NONPARAMETRIC TESTS FOR POSITIVE QUADRANT DEPENDENCE

Michel Denuit

Institut des Sciences Actuarielles & Institut de Statistique

Université Catholique de Louvain

Olivier Scaillet

HEC Genève and FAME

Université de Genève

corresponding author:

Bd Carl Vogt, 102

CH - 1211 Genève 4, Suisse

scaillet@hec.unige.ch

++ 41 22 379 88 16

This version: March 2004 (first version : January 2001)

Abstract

We consider distributional free inference to test for positive quadrant dependence, i.e. for the probability that two variables are simultaneously small (or large) being at least as great as it would be were they dependent. Tests for its generalization to higher dimensions, namely positive orthant dependence, are also analyzed. We propose two types of testing procedures. The first procedure is based on the specification of the dependence concepts in terms of distribution functions, while the second procedure exploits the copula representation. For each specification a distance test and an intersection-union test for inequality constraints are developed for time dependent data. An empirical illustration is given for US insurance claim data, where we discuss practical implications for the design of reinsurance treaties. Another application concerns detection of positive quadrant dependence between the HFR and CSFB/Tremont market neutral hedge fund indices and the S&P500 index.

Key words and phrases: Nonparametric, Positive Quadrant Dependence, Copula, Inequality Constraint Test, Risk Management.

JEL Classification: C12, D81, G10, G21, G22.

1 Introduction

The development and analysis of quantitative models for losses in large portfolios of insurance contracts or financial assets has been an area of interest for practitioners, regulators and academics for several years. These models aim at capturing the losses due to default events or adverse movements of asset prices. In fact, most financial institutions are now routinely using risk management systems to adequately control their risks or to suitably allocate their capital. This has been driven by either internal requirements (efficient use of capital invested by shareholders, development of new business lines) or external constraints (Capital Adequacy Requirement of the Basle Committee on Banking Supervision, prudential rules imposed by European or American regulators on financial institutions). Clearly, the dependence between financial instruments materially affects risk measures and asset allocations resulting from optimal portfolio selections. The analysis of the dependence structure cannot be neglected and reveals much of the danger associated to a given position.

Unfortunately, contemporary techniques too often revolve around the use of linear correlation to describe a dependence between risks and implicitly assume normally distributed risks (mainly for mathematical convenience). But what does positive correlation really mean? In a normal world, positive correlation entails strong positive dependence notions, see Tong (1990). However, as illustrated by Embrechts, NcNeil and Straumann (2000), the dependence properties of the normal world do not hold in a non-normal world. Modern risk management calls for an understanding of stochastic dependence going beyond simple linear correlation. In that respect, dependence concepts like comonotonicity, multivariate total positivity, conditional increasingness in sequence, association and positive quadrant dependence (and its multivariate extensions) are of prime importance, and should be understood and used by practitioners.

In the management of large portfolios, the main risk is the joint occurrence of a number of default events or the simultaneous downside evolution of prices. A better knowledge of the dependence between financial assets or claims is crucial to assess the risk of loss clustering. This clustering behavior can be described by a useful concept known as positive quadrant dependence (PQD) for bivariate distributions (Lehmann (1966)) and positive orthant dependences (POD) for more than two risks. This type of dependence tells us how two, or more, random variables behave together when they are simultaneously small (or large). More precisely two random variables are PQD if the probability that they are simultaneously small is at least as great as it would be were they independent. The aim of this paper is to provide procedures relevant to testing for the presence or not of a PQD behavior in the data.

One of the main interests in this dependence structure is that it allows the risk manager to directly compare the sum of PQD random variables with the corresponding sum under the independence assumption (see the appendix for further details). The comparison is in the sense of different stochastic orderings expressing the common preferences of rational decision-makers (in the framework of the classical von Neuman-Morgenstern expected utility theory, as well as in other theories, cf. Yaari (1987)). Inferring that two claims are PQD, regardless of the strength of this dependence, immediately allows us to infer the underestimation of most insurance premiums involving a portfolio of these two claims if the independence assumption is made instead. In a financial setting the same holds true but for risk measures and derivative prices related to a portfolio of two PQD financial assets.

A related notion, namely asymptotic dependence, is empirically analyzed and tested in Poon, Rockinger and Tawn (2003). It exactly corresponds to the PQD concept but for loss probabilities, resp. loss levels, tending to zero, resp. minus infinity [1]. These authors discuss the use of asymptotic dependence concepts in a number of financial applications such as portfolio selection, risk management, Sharpe ratio targeting, hedging, option valuation and credit risk analysis. Most of their discussion remains valid in the PQD case but for less extreme risks, and we refer to their paper to get further substantial justification for the use of dependence tests in resolving interesting financial hypotheses. Note however that we prefer the PQD concept over the asymptotic independence notion for several reasons. First, two asymptotically independent variables may still exhibit a PQD behavior (cf. the notorious case of a Gaussian copula), and thus PQD should be the primary object of focus. Even risks which are far from extreme can lead to severe damages. Second, asymptotic independence is essentially a bivariate concept, and has not yet been characterized in higher dimensions. Third, current available inference for asymptotic independence has been developed in an i.i.d. context, and has not yet been rigorously justified in a time series context.

Finally let us emphasize that financial security systems are generally complex, and their outcomes usually involve several dimensions. Describing relationships among different dimensions is a basic technique for explaining the behavior of risk control mechanisms to concerned business and public policy decision-makers. In that respect, copula functions can be of great use for risk managers and actuaries. The concept of "copulas" or "copula functions" as named by Sklar (1959) originates in the context of probabilistic metric spaces. The idea behind this concept is the following: for multivariate distributions, the univariate marginals and the dependence structure can be separated and the latter may be represented by a copula. The word copula, resp. copulare, is a latin noun, resp. verb, that means "bond", resp. "to connect" or "to join". The term copula is used in grammar and logic to describe that part of a proposition which connects the subject and predicate. In statistics, it now describes the function that "joins" one-dimensional distribution functions to form multivariate ones, and may serve to characterize dependence concepts such as PQD and POD (see Nelsen (1999) and Joe (1997) for definitions and further details). Specification of PQD and POD hypotheses can thus be made in terms of distribution functions (specification in terms of loss levels) or copulas (specification in terms of probability levels), and this will lead to different inference procedures. Note that some regulators in the banking industry think in terms of probability levels in place of loss levels when assessing financial risks. Credit rating analysis is also directly linked to default probability levels and not loss levels. This explains why we develop both types of inference.

The paper is organized as follows. In Section 2, we start with the definition of PQD before presenting its multivariate extensions. In Section 3 we describe the null and alternative hypotheses we are interested in, and develop testing procedures in a time series setting. These procedures are closely related to the inference tools for traditional first order and second order stochastic dominance (and their phrasing in terms of utility functions), which also rely on distance and intersection-union tests for inequality constraints (see Davidson and Duclos (2000) and the references therein). They are of a nonparametric nature and thus avoid misspecification problems and distortions which could be associated with parametric approaches (see Fermanian and Scaillet (2003, 2004) for several examples related to copula modeling). Two empirical illustrations are given in Section 4. We first consider US insurance claim data, and provide a comparison of premiums computed under different dependence assumptions. This allows us to discuss effect of PQD on the pricing of reinsurance treaties. In a second application we examine whether there is PQD between the HFR and CSFB/Tremont market neutral hedge fund indices and the S&P500 index. Section 5 concludes. Proofs are gathered in the appendix. In the latter we also give further illustrations of the interest in positive dependence notions with the help of various stochastic inequalities. We provide some relevant examples coming from insurance and finance. We believe that the examples and results developed in the appendix help to clarify the link between the PQD concept and stochastic dominance (as well as their economic interpretation in terms of a utility function) as well as why PQD is important in a number of actuarial and financial applications.

2 Dependence notions

2.1 Positive quadrant dependence

The concept of positive quadrant dependence (PQD) is introduced in Lehmann (1966) and describes how two random variables behave together when they are simultaneously small (or large). As already mentioned, joint occurrence of large losses or very negative returns is of particular interest in risk management.

Formally, two random variables Y_1 and Y_2, or the random pair $\boldsymbol{Y} = (Y_1, Y_2)'$, are said to be positively quadrant dependent if, for all $(y_1, y_2) \in \mathbb{R}^2$,

$$P[Y_1 \leq y_1, Y_2 \leq y_2] \geq P[Y_1 \leq y_1]P[Y_2 \leq y_2]. \tag{2.1}$$

This states that two random variables are PQD if the probability that they are simultaneously small is at least as great as it would be were they independent. Of course, (2.1) is equivalent to

$$P[Y_1 > y_1, Y_2 > y_2] \geq P[Y_1 > y_1]P[Y_2 > y_2] \tag{2.2}$$

which enjoys a similar interpretation (with "small" replaced with "large") [2].

Since "positive" refers to a comparison with independence, let $\boldsymbol{Y}^{\perp}$ denote an independent version of the random vector $\boldsymbol{Y}$, that is, $\boldsymbol{Y}$ and $\boldsymbol{Y}^{\perp}$ have identical univariate marginals and $\boldsymbol{Y}^{\perp}$ has independent components. Considering (2.1)-(2.2), PQD appears as a comparison of the joint distribution of $\boldsymbol{Y}$ to that of $\boldsymbol{Y}^{\perp}$. It can thus be considered as a special case of comparisons of pairs of bivariate distributions with identical marginals. This yields the concordance order introduced by Yanagimoto and Okamoto (1969) and further studied by Tchen (1980) and Kimeldorf and Sampson (1987) (see Cebrian, Denuit and Scaillet (2004) for testing procedures). PQD is in particular satisfied when random variables are regression dependent (see Dachraoui and Dionne (2003) for definition and use of this dependence concept for optimal portfolio selection in presence of dependent risky assets).

Clearly, Y_1 and Y_2 are PQD if, and only if, $g_1(Y_1)$ and $g_2(Y_2)$ are PQD for any increasing functions g_1 and g_2. This indicates that PQD is a property of the underlying copula and is not influenced by the marginals (see Joe (1997) for a proof). Inequality (2.1) can then also be written in terms of the copula C of the two random variables, since (2.1) is equivalent to the condition that, for all $\boldsymbol{u} \in [0, 1]^2$,

$$C(u_1, u_2) \geq C^{\perp}(u_1, u_2) = u_1 u_2, \tag{2.3}$$

where $C^{\perp}$ is called the independence copula.

Remark that bivariate copulas are directly linked with scale invariant measures of association such as Kendall's τ and Spearman's ρ through:

$$\begin{aligned}\tau &= 1-4\int_0^1\int_0^1 \frac{\partial C(u_1,u_2)}{\partial u_1}\frac{\partial C(u_1,u_2)}{\partial u_2}du_1du_2,\\ \rho &= 12\int_0^1\int_0^1 C(u_1,u_2)du_1du_2-3,\end{aligned}$$

which are obviously equal to zero when $C(u_1,u_2)=C^{\perp}(u_1,u_2)$.

