continuously a whole range of dependence between the lower Fréchet bound copula $\max(u + v - 1, 0)$, the independent copula uv , and the upper Fréchet bound copula $\min(u, v)$. Such families are called *inclusive* or *comprehensive*. The extensive list by Nelsen(1999), p. 96, contains only two one-parameter inclusive families of copulas, namely those by Frank(1979) and Clayton(1978) also considered above. Another one, considered first in Hürlimann(2000b), is described in the next Section.

4. The linear Spearman copula.

Consider the *linear Spearman copula*, which is defined as follows. For $q \in [0, 1]$ one has

$$C_q(u, v) = \begin{cases} [u + q(1 - u)] \cdot v, & v \leq u, \\ [v + q(1 - v)] \cdot u, & v > u, \end{cases} \quad (4.1)$$

and for $q \in [-1, 0]$ one has

$$C_q(u, v) = \begin{cases} (1 + q) \cdot u, & u + v < 1, \\ uv + q \cdot (1 - u) \cdot (1 - v), & u + v \geq 1. \end{cases} \quad (4.2)$$

For $q \in [0, 1]$ this copula is family B11 in Joe(1997), p. 148. It represents a mixture of perfect dependence and independence. If X and Y are uniform(0,1), $Y = X$ with probability q and Y is independent of X with probability $1 - q$, then (X, Y) has the linear Spearman copula. This distribution has been first considered by Konijn(1959) and motivated in Cohen(1960) along Cohen's kappa statistic (see Hutchinson and Lai(1990), Section 10.9). For the extended copula, the chosen nomenclature *linear* refers to the piecewise linear sections of this copula, and *Spearman* refers to the fact that the *grade*

correlation coefficient \mathbf{r}_s by Spearman(1904) coincides with the parameter \mathbf{q} . This follows from the calculation

$$\mathbf{r}_s = 12 \cdot \int_0^1 \int_0^1 [C_q(u, v) - uv] dudv = \mathbf{q}, \quad (4.3)$$

where a proof of the integral representation is given in Nelsen(1991). The linear Spearman copula, which leads to the so-called *linear Spearman bivariate distribution*, has a singular component, which according to Joe should limit its field of applicability. Despite of this it has many interesting and important properties and is suitable for analytical computation.

For the reader's convenience, let us describe first two extremal properties. Kendall's \mathbf{t} for this copula equals using Nelsen(1991) :

$$\mathbf{t} = 1 - 4 \cdot \int_0^1 \int_0^1 \frac{\partial}{\partial u} C_q(u, v) \cdot \frac{\partial}{\partial v} C_q(u, v) dudv = \frac{1}{3} \mathbf{r}_s \cdot [2 + \text{sgn}(\mathbf{r}_s) \mathbf{r}_s]. \quad (4.4)$$

Invert this to get

$$\mathbf{r}_s = \begin{cases} -1 + \sqrt{1 + 3\mathbf{t}}, & \mathbf{t} \geq 0, \\ 1 - \sqrt{1 - 3\mathbf{t}}, & \mathbf{t} \leq 0. \end{cases} \quad (4.5)$$

Relate this to the convex two-parameter copula by Fréchet(1958) defined by

$$C_{a,b}(u, v) = \mathbf{a} \cdot C_{-1}(u, v) + (1 - \mathbf{a} - \mathbf{b}) \cdot C_0(u, v) + \mathbf{b} \cdot C_1(u, v), \quad \mathbf{a}, \mathbf{b} \geq 0, \quad \mathbf{a} + \mathbf{b} \leq 1. \quad (4.6)$$

Since $\mathbf{r}_s = \mathbf{a} - \mathbf{b}$ and $\mathbf{t} = \frac{\mathbf{a} - \mathbf{b}}{3} (2 + \mathbf{a} + \mathbf{b})$ for this copula, one has the inequalities

$$\begin{aligned} \mathbf{t} \leq \mathbf{r}_s \leq -1 + \sqrt{1 + 3\mathbf{t}}, & \quad \mathbf{t} \geq 0, \\ 1 - \sqrt{1 - 3\mathbf{t}} \leq \mathbf{r}_s \leq \mathbf{t}, & \quad \mathbf{t} \leq 0. \end{aligned} \quad (4.7)$$

The linear Spearman copula satisfies the following extremal property. For $\mathbf{t} \geq 0$ the upper bound for \mathbf{r}_s in Fréchet's copula is attained by the linear Spearman copula, and for $\mathbf{t} \leq 0$ it is the lower bound, which is attained.

In case $\mathbf{t} \geq 0$ a second more important extremal property holds, which is related to a conjectural statement. Recall that Y is *stochastically increasing* on X , written $SI(Y|X)$, if $\Pr(Y > y|X = x)$ is a nondecreasing function of x for all y . Similarly, X is *stochastically increasing* on Y , written $SI(X|Y)$, if $\Pr(X > x|Y = y)$ is a nondecreasing function of y for all x . (Note that Lehmann(1966) speaks instead of *positive regression dependence*). If X and Y are continuous random variables with copula $C(u, v)$, then one has the equivalences (Nelsen(1999), Theorem 5.2.10) :

$$SI(Y|X) \Leftrightarrow \frac{\partial}{\partial u} C(u, v) \text{ is nonincreasing in } u \text{ for all } v \quad (4.8)$$

$$SI(X|Y) \Leftrightarrow \frac{\partial}{\partial v} C(u, v) \text{ is nonincreasing in } v \text{ for all } u \quad (4.9)$$

