NONPARAMETRIC TESTS FOR POSITIVE QUADRANT DEPENDENCE

Michel Denuit
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve, Belgium
denuit@stat.ucl.ac.be

Olivier Scaillet
IRES, Département des Sciences Economiques,
and Institut d'Administration et de Gestion
Université Catholique de Louvain
Place Montesquieu, 3
B-1348 Louvain-la-Neuve, Belgium
scaillet@ires.ucl.ac.be

This version: April 2001 (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 generalisation in higher dimensions, namely positive orthant dependences, are also analysed. 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 depending on the definition of null and alternative hypotheses. An empirical illustration is given for US and Danish insurance claim data. Practical implications for the design of reinsurance treaties are also discussed.

Key words and phrases: Nonparametric, Stochastic Ordering, Positive Quadrant Dependence, Positive Orthant Dependence, Copula, Inequality Constraint Test, Risk Management, Loss Severity Distribution.

JEL Classification: C12, D81, G10, G21, G22.
AMS 2000 Subject Classification: 60E15, 62G10, 62G30, 62P05, 91B28, 91B30.

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 purpose to capture 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 impulsed 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.

In actuarial science, the increasing complexity of insurance and reinsurance products has lead to increased interest in the modeling of dependent risks (think of the emergence of multi-line products which require sophisticated risk evaluation mechanisms, see PINQUET (1998) for simple and powerful models applicable to packaged products in car insurance). Also, Dynamic Financial Analysis (see KAUFMANN, GADMER and KLETT (1999) for an introduction) provides actuaries with an integrated risk management technique. Its main characteristic is to integrate the investment and underwriting risk to which an insurer is exposed. DFA necessitates a model that combines information on marginal distributions together with ideas on interdependencies.

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 the normal world, positive correlation entails strong positive dependence notions, see TONG (1990). However, as illustrated by EMBRECHTS et al. (2000), dependence properties of the normal world no more hold in the 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 correctly be understood by practitioners.

In the management of large portfolios, the main risk is the occurrence of many joint default events or simultaneous downside evolution of prices. A deep knowledge of the dependence between financial assets or claims is crucial to better assess this risk of loss clustering and therefrom achieve a performant risk management in finance and insurance. This clustering behaviour can be described by a useful concept known as positive quadrant dependence (PQD) for bivariate distributions (LEHMANN (1966)) and positive orthant dependences (POD) for dimensions higher than two. 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.

One of the main interest of this dependence structure is that it allows the risk manager to compare the sum of PQD random variables with the corresponding sum under the independence assumption. 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, like YAARI (1987)'s one).

Finally let us remark 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 usefulness 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 is a latin noun that means "couple", and 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 "couples" one-dimensional distribution functions to form multivariate ones, and may serve to characterize dependence concepts such as PQD and POD.

The paper is organized as follows. In Section 2, we review several stochastic order relations. In Section 3, we recall the definition of copula functions, as well as the classical Sklar's representation theorem for multivariate distributions. Specification of hypotheses in terms of distribution functions or copulas will lead to different inferences. Section 4 gives the definition of PQD and of some of its multivariate extensions. In Section 5, we illustrate the interest of these positive dependence notions with the help of various useful stochastic inequalities. We provide some relevant examples coming from measurement of inequality and poverty, as well as from life insurance. In Section 6 we describe the null and alternative hypotheses we are interested in, and develop testing procedures for such purpose. These procedures are closely related to the inference tools for traditional first order and second order stochastic dominance, which also rely on distance and intersection-union tests for inequality constraints (see DAVIDSON and DUCLOS (2000) and the references therein). An empirical illustration on US and Danish insurance claim data is proposed in Section 7. Therein we provide a comparison of premiums computed under different dependence assumptions, and discuss effect of PQD on the pricing of reinsurance treaties. Section 8 concludes. Proofs are gathered in an appendix.

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 Stochastic orderings

Stochastic orderings are binary relations defined on classes of probability distributions. They aim to mathematically translate intuitive ideas like "being larger" or "being more variable" for random quantities. They thus extend the classical mean-variance approach to compare riskiness.

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 book by Shaked and Shanthikumar (1994). For a rigorous treatment of $\preceq_{k-\text{icv}}$, see Rolski (1976) and Fishburn (1976).

We summarize hereafter the main characterizations of $\preceq_{\text{d}}$, $\preceq_{\text{icv}}$ and $\preceq_{\text{cv}}$. Let F_1 and F_2 be the respective cdf's for Y_1 and Y_1. Let F_1^{-1} and F_2^{-1} denote the corresponding quantile transformations, defined as

$$F_j^{-1}(p) = \inf\{x \in \mathbb{R} | F_j(x) \geq p\}, \;\; p \in [0,1], \;\; j = 1, 2.$$

Theorem 2.1.