2.2 Positive orthant dependences

The bivariate notion of PQD has been generalized to higher dimensions in several ways, see e.g. Newman (1984). We consider here positive orthant dependences.

Positive orthant dependence offers a direct extension of PQD in three or more dimensions where orthants replace quadrants. This yields the following definitions, directly inspired from (2.1) and (2.2). A random vector $\boldsymbol{Y}=(Y_1,...,Y_n)'$ is said to be positively lower orthant dependent (PLOD, in short) when the inequalities

$$P[\boldsymbol{Y}\leq \boldsymbol{y}]\geq P[\boldsymbol{Y}^{\perp}\leq \boldsymbol{y}]=\prod_{i=1}^{n}P[Y_i\leq y_i] \tag{2.4}$$

hold for any $\boldsymbol{y}\in\mathbb{R}^n$. It is said to be positively upper orthant dependent (PUOD, in short) when the inequalities

$$P[\boldsymbol{Y}> \boldsymbol{y}]\geq P[\boldsymbol{Y}^{\perp}> \boldsymbol{y}]=\prod_{i=1}^{n}P[Y_i> y_i] \tag{2.5}$$

hold for any $\boldsymbol{y}\in\mathbb{R}^n$. Of course, (2.4) and (2.5) are no more equivalent when $n\geq 3$.

Intuitively, (2.5) means that $Y_1,Y_2,\ldots,Y_n$ are more likely to simultaneously have large values, compared with a vector of independent random variables with the same corresponding univariate marginals. Inequality (2.4) is similarly interpreted. When (2.4) and (2.5) hold together, then $\boldsymbol{Y}$ is said to be positively orthant dependent (POD, in short). In particular POD is fulfilled when variables are associated (see Milgrom and Weber (1982) for definition and use of the association concept in auction theory).

In terms of the copula C associated to the random vector $\boldsymbol{Y}$, (2.4) can be written as

$$C(\boldsymbol{u})\geq \prod_{j=1}^{n}u_j, \tag{2.6}$$

and (2.5) as

$$\overline{C}(\boldsymbol{u}) \geq \prod_{j=1}^{n}(1-u_j), \tag{2.7}$$

for all $\boldsymbol{u} \in [0,1]^n$, where $\overline{C}$ denotes the survival copula associated with C (see Nelsen (1999) for a definition).

Finally let us remark that other dependence concepts such as negative quadrant dependence (NQD) and negative orthant dependence (NOD) may also be defined by reversing the sense of one, or all inequalities in (2.1) and (2.4) (see Nelsen (1999)). Testing procedures similar to ours may easily be developed for these cases. We focus hereafter on PQD and PLOD, and not on NQD and NOD, since we believe that the former notions are the most relevant in standard risk management applications with long positions. Nevertheless, the other concepts could also be of interest for other applications as, for instance, when determining whether a risk tends to hedge another one.

3 Hypotheses testing

We consider a strictly stationary process $\{\boldsymbol{Y}_t, t \in \mathbb{Z}\}$ taking values in $\mathbb{R}^n$. The observations consist in a realization of $\{\boldsymbol{Y}_t;\ t = 1, ..., T\}$. These data may correspond to either observed individual losses on n insurance contracts, the amounts of claims reported by a given policy holder on n different guarantees in a multiline product or observed returns of n financial assets. For inference we will also need to assume that the process is strong mixing (α-mixing) with mixing coefficients α_t such that $\alpha_T = O(T^{-a})$ for some $a > 1$ as $T \to \infty$ (see Doukhan (1994) for relevant definition and examples).

We denote by $f(\boldsymbol{y})$, $F(\boldsymbol{y})$, the pdf and cdf of $\boldsymbol{Y} = (Y_1, ..., Y_n)'$ at point $\boldsymbol{y} = (y_1, ..., y_n)'$. The marginal pdf and cdf of each element Y_j at point y_j, $j = 1, ..., n$, will be written $f_j(y_j)$, and $F_j(y_j)$, respectively.

Hereafter we develop two testing methods. The first is based on a specification in terms of distribution functions (specification in terms of loss levels), while the second relies on copulas (specification in terms of probability levels).

3.1 Inference based on distribution functions

Let us start with the definition (2.4) of PLOD. Obviously we get that (2.4) can be rewritten in terms of cdfs as $F(\boldsymbol{y}) \geq \prod_{j=1}^{n} F_j(y_j)$. As in traditional stochastic dominance tests we use a version of the conditions defining PQD and PLOD on a predetermined grid, and only consider a fixed number of distinct points, say d points $\boldsymbol{y}_i = (y_{i1}, ..., y_{in})'$ in $\mathbb{R}^n$, $i = 1, ..., d$. In actuarial science, these points will cover the whole range of possible losses. The direct insurer may desire resorting to a truncated distribution when reinsurance has been bought, while the reinsurer may want to restrict its attention to the conditional distribution of excesses over a high threshold. If special attention is paid to the joint occurrence of larges losses, the grid ought to be refined in these regions.

We define $D_F^i = F(\boldsymbol{y}_i) - \prod_{j=1}^{n} F_j(y_{ij})$, and $\boldsymbol{D}_F = (D_F^1, ..., D_F^d)'$. The null hypothesis of a test for PLOD may be written as

$$H_F^0 = \{\boldsymbol{D}_F : \boldsymbol{D}_F \geq 0\},$$

and we take as alternative hypothesis:

$$H_F^1 \quad = \quad \{\boldsymbol{D}_F : \boldsymbol{D}_F \text{ unrestricted }\}.$$

To examine these hypotheses we will use the usual distance tests for inequality constraints, initiated in the multivariate one-sided hypothesis literature for positivity of the mean (Bartholomew (1959a,b)).

We may also consider a test for non-PLOD based on the null hypothesis:

$$\bar{H}_F^0 = \{\boldsymbol{D}_F : D_F^i \leq 0 \text{ for some } i\},$$

and the alternative hypothesis:

$$\bar{H}_F^1 = \{\boldsymbol{D}_F : D_F^i > 0 \text{ for all } i\}.$$

These hypotheses will be tested through intersection-union tests based on the minimum of a t-statistic.

Both testing procedures will be built from the empirical counterpart $\hat{D}_F^i$ of D_F^i obtained by substituting the empirical distributions for the unknown distributions. The joint and individual empirical distributions are given by

$$\hat{F}(\boldsymbol{y}_i) = \frac{1}{T}\sum_{t=1}^{T}\prod_{j=1}^{n}\mathbb{I}\{Y_{jt} \leq y_{ij}\}, \qquad i = 1, ..., d, \tag{3.1}$$

$$\hat{F}_j(y_{ij}) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}\{Y_{jt} \leq y_{ij}\}, \qquad i = 1, ..., d,\ j = 1, ..., n. \tag{3.2}$$

The following proposition gives the asymptotic distribution of $\hat{\boldsymbol{D}}_F$ for time dependent data.

Proposition 3.1. *The random vector $\sqrt{T}(\hat{\boldsymbol{D}}_F - \boldsymbol{D}_F)$ converges in distribution to a d-dimensional normal random variable with mean zero and covariance matrix $\boldsymbol{V}_F$ whose elements for $k, l = 1, ..., d$ are*

$$\begin{aligned} v_{F,kl} = & \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{\boldsymbol{Y}_0 \leq \boldsymbol{y}_k\}, \mathbb{I}\{\boldsymbol{Y}_t \leq \boldsymbol{y}_l\}] \\ & - \sum_{h=1}^{n}[\prod_{j=1,j\neq h}^{n} F_j(y_{kj})] \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{Y_{h0} \leq y_{kh}\}, \mathbb{I}\{\boldsymbol{Y}_t \leq \boldsymbol{y}_l\}] \\ & - \sum_{h'=1}^{n}[\prod_{j=1,j\neq h'}^{n} F_j(y_{lj})] \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{\boldsymbol{Y}_0 \leq \boldsymbol{y}_k\}, \mathbb{I}\{Y_{h't} \leq y_{lh'}\}] \\ & + \sum_{h=1}^{n}\sum_{h'=1}^{n}[\prod_{j=1,j\neq h}^{n} F_j(y_{kj})][\prod_{j=1,j\neq h'}^{n} F_j(y_{lj})] \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{Y_{h0} \leq y_{kh}\}, \mathbb{I}\{Y_{h't} \leq y_{lh'}\}]. \end{aligned}$$

In the i.i.d. case the asymptotic covariance matrix $\boldsymbol{V}_F$ can be simplified. Let us define $\boldsymbol{y}_{k\wedge l} = (y_{k1} \wedge y_{l1}, ..., y_{kn} \wedge y_{ln})'$ where $a \wedge b = \min(a, b)$. We further denote: $\boldsymbol{y}_{k\wedge_j l} = (y_{k1}, ..., y_{kj} \wedge y_{lj}, ..., y_{kn})'$, and $F_{j_1 j_2}(\boldsymbol{y}_{k\wedge l}) = P[Y_{j_1} \leq (y_{kj_1} \wedge y_{lj_1}), Y_{j_2} \leq (y_{kj_2} \wedge y_{lj_2})]$, $j_1, j_2 = 1, ..., n$, $j_1 \neq j_2$. Then the covariance matrix $\boldsymbol{V}_F$ has elements

$$v_{F,kl} = \boldsymbol{b}_k' \boldsymbol{A}_{kl} \boldsymbol{b}_l, \qquad k, l = 1, ..., d,$$

where

$$\boldsymbol{b}_i = \begin{pmatrix} 1 & -\prod_{j=1,j\neq 1}^{n} F_j(y_{ij}) & \dots & -\prod_{j=1,j\neq n}^{n} F_j(y_{ij}) \end{pmatrix}', \qquad i = 1, ..., d,$$

and $\boldsymbol{A}_{kl}$ is equal to

$$\begin{pmatrix} F(\boldsymbol{y}_{k\wedge l})-F(\boldsymbol{y}_k)F(\boldsymbol{y}_l) & F(\boldsymbol{y}_{k\wedge_1 l})-F(\boldsymbol{y}_k)F_1(y_{l1}) & \dots & F(\boldsymbol{y}_{k\wedge_n l})-F(\boldsymbol{y}_k)F_n(y_{ln}) \\ F(\boldsymbol{y}_{l\wedge_1 k})-F(\boldsymbol{y}_l)F_1(y_{k1}) & F_1(y_{k\wedge l,1})-F_1(y_{k1})F_1(y_{l1}) & \dots & F_{1n}(\boldsymbol{y}_{k\wedge l})-F_1(y_{k1})F_n(y_{ln}) \\ \vdots & \vdots & \ddots & \vdots \\ F(\boldsymbol{y}_{l\wedge_n k})-F(\boldsymbol{y}_l)F_n(y_{kn}) & F_{1n}(\boldsymbol{y}_{k\wedge l})-F_1(y_{l1})F_n(y_{kn}) & \dots & F_n(y_{k\wedge l,n})-F_n(y_{kn})F_n(y_{ln}) \end{pmatrix}.$$

A consistent estimate $\hat{\boldsymbol{V}}_F$ of $\boldsymbol{V}_F$ can be obtained by replacing the unknown distribution F by its empirical counterpart $\hat{F}$.