The *Hutchinson-Lai conjecture* consists of the following statement. If (X, Y) satisfies the properties (4.8), then Spearman's r_s satisfies the inequalities

$$-1 + \sqrt{1 + 3t} \leq r_s \leq \min \left\{ \frac{3}{2}t, 2t - t^2 \right\} \quad (4.10)$$

The upper bound $2t - t^2$ is attained for the one-parameter copula introduced by Kimeldorf and Sampson(1975b) (see also Hutchinson and Lai(1990), Section 13.7). The lower bound is attained by the linear Spearman copula, as shown already by Konijn(1959), p. 277. Alternatively, if the conjecture holds, the maximum value of Kendall's t by given r_s is attained for the linear Spearman copula. Note that the upper bound $r_s \leq \frac{3}{2}t$ has been disproved recently by Nelsen(1999), Exercise 5.36. The remaining conjecture $-1 + \sqrt{1 + 3t} \leq r_s \leq 2t - t^2$ is still unsettled.

As an important modelling characteristic, let us show that the linear Spearman copula leads to a simple tail dependence structure, which is of interest when extreme values are involved. Recall that the *coefficient of (upper) tail dependence* of a couple (X, Y) is defined by

$$I = I(X, Y) = \lim_{a \rightarrow 1^-} \Pr(Y > Q_Y(\mathbf{a}) | X > Q_X(\mathbf{a})), \quad (4.11)$$

provided a limit I in $[0, 1]$ exists ($Q_X(u) = \inf \{x | \Pr(X \leq x) \geq u\}$ denotes a quantile function of X). If $I \in (0, 1]$ then the couple (X, Y) is called *asymptotically dependent* (in the upper tail) while if $I = 0$ one speaks of *asymptotic independence*. Tail dependence is an asymptotic property of the copula. Its calculation follows easily from the relation

$$\begin{aligned} & \Pr(Y > Q_Y(\mathbf{a}) | X > Q_X(\mathbf{a})) \\ &= \frac{1 - \Pr(X \leq Q_X(\mathbf{a})) - \Pr(Y \leq Q_Y(\mathbf{a})) + \Pr(X \leq Q_X(\mathbf{a}), Y \leq Q_Y(\mathbf{a}))}{1 - \Pr(X \leq Q_X(\mathbf{a}))}. \end{aligned} \quad (4.12)$$

For a linear Spearman couple one obtains

$$I(X, Y) = \lim_{a \rightarrow 1^-} \frac{1 - 2a + C_q(\mathbf{a}, \mathbf{a})}{1 - a} = \lim_{a \rightarrow 1^-} (1 - a + qa) = q. \quad (4.13)$$

Therefore, unless X and Y are independent, a linear Spearman couple is always asymptotically dependent. This is a desirable property in insurance and financial modelling, where data tend to be dependent in their extreme values. In contrast to this, the ubiquitous Gaussian copula yields always asymptotic independence, unless perfect correlation holds (Sibuya(1961), Resnick(1987), Chap. 5, Embrechts et al.(1998), Section 4.4).

5. Bivariate copula fitting.

Our view of multivariate statistical modelling is that of Joe(1997), Section 1.7 :

“Models should try to capture important characteristics, such as the appropriate density shapes for the univariate margins and the appropriate dependence structure, and otherwise be as simple as possible.”

To fulfill this, a bivariate model should satisfy some desirable properties (Joe(1997), Section 4.1, Klugman and Parsa(1999)) :

- a) The bivariate distribution and/or density should preferably have a closed-form representation, at least numerical evaluation should be possible.
- b) The marginal distributions and/or densities should belong to the same parametric family and numerical evaluation should be possible.
- c) A parameter of the model should describe the dependence between the margins and cover a wide range of dependence.
- d) The model should be sufficiently flexible to fit the available data.

Modelling bivariate cumulative returns with the copulas of Sections 3 and 4 yields a closed-form representation for the bivariate distribution, which can be numerically evaluated provided the margins follow $NI\Gamma(\mathbf{m}, c, 3)$ distributions. Therefore the properties a) and b) are fulfilled. The parameter \mathbf{q} in our copulas, which models the dependence, covers the relevant range of positive dependence lying between the independent copula and the upper Fréchet bound copula, hence c) is satisfied. It remains to analyze point d), that is the effective fitting of the chosen copulas to actual data. Statistical inference is done by specifying the estimation method and a bivariate goodness-of-fit test.

To estimate the 5 parameters $(\mathbf{m}_X, c_X, \mathbf{m}_Y, c_Y, \mathbf{q})$ of a copula-based bivariate model $F(x, y) = C_q[F_X(x), F_Y(y)]$, where $F_X(x), F_Y(x)$ are $NI\Gamma(\mathbf{m}_X, c_X, 3), NI\Gamma(\mathbf{m}_Y, c_Y, 3)$ distributed, we apply a method close in spirit to the method of *inference function for margins* or *IFM method* studied in McLeish and Small(1988), Xu(1996), and Joe(1997), Section 10.1. A simple case of the IFM method consists of doing two separate maximum likelihood estimations of the univariate marginal distributions, followed by an optimization of the bivariate likelihood as a function of the dependence parameter. This procedure is computationally simpler and less time-consuming compared with a simultaneous estimation of all parameters from the bivariate likelihood. A statistical theory of the IFM method is developed in Joe(1997), Section 10.1. Proceeding similarly, we perform two separate maximum likelihood estimations of the univariate margins, followed by an estimation of the dependence parameter. However, we do not maximize the bivariate likelihood, except for comparison purposes in the Tables 5.1 and 5.2. Instead, we determine the dependence parameter, which maximizes the p-value (respectively minimizes the bivariate chi-square statistic) of a bivariate version of the usual Pearson goodness-of-fit test. We note that an application of the uniform test by Quesenberry(1986) for testing Y given X and X given Y , as proposed in Klugman and Prasa(1999), Section 5, is only satisfactory in one way. Testing the stock return Y given the index return X is accepted, while testing the index return X given the stock return Y is rejected.