(i) $Y_1 \preceq_d Y_2 \Leftrightarrow F_1(x) \geq F_2(x)$ *for all* x*;*

(ii) $Y_1 \preceq_d Y_2 \Leftrightarrow F_1^{-1}(p) \leq F_2^{-1}(p)$ *for all* $p \in (0,1)$*;*

(iii) $Y_1 \preceq_{icv} Y_2 \Leftrightarrow \int_{-\infty}^{x} F_1(u)\,du \geq \int_{-\infty}^{x} F_2(u)\,du$ *for all* x*;*

(iv) $Y_1 \preceq_{icv} Y_2 \Leftrightarrow \int_0^p F_1^{-1}(u)\,du \leq \int_0^p F_2^{-1}(u)\,du$ *for all* $p \in (0,1)$*;*

(v) $Y_1 \preceq_{icv} Y_2 \Leftrightarrow$ *there exists a random variable* Z *such that* $Y_1 \preceq_{cv} Z \preceq_d Y_2$*;*

(vi) $Y_1 \preceq_{k-icv} Y_2 \Leftrightarrow$

$$\int_{-\infty}^{x}\int_{-\infty}^{x_{k-1}}\cdots\int_{-\infty}^{x_2} F_1(x_1)\,dx_1dx_2\cdots dx_{k-1}$$
$$\geq \int_{-\infty}^{x}\int_{-\infty}^{x_{k-1}}\cdots\int_{-\infty}^{x_2} F_2(x_1)\,dx_1dx_2\cdots dx_{k-1} \quad \textit{for all} \quad x,$$

provided the integrals are finite (the integrals are finite if F_1 and F_2 have finite $(k-1)$th moments).

Statistical inference for $\preceq_{\mathrm{d}}$, $\preceq_{\mathrm{icv}}$ and $\preceq_{k-\mathrm{icv}}$ is investigated in vast details in Davidson and Duclos (2000), where connections with economic and social welfare in different populations are explicited. Different approaches are possible. Either empirical analogs of the iterated integrals of the cdf's are computed and compared at every couple of observations. This is the route followed e.g. by McFadden (1989) or Kaur, Prakasa Rao and Singh (1994). Or, which is more often the case, a predetermined grid of a much smaller number of points (typically, quantiles of the underlying distributions) is used. In the latter case, the stochastic ranking implies a set of inequalities $D^i \geq 0$ for $i = 1, 2, \ldots, d$, where d is the number of points in the predetermined grid. This will be the foundation for the statistical tests discussed in Section 6. Before moving to that point we need to present a very powerful theorem due to Sklar (1959).

3 Sklar's representation for multivariate distributions

We consider a setting made of i.i.d. observations $\{\boldsymbol{Y}_t;\ t = 1, ..., T\}$ of a random vector $\boldsymbol{Y}$ taking values in $\mathbb{R}^n$. 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.

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. As already pointed out, how the joint distribution F is "coupled" to its univariate margins F_j, can be described by a copula. While the joint distribution F provides complete information concerning the behaviour of $\boldsymbol{Y}$, copulas allow to separate dependence and marginal behaviour of the elements constituting $\boldsymbol{Y}$. Before defining formally a copula, we would like to refer the reader to Nelsen (1999) and Joe (1997) for more extensive theoretical treatments.

A n-dimensional copula C is simply (the restriction to $[0,1]^n$ of) an n-dimensional cdf with unit uniform marginals. The reason why a copula is useful in revealing the link between the joint distribution and its margins transpires from the following theorem.

Theorem 3.1. (Sklar's Theorem)
Let F be an n-dimensional cdf with margins $F_1, ..., F_n$. Then there exists an n-copula C such that for all $\boldsymbol{y}$ in $\mathbb{R}^n$,

$$F(\boldsymbol{y}) = C\big(F_1(y_1), ..., F_n(y_n)\big). \tag{3.1}$$

If $F_1, ..., F_n$ are all continuous, then C is uniquely defined. Otherwise, C is uniquely determined on range $F_1 \times ... \times$ range F_n. Conversely, if C is an n-copula and $F_1, ..., F_n$ are cdf's, then the function F defined by (3.1) is an n-dimensional cdf with margins $F_1, ..., F_n$.

Although copulas constitute a less well-known approach to describing dependence than correlation, they offer the best understanding of the general concept of dependency. In particular, copulas share the nice property that strictly increasing transformations of the underlying random variables result in the transformed variables having the same copula (what is not true for linear correlation).

As an immediate corollary of Sklar's Theorem, we have

$$C(\boldsymbol{u}) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)) \tag{3.2}$$

for any $\boldsymbol{u} \in [0,1]^n$. From Expression (3.2), we may observe that the dependence structure embodied by the copula can be recovered from the knowledge of the joint cdf F and its margins F_j. This will be used later to deliver empirical estimates of copulas.

Note that independence between random variables can be characterized through copulas. Indeed, n random variables are independent if, and only if, their copula is $C(\boldsymbol{u}) = C^{\perp}(\boldsymbol{u}) = \prod_{j=1}^{n} u_j$, for all $\boldsymbol{u} \in [0,1]^n$. $C^{\perp}$ is further referred to as the independence copula. This characterization of independence is the starting point for the analysis of positive quadrant dependence through the use of copulas, which we wish to develop now.

4 Dependence notions

4.1 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 occurence of large losses or very negative returns is of particular interest in risk management.