3.2 Inference based on copulas

Let us now proceed with the analogous quantities when we use copulas, and take d points $\boldsymbol{u}_i = (u_{i1}, ..., u_{in})'$, with $u_{ij} \in (0,1)$, $i = 1, ..., d$, $j = 1, \dots, n$. The d points correspond here to probability levels instead of return or loss levels. We may then define $D_C^i = C(\boldsymbol{u}_i) - \prod_{j=1}^n u_{ij}$, and $\boldsymbol{D}_C = (D_C^1, ..., D_C^d)'$. As in the previous lines we may consider the null hypothesis for a test for PLOD:

$$H_C^0 = \{\boldsymbol{D}_C : \boldsymbol{D}_C \geq 0\},$$

together with the alternative hypothesis:

$$H_C^1 = \{\boldsymbol{D}_C : \boldsymbol{D}_C \text{ unrestricted }\},$$

while the test for non-PLOD can be based on the null hypothesis:

$$\bar{H}_C^0 = \{\boldsymbol{D}_C : D_C^i \leq 0 \text{ for some } i\},$$

with

$$\bar{H}_C^1 = \{\boldsymbol{D}_C : D_C^i > 0 \text{ for all } i\},$$

as alternative hypothesis.

We assume hereafter that all cdfs are continuous, and that the cdf F_j of Y_{jt}, is such that the equation $F_j(y) = u_{ij}$ admits a unique solution denoted by ζ_{ij}, $i = 1, ..., d$, $j = 1, ..., n$, while $f_j(\zeta_{ij}) > 0$ at each quantile ζ_{ij}.

In view of the relationship $C(\boldsymbol{u}) = F(F_1^{-1}(u_1), ..., F_n^{-n}(u_n))$, we may think of estimating $C(\boldsymbol{u}_i) = F(\boldsymbol{\zeta}_i)$ by $\hat{C}(\boldsymbol{u}_i) = \hat{F}(\hat{\boldsymbol{\zeta}}_i)$ where $\hat{\boldsymbol{\zeta}}_i = (\hat{\zeta}_{i1}, ..., \hat{\zeta}_{in})'$ is made of the empirical univariate quantiles $\hat{\zeta}_{ij}$. The main difference when compared with (3.1) is that the levels are no longer given deterministic values, but quantiles estimated on the basis of sample information, and thus random quantities. As we will see in a moment this slightly complicates matters, but one often prefers (or is imposed, for instance, by regulators) to work with predetermined probability levels instead of loss levels.

Let us put $\boldsymbol{\zeta}_{k\wedge_j l} = (\zeta_{k1}, ..., \zeta_{kj} \wedge \zeta_{lj}, ..., \zeta_{kn})'$, $u_{k\wedge l,j} = (u_{kj} \wedge u_{lj})$, and $F_{j_1 j_2}(\boldsymbol{\zeta}_{k\wedge l}) = P[Y_{j_1} \leq (\zeta_{kj_1} \wedge \zeta_{lj_1}), Y_{j_2} \leq (\zeta_{kj_2} \wedge \zeta_{lj_2})]$, $j_1, j_2 = 1, ..., n$, $j_1 \neq j_2$. Then the following proposition gives the asymptotic distribution of $\hat{\boldsymbol{D}}_C$.

Proposition 3.2. *The random vector* $\sqrt{T}(\hat{\boldsymbol{D}}_C - \boldsymbol{D}_C)$ *converges in distribution to a d-dimensional normal random variable with mean zero and covariance matrix* $\boldsymbol{V}_C$ *whose elements for* $k, l = 1, ..., d$ *are*

$$\begin{aligned} v_{C,kl} &= \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{\boldsymbol{Y}_0 \leq \boldsymbol{\zeta}_k\}, \mathbb{I}\{\boldsymbol{Y}_t \leq \boldsymbol{\zeta}_l\}] \\ &\quad - \sum_{h=1}^{n} \frac{\frac{\partial F(\boldsymbol{\zeta}_k)}{\partial x_h}}{f_h(\zeta_{kh})} \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{Y_{h0} \leq \zeta_{kh}\}, \mathbb{I}\{\boldsymbol{Y}_t \leq \boldsymbol{\zeta}_l\}] \\ &\quad - \sum_{h'=1}^{n} \frac{\frac{\partial F(\boldsymbol{\zeta}_l)}{\partial x_{h'}}}{f_{h'}(\zeta_{lh'})} \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{\boldsymbol{Y}_0 \leq \boldsymbol{\zeta}_k\}, \mathbb{I}\{Y_{h't} \leq \zeta_{lh'}\}] \\ &\quad + \sum_{h=1}^{n} \sum_{h'=1}^{n} \frac{\frac{\partial F(\boldsymbol{\zeta}_k)}{\partial x_h}}{f_h(\zeta_{kh})} \frac{\frac{\partial F(\boldsymbol{\zeta}_l)}{\partial x_{h'}}}{f_{h'}(\zeta_{lh'})} \sum_{t\in\mathbb{Z}} Cov[\mathbb{I}\{Y_{h0} \leq \zeta_{kh}\}, \mathbb{I}\{Y_{h't} \leq \zeta_{lh'}\}]. \end{aligned}$$

As in the previous case the asymptotic covariance matrix $\boldsymbol{V}_C$ can be simplified in the i.i.d. case. We get for its elements:

$$v_{C,kl} = \boldsymbol{b}_k' \boldsymbol{A}_{kl} \boldsymbol{b}_l, \qquad k, l = 1, ..., d,$$

where

$$\boldsymbol{b}_i = \begin{pmatrix} 1 & \frac{-\frac{\partial F(\boldsymbol{\zeta}_i)}{\partial x_1}}{f_1(\zeta_{i1})} & ... & \frac{-\frac{\partial F(\boldsymbol{\zeta}_i)}{\partial x_n}}{f_n(\zeta_{in})} \end{pmatrix}', \qquad i = 1, ..., d,$$

and

$$\boldsymbol{A}_{kl} = \begin{pmatrix} F(\boldsymbol{\zeta}_{k\wedge l}) - F(\boldsymbol{\zeta}_k)F(\boldsymbol{\zeta}_l) & F(\boldsymbol{\zeta}_{k\wedge_1 l}) - F(\boldsymbol{\zeta}_k)u_{l1} & \dots & F(\boldsymbol{\zeta}_{k\wedge_n l}) - F(\boldsymbol{\zeta}_k)u_{ln} \\ F(\boldsymbol{\zeta}_{l\wedge_1 k}) - F(\boldsymbol{\zeta}_l)u_{k1} & u_{k\wedge l,1} - u_{k1}u_{l1} & \dots & F_{1n}(\boldsymbol{\zeta}_{k\wedge l}) - u_{k1}u_{ln} \\ \vdots & \vdots & \ddots & \vdots \\ F(\boldsymbol{\zeta}_{l\wedge_n k}) - F(\boldsymbol{\zeta}_l)u_{kn} & F_{1n}(\boldsymbol{\zeta}_{k\wedge l}) - u_{l1}u_{kn} & \dots & u_{k\wedge l,n} - u_{kn}u_{ln} \end{pmatrix}.$$

Note that the asymptotic covariance matrix $\boldsymbol{V}_C$ involves derivatives of F and the univariate densities f_j. These quantities may be estimated by standard kernel methods (see e.g. Scott (1992)) in order to deliver a consistent estimate $\hat{\boldsymbol{V}}_C$ of $\boldsymbol{V}_C$. For example we may take a Gaussian kernel and different bandwidth values h_j in each dimension, which leads to:

$$\begin{aligned} \frac{\partial \hat{F}(\widehat{\boldsymbol{\zeta}}_i)}{\partial x_j} &= (Th_j)^{-1} \sum_{t=1}^{T} \varphi\left(\frac{Y_{jt} - \hat{\zeta}_{ij}}{h_j}\right) \prod_{l \neq j}^{n} \Phi\left(\frac{Y_{lt} - \hat{\zeta}_{il}}{h_l}\right), \\ \hat{f}_j(\hat{\zeta}_{ij}) &= (Th_j)^{-1} \sum_{t=1}^{T} \varphi\left(\frac{Y_{jt} - \hat{\zeta}_{ij}}{h_j}\right), \end{aligned}$$

where φ and Φ denote the pdf and cdf of a standard Gaussian variable. In the empirical section of the paper, we opt for the standard choice (rule of thumb) for the bandwidths h_j, that is $1.05T^{-1/5}$ times the estimated standard deviation of Y_j.

3.3 Testing procedures

The distributional results of Propositions 3.1 and 3.2 are the building blocks of the testing procedures. The first testing procedure considers H_0^F (resp. H_0^C) against H_1^F (resp. H_1^C) and makes use of distance tests. It will be relevant when one or more components of $\hat{\boldsymbol{D}}_K$, $K = F, C$, are found to be negative (in such a case one wants to know whether this invalidates PLOD).

Let $\tilde{\boldsymbol{D}}_K$, $K = F, C$, be the solution of the constrained quadratic minimization problem:

$$\inf_{\boldsymbol{D}} T(\boldsymbol{D} - \hat{\boldsymbol{D}}_K)' \hat{\boldsymbol{V}}_K^{-1} (\boldsymbol{D} - \hat{\boldsymbol{D}}_K) \qquad s.t. \qquad \boldsymbol{D} \geq 0, \tag{3.3}$$

where $\hat{\boldsymbol{V}}_K$ is a consistent estimate of $\boldsymbol{V}_K$, and put

$$\hat{\xi}_K = T(\tilde{\boldsymbol{D}}_K - \hat{\boldsymbol{D}}_K)' \hat{\boldsymbol{V}}_K^{-1} (\tilde{\boldsymbol{D}}_K - \hat{\boldsymbol{D}}_K).$$

Roughly speaking, $\tilde{\boldsymbol{D}}_K$ is the closest point to $\hat{\boldsymbol{D}}_K$ under the null in the distance measured in the metric of $\hat{\boldsymbol{V}}_K$, and the test statistic $\hat{\xi}_K$ is the distance between $\tilde{\boldsymbol{D}}_K$ and $\hat{\boldsymbol{D}}_K$. The idea is to reject H_0^K when this distance becomes too large.

The asymptotic distribution of $\hat{\xi}_K$ under the null (see e.g. Gouriéroux, Holly and Monfort (1982), Kodde and Palm (1986), Wolak (1989a,b)) is such that for any positive x:

$$P[\hat{\xi}_K \geq x] = \sum_{i=1}^{d} P[\chi_i^2 \geq x] w(d, d-i, \hat{\boldsymbol{V}}_K) + o(1),$$

where the weight $w(d, d-i, \hat{\boldsymbol{V}}_K)$ is the probability that $\tilde{\boldsymbol{D}}_K$ has exactly $d-i$ positive elements.