Let us describe in more details our estimation of the dependence parameter \mathbf{q} . We assume the parameters $\mathbf{m}_X, c_X, \mathbf{m}_Y, c_Y$ have been determined using two separate maximum likelihood estimations. Given $n+1$ daily observations I_i of a market index and $n+1$ daily prices S_i of a stock in the index family, let $X_i = \frac{I_{i+1}}{I_i}, Y_i = \frac{S_{i+1}}{S_i}, i = 1, \dots, n$ be the daily cumulative returns used for statistical fitting.

First, we want to validate the estimation of the marginal distributions using two separate univariate Pearson chi-square tests. Following Hürliemann(2000a), Section 6, the raw data X_i is regrouped into 6 classes $(v_0, v_1], (v_1, v_2], \dots, (v_5, v_6]$, where the boundaries v_i 's are chosen such that the number of observations I_1, I_2, \dots, I_6 in the corresponding classes are as much symmetrically distributed as possible ($v_3 = 1$ appears adequate). This arrangement is motivated by the fact that the $NIF(\mathbf{m}, c, 3)$ is a symmetric distribution. The data Y_i is regrouped into 6 similar classes $(w_0, w_1], (w_1, w_2], \dots, (w_5, w_6]$ with observations h_1, h_2, \dots, h_6 in each class. One obtains two separate chi-square statistics

$$\mathbf{c}_X^2 = \sum_{i=1}^6 \frac{(I_i - nf_i)^2}{nf_i}, \quad \mathbf{c}_Y^2 = \sum_{i=1}^6 \frac{(h_i - ng_i)^2}{ng_i}, \quad (5.1)$$

$$f_i = \frac{F_X(v_i) - F_X(v_{i-1})}{1 - F_X(v_0)}, \quad g_i = \frac{F_Y(w_i) - F_Y(w_{i-1})}{1 - F_Y(w_0)}.$$

In the examples below we report the p-values p_X, p_Y corresponding to $\mathbf{c}_X^2, \mathbf{c}_Y^2$ for a chi-square distribution with 3 degrees of freedom. If these values are sufficiently high, the marginal distributions have been fitted in a satisfactory way.

Next, to create a meaningful bivariate chi-square statistic, we look at the number of observations $z_{i,j}$ in the 36 two-dimensional intervals $(v_{i-1}, v_i] \times (w_{j-1}, w_j]$, $i, j = 1, \dots, 6$. By Moore(1978/86) (Klugman et al.(1998), p. 121) regroup these intervals in 10 larger rectangular interval classes as follows. Recommended is an expected frequency of at least 1% in each class and a 5% expected frequency in 80% of the classes. With $n = 250$ daily observations over an approximate one-year period the following rectangular regrouping in 10 classes $C_k, k = 1, \dots, 10$, appears adequate (at least in our examples) :

| | w_0 | w_1 | w_2 | w_3 | w_4 | w_5 | w_6 |
|-------|-------|-------|-------|----------|-------|-------|-------|
| v_0 | | | | | | | |
| v_1 | C_1 | C_1 | C_1 | C_1 | C_1 | C_1 | C_1 |
| v_2 | C_1 | C_3 | C_5 | C_5 | C_5 | C_5 | C_2 |
| v_3 | C_1 | C_6 | C_7 | C_8 | C_5 | C_5 | C_2 |
| v_4 | C_1 | C_6 | C_9 | C_{10} | C_5 | C_5 | C_2 |
| v_5 | C_1 | C_6 | C_6 | C_6 | C_4 | C_4 | C_2 |
| v_6 | C_2 | C_2 | C_2 | C_2 | C_2 | C_2 | C_2 |

Our bivariate chi-square statistic is constructed from this configuration as follows. Consider the fitted or expected number of observations $f_{i,j}$ in each of the 36 two-dimensional intervals $(v_{i-1}, v_i] \times (w_{j-1}, w_j]$ given by

$$f_{i,j} = n \cdot [F(v_i, w_j) - F(v_{i-1}, w_j) - F(v_i, w_{j-1}) + F(v_{i-1}, w_{j-1})], \quad (5.2)$$

$$i, j = 1, \dots, 6, \quad F(x, y) = C_q [F_X(x), F_Y(y)]$$

Through summation of $z_{i,j}$'s, respectively $f_{i,j}$'s, one obtains the number of observations O_k , respectively the expected number of observations E_k , in each class C_k , $k = 1, \dots, 10$. The bivariate chi-square statistic is then defined by

$$c^2 = \sum_{k=1}^{10} \frac{(O_k - E_k)^2}{E_k}. \quad (5.3)$$

It is now possible to obtain for each copula numerical values of \mathbf{q} , which minimize c^2 respectively maximize the bivariate p-value corresponding to c^2 for a chi-square distribution with 4 degrees of freedom. For comparisons the value of the bivariate negative log-likelihood is also of interest. It is defined by

$$-\ln L = n \cdot \ln[1 - F_X(v_0) - F_Y(w_0) + F(v_0, w_0)] - \sum_{i=1}^6 \sum_{j=1}^6 z_{i,j} \cdot \ln[F(v_i, w_j) - F(v_{i-1}, w_j) - F(v_i, w_{j-1}) + F(v_{i-1}, w_{j-1})]. \quad (5.4)$$