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

$$P[\boldsymbol{Y} \leq \boldsymbol{y}] \geq P[\boldsymbol{Y}^{\perp} \leq \boldsymbol{y}] = P[Y_1 \leq y_1]P[Y_2 \leq y_2]. \tag{4.1}$$

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, (4.1) is equivalent to

$$P[\boldsymbol{Y} > \boldsymbol{y}] \geq P[\boldsymbol{Y}^{\perp} > \boldsymbol{y}] = P[Y_1 > y_1]P[Y_2 > y_2] \tag{4.2}$$

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

Considering (4.1)-(4.2), 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). PQD is in particular satisfied when random variables are regression dependent (see Dachraoui and Dionne (2000) 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 shows that PQD is a property of the underlying copula and is not influenced by the marginals. Inequality (4.1) can then also be written in terms of the copula C of the two random variables, since (4.1) is equivalent to the condition that, for all $\boldsymbol{u} \in [0,1]^2$,

$$C(\boldsymbol{u}) \geq C^{\perp}(\boldsymbol{u}) = u_1 u_2. \tag{4.3}$$

4.2 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 dependencies.

Positive orthant dependences offer nice extensions of PQD: in three or more dimensions, orthants are substituted for quadrants. This yields the following definitions, directly inspired from (4.1) ands (4.2). A random vector $\boldsymbol{Y}$ is said to be positively lower orthant dependent (PLOD, in short) when the inequalities

$$P[\boldsymbol{Y} \leq \boldsymbol{y}] \geq P[\boldsymbol{Y}^{\perp} \leq \boldsymbol{y}] = \prod_{i=1}^{n} P[Y_i \leq y_i] \tag{4.4}$$

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

$$P[\boldsymbol{Y} > \boldsymbol{y}] \geq P[\boldsymbol{Y}^{\perp} > \boldsymbol{y}] = \prod_{i=1}^{n} P[Y_i > x_i] \tag{4.5}$$

hold for any $\boldsymbol{y} \in \mathbb{R}^n$. Of course, (4.4) and (4.5) are no more equivalent when $n \geq 3$.

Intuitively, (4.5) means that $Y_1, Y_2, \ldots, Y_n$ are more likely simultaneously to have large values, compared with a vector of independent rv's with the same corresponding univariate marginals. Inequality (4.4) is similarly interpreted. When (4.4) and (4.5) simultaneously hold, then $\boldsymbol{Y}$ is said to be positively orthant dependent (POD, in short). POD is in particular 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 $\boldsymbol{Y}$, (4.4) can be written as

$$C(\boldsymbol{u}) \geq \prod_{j=1}^{n} u_j, \tag{4.6}$$

and (4.5) as

$$\overline{C}(\boldsymbol{u}) \geq \prod_{j=1}^{n} (1 - u_j), \tag{4.7}$$

for all $\boldsymbol{u} \in [0,1]^n$, where $\overline{C}$ denotes the survival copula associated with C.

Hence, PQD and PLOD may be characterized in terms of either cdf's or copulas, and thus may be checked, once cdf's or copulas are empirically known. In Section 6 we develop inference tools for that purpose.

Finally let us note that other dependence concepts such as negative quadrant dependence (NQD) and negative orthant dependences (NOD) may also be defined by reversing the sense of one, or all inequalities in (4.1) and (4.4) (see Nelsen (1999)). Testing procedures similar to ours may easily be developed for these cases. We focus hereafter to PQD and PLOD since we believe that they are the most relevant dependence notions in standard risk management applications. Nevertheless, the other concepts could also be of interest for other applications as, for instance, to determine whether a risk tends to hedge another one.

5 Applications of positive dependence notions

In the next lines, we illustrate the practical relevance of the positive dependence notions in measurement of inequality and poverty as well as life insurance.

5.1 PQD

One of the main interest of PQD is for comparison with random couples with identical marginals but independent components. This comes from the following result of which we provide a short proof in appendix. It is a straightforward adaptation of the result of Dhaene and Goovaerts (1996) established in the convex actuarial setting.

Proposition 5.1. *If Y_1 and Y_2 are PQD, then $Y_1 + Y_2 \preceq_{cv} Y_1^{\perp} + Y_2^{\perp}$.*

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. The quantity $E(Y_1 + Y_2 - \kappa)_+$ is referred to as the stop-loss premium relating to $Y_1 + Y_2$ in actuarial science (κ is called the deductible). In finance, when appropriately discounted, it can be regarded as the price of a basket option with Y_1 and Y_2 as underlying assets and κ as strike price. The convenient assumption of independence may thus lead to serious underpricing of insurance premiums or option prices. This will be confirmed by the empirical results of this paper.

The Lorenz order is defined by means of pointwise comparison of Lorenz curves. The latter is used in economics to measure the inequality of incomes (see Beach and Davidson (1983), Dardanoni and Forcina (1999) for related inference). More precisely, let Y be a non-negative random variable with cdf F. The Lorenz curve L associated with Y is then defined by

$$L(p) = \frac{1}{EY}\int_{t=0}^{p} F^{-1}(u)du, \;\; p \in [0,1].$$

When Y represents the income of the individuals in some population, L maps $p \in [0,1]$ to the proportion of the total income of the population which accrues to the poorest $100p$ % of the population.