Computation of the solution $\tilde{\boldsymbol{D}}_K$ can be performed by a numerical optimization routine for constrained quadratic programming problems available in most statistical software. Closed form solution for the weights are available for $d \leq 4$ (Kudo (1963)). For higher dimensions one usually relies on a simple Monte Carlo technique as advocated in Gouriéroux, Holly and Monfort (1982) (see also Wolak (1989a)). Indeed it is enough to draw a given large number of realizations of a multivariate normal with mean zero and covariance matrix $\hat{\boldsymbol{V}}_K$. Then use these realizations as $\hat{\boldsymbol{D}}_K$ in the above minimization problem (3.3), compute $\tilde{\boldsymbol{D}}_K$, and count the number of elements of the vector greater than zero. The proportion of draws such that $\tilde{\boldsymbol{D}}_K$ has exactly $d-i$ elements greater that zero gives a Monte Carlo estimate of $w(d, d-i, \hat{\boldsymbol{V}}_K)$. If one wishes to avoid this computational burden, the upper and lower bound critical values of Kodde and Palm (1986) can be adopted.

Let us now turn our attention to the second procedure aimed to test $\bar{H}_0^F$, resp. $\bar{H}_0^C$, against $\bar{H}_1^F$, resp. $\bar{H}_1^C$, and relying on the intersection-union principle. It will be used when all the components of $\hat{\boldsymbol{D}}_K$ are found to be positive. The question is then whether this suffices to ensure PLOD.

Let $\hat{\gamma}_K^i = \sqrt{T}\hat{D}_K^i / \sqrt{\hat{v}_{K,ii}}$, $K = F, C$. Then under $\bar{H}_K^0$, the limit of $P[\inf \hat{\gamma}_K^i > z_{1-\alpha}]$ will be less and exactly equal to α if $D_K^i = 0$ for a given i and $D_K^l > 0$ for $l \neq i$, while its limit is one under $\bar{H}_K^1$. Hence the test consisting of rejecting $\bar{H}_K^0$ when $\inf \hat{\gamma}_K^i$ is above the $(1-\alpha)$-quantile $z_{1-\alpha}$ of a standard normal distribution has an upper bound α on its asymptotic size (see e.g. Howes (1993), Kaur, Prakasa Rao and Singh (1994)).

Power issues are extensively discussed by Dardanoni and Forcina (1999) (see also the comments in Davidson and Duclos (2000)). First, approaches based on distance tests exploit the covariance structure, and are thus expected to achieve better power properties relative to approaches, such as ones based on t-statistics, that do not account for it. In a set of Monte Carlo experiments for standard stochastic dominance and nondominance tests, they find that, indeed, distance tests are worth the extra amount of computational work (see also Tse and Zhang (2003) for further Monte Carlo evidence). Since the form of our tests is very similar, we expect their results to hold in our setting as well. Second, it is possible that nonrejection of the null of dominance, here PLOD, by distance tests occurs along with the nonrejection of the null of nondominance, here non-PLOD, by intersection-union tests. This is due to the highly conservative nature of the latter, and will typically occur in our setting if $\hat{\boldsymbol{D}}_K$ is close enough to zero for a number of coordinates. Finally, Barrett and Donald (2003) report that distance tests for stochastic dominance do not seem to suffer from size distortion problems in samples as small as $T = 50$ [3]. They also confirm the nice behavior of distance tests in terms of power.

4 Empirical illustrations

This section illustrates the implementation of the testing procedures described in the previous section. We provide two empirical applications; the first concerns the detection of PQD in US insurance claim data, and its effect on premium valuation, the second is devoted to hedge fund and stock index data.

4.1 US Losses and ALAE's

Various processes in casualty insurance involve correlated pairs of variables. A prominent example is the loss and allocated loss adjustment expenses (ALAE, in short) on a single claim. Here ALAE are type of insurance company expenses that are specifically attributable to the settlement of individual claims such as lawyers' fees and claims investigation expenses. The joint modeling in parametric settings of those two variables has been examined by Frees and Valdez (1998), and Klugman and Parsa (1999). The data used in these empirical studies were collected by the US Insurance Services Office, and comprise general liability claims randomly chosen from late settlement lags. Frees and Valdez (1998) choose the Pareto distribution to model the margins and select Gumbel and Frank copulas (on the basis of a graphical procedure suitable for Archimedean copulas). Both models express PQD by their estimated parameter values. Klugman and Parsa (1999) opt for the Inverse Paralogistic for the losses and for the Inverse Burr for ALAE's. They use Frank's copula. Again, the estimated value of the dependence parameter entails PQD for losses and ALAE's. In the following we rely on a nonparametric approach to assess PQD. This assessment has many implications in insurance, for example, for the computation of reinsurance premiums (where the sharing of expenses between the ceding company and the reinsurer has to be decided on) and for the determination of the expense level for a given loss level (for reserving an appropriate amount to cover future settlement expenses).

The data consist in $T = 1,466$ uncensored observed values of the pair (LOSS,ALAE). Some summary statistics are gathered in Table 1. The estimated values for Pearson's r, Kendall's τ and Spearman's ρ are 0.3805, 0.3067 and 0.4437, respectively. All of them are judged significantly positive at 1%. Because some very high values of the variables are contained in the data set, we will work on a logarithmic scale. This will not alter the results of our analysis since (LOSS,ALAE) PQD $\Leftrightarrow$ (log(LOSS),log(ALAE)) PQD. Note that the transformation of the margins results in a new Pearson's r (linear correlation coefficient) of 0.4313, while Kendall's and Spearman's values are left unchanged. These are not affected by strictly increasing transformation of the variables.

please insert Table 1

Figure 1 shows the kernel estimator of the bivariate pdf of the pair (log(LOSS),log(ALAE)), together with its contour plot. This estimation relies on a product of Gaussian kernels and bandwidth values selected by the standard rule of thumb (Scott (1992)). The graphs obviously suggest strong positive dependence between both variables.

In order to test whether PQD holds on the whole observation domain, we take 49 points coming from the equally spaced grid $\{6, 7, \ldots, 12\} \times \{6, 7, \ldots, 12\}$. This leads to a vector $\hat{\boldsymbol{D}}_F$ with only one negative component -0.00015. We wish to check whether this invalidates PQD or not. The distance between $\tilde{\boldsymbol{D}}_F$ and $\hat{\boldsymbol{D}}_F$ is found to be 2.4×10^{-7}. Lower bounds on the critical values obtained by Kodde and Palm (1986) are given in Table 2 for different levels α. Note that they do not depend on the grid size d. In view of these bounds we do not reject the null of PQD at any reasonable confidence level.

please insert Table 2

Let us now consider a positive dependence, but only in the upper tails[4]. We take the grid $\{10, 10.3, 10.6, 11, 11.3, 11.6, 12\} \times \{10, 10.3, 10.6, 11, 11.3, 11.6, 12\}$. All 49 components of $\hat{\boldsymbol{D}}_F$ are strictly positive, which means that H_0^F is automatically not rejected. The intersection-union test may then be used to know whether the data exhibit PQD. We get $\min \hat{\gamma}_F^i = 0.10081$ which does not allow us to reject $\overline{H}_0^F$ in favor of PQD. This non rejection is due to the closeness of $\hat{\boldsymbol{D}}_F$ to zero for a large number of coordinates. This point has already been discussed at the end of Section 3.

Let us now turn to copula based tests. For the $\boldsymbol{u}_i$'s, we take the 81 deciles of the grid $\{0.1, 0.2, \ldots, 0.9\} \times \{0.1, 0.2, \ldots, 0.9\}$. All components of the corresponding $\hat{\boldsymbol{D}}_C$ are positive, so that H_0^C cannot be rejected. For the intersection-union test, we obtain $\min \hat{\gamma}_C^i = 0.94894$ which does not allow us to reject $\overline{H}_0^C$ in favor of $\overline{H}_1^C$. If we focus on the tails, taking the high percentiles in $\{0.91, 0.92, \ldots, 0.99\} \times \{0.91, 0.92, \ldots, 0.99\}$, we get that all components of $\hat{\boldsymbol{D}}_C$ are again positive resulting in the non-rejection of H_0^C. Further, $\min \hat{\gamma}_C^i = 0.6983$, so that $\overline{H}_0^C$ is not rejected, either.

It has to be pointed out that the choice of the bandwidth has very little impact on the values of the test statistics. They have been computed with half, twice and three times the standard choice, and this has only resulted in small variations.

Let us now discuss the practical implications of the presence of PQD in the previous data. We look at the impact on premium valuation in reinsurance treaties. We consider a reinsurance treaty on a policy with unlimited liability and insurer's retention R. Assuming a prorata sharing of expenses, the reinsurer's payment for a given realization of (LOSS,ALAE) is described by the function

$$g(\text{LOSS,ALAE}) = \begin{cases} 0 \text{ if LOSS} \leq R, \\ \text{LOSS} - R + \frac{\text{LOSS-}R}{\text{LOSS}}\text{ALAE if LOSS} > R. \end{cases}$$

The pure premium relating to this reinsurance treaty is

$$\pi = E[g(\text{LOSS,ALAE})].$$

The results in Table 3 provide the premiums the reinsurer would have assessed to cover costs of losses and expenses according to various insurer's retention. Three situations have been considered:

1. the first one assumes independence, i.e.
$$\hat{\pi} = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{t'=1}^{T} g(\text{LOSS}_t, \text{ALAE}_{t'});$$
2. the second one takes into account the dependence expressed by the data, i.e.
$$\hat{\pi} = \frac{1}{T}\sum_{t=1}^{T} g(\text{LOSS}_t, \text{ALAE}_t);$$
3. and the last one resorts to the classical comonotonic approximation for (LOSS,ALAE), i.e.
$$\hat{\pi} = \frac{1}{T}\sum_{t=1}^{T} g\left(\text{LOSS}_t, \hat{F}_2^{-1}\left(\hat{F}_1(\text{LOSS}_t)\right)\right).$$

We see that substantial mispricing could result from the independence hypothesis (as predicted by the theory developed in the appendix), while the comononotic approximation is too conservative. We see that independence generates lower premiums than those suggested by the data, which themselves are smaller than those based on the comonotonic assumption (as they are theoretically bound to be under PQD).

please insert Table 3

4.2 Hedge fund and stock indices

This second empirical illustration concerns data on the HFR and CSFB/Tremont market neutral hedge fund indices and the S&P500 index. We consider returns recorded monthly from 31/01/1994 to 31/10/2003, i.e. 118 observations. Figure 2 shows bivariate scatterplots of the data. Table 4 gathers the summary statistics for these data. The estimated values for Pearson's r, Kendall's τ and Spearman's ρ are 0.2786, 0.1889 and 0.2743 for the pair (HFR,CSFB/Tremont). All of them are judged significantly positive at 1%. We can already observe that the amount of dependence between the two indices is not as high as could have been expected. In fact their respective composition does not span exactly the same hedge fund universe, and this explains the observed differences. For the pair (CSFB/Tremont,S&P500), we get r=0.3952, τ=0.2543 and ρ=0.3728, which are again significant at 1%. On the contrary we get r=0.1406, τ=0.0630 and ρ=0.0875 for the pair (HFR,S&P500), and these values cannot be taken as significantly different from zero at 1%. This second index seems to be more "market neutral" than the first one, and is often preferred by practitioners as a benchmark for such strategies. We can remark that both indices exhibit a low linear dependence (low beta) with the S&P500 index because of the "neutral" target of the hedge funds composing them. However these indices may exhibit a strong nonlinear dependence. If so, this should be revealed by a PQD behavior. Figure 3 shows the three series, while Figure 4 displays the estimated autocorrelation coefficients for various lags. It can be seen that the monthly returns have a rather low degree of autocorrelation.