Following the IFM method described above, one can obtain numerical values of \mathbf{q} , which minimize $-\ln L$ under a p-value of at least 5% (the model should not be rejected at this significance level). Our examples show that the IFM method reduces the p-value of c^2 sometimes rather drastically. For this reason, we prefer the proposed bivariate minimum chi-square or maximum p-value estimation method.

In our examples, we start with a comparison of fit based on the IFM method. Tables 5.1 and 5.2 report copula fitting results for daily cumulative returns between the SMI index and a stock in the index family for the one-year period between September 29, 1998 and September 24, 1999. For the Credit Suisse Group stock the Gumbel-Hougaard copula maximizes the bivariate log-likelihood, the linear Spearman copula yields the highest p-value. For the Novartis stock the Clayton copula maximizes the bivariate log-likelihood, the Frank copula yields the highest p-value.

Table 5.1 : IFM method for Credit Suisse Group stock

| location-scale parameters : $m_x = 1.000656016$, $c_x = 0.030403482$, $m_y = 1.001624799$, $c_y = 0.051511031$. | | | | | |
|--|--------------|---------|-------------------|-------|-------|
| model | \mathbf{q} | p-value | $(-\ln L)_{\min}$ | p_x | p_y |
| Gumbel-Hougaard | 2.65 | 13.5 | 599.4 | 74.9 | 84.9 |
| Galambos | 1.97 | 13.1 | 600.3 | 74.9 | 84.9 |
| Hüsler-Reiss | 2.48 | 5 | 602.9 | 74.9 | 84.7 |
| Frank | 9.25 | 16.7 | 604.4 | 75.6 | 82.9 |
| Cuadras-Augé | 0.665 | 31.1 | 610.8 | 75.9 | 83.5 |
| Linear Spearman | 0.52 | 66.7 | 614.2 | 76.6 | 83.4 |
| Clayton | 3.58 | 5 | 629.5 | 76.4 | 85.1 |

Table 5.2 : IFM method for Novartis stock

| location-scale parameters : $m_x = 1.000656016$, $c_x = 0.030403482$, $m_y = 1.000225799$, $c_y = 0.035275623$. | | | | | |
|--|-------|---------|-------------------|-------|-------|
| model | q | p-value | $(-\ln L)_{\min}$ | p_X | p_Y |
| Clayton | 2.25 | 30.7 | 592.8 | 75.4 | 93.8 |
| Frank | 6.9 | 32.5 | 593 | 75.6 | 93.4 |
| Gumbel-Hougaard | 2.23 | 11.6 | 594.9 | 75.1 | 93.4 |
| Galambos | 1.57 | 11 | 595.1 | 75 | 93.3 |
| Hüsler-Reiss | 2.03 | 9.5 | 596 | 75.1 | 93.3 |
| Cuadras-Augé | 0.619 | 5.5 | 611.5 | 75.9 | 93.5 |
| Linear Spearman | 0.474 | 25.7 | 612.7 | 76.7 | 93.6 |

The Tables 5.3 to 5.8 report comparisons of fit using the bivariate minimum chi-square estimation method for 6 pairs of daily cumulative returns between the SMI index and a stock in the index family for the same one-year period. Except for the Nestlé stock, whose fitted marginal distribution is rejected, we find in most cases satisfactory fits, in particular for the SMI stocks Credit Suisse Group, Sulzer, Novartis and UBS. The rejected fit of the Nestlé marginal distribution may be due to the fact that the extreme values of the daily returns were rather moderate between -6% and 6%. The fit of the Swisscom stock is not rejected but less satisfactory. Rather surprisingly, the highest bivariate p-value is obtained for the very tractable linear Spearman copula (except for the Swisscom stock). Except for the Cuadras-Augé (UBS) and Clayton copulas (Credit Suisse Group and Swisscom), the p-values seem sufficiently high. In our view the linear Spearman, Frank and Gumbel-Hougaard copulas provide the best overall fits for the analyzed pairs.