Consider two non-negative random variables Y_1 and Y_2 with finite expectations. Then, Y_1 is said to be smaller than Y_2 in the Lorenz order, henceforth denoted by $Y_1 \preceq_{\text{Lorenz}} Y_2$, when $L_1(p) \geq L_2(p)$ for all $p \in [0,1]$. When $Y_1 \preceq_{\text{Lorenz}} Y_2$ holds, Y_1 does not exhibit more inequality in the Lorenz sense than does Y_2. A standard reference for $\preceq_{\text{Lorenz}}$ is Arnold (1987).

From Theorem 2.1(iv), we see that provided $EY_1 = EY_2$, $Y_1 \preceq_{\text{Lorenz}} Y_2 \Leftrightarrow Y_2 \preceq_{\text{cv}} Y_1$. Hence, if Y_1 and Y_2 are PQD then $Y_1^{\perp} + Y_2^{\perp} \preceq_{\text{Lorenz}} Y_1 + Y_2$ in virtue of Proposition 5.1. Let us give an interpretation of this stochastic inequality. Let Y_1 (resp. Y_2) denote the husbands' (resp. wives') income in some population. Saying that Y_1 and Y_2 are PQD means that, as the saying goes, "birds of a feather flock together": men and women earning large (resp. small) salaries tend to be associated. Such a population exhibits more inequality in the Lorenz sense than a population where spouses' earnings are independent. This type of inequality measurement based on PQD may also be applied to total household incomes in two countries instead of husbands' and wives' incomes in one country.

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

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

which readily follow from (4.1)-(4.2). 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 analysed by DENUIT and CORNET (1999).

5.2 PLOD

Proposition 5.1 no more holds if PLOD is substituted for PQD. Rather, the following result holds true.

Proposition 5.2. *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.*

Comparing Propositions 5.1 and 5.2, we see that $\preceq_{\text{cv}}$ is replaced with $\preceq_{n-\text{icv}}$ in dimension n. Besides, as it can be seen from Proposition 5.2, we only get weaker orderings in higher dimensions. To get $\preceq_{\text{cv}}$ as in Proposition 5.1, we need another dependence notion callled positive cumulative dependence (PCD in short) and 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 5.3. *Provided $\boldsymbol{Y}$ is PCD, the stochastic inequality $\sum_{i=1}^{n} Y_i \preceq_{cv} \sum_{i=1}^{n} Y_i^{\perp}$ holds.*

From (4.4) and (4.5), it is easy to get the following result that reinforces a stochastic inequality obtained by BACCELLI and MAKOWSKI (1989).

Proposition 5.4. *Let S be a subset of $\{1, 2, \ldots, n\}$. Provided $\boldsymbol{Y}$ 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 interest of Proposition 5.4 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:

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

and

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

From Proposition 5.4, it follows that

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

which in turn implies that

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

where the superscript "$\perp$" is used to indicate that the annuity is based on $T^{\perp}_{(x_1,x_2,\dots,x_n)}$ or $T^{\perp}_{\overline{(x_1,x_2,\dots,x_n)}}$. This means that for POD remaining lifetimes, the independence assumption (while leaving the marginal cdf's 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.

6 Hypotheses testing

Now that the relevant theoretical concepts and applications have been presented, we may turn our attention to inference. We develop two testing methods. The first one is based on a specification in terms of distribution functions, while the second one relies on copulas.

6.1 Inference based on distribution functions

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. 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.

Let us start with the definition (4.4) of PLOD in terms of cdf's, and take d points $\boldsymbol{y}_i = (y_{i1}, \dots, y_{in})'$ in $\mathbb{R}^n$. We define $D^i_F = F(\boldsymbol{y}_i) - \prod_{j=1}^n F_j(y_{ij})$, and $\boldsymbol{D}_F = (D^1_F, \dots, D^d_F)'$. The null hypothesis of a test for PLOD may be written as

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

and we take as alternative hypothesis:

$$H^1_F \;\; = \;\; \{\boldsymbol{D}_F : \boldsymbol{D}_F \text{ unrestricted }\}.$$

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

$$\hat{F}(\boldsymbol{y}_i) = \frac{1}{T}\sum_{t=1}^{T}\prod_{j=1}^{n}\mathbb{I}[Y_{jt} \leq y_{ij}], \qquad i = 1, ..., d, \tag{6.1}$$

$$\hat{F}_j(y_{ij}) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}[Y_{jt} \leq y_{ij}], \qquad i = 1, ..., d,\ j = 1, ..., n. \tag{6.2}$$

Let us define $\boldsymbol{y}_{k\wedge l} = (y_{k1} \wedge y_{l1}, ..., y_{kn} \wedge y_{ln})'$ where $a \wedge b = \min(a, b)$. Then the following proposition gives the asymptotic distribution of $\widehat{\boldsymbol{D}_F}$.

Proposition 6.1. *The random vector $\sqrt{T}(\widehat{\boldsymbol{D}_F} - \boldsymbol{D}_F)$ converges in distribution to a d-dimensional normal random variable with mean zero and covariance matrix $\boldsymbol{V}_F$ whose elements are*

$$v_{F,kl} = F(\boldsymbol{y}_{k\wedge l}) - F(\boldsymbol{y}_k)F(\boldsymbol{y}_l), \qquad k, l = 1, ..., d.$$

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

6.2 Inference based on copulas

Let us now proceed with the analoguous quantities when we use copulas, and take d points $\boldsymbol{u}_i = (u_{i1}, ..., u_{in})'$, with $u_{ij} \in (0, 1)$, $i = 1, ..., d$, $j = 1, \ldots, n$.