please insert Table 4

please insert Figure 2

please insert Figure 3

please insert Figure 4

We only provide the results for the copula based tests since the regulatory environment for risk management in banks is specified in terms of probability levels instead of loss levels. For the $\boldsymbol{u}_i$'s, because of the small amount of data points, we take the grid $\{0.2, 0.4, 0.6, 0.8\} \times \{0.2, 0.4, 0.6, 0.8\}$. All components of the corresponding $\hat{\boldsymbol{D}}_C$ are positive for the pairs (HFR,CSFB/Tremont) and (CSFB/Tremont,S&P500) (but not for the pair (HFR,S&P500)), so that H_0^C cannot be rejected in these two cases. For the intersection-union test, we obtain $\min \hat{\gamma}_C^i = 0.2626$ for the couple (HFR,CSFB/Tremont) and $\min \hat{\gamma}_C^i = 0.6941$ for the pair (CSFB/Tremont,S&P500) so that we do not reject $\overline{H}_0^C$ in favor of $\overline{H}_1^C$ either. Some of the components of $\hat{\boldsymbol{D}}_C$ are negative for the pair (HFR,S&P500). Therefore, we compute $\hat{\xi}_C = 0.0025$ and we do not reject the null hypothesis of PQD.

Note that the PQD behavior directly implies (cf. the theoretical results of the appendix) that the price of a call option on a basket made of one of the market neutral hedge fund index and the S&P500 index will be underestimated if we use a zero correlation as input (independence) in the Black-Scholes model (even if the assumption of normally distributed returns is true for the margins).

5 Concluding remarks

In this paper we have analyzed simple distributional free inference for positive quadrant and positive lower orthant dependences. The various testing procedures have proven to be empirically relevant to the analysis of dependences among insurance and financial data. In particular they suggest the strong PQD nature of these data. Hence they complement ideally the existing battery of inference tools dedicated to joint risk analysis. They should further help to achieve a better understanding and design of insurance contracts as well as option contracts in terms of premium valuation.

APPENDIX

A Proof of Proposition 3.1

Let us consider the second term $\prod_{j=1}^{n} \hat{F}_j(y_{ij})$ of $\hat{D}_F^i$. It can be approximated by

$$\prod_{j=1}^{n} \hat{F}_j(y_{ij}) = \prod_{j=1}^{n} F_j(y_{ij}) + \sum_{h=1}^{n} \left\{ \prod_{j=1, j\neq h}^{n} \bar{F}_j(y_{ij}) \right\} (\hat{F}_h(y_{ih}) - F_h(y_{ih})) + o_p(T^{-1/2}),$$

where $\bar{F}_j(y_{ij})$ is a mean value located between $\hat{F}_j(y_{ij})$ and $F_j(y_{ij})$. Since we know that the empirical process $\sqrt{T}(\hat{F} - F)$ tends weakly to a centered Gaussian process $\mathbb{G}$ in the space of a.s. bounded function of $\mathbb{R}^n$ (see e.g. Rio (2000)), and that the covariance function of $\mathbb{G}$ is given by

$$\text{Cov}[\mathbb{G}(\boldsymbol{y}_k), \mathbb{G}(\boldsymbol{y}_l)] = \sum_{t\in\mathbb{Z}} \text{Cov}[\mathbb{I}\{\boldsymbol{Y}_0 \leq \boldsymbol{y}_k\}, \mathbb{I}\{\boldsymbol{Y}_t \leq \boldsymbol{y}_l\}],$$

the stated result follows after computation of all covariance terms.

B Proof of Proposition 3.2

Let $M = \{\mathbb{I}\{\,\cdot \leq x_1\}...\mathbb{I}\{\,\cdot \leq x_n\} : x_j \in \mathbb{R}, j = 1, ..., n\}$. Since M satisfies Pollard's entropy condition for some finite constant taken as envelope, the sequence

$$\left\{ \hat{F}(\boldsymbol{x}) = T^{-1} \sum_{t=1}^{T} \prod_{j=1}^{n} \mathbb{I}\{Y_{jt} \leq x_j\} : T \geq 1 \right\}$$

is stochastically differentiable at $\boldsymbol{\zeta}_i$ with random derivative $(d \times 1)$-vector $D\hat{F}(\boldsymbol{\zeta}_i)$ (see e.g. Pollard (1985), Andrews (1989,1999) for definition, use and check of stochastic differentiability). It means that we have the approximation:

$$\hat{F}(\hat{\boldsymbol{\zeta}}_i) = \hat{F}(\boldsymbol{\zeta}_i) + D\hat{F}(\bar{\boldsymbol{\zeta}}_i)'(\hat{\boldsymbol{\zeta}}_i - \boldsymbol{\zeta}_i) + o_p(T^{-1/2}),$$

where $\bar{\boldsymbol{\zeta}}_i$ is a mean value located between $\hat{\boldsymbol{\zeta}}_i$ and $\boldsymbol{\zeta}_i$.
Similarly we get the approximations:

$$\hat{F}_j(\hat{\zeta}_{ij}) = \hat{F}_j(\zeta_{ij}) + D\hat{F}_j(\bar{\zeta}_{ij})(\hat{\zeta}_{ij} - \zeta_{ij}) + o_p(T^{-1/2}).$$

Combining these approximations and using $F_j(\zeta_{ij}) = u_{ij} = \hat{F}_j(\hat{\zeta}_{ij})$ leads to

$$\hat{F}(\hat{\boldsymbol{\zeta}}_i) = \hat{F}(\boldsymbol{\zeta}_i) - D\hat{F}(\bar{\boldsymbol{\zeta}}_i)'\text{diag}\, S_i + o_p(T^{-1/2}),$$

where S_i is the stack of $(\hat{F}_j(\zeta_{ij}) - u_{ij})/D\hat{F}_j(\bar{\zeta}_{ij})$, $j = 1, ..., n$, and diag S_i is the diagonal matrix built from this stack.

Using the convergence in probability of $D_j\hat{F}(\bar{\boldsymbol{\zeta}}_i)$ to $\partial F(\boldsymbol{\zeta}_i)/\partial x_j$, $j = 1, ..., n$, and $D_j\hat{F}_j(\bar{\zeta}_{ij})$ to $f_j(\zeta_{ij})$, we may deduce the given result from the behaviour of the empirical process $\sqrt{T}(\hat{F} - F)$ at the limit and computations of all covariance terms.

C Applications of positive dependence notions

C.1 PQD

Let us define the following utility classes. Let U_1 contain all non-decreasing utilities $u : \mathbb{R} \to \mathbb{R}$. Let U_2 be the restriction of U_1 to its concave elements. More generally, for $k \geq 3$, let U_k be $(k-1)$ times continuously differentiable utility functions u such that $\lim_{x\to+\infty} u(x) \equiv u(+\infty)$ is finite, $\lim_{x\to+\infty} u^{(j)}(x) = 0$ for $j = 1, \ldots, k-1$ and $(-1)^{k-1}u^{(k-1)}$ is non-decreasing.

In what follows, we assume that decision-makers maximize a von Neumann-Morgenstern expected utility (but we mention that results involving U_1 and U_2 still hold in dual theories for choice under risk, see e.g. Denuit, Dhaene and Van Wouwe (1999) for further information).

Let Y_1 and Y_2 be two random variables such that $Eu(Y_1) \leq Eu(Y_2)$ holds for all $u \in U_1$ (resp. $u \in U_2$), provided the expectations exist. Then Y_1 is said to be smaller than Y_2 in the stochastic dominance (resp. increasing concave order), denoted as $Y_1 \preceq_{\text{d}} Y_2$ (resp. $Y_1 \preceq_{\text{icv}} Y_2$). From the very definitions of $\preceq_{\text{d}}$ and $\preceq_{\text{icv}}$, we see that these stochastic orderings express the common preferences of the classes of profit-seeking decision-makers, and of profit-seeking risk-averters, respectively. This provides an intuitive meaning to rankings in the $\preceq_{\text{d}}$- or $\preceq_{\text{icv}}$-sense.

If $Y_1 \preceq_{\text{icv}} Y_2$ and $EY_1 = EY_2$, then we write $Y_1 \preceq_{\text{cv}} Y_2$. In this case $Eu(Y_1) \leq Eu(Y_2)$ for all the concave utilities u, so that Y_2 is preferred over Y_1 by all risk-averters. Furthermore, if $Eu(Y_1) \leq Eu(Y_2)$ for all $u \in U_k$, provided the expectations exist, then Y_1 is said to be smaller than Y_2 in the k-increasing concave order, denoted as $Y_1 \preceq_{k-\text{icv}} Y_2$. By convention we assume that $\preceq_{k-\text{icv}}$ reduces to $\preceq_{\text{icv}}$ and $\preceq_{\text{d}}$ for $k = 2$ and $k = 1$, respectively.

For a more detailed exposition of stochastic orderings, see e.g. the review papers by Kroll and Levy (1980) and Levy (1992), the classified bibliography by Mosler and Scarsini (1993) and the books by Shaked and Shanthikumar (1994) and Mari and Kotz (2001). For a rigorous treatment of $\preceq_{k-\text{icv}}$, see Rolski (1976) and Fishburn (1976).

As already mentioned in the introduction, one of the main interests in PQD arises from the comparison with random pairs with identical marginals but independent components. This comes from the following result for which we provide a short proof. It is a straightforward adaptation of the result in Dhaene and Goovaerts (1996) established in the convex actuarial setting.

Proposition C.1. *If Y_1 and Y_2 are PQD, then $Y_1+Y_2 \preceq_{cv} Y_1^{\perp}+Y_2^{\perp}$, where $Y_1^{\perp}+Y_2^{\perp}$ is the sum of the components of the independent version $\mathbf{Y}^{\perp}$ of $\mathbf{Y}$ defined in* (2.1).