Table 5.3 : Bivariate maximum p-value method for Credit Suisse Group stock

| location-scale parameters : $m_x = 1.000656016$, $c_x = 0.030403482$, $m_y = 1.001624799$, $c_y = 0.051511031$. | | | | | |
|--|-------|--------------------|----------|-------|-------|
| model | q | $(p-value)_{\max}$ | $-\ln L$ | p_X | p_Y |
| Linear Spearman | 0.495 | 71.8 | 614.5 | 76.6 | 83.4 |
| Cuadras-Augé | 0.64 | 36.9 | 611.3 | 76 | 83.3 |
| Gumbel-Hougaard | 3.05 | 28.8 | 601.4 | 74.7 | 85.4 |
| Galambos | 2.37 | 28.2 | 602.8 | 74.6 | 85.5 |
| Frank | 10.6 | 26.4 | 605.7 | 75.5 | 83 |
| Hüsler-Reiss | 3.15 | 26.4 | 609.5 | 74.5 | 85.5 |
| Clayton | 3.83 | 5.7 | 632.2 | 76.4 | 85.1 |

Table 5.4 : Bivariate maximum p-value method for Novartis stock

| location-scale parameters : $m_x = 1.000656016$, $c_x = 0.030403482$, $m_y = 1.000225799$, $c_y = 0.035275623$. | | | | | |
|--|-------|----------------------|----------|-------|-------|
| model | q | $(p - value)_{\max}$ | $-\ln L$ | p_X | p_Y |
| Linear Spearman | 0.399 | 56.4 | 614.3 | 76.6 | 93.5 |
| Frank | 6.9 | 32.5 | 593 | 75.6 | 93.4 |
| Clayton | 2.18 | 31.1 | 592.8 | 75.3 | 93.8 |
| Cuadras-Augé | 0.547 | 15.6 | 612.9 | 76 | 93.4 |
| Gumbel-Hougaard | 2.23 | 11.6 | 594.9 | 75.1 | 93.4 |
| Galambos | 1.56 | 11.1 | 595.1 | 75 | 93.3 |
| Hüsler-Reiss | 2.17 | 10.3 | 596.5 | 75 | 93.3 |

Table 5.5 : Bivariate maximum p-value method for UBS stock

| location-scale parameters : $m_x = 1.000656016$, $c_x = 0.030403482$, $m_y = 1.0016621$, $c_y = 0.0477182$ | | | | | |
|--|-------|----------------------|----------|-------|-------|
| model | q | $(p - value)_{\max}$ | $-\ln L$ | p_X | p_Y |
| Linear Spearman | 0.403 | 38 | 629.6 | 77.1 | 52.3 |
| Clayton | 2.42 | 25.3 | 648.2 | 76.7 | 49.6 |
| Frank | 7.43 | 14.5 | 633 | 76.7 | 54.2 |
| Gumbel-Hougaard | 2.38 | 14.1 | 626.9 | 76.8 | 52.6 |
| Galambos | 1.69 | 14 | 629 | 76.8 | 52.6 |
| Hüsler-Reiss | 2.34 | 13.8 | 638.6 | 76.8 | 52.5 |
| Cuadras-Augé | 0.55 | 5.5 | 634 | 76.7 | 54 |

Table 5.6 : Bivariate maximum p-value method for Nestlé stock

| location-scale parameters : $m_x = 1.000656016$, $c_x = 0.030403482$, $m_y = 1.00096518$, $c_y = 0.03362619$ | | | | | |
|--|-------|----------------------|----------|-------|-------|
| model | q | $(p - value)_{\max}$ | $-\ln L$ | p_X | p_Y |
| Cuadras-Augé | 0.5 | 34 | 604.1 | 79.1 | 0 |
| Gumbel-Hougaard | 2.11 | 32.1 | 579.9 | 80.8 | 0 |
| Linear Spearman | 0.377 | 31.5 | 606.1 | 78.8 | 0 |
| Galambos | 1.42 | 31.2 | 579.5 | 80.8 | 0 |
| Hüsler-Reiss | 2.01 | 29.7 | 578.9 | 80.9 | 0 |
| Frank | 6.34 | 29.3 | 584.9 | 75.6 | 0.1 |
| Clayton | 2.03 | 2.5 | 592.2 | 75.8 | 0 |

Table 5.7 : Bivariate maximum p-value method for Sulzer stock

| location-scale parameters : $m_x = 1.000656016, c_x = 0.030403482,$ $m_y = 1.0014062, c_y = 0.0466025$ | | | | | |
|---|-------|----------------------|----------|-------|-------|
| model | q | $(p - value)_{\max}$ | $-\ln L$ | p_X | p_Y |
| Linear Spearman | 0.264 | 69.7 | 649.1 | 76.8 | 46.7 |
| Gumbel-Hougaard | 1.77 | 64.9 | 636.9 | 77.1 | 47.9 |
| Galambos | 1.09 | 64.9 | 637.8 | 77.1 | 47.9 |
| Hüsler-Reiss | 1.62 | 64.8 | 639.9 | 77 | 48 |
| Frank | 4.72 | 50.8 | 641.8 | 76.3 | 48.7 |
| Cuadras-Augé | 0.382 | 46.2 | 649.9 | 76.7 | 47.8 |
| Clayton | 1.45 | 24.3 | 634.9 | 76.2 | 42.7 |

Table 5.8 : Bivariate maximum p-value method for Swisscom stock

| location-scale parameters : $m_x = 1.000656016, c_x = 0.030403482,$ $m_y = 1.00071313, c_y = 0.05098643$ | | | | | |
|---|-------|----------------------|----------|-------|-------|
| model | q | $(p - value)_{\max}$ | $-\ln L$ | p_X | p_Y |
| Cuadras-Augé | 0.151 | 15.9 | 700.6 | 75.6 | 18.2 |
| Frank | 2.09 | 12.8 | 695.7 | 75.3 | 18.1 |
| Linear Spearman | 0.094 | 9.1 | 703.7 | 75.2 | 18.3 |
| Gumbel-Hougaard | 1.22 | 8.5 | 696.3 | 75.7 | 18 |
| Galambos | 0.47 | 7.9 | 696.6 | 75.9 | 18 |
| Hüsler-Reiss | 0.84 | 7.7 | 697 | 75.9 | 17.9 |
| Clayton | 0.36 | 2.5 | 701.9 | 72 | 18.4 |

6. Covariance estimation with the linear Spearman copula.

To present a practical illustration of our estimation results, it is useful to compare standard empirical values of the linear correlation coefficient with the estimated linear correlation coefficient from a fitted copula. This is of interest because for heavy tailed data the standard product-moment correlation estimator has a very bad performance, and there is a need for more robust estimators (e.g. Lindskog(2000)). In view of the importance of correlation measures in modern finance (mean-variance portfolio theory, CAPM models), this is a highly relevant issue.