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 cdf 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 ζ_{ij}, $i = 1, ..., d$, $j = 1, ..., n$, while $f_j(\zeta_{ij}) > 0$ at each quantile ζ_{ij}.

In view of (3.2) 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 (6.1) is that the levels are no more 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 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 $\widehat{\boldsymbol{D}_C}$.

Proposition 6.2. *The random vector $\sqrt{T}(\widehat{\boldsymbol{D}_C} - \boldsymbol{D}_C)$ converges in distribution to a d-dimensional normal random variable with mean zero and covariance matrix $\boldsymbol{V}_C$ whose elements are*

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

where

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

and

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

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 $\widehat{\boldsymbol{V}_C}$ 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 bandwiths h_j, that is $1.05T^{-1/5}$ times the estimated standard deviation of Y_j.

6.3 Testing procedure

The distributional results of Propositions 6.1 and 6.2 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 several components of $\widehat{\boldsymbol{D}_K}$ are found to be negative (in such a case one wants to know whether this invalidates PLOD).

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

$$\inf_{\boldsymbol{D}} T(\boldsymbol{D} - \widehat{\boldsymbol{D}_K})' \widehat{\boldsymbol{V}_K}^{-1} (\boldsymbol{D} - \widehat{\boldsymbol{D}_K}) \qquad s.t. \qquad \boldsymbol{D} \geq 0, \tag{6.3}$$

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

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

Roughly speaking, $\widetilde{\boldsymbol{D}_K}$ is the closest point to $\widehat{\boldsymbol{D}_K}$ under the null in the distance measured in the metric of $\widehat{\boldsymbol{V}}_K$, and the test statistic $\hat{\xi}_K$ is the distance between $\widetilde{\boldsymbol{D}_K}$ and $\widehat{\boldsymbol{D}_K}$. 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:

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

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

Computation of the solution $\widetilde{\boldsymbol{D}}_K$ can be performed by a numerical optimisation 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 realisations of a multivariate normal with mean zero and covariance matrix $\widehat{\boldsymbol{V}}_K$. Then use these realisations as $\widehat{\boldsymbol{D}_K}$ in the above minimisation problem (6.3), compute $\widetilde{\boldsymbol{D}}_K$, and count the number of elements of the vector greater than zero. The proportion of draws such that $\widetilde{\boldsymbol{D}}_K$ has exactly $d - i$ elements greater that zero gives a Monte Carlo estimate of $w(d, d-i, \widehat{\boldsymbol{V}_K})$. 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 testing 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 $\widehat{\boldsymbol{D}}_K$ 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 or equal to α, 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 the asymptotic size and is consistent (see e.g. Howes (1993), Kaur, Prakasa Rao and Singh (1994)).

Power issues are extensively studied for stochastic dominance and nondominance tests in Dardanoni and Forcina (1999) (see also the comments in Davidson and Duclos (2000)). They carry over to our case. 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, they find that, indeed, distance tests are worth the extra amount of computational work. 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 $\widehat{\boldsymbol{D}_K}$ is close enough to zero for a number of coordinates.

7 Empirical illustrations

This section illustrates the implementation of the testing procedures described in the previous section. We provide two empirical applications to insurance. They concern the detection of PQD in US and Danish insurance claim data, and its effect on premium valuation.

7.1 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 modelling 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 choosen 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 7.1. 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.

	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 7.1: Summary statistics for variables LOSS and ALAE.

Figure 7.7.1 shows the kernel estimator of the bivariate pdf of the couple (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 $\widehat{\boldsymbol{D}_F}$ with only one negative component -0.0002. We wish to check thanks to the distance test whether this invalidates PQD or not. The distance between $\widehat{\boldsymbol{D}_F}$ and $\widetilde{\boldsymbol{D}_F}$ is found to be 6.5×10^{-12}. Lower bounds on the critical values obtained by Kodde and Palm (1986) are given in Table 7.2 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.

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 7.2: Lower bounds on critical values for the distance test.

Let us now consider a positive dependence, but only in the upper tails. 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 $\widehat{\boldsymbol{D}_F}$ 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 $\min\hat{\gamma}_F^i=0.1553$ which does not allow us to reject $\overline{H}_0^F$ in favor of PQD. This non rejection is due to the closeness of $\widehat{\boldsymbol{D}_F}$ to zero for a large number of coordinates. This point has already been discussed at the end of Section 6.