Proof: Let us first note that

$$\int_{-\infty}^{x} P[Y_1+Y_2 \leq t]dt = \Big[tP[Y_1+Y_2 \leq t]\Big]_{-\infty}^{x} - \int_{-\infty}^{x} tdP[Y_1+Y_2 \leq t] = E(x-Y_1-Y_2)_+.$$

So, we want to show that the inequality $E(x - Y_1 - Y_2)_+ \geq E(x - Y_1^{\perp} - Y_2^{\perp})_+$ holds for any real constant x. Now, let us express $E(x - Y_1 - Y_2)_+$ in terms of the joint cdf of $\boldsymbol{Y}$. Note that

$$\int_{-\infty}^{x} \mathbb{I}[y_1 \leq t, y_2 \leq x - t]dt = \int_{-\infty}^{x} \mathbb{I}[y_1 \leq t \leq x - y_2]dt = (x - y_1 - y_2)_+$$

whence it follows that

$$E(x - Y_1 - Y_2)_+ = \int_{-\infty}^{x} P[Y_1 \leq t, Y_2 \leq x - t]dt.$$

Finally,

$$\begin{aligned} &E(x - Y_1 - Y_2)_+ - E(x - Y_1^{\perp} - Y_2^{\perp})_+ \\ &\quad = \int_{-\infty}^{x} \Big\{P[Y_1 \leq t, Y_2 \leq x - t] - P[Y_1 \leq t]P[Y_2 \leq x - t]\Big\}dt, \end{aligned}$$

where the integrand $\{\ldots\}$ is non-negative provided Y_1 and Y_2 are PQD, which ends the proof. Q.E.D.

This means that when PQD holds, every risk-averter agrees to say that $Y_1 + Y_2$ is less favorable than the corresponding sum under independence. Consequently, most insurance premiums and risk measures will be larger for $X_1 + X_2$ than for $X_1^{\perp} + X_2^{\perp}$ (since the principles used to calculate such quantities are in accordance with the common preferences of risk-averters). For instance, since the function $x \mapsto -(x-\kappa)_+$, with $(\cdot)_+ = \max\{0, \cdot\}$, is concave for any $\kappa \in \mathbb{R}$, the inequality $E(Y_1^{\perp} + Y_2^{\perp} - \kappa)_+ \leq E(Y_1 + Y_2 - \kappa)_+$ holds true for all κ. In actuarial science the quantity $E(Y_1 + Y_2 - \kappa)_+$ is referred to as the stop-loss premium related to the portfolio $Y_1 + Y_2$ (κ is called the deductible). In finance, when appropriately discounted, it can be regarded as the price of a basket option with $Y_1 + Y_2$ as underlying asset portfolio and κ as strike price. The convenient assumption of independence may thus lead to serious underpricing of insurance premiums or option prices. This type of underpricing has been illustrated in the empirical section of the paper.

Let us now provide an application of PQD in life insurance. Standard actuarial theory of multiple life insurance traditionally postulates the independence for the remaining lifetimes in order to evaluate the amount of premium relating to an insurance contract involving multiple lives. Nevertheless, this hypothesis obviously relies on computational convenience rather than realism. A fine example of possible dependence among insured persons is certainly a contract issued to a married couple. In such a case, the actuary has to wonder whether the independence assumption is reasonable and whether it would not be wiser to build an appropriate price list incorporating possible effects of a dependence among time-until-death random variables.

Specifically, let T_{x_1} (resp. T_{x_2}) be the husband's (resp. wife's) lifetime, where x_1 (resp. x_2) stands for the age of the husband (resp. wife) at the start of the contract. In light of clinical studies, the PQD assumption for T_{x_1} and T_{x_1} seems reasonable. This has been empirically investigated using official Belgian statistics by Denuit and Cornet (1999) in a Markovian parametric setting. Of course, the statistical tests developed in this paper are useful in that respect, since they avoid the parametric assumption often made in actuarial studies, namely a Gompertz-Makeham distribution for the remaining lifetimes.

For insurance policies sold to married couples, PQD for T_{x_1} and T_{x_2} allows the actuary to know whether the independence assumption generates implicit safety loading or, on the contrary, leads to insufficient premium amounts. Indeed, this simply comes from the fact that the PQD assumption for T_{x_1} and T_{x_2} ensures that

$$\min\{T_{x_1}^{\perp}, T_{x_2}^{\perp}\} \preceq_{\mathrm{d}} \min\{T_{x_1}, T_{x_2}\} \text{ and } \max\{T_{x_1}, T_{x_2}\} \preceq_{\mathrm{d}} \max\{T_{x_1}^{\perp}, T_{x_2}^{\perp}\}$$

which readily follow from (2.1)-(2.2). Now, let us consider annuities (i.e. contractual guarantees that promise to provide periodic income over the lifetimes of individuals). The n-year last-survivor (resp. joint-life) annuity pays \$ 1 at the end of the years $1, 2, \ldots, n$ as long as either spouse survives (resp. both spouses survive). The net present value of the insurer's payments are obviously increasing functions of $\max\{T_{x_1}, T_{x_2}\}$ for the last-survivor annuity and of $\min\{T_{x_1}, T_{x_2}\}$ for the joint-life annuity. The net single premium corresponding to the last-survivor (resp. joint-life) annuity is denoted $a_{\overline{(x_1x_2)};\overline{n}|}$ (resp. $a_{(x_1x_2);\overline{n}|}$); it is simply the mathematical expectation of net present value of the insurer's payments (see Gerber (1995) for further details on actuarial notations and concepts). Let us denote as $a^{\perp}_{\overline{(x_1x_2)};\overline{n}|}$ and $a^{\perp}_{(x_1x_2);\overline{n}|}$ the corresponding premiums computed on the basis of the independence assumption for the remaining lifetimes. In case T_{x_1} and T_{x_2} are PQD then $a_{\overline{(x_1x_2)};\overline{n}|} \leq a^{\perp}_{\overline{(x_1x_2)};\overline{n}|}$ and $a_{(x_1x_2);\overline{n}|} \geq a^{\perp}_{(x_1x_2);\overline{n}|}$ hold true. Similar conclusions can be obtained for most standard life insurance contracts making the PQD assumption of paramount importance. This has been pointed out by Norberg (1989) and further analyzed by Denuit and Cornet (1999).

In a banking context, one area which is very close to insurance risk is operational risk. Knowing that lifetimes on two different types of operational risk (for example, fraud and IT crashes) are PQD may help to decide whether reserve funds built as if they were independent are high enough to absorb operational losses.

C.2 PLOD

Proposition C.1 no longer holds if PLOD is substituted for PQD. Rather, the following result holds true.

Proposition C.2. *Provided* $\boldsymbol{Y}$ *is PLOD, the stochastic inequality* $\sum_{i=1}^n Y_i \preceq_{n-icv} \sum_{i=1}^n Y_i^\perp$ *holds.*

Proof: Let $u \in U_n$. Then, invoking integration by parts yields

$$\begin{aligned} Eu\left(\sum_{i=1}^n Y_i\right) &= \int \ldots \int_{\boldsymbol{y}\in\mathbb{R}^n} u\left(\sum_{i=1}^n y_i\right) dP[\boldsymbol{Y} \leq \boldsymbol{y}] \\ &= u(+\infty) + (-1)^n \int \ldots \int_{\boldsymbol{y}\in\mathbb{R}^n} P[\boldsymbol{Y} \leq \boldsymbol{y}] du^{(n-1)}\left(\sum_{i=1}^n y_i\right) dy_1 \ldots dy_{n-1}. \end{aligned}$$

Provided $\boldsymbol{Y}$ is PLOD, we get

$$Eu\left(\sum_{i=1}^n Y_i^\perp\right) - Eu\left(\sum_{i=1}^n Y_i\right)$$

$$= (-1)^n \int \ldots \int_{\boldsymbol{y}\in\mathbb{R}^n} \Big\{P[\boldsymbol{Y}^\perp \leq \boldsymbol{y}] - P[\boldsymbol{Y} \leq \boldsymbol{y}]\Big\} du^{(n-1)}\left(\sum_{i=1}^n y_i\right) dy_1 \ldots dy_{n-1} \geq 0,$$

for any $u \in \mathbb{U}_n$, whence the announced result follows. Q.E.D.

Comparing Propositions C.1 and C.2, we see that $\preceq_{\text{cv}}$ is replaced with $\preceq_{n-\text{icv}}$ in dimension n. Besides, as can be seen from Proposition C.2, we only get weaker orderings in higher dimensions. To get $\preceq_{\text{cv}}$ as in Proposition C.1, we need another dependence notion called positive cumulative dependence (PCD in short) which is defined as follows: the random variables $Y_1, Y_2, \ldots, Y_n$ are PCD if the random couples $(\sum_{i=1}^{j-1} Y_i, Y_j)$ are PQD for $j = 2, 3, \ldots, n$.

The following result is inspired from Denuit, Dhaene and Ribas (2001).

Proposition C.3. *Provided* $\boldsymbol{Y}$ *is PCD, the stochastic inequality* $\sum_{i=1}^n Y_i \preceq_{cv} \sum_{i=1}^n Y_i^\perp$ *holds.*

Proof: We proceed by recurrence. From Proposition C.1, we know that provided $\boldsymbol{Y}$ is PCD,

$$Y_1 + \ldots + Y_k + Y_{k+1} = (Y_1 + \ldots + Y_k) + Y_{k+1} \preceq_{\text{cv}} (Y_1 + \ldots + Y_k)^\perp + Y_{k+1}^\perp.$$

Next, by induction hypothesis,

$$Y_1 + \ldots + Y_k \preceq_{\text{cv}} Y_1^\perp + \ldots + Y_k^\perp.$$

Since the concave order is closed under convolution (i.e. given three independent random variables A, B and C, $A \preceq_{\text{cv}} B \Rightarrow A + C \preceq_{\text{cv}} B + C$; see e.g. Shaked and Shanthikumar (1994) for details) we then have that

$$(Y_1 + \ldots + Y_k)^\perp + Y_{k+1}^\perp \preceq_{\text{cv}} Y_1^\perp + \ldots + Y_k^\perp + Y_{k+1}^\perp.$$

Now, invoking the transitivity of the $\preceq_{\text{cv}}$ order relation yields

$$Y_1 + \ldots + Y_k + Y_{k+1} \preceq_{\text{cv}} Y_1^\perp + \ldots + Y_k^\perp + Y_{k+1}^\perp,$$

which yields the stated result. Q.E.D.

From (2.4) and (2.5), it is easy to get the following result that reinforces a stochastic inequality obtained by Baccelli and Makowski (1989).