Extreme, synchronized rises and falls of indices and stocks occur infrequently, but more often than what is predicted by a bivariate normal model. Since the linear Spearman copula has a simple tail dependence structure capable to model extreme values (end of Section 4) and our copula fitting was quite satisfactory, we focus on the evaluation of the covariance for this copula. The following general result has been first derived in Hürlimann(2000b).

Theorem 6.1. Let (X, Y) be distributed as $F(x, y) = C_q[F_X(x), F_Y(y)]$, where $C_q(u, v)$ is the linear Spearman copula, and the continuous and strictly increasing marginal distributions

are defined on the open supports (a_x, b_x) , (a_y, b_y) . For an arbitrary differentiable function $\mathbf{y}(y)$, assume the following regularity assumption holds :

$$\begin{aligned} \lim_{y \rightarrow a_y} \mathbf{y}(y)F_Y(y) &= 0, \\ \lim_{y \rightarrow b_y} \mathbf{y}(y) \left(E[X|F_Y(y)] - \int_{a_y}^y F_X^{-1}[F_Y^q(y)] dF_Y(y) \right) &= 0. \end{aligned} \quad (\text{RA})$$

Then one has the covariance formula

$$\text{Cov}[X, \mathbf{y}(Y)] = \text{sgn}(\mathbf{q})\mathbf{q} \cdot E\left[\left(F_X^{-1}[F_Y^q(Y)] - E[X]\right) \cdot \mathbf{y}(Y)\right], \quad (6.1)$$

where one sets

$$F_Y^q(y) = \begin{cases} F_Y(y), & \mathbf{q} \geq 0, \\ \bar{F}_Y(y), & \mathbf{q} < 0, \end{cases} \quad (6.2)$$

with the abbreviation $\bar{F}_Y(y) = 1 - F_Y(y)$.

Proof. Let us first derive the regression function $E[X|Y=y]$. The conditional distribution of X given $Y=y$ equals for $\mathbf{q} \geq 0$

$$F(x|y) = \frac{\partial C_{\mathbf{q}}}{\partial v}[F_X(x), F_Y(y)] = \begin{cases} F_X(x) + \mathbf{q}\bar{F}_X(x), & x \geq F_X^{-1}[\bar{F}_Y(y)] \\ (1-\mathbf{q})F_X(x), & x < F_X^{-1}[F_Y(y)] \end{cases} \quad (6.3)$$

and for $\mathbf{q} < 0$

$$F(x|y) = \begin{cases} (1+\mathbf{q})F_X(x), & x < F_X^{-1}[\bar{F}_Y(y)] \\ F_X(x) - \mathbf{q}\bar{F}_X(x), & x \geq F_X^{-1}[F_Y(y)] \end{cases} \quad (6.4)$$

Through calculation one obtains the regression formula

$$E[X|y] = \int_0^\infty [1 - F(x|y)] dx - \int_{-\infty}^0 F(x|y) dx \quad (6.5)$$

$$= \begin{cases} E[X] - \mathbf{q} \cdot (E[X] - F_X^{-1}[F_Y(y)]), & \mathbf{q} \geq 0, \\ E[X] + \mathbf{q} \cdot (E[X] - F_X^{-1}[\bar{F}_Y(y)]), & \mathbf{q} \leq 0, \end{cases} \quad (6.6)$$

which is a weighted average of the mean $E[X]$ and the quantile $F_X^{-1}[F_Y(y)]$ respectively $F_X^{-1}[\bar{F}_Y(y)]$. To obtain the stated covariance formula, one uses the well-known formula by Hoeffding(1940) and Lehmann(1966), Lemma 2, to get the expression

$$\begin{aligned}
\text{Cov}[X, \mathbf{y}(Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(x, y) - F_X(x)F_Y(y)] \mathbf{y}'(y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_Y(y) [F(x|Y \leq y) - F_X(x)] \mathbf{y}'(y) dx dy \\
&= \int_{-\infty}^{\infty} F_Y(y) \left[\int_{-\infty}^0 F(x|Y \leq y) dx - \int_{-\infty}^0 F_X(x) dx + \int_0^{\infty} \bar{F}_X(x) dx - \int_0^{\infty} [1 - F(x|Y \leq y)] dx \right] \mathbf{y}'(y) dy \\
&= \int_{-\infty}^{\infty} (E[X] - E[X|Y \leq y]) F_Y(y) \mathbf{y}'(y) dy.
\end{aligned} \tag{6.7}$$

Furthermore, from (6.6) one obtains

$$E[X] - E[X|Y \leq y] = \text{sgn}(\mathbf{q}) \mathbf{q} \cdot (E[X] - E[F_X^{-1}[F_Y^q(Y)]|Y \leq y]). \tag{6.8}$$