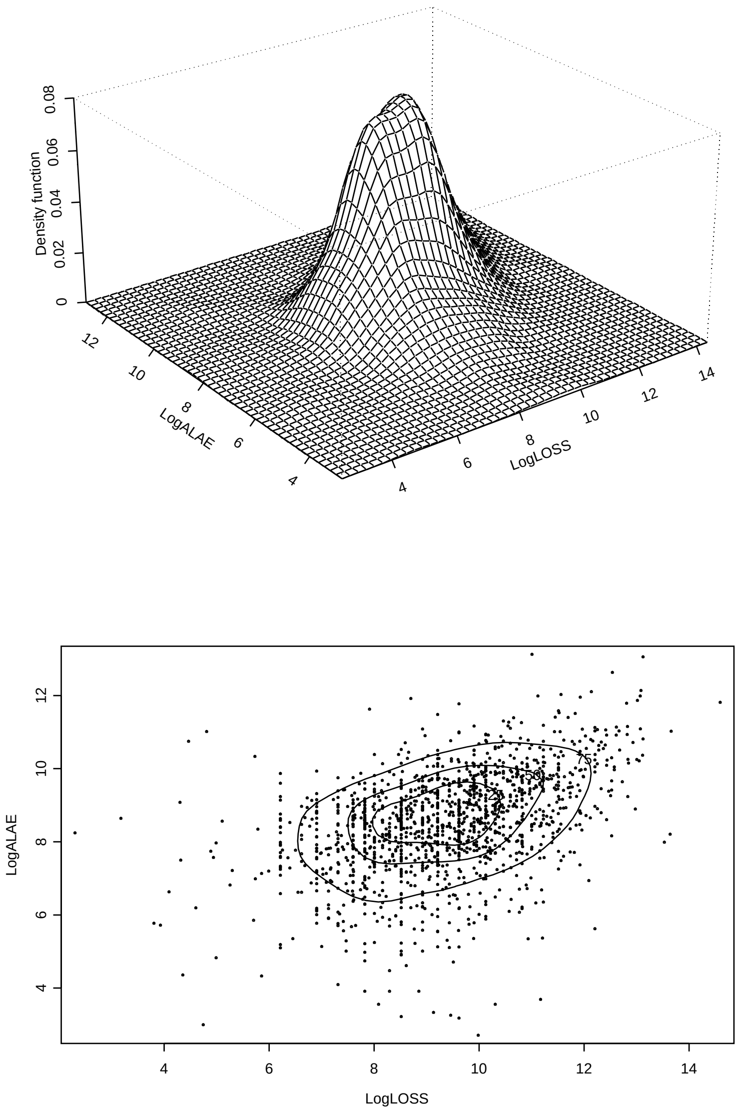

Figure 7.7.1: Kernel estimation of the bivariate pdf for (log(LOSS),log(ALAE)).

Let us now turn to copula based tests. For the $\boldsymbol{u}_i$'s, we take the 81 deciles of the grid $\{0.1, 0.2, \dots, 0.9\} \times \{0.1, 0.2, \dots, 0.9\}$. All components of the corresponding $\widehat{\boldsymbol{D}}_C$ are positive, so that H_0^C cannot be rejected. For the intersection-union test, we obtain $\min \widehat{\gamma}_C^i = 0.9489$ 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 $\{0.91, 0.92, \dots, 0.99\} \times \{0.91, 0.92, \dots, 0.99\}$, we get that all components of $\widehat{\boldsymbol{D}}_C$ are again positive resulting in the non-rejection of H_0^C. Further, $\min \widehat{\gamma}_C^i = 0.6983$, so that $\overline{H}_0^C$ is not rejected, either.

It has to be pointed out that the choice of the bandwith 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 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 7.3 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.

$$\widehat{\pi} = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} g(\text{LOSS}_t, \text{ALAE}_{t'});$$

2. the second one takes into account the dependence expressed by the data, i.e.

$$\widehat{\pi} = \frac{1}{T} \sum_{t=1}^{T} g(\text{LOSS}_t, \text{ALAE}_t);$$

3. and the last one resorts to the classical comonotonic approximation for (LOSS,ALAE), i.e.

$$\widehat{\pi} = \frac{1}{T} \sum_{t=1}^{T} g\left(\text{LOSS}_t, \widehat{F}_2^{-1}\left(\widehat{F}_1(\text{LOSS}_t)\right)\right).$$

We see that substantial mispricing could result from the independence hypothesis, 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).

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 7.3: Pure premiums for a reinsurance treaty with retention R.

7.2 Danish fire losses

These data comprise losses over one million Danish Krone for the years 1980-1989. Loss figures are classified as damage to buildings (variable Buildings), and damage to furniture and personal property (variable Contents), and consist in $T = 1,485$ data points. Total losses made of the sum of both losses have been previously studied in the context of tail analysis by Embrechts, Kluppelberg and Mikosch (1997), McNeil (1997), and Scaillet (2000).

Table 7.4 gives summary statistic for the variables "Buildings" and "Contents". The estimated values for Pearson's r, Kendall's τ and Spearman's ρ are 0.5362, 0.0741 and 0.1385, respectively. All of them are significantly positive. Because variables take sometime very high values, we again decide to work on a logarithmic scale. After this logarithmic transformation, the new Pearson's r is 0.2315.

	Buildings	Contents
Mean	1,731,012	1,391,979
Std Dev.	2,842,519	3,776,137
Skew.	12.102	9.068
Kurt.	217.587	122.316
Min	25,000	10,000
Max	65,000,000	72,500,000
1st Quart.	800,000	250,000
Median	1,100,000	430,000
3rd Quart.	1,775,000	1,000,000

Table 7.4: Summary statistics for variables Buildings and Contents.

Figure 7.7.2 displays a kernel estimate of the bivariate pdf of the couple (log(Buildings),log(Contents)), as well as its associated contour plot. Again positive dependence is expected in light of these graphs. The shape of the dependence is however different than for losses and ALAE's.