Proposition C.4. *Let S be a subset of $\{1, 2, \ldots, n\}$. Provided $\mathbf{Y}$ is POD, the stochastic inequalities $\min_{i \in S} Y_i^\perp \preceq_d \min_{i \in S} Y_i$ and $\max_{i \in S} Y_i \preceq_d \max_{i \in S} Y_i^\perp$ both hold.*

Let us illustrate the importance of Proposition C.4 in life insurance. Consider n individuals aged $x_1, x_2, \ldots, x_n$, respectively, with remaining lifetimes T_{x_1}, T_{x_2}, $\ldots$, T_{x_n}, respectively. The joint life status $(x_1, x_2, \ldots, x_n)$ exists as long as all individual statuses exist. This status has remaining lifetime:

$$T_{(x_1, x_2, \ldots, x_n)} = \min\{T_{x_1}, T_{x_2}, \ldots, T_{x_n}\}.$$

The last survivor status $\overline{(x_1, x_2, \ldots, x_n)}$ exists as long as at least one of the individual status is alive. Its remaining lifetime is given by

$$T_{\overline{(x_1, x_2, \ldots, x_n)}} = \max\{T_{x_1}, T_{x_2}, \ldots, T_{x_n}\}.$$

Let us now assume that $\mathbf{T} = (T_{x_1}, T_{x_2}, \ldots, T_{x_n})$ is POD. Let us also introduce the following straightforward notation:

$$T_{(x_1, x_2, \ldots, x_n)}^\perp = \min\left\{T_{x_1}^\perp, T_{x_2}^\perp, \ldots, T_{x_n}^\perp\right\}$$

and

$$T^{\perp}_{\overline{(x_1,x_2,\ldots,x_n)}} = \max\left\{T^{\perp}_{x_1}, T^{\perp}_{x_2}, \ldots, T^{\perp}_{x_n}\right\}.$$

From Proposition C.4, it follows that

$$T^{\perp}_{(x_1,x_2,\ldots,x_n)} \preceq_{\mathrm{d}} T_{(x_1,x_2,\ldots,x_n)} \text{ and } T_{\overline{(x_1,x_2,\ldots,x_n)}} \preceq_{\mathrm{d}} T^{\perp}_{\overline{(x_1,x_2,\ldots,x_n)}},$$

which in turn implies that

$$a^{\perp}_{(x_1,x_2,\ldots,x_n);\overline{n|}} \leq a_{(x_1,x_2,\ldots,x_n);\overline{n|}} \text{ and } a_{\overline{(x_1,x_2,\ldots,x_n)};\overline{n|}} \leq a^{\perp}_{\overline{(x_1,x_2,\ldots,x_n)};\overline{n|}},$$

where the superscript "$\perp$" is used to indicate that the annuity is based on $T^{\perp}_{(x_1,x_2,\ldots,x_n)}$ or $T^{\perp}_{\overline{(x_1,x_2,\ldots,x_n)}}$. This means that for POD remaining lifetimes, the independence assumption (while leaving the marginal cdfs unchanged) leads to an underestimation of the net single premium (and reserves) of a joint life annuity. The opposite conclusion holds for the last survivor annuity. Similar conclusions can be drawn for endowment and whole life insurance.

Finally let us stress that similar contracts have recently been proposed in finance in the context of over-the-counter credit derivatives. For example the writer of the so-called "first-to-default" contract have to pay to the buyer a given amount of dollars contingent to the first time of an observed default among a given set of names. In that setting assuming independence between time-until-default will also lead to price underestimation.

Notes

[1]It is worth mentioning that we depart from the actuarial literature by assigning a negative sign to losses in this paper. This is in line with the agreement in force in finance for asset returns.

[2]Remark that the inequality (2.2) can also be restated in terms of a conditional probability: $P[Y_1 > y_1 | Y_2 > y_2] \geq P[Y_1 > y_1]$, as in the definition of asymptotic independence.

[3]Note that we usually face large sample sizes in finance and insurance, and thus we do not expect size distortion to be a major issue.

[4]In other applications where the number of observations available in tails is small because of limited data sets, one should be cautious about performance of tests focusing on these regions.

Acknowledgements

The authors wish to express their deep gratitude to René Garcia, the Associate Editor and the two anonymous referees who have corrected some mistakes in the original manuscript and whose numerous remarks and suggestions have considerably improved the text.

We would like to thank Professors Frees and Valdez for kindly providing the Loss-ALAE data, which were collected by the US Insurance Services Office (ISO). We are grateful to Professor Irène Gijbels for her comments as well as to participants at the Bachelier world congress 2002, LAMES 2002, Journées ASU 2002, and seminars at CREST, HEC Paris, HEC Lausanne, LSE, GREMAQ, Zurich University, for their remarks. We also thank Oana Purcaru for checking the Splus programs.

The second author receives support by the Swiss National Science Foundation through the National Center of Competence: Financial Valuation and Risk Management. Part of this research was done when he was visiting THEMA and IRES.

References

[1] Andrews, D. (1994). "Empirical Process Methods in Econometrics." Handbook of Econometrics, Vol. IV, eds R. Engle and D. McFadden, 2247-2294.

[2] Andrews, D. (1999). "Estimation when a Parameter is on a Boundary." Econometrica 67, 1341-1384.

[3] Baccelli, F., and A. Makowski. (1989). "Multidimensional Stochastic Ordering and Associated Random Variables." Operations Research 37, 478-487.

[4] Barrett, G., and S. Donald. (2003). "Consistent Tests for Stochastic Dominance." Econometrica 71, 71-104.

[5] Bartholomew, D. (1959a). "A Test of Homogeneity for Ordered Alternatives, I." Biometrika 46, 36-48.

[6] Bartholomew, D. (1959b). "A Test of Homogeneity for Ordered Alternatives, II." Biometrika 46, 328-335.

[7] Beach, C., and R. Davidson. (1983). "Distribution-Free Statistical Inference with Lorenz Curves and Income Shares." Review of Economic Studies 50, 723-735.

[8] Cebrian, A., M. Denuit, and O. Scaillet. (2004). "Testing for Concordance Ordering." FAME DP 41; forthcoming in ASTIN Bulletin.

[9] Dachraoui, K., and G. Dionne. (2003). "Conditions Ensuring the Separability of Asset Demand for All Risk Averse Investors." Working Paper, HEC Montréal.

[10] Dardanoni, V., and A. Forcina. (1999). "Inference for Lorenz Curve Orderings." Econometrics Journal 2, 48-74.

[11] Davidson, R., and J.-Y. Duclos. (2000). "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality." Econometrica 68, 1435-1464.

[12] Denuit, M. and A. Cornet. (1999). "Premium Calculation with Dependent Time-until-Death Random Variables: The Widow's Pension." Journal of Actuarial Practice 7, 147-180.

[13] Denuit, M., J. Dhaene and C. Ribas. (2001). "Does positive dependence between individual risks increase stop-loss premiums?", Insurance: Mathematics and Economics 28, 305-308.

[14] Denuit, M., J. Dhaene, and M. Van Wouwe. (1999). "The Economics of Insurance: a Review and Some Recent Developments." Bulletin of the Swiss Association of Actuaries 2, 137-175.

[15] Dhaene, J., and M. Goovaerts. (1996). "Dependency of Risks and Stop-Loss Order." ASTIN Bulletin 26, 201-212.

[16] Doukhan, P. (1994). Mixing: Properties and Examples, Lecture Notes in Statistics, Springer-Verlag, New York.

[17] Embrechts, P., A. McNeil, and D. Straumann. (2000). "Correlation and Dependency in Risk Management: Properties and Pitfalls." Risk Management: Value at Risk and Beyond, eds Dempster M. and Moffatt H., Cambridge University Press, Cambridge.

[18] Fermanian, J.-D., and O. Scaillet. (2003). "Nonparametric estimation of copulas for time series." Journal of Risk 5, 25-54.

[19] Fermanian, J.-D., and O. Scaillet. (2004). "Some statistical pitfalls in copula modelling for financial applications." HEC Geneva working paper 2004-2.

[20] Fishburn, P. (1980). "Continua of Stochastic Dominance Relations for Unbounded Probability Distributions." Journal of Mathematical Economics 7, 271-285.

[21] Frees, E., and E. Valdez. (1998). "Understanding Relationships Using Copulae." North American Actuarial Journal 2, 1-25.

[22] Gerber, H. (1995). Life Insurance Mathematics, Springer, Berlin.

[23] Gouriéroux, C., A. Holly, and A. Monfort. (1982). "Likelihood Ratio, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters." Econometrica 50, 63-80.

[24] Howes, S. (1993). "Inferring Population Rankings from Sample Data." STICERD, discussion paper.

[25] Joe, H. (1997). Multivariate Models and Dependence Concepts. Chapman & Hall, London.

[26] Kaur, A., B. Prakasa Rao, and H. Singh. (1994). "Testing for Second-Order Dominance of Two Distributions." Econometric Theory 10, 849-866.

[27] Kimeldorf, G. and A. Sampson. (1987). "Positive Dependence Orderings." Annals of the Institute of Statistical Mathematics 39, 113-128.

[28] Klugman, S., and R. Parsa. (1999). "Fitting Bivariate Loss Distributions with Copulas." Insurance Mathematics and Economics 24, 139-148.

[29] Kodde, D., and F. Palm. (1986). "Wald Criteria for Jointly Testing Equality and Inequality Restrictions." Econometrica 54, 1243-1248.

[30] Kroll, Y., and H. Levy. (1980). "Stochastic Dominance Criteria : a Review and Some New Evidence." Research in Finance, Vol. II, p. 263-227, JAI Press, Greenwich.

[31] Kudo, A., (1963). "A Multivariate Analogue for the One-Sided Test." Biometrika 50, 403-418.

[32] Lehmann, E. (1966). "Some Concepts of Dependence." Annals of Mathematical Statistics 37, 1137-1153.

[33] Levy, H. (1992). "Stochastic Dominance and Expected Utility : Survey and Analysis." Management Science 38, 555-593.

[34] Mari, D., and S. Kotz. (2001). Correlation and Dependence. Imperial College Press, London.

[35] Milgrom, P., and R. Weber. (1982). "A Theory of Auctions and Competitive Bidding." Econometrica 50, 1089-1122.

[36] Mosler, K., and M. Scarsini. (1993). Stochastic Orders and Applications, a Classified Bibliography. Springer-Verlag, Berlin.

[37] Nelsen, R. (1999). An Introduction to Copulas. Lecture Notes in Statistics, Springer-Verlag, New-York.

[38] Newman, C. (1984). "Asymptotic Independence and Limit Theorems for Positively and Negatively Dependent Random Variables." Inequalities in Statistics and Probability, IMS Lecture Notes 5, 127-140.

[39] Norberg, R. (1989). "Actuarial Analysis of Dependent Lives." Bulletin de l'Association Suisse des Actuaires, 243-254.

[40] Pollard, D. (1985). "New Ways to Prove Central Limit Theorems." Econometric Theory 1, 295-314.

[41] Poon, S., M. Rockinger, and J. Tawn. (2003). "Extreme-Value Dependence in Financial Markets: Diagnostics, Models and Financial Implications." forthcoming in Review of Financial Studies.

[42] Rio, E. (2000). Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants. SMAI, Springer, France.

[43] Rolski, T. (1976). "Order Relations in the Set of Probability Distribution Functions and their Applications in Queueing Theory." Dissertationes Mathematicae 132, 5-47.

[44] Scott, D. (1992). "Multivariate Density Estimation: Theory, Practice and Visualisation." John Wiley & Sons, New-York.

[45] Shaked, M. and J. Shanthikumar (1994). Stochastic Orders and their Applications. Academic Press, New York.

[46] Sklar, A. (1959). "Fonctions de Répartition à n Dimensions et leurs Marges." Publ. Inst. Stat. Univ. Paris 8, 229-231.

[47] Tchen, A. (1980). "Inequalities for Distributions with Given Marginals." The Annals of Probability 8, 814-827.

[48] Tong, Y. (1990). The Multivariate Normal Distribution. Springer, New York.