Inserted in (6.7), a partial integration yields

$$\begin{aligned}
\text{Cov}[X, \mathbf{y}(Y)] &= \text{sgn}(\mathbf{q}) \mathbf{q} \cdot \int_{a_Y}^{b_Y} \left(E[X] F_Y(y) - \int_{a_Y}^y F_X^{-1}[F_Y^q(y)] dF_Y(y) \right) \mathbf{y}'(y) dy \\
&= \text{sgn}(\mathbf{q}) \mathbf{q} \cdot \left(\mathbf{y}(y) \left[E[X] F_Y(y) - \int_{a_Y}^y F_X^{-1}[F_Y^q(y)] dF_Y(y) \right]_{a_Y}^{b_Y} \right. \\
&\quad \left. + \int_{a_Y}^{b_Y} \mathbf{y}(y) (F_X^{-1}[F_Y^q(y)] - E[X]) dF_Y(y) \right)
\end{aligned} \tag{6.9}$$

which implies (6.1) by the regularity assumption. \diamond

The application of this result to margins from a symmetric location-scale family is simple.

Corollary 6.1. Under the assumptions from Theorem 6.1 suppose that

$$F_X(x) = F_Z\left(\frac{x - \mathbf{m}_X}{c_X}\right), \quad F_Y(y) = F_Z\left(\frac{y - \mathbf{m}_Y}{c_Y}\right), \quad \mathbf{m}_X = E[X], \quad \mathbf{m}_Y = E[Y], \quad \text{and} \quad \bar{F}_Z(-z) = F_Z(z).$$

Then one has

$$\text{Cov}[X, \mathbf{y}(Y)] = \mathbf{q} \frac{c_X}{c_Y} \text{Cov}[Y, \mathbf{y}(Y)]. \tag{6.10}$$

Proof. The result follows from (6.1) noting that $F_X^{-1}[F_Y^q(y)] = \mathbf{m}_X + \text{sgn}(\mathbf{q}) \frac{c_X}{c_Y} (y - \mathbf{m}_Y)$. \diamond

Example 6.1.

In case $X \sim N\Gamma(\mathbf{m}_X, c_X, \mathbf{a})$, $Y \sim N\Gamma(\mathbf{m}_Y, c_Y, \mathbf{a})$, $\mathbf{a} > 1$, one has $c_X = \sqrt{2(\mathbf{a} - 1)\text{Var}[X]}$, $c_Y = \sqrt{2(\mathbf{a} - 1)\text{Var}[Y]}$. For $\mathbf{y}(y) = y$ the regularity assumption (RA) holds, hence $\text{Cov}[X, Y] = \mathbf{q} \sqrt{\text{Var}[X]\text{Var}[Y]}$. In this special situation Spearman's correlation coefficient \mathbf{q} coincides with Pearson's linear correlation coefficient. It is therefore possible to compare the standard product-moment correlation estimator with the estimated \mathbf{q} obtained from the linear Spearman copula fitting. Results for some stocks from the SMI index family are found in Table 6.1.

Table 6.1 : linear correlation estimators for the linear Spearman copula

| Stock | Product-moment estimator | Fitted estimator |
|---------------------|--------------------------|------------------|
| Credit Suisse Group | 0.829 | 0.495 |
| UBS | 0.708 | 0.403 |
| Nestlé | 0.777 | 0.377 |
| Novartis | 0.798 | 0.399 |
| Sulzer | 0.663 | 0.264 |
| Swisscom | 0.285 | 0.094 |

The discrepancy between both estimators is considerable. However, on a relative scale both estimators rank the strength of dependence similarly.

References.

- Benes, V. and J. Stepen (1997).* (Ed.) Distributions with Given Marginals and Moment Problems. Kluwer Academic Publishers.
- Blattberg, R.C. and N.J. Gonedes (1974).* A comparison of the stable and Student distributions as statistical models for stock prices. *Journal of Business* 47, 244-80.
- Bouyé, E., A. Nikeghbali, G. Riboulet and T. Roncalli (2000).* Copulas for Finance. A Reading Guide and some Applications. Preprint.
- Carriere, J. (1994).* Dependent decrement theory. *Transactions of the Society of Actuaries* XLVI, 45-74.
- Clayton, D. (1978).* A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65, 141-151.
- Cohen, J. (1960).* A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, 37-46.
- Cook, R.D. and M.E. Johnson (1981).* A family of distributions for modeling non-elliptically symmetric multivariate data. *Journal of the Royal Statist. Society, Ser. B* 43, 210-218.
- Cuadras, C.M. and J. Augé (1981).* A continuous general multivariate distribution and its properties. *Communications in Statistics - Theory and Methods, Ser. A* 10(4), 339-53.
- Dall'Aglio, G., Kotz, S. and G. Salinetti (1991).* (Ed.) *Advances in Probability Distributions with Given Marginals.* Kluwer Academic Publishers.
- Embrechts, P., McNeil, A. and D. Straumann (1998).* Correlation and dependency in Risk Management : properties and pitfalls. Preprint ETHZ, available from <http://www.math.ethz.ch/~embrechts>.
- Frank, M.J. (1979).* On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Mathematicae* 19, 194-226.
- Fréchet, M. (1958).* Remarques au sujet de la note précédente. *Comptes Rendus de l'Académie des Sciences de Paris, Série I Math.* 246, 2719-2220.
- Frees, E.W., Carriere, J., and E.A. Valdez (1996).* Annuity valuation with dependent mortality. *Journal of Risk and Insurance* 63, 229-261.
- Frees, E.W., and E.A. Valdez (1998).* Understanding relationships using copulas. *North American Actuarial Journal* 2, 1-25.
- Galambos, J. (1975).* Order statistics of samples from multivariate distributions. *Journal of the American Statistical Association* 70, 674-680.
- Gumbel, E.J. (1960).* Distributions des valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris* 9, 171-173.