Taking the equally spaced grid $\{11, 12, \ldots, 17\} \times \{11, 12, \ldots, 17\}$ over the whole observation domain we get 15 negative components for $\widehat{\boldsymbol{D}_F}$. The distance between $\widetilde{\boldsymbol{D}_F}$ and $\widehat{\boldsymbol{D}_F}$ is found to be 0.33. Hence we do not reject the null of PQD at any reasonable confidence level (see the bounds of Table 7.2).

Let us now examine the presence of positive dependence in the upper tails. We take the grid $\{15, 15.3, 15.6, 16, 16.3, 16.6, 17\} \times \{15, 15.3, 15.6, 16, 16.3, 16.6, 17\}$. All components of $\widehat{\boldsymbol{D}_F}$ are positive. We have $\min \hat{\gamma}_F^i = 1.015$ in the intersection-union test. This does not allow

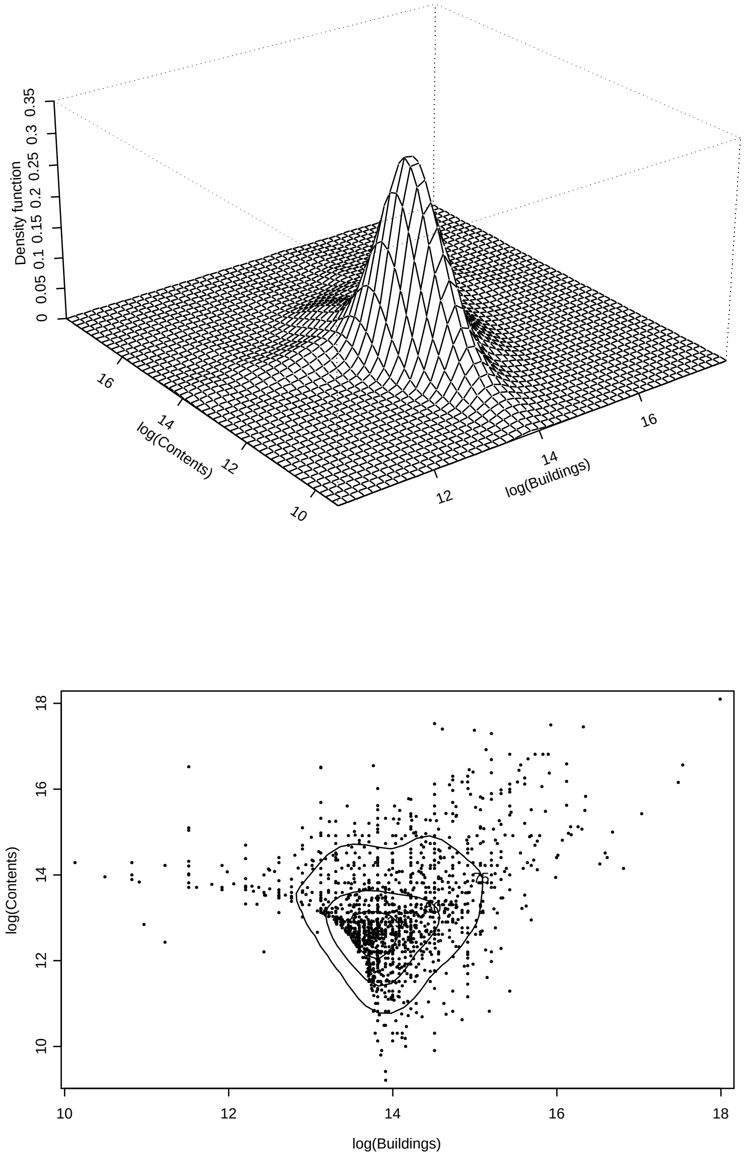

Figure 7.7.2: Kernel estimation of the bivariate pdf for (log(Buildings),log(Contents)).

us to reject $\overline{H}_0^F$ in favor of PQD (this is again due to closeness of $\widehat{\boldsymbol{D}_F}$ to zero for a large number of coordinates).

For the copula based inference procedures, we take the decile grid $\{0.1, 0.2, \ldots, 0.9\} \times \{0.1, 0.2, \ldots, 0.9\}$ as before. The resulting vector $\widehat{\boldsymbol{D}}_C$ exhibits 24 negative components, with a minimum of -0.0533. The question is: does this invalidate PQD? The distance between $\widetilde{\boldsymbol{D}}_C$ and $\widehat{\boldsymbol{D}}_C$ in the metric induced by $\widehat{\boldsymbol{V}}_C$ is 0.28. On the basis of the values in Table 7.2, this does not imply the rejection of H_0^C in favor of H_1^C. If we concentrate on the right tails, and use the percentile grid $\{0.91, 0.92, \ldots, 0.99\} \times \{0.91, 0.92, \ldots, 0.99\}$, all components of $\widehat{\boldsymbol{D}}_C$ are positive. This implies non-rejection of H_0^C. Further, $\min \widehat{\gamma}_C^i = 0.3402$, so that $\overline{H}_0^C$ is not rejected, either. As it can be observed from the results in both empirical illustrations, the intersection-union test is extremely conservative and seems to lead to sparse rejection.