[49] Tse, Y., and X. Zhang. (2003). "A Monte Carlo Investigation of Some Tests for Stochastic Dominance." forthcoming in Journal of Statistical Computation and Simulation.

[50] Wolak, F. (1989a) "Testing Inequality Constraints in Linear Econometric Models." Journal of Econometrics 41, 205-235.

[51] Wolak, F. (1989b). "Local and Global of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models." Econometric Theory 5, 1-35.

[52] Yaari, M. (1987). "The Dual Theory of Choice under Risk." Econometrica 55, 95-115.

[53] Yanagimoto, Y., and M. Okamoto. (1969). "Partial orderings of permutations and monotonicity of a rank correlation statistic." Annals of the Institute of Statistical Mathematics 21, 489–506.

	LOSS	ALAE
Mean	37,109.58	12,017.47
Std Dev.	92,512.80	26,712.35
Skew.	10.95	10.07
Kurt.	209.62	152.39
Min	10.00	15.00
Max	2,173,595.00	501,863.00
1st Quart.	3,750.00	2,318.25
Median	11,048.50	5,420.50
3rd Quart.	32,000.00	12,292.00

Table 1: Summary statistics for variables LOSS and ALAE.

α	25%	10%	5%	2.5%	1%	0.5%	0.1%
Lower bound	0.455	1.642	2.706	3.841	5.412	6.635	9.500

Table 2: Lower bounds on critical values for the distance test.

R	10,000	50,000	100,000	500,000	1,000,000
indep.	33,308.9054	19,108.3604	12,402.7515	1,800.9984	804.9684
dep.	36,765.8687	21,227.8071	13,801.1927	1,875.0277	850.1686
comon.	38,962.6734	23,271.1908	15,407.7782	2,308.0139	985.3801

Table 3: Pure premiums for a reinsurance treaty with retention R.

HFR	
Mean	0.0070372
Std Dev.	0.0094661
Skew.	0.2279256
Kurt.	0.4175570
Min	-0.0167000
Max	0.0359000
1st Quart.	0.0015250
Median	0.0063000
3rd Quart.	0.01.2975
CSFB/Tremont	
Mean	0.0085466
Std Dev.	0.0089309
Skew.	0.2007868
Kurt.	0.1991623
Min	-0.0115000
Max	0.0032600
1st Quart.	0.0028250
Median	0.0081000
3rd Quart.	0.0142750
S&P500	
Mean	0.0068819
Std Dev.	0.0462238
Skew.	-0.7155344
Kurt.	0.5972095
Min	-0.1575860
Max	0.0923238
1st Quart.	-0.0220540
Median	0.0122834
3rd Quart.	0.0391559

Table 4: Summary statistics for HFR, CSFB/Tremont and S&P500.

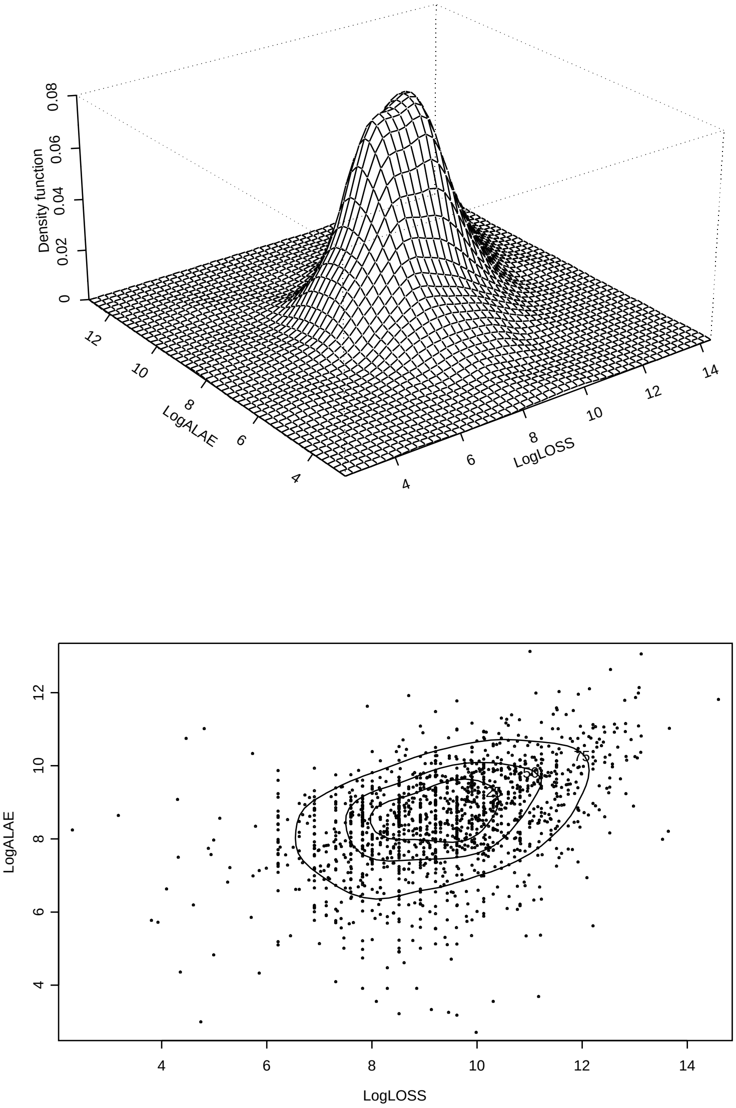

Figure 1: Kernel estimation of the bivariate pdf for (log(LOSS),log(ALAE)).

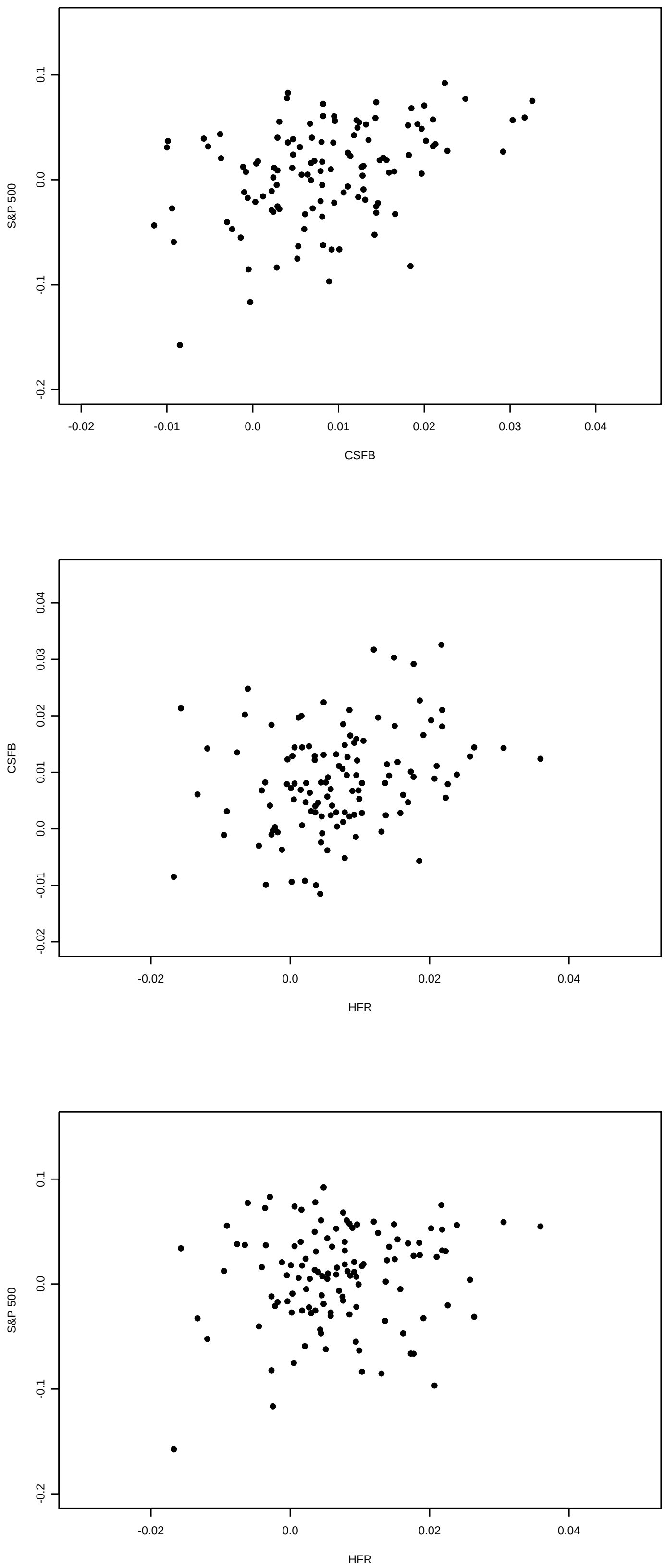

Figure 2: Bivariate scatterplots of the data.

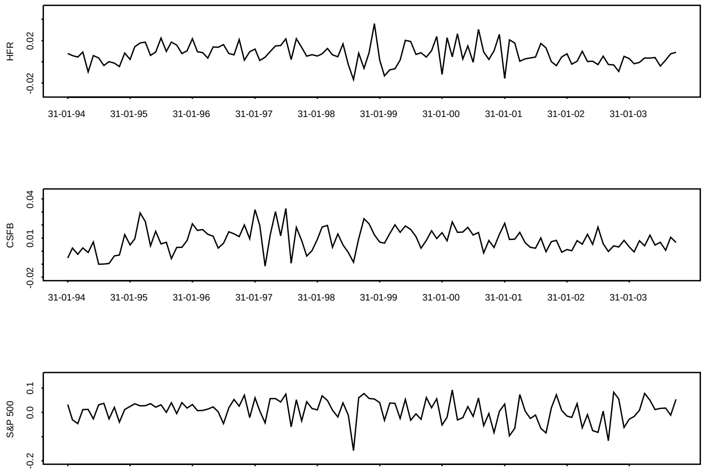

Figure 3: Plots of the time series.

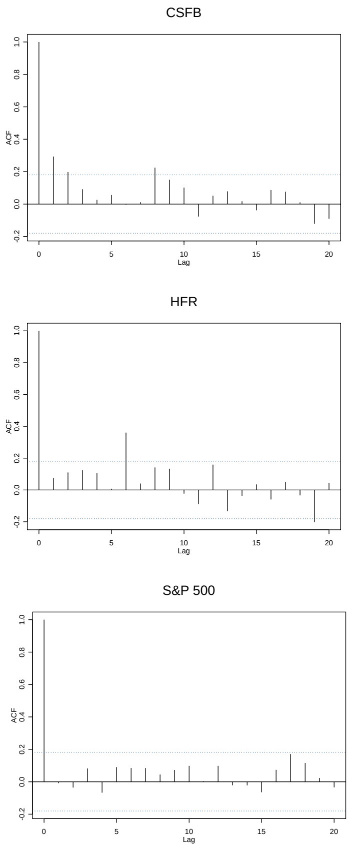

Figure 4: Autocorrelograms of the time series.