- Hoefding, W.* (1940). Massstabinvariante Korrelations-theorie. Schriften des Mathematischen Instituts der Universität Berlin 5, 181-233.
- Hogg, R. and S. Klugman* (1984). Loss Distributions. John Wiley. New York.
- Hougaard, P.* (1986). A class of multivariate failure time distributions. *Biometrika* 73, 671-678.
- Hürlimann, W.* (1995). Is there a rational evidence for an infinite variance asset pricing model ? Proceedings of the 5th International AFIR Colloquium.
- Hürlimann, W.* (2000a). Financial data analysis with two symmetric distributions. First Prize Gunnar Benktander ASTIN Award Competition 2000. *ASTIN Bulletin* 31, 187-211.
- Hürlimann, W.* (2000b). A Spearman multivariate distribution with fixed margins (I) theory (II) applications. 4th International Congress Insurance : Mathematics and Economics.
- Hüsler, J. and R.-D. Reiss* (1989). Maxima of normal random vectors : between independence and complete dependence. *Statistics and Probability Letters* 7, 283-286.
- Hutchinson, T.P. and C.D. Lai* (1990). *Continuous Bivariate Distributions, Emphasizing Applications*. Rumsby, Sydney, Australia.
- Joe, H.* (1997). *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability 73. Chapman and Hall, London.
- Kimeldorf, G. and A. Sampson* (1975a). Uniform representations of bivariate distributions. *Communications in Statistics* 4, 617-627.
- Kimeldorf, G. and A. Sampson* (1975b). One-parameter families of bivariate distributions with fixed marginals. *Communications in Statistics* 4, 293-301.
- Klugman, S., Panjer, H. and G. Willmot* (1998). *Loss Models. From Data to Decisions*. John Wiley, New York.
- Klugman, S.A. and R. Prasa* (1999). Fitting bivariate loss distributions with copulas. *Insurance : Mathematics and Economics* 24, 139-48.
- Kon, S.J.* (1984). Models of stock returns – a comparison. *Journal of Finance* 39, 147-65.
- Konijn, H.S.* (1959). Positive and negative dependence of two random variables. *Sankhyà* 21, 269-80.
- Lehmann, E.L.* (1966). Some concepts of dependence. *Annals of Mathematical Statistics* 37, 1137-53.
- Li, D.X.* (1999). On default correlation : a copula function approach. RiskMetrics Group. Working Paper.
- Lindskog, F.* (2000). Linear correlation estimation. Preprint RiskLab ETH Zürich, available from <http://www.math.ethz.ch/~lindskog>.
- McLeish, D.L. and C.G. Small* (1988). *The Theory and Applications of Statistical Inference Functions*. Lecture Notes in Statistics 44. Springer-Verlag, New York.
- Moore, D.* (1978). Chi-square tests. In Hogg, R. (Ed.). *Studies in Statistics*, vol. 19. Mathematical Association of America, 453-63.
- Moore, D.* (1986). Tests of chi-squared type. In D'Agostino, R. and M. Stephens (Eds.). *Goodness-of-Fit Techniques*. Marcel Dekker, New York, 63-95.
- Nelsen, R.B.* (1991). Copulas and association. In Dall'Aglio, G., Kotz, S. and G. Salinetti (Eds). *Advances in Probability Distributions with Given Marginals*. Kluwer, Dordrecht, The Netherlands, 51-74.
- Nelsen, R.B.* (1999). *An Introduction to Copulas*. Springer-Verlag.
- Praetz, P.D.* (1972). The distribution of share price changes. *Journal of Business* 45, 49-65.
- Quesenberry, C.* (1986). Some transformation methods in goodness-of-fit. In : D'Agostino, R., and M. Stephens (Eds). *Goodness-of-Fit Techniques*. Marcel Dekker, New York, 235-277.
- Resnick, S.I.* (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.

- Rüschendorf, L., Schweizer, B. and M.D. Taylor* (1996). (Ed.) Distributions with Fixed Marginals and Related Topics. Institute of Mathematical Statistics, Hayward.
- Schwartz, G.* (1978). Estimating the dimension of a model. *Annals of Statistics* 6, 461-64.
- Schweizer, B. and A. Sklar* (1983). Probabilistic Metric Spaces. North-Holland, New York.
- Sibuya, M.* (1961). Bivariate extreme statistics. *Annals of Mathematical Statistics* 11, 195-210.
- Sklar, A.* (1959). Fonctions de répartition et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris* 8, 229-31.
- Spearman, C.* (1904). The proof and measurement of association between two things. *American Journal of Psychology* 15, 72-101.
- Taylor, S.J.* (1992). *Modeling Financial Time Series* (3rd reprint). John Wiley.
- Xu, J.J.* (1996). Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data. Ph.D. thesis, Department of Statistics, University of British Columbia.