To end this empirical section, we propose to analyse the effect of PQD on stop-loss reinsurance premiums. We examine several deductibles κ, and estimate $E(\text{Buildings+Contents} - \kappa)_+$. Again, we consider three situations: independence, actual dependence and comonotonicity. Stop-loss premiums are displayed in Table 7.5. As expected, they are listed in ascending order. The premium computed under actual dependence considerably exceeds the price under independence. This is especially true for large values of κ, reflecting the strong positive dependence in the tails. Besides the comonotonic assumption delivers too heavy premiums.

κ ($\times 10^6$)	5	7.5	10	12.5	15	17.5	20
indep.	782,667	556,565	420,241	328,537	265,217	219,057	185,606
dep.	919,680	698,131	550,809	439,888	356,078	297,493	254,729
comon.	1,046,330	802,787	652,342	545,302	461,813	397,181	348,361

Table 7.5: Stop-loss premiums for different deductibles κ.

8 Concluding remarks

In this paper we have analysed 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 dependencies among US and Danish insurance claim 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, and should help to achieve a better design of insurance contracts in terms of premium valuation.

APPENDIX

A Proof of Proposition 5.1

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,

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

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

B Proof of Proposition 5.2

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

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

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

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

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

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

C Proof of Proposition 5.3

We proceed by recurrence. From Proposition 5.1, we know that provided $\boldsymbol{Y}$ is PCD, $Y_1 + Y_2 \preceq_{\text{cv}} Y_1^{\perp} + Y_2^{\perp}$. Assume now that $Y_1 + \ldots + Y_k \preceq_{\text{cv}} Y_1^{\perp} + \ldots + Y_k^{\perp}$ holds true. Then, $Y_1 + \ldots + Y_k \preceq_{\text{cv}} Y_1^{\perp} + \ldots + Y_k^{\perp}$ is also valid since $(Y_1 + \ldots + Y_k, Y_{k+1})$ is PQD.

D Proof of Proposition 6.1

From standard properties of the empirical distribution, the first term of the difference $\hat{D}_F^i = \hat{F}(\boldsymbol{y}_i) - \prod_{j=1}^n \hat{F}_j(y_{ij})$ is of order $T^{-1/2}$, while the second term involves a product of order $T^{-n/2}$. This means that only the first term contributes to the asymptotic distribution. The stated result is then a direct consequence of the central limit theorem and a simple computation of the asymptotic covariance:

$$\lim_{T\to\infty} T\,\text{Cov}\,(\hat{F}(\boldsymbol{y}_k), \hat{F}(\boldsymbol{y}_l)) \;=\; F(y_{k1} \wedge y_{l1}, ..., y_{kn} \wedge y_{ln}) - F(\boldsymbol{y}_k)F(\boldsymbol{y}_l).$$

E Proof of Proposition 6.2

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:

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

where $\bar{\boldsymbol{\zeta}}_i$ is a mean value located between $\widehat{\boldsymbol{\zeta}}_i$ 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

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

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 stated result from the central limit theorem and computation of several covariance terms such as:

$$\lim_{T\to\infty} T\,\mathrm{Cov}\,[\hat{F}(\boldsymbol{\zeta}_k), \hat{F}_j(\zeta_{lj})] \;=\; F(\zeta_{k1}, ..., \zeta_{kj} \wedge \zeta_{lj}, ..., \zeta_{kn}) - F(\boldsymbol{\zeta}_k)F_j(\zeta_{lj}),$$

$$\lim_{T\to\infty} T\,\mathrm{Cov}\,[\hat{F}_j(\zeta_{kj}), \hat{F}_j(\zeta_{lj})] \;=\; F_j(\zeta_{kj} \wedge \zeta_{lj}) - F_j(\zeta_{kj})F_j(\zeta_{lj}),$$

$$\lim_{T\to\infty} T\,\mathrm{Cov}\,[\hat{F}_{j_1}(\zeta_{kj_1}), \hat{F}_{j_2}(\zeta_{lj_2})] \;=\; F_{j_1j_2}(\zeta_{kj_1} \wedge \zeta_{lj_1}, \zeta_{kj_1} \wedge \zeta_{lj_2}) - F_{j_1}(\zeta_{kj_1})F_{j_2}(\zeta_{lj_2}).$$

Acknowledgements

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 also thank Professors Embrechts and McNeil for providing the Danish data on fire insurance losses, which were collected by Mette Rytgaard at Copenhagen Re. We are grateful to Professor Irène Gijbels for her comments.

The first author acknowledges financial support of the Belgian Government under "Projet d'Actions de Recherche Concertées" (No. 98/03-217). The second author receives support from the Belgian Program on Interuniversity Poles of Attraction (PAI nb. P4/01). Part of this research was done when he was visiting THEMA.

Downloadable at www.econ.ucl.ac.be/IRES/CSSSP/home_pa_pers/scaillet/scaill.htm

References

[1] Andrews, D. (1994): "Empirical Process Methods in Econometrics", in *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] Arnold, B. (1987): *Majorization and the Lorenz Order: A Brief Introduction*, Springer-Verlag, New York, NY.

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

[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] Dachraoui, K. and G. Dionne (2000): "Optimal Financial Portfolio and Dependence of Risky Assets", Working Paper 00-12, HEC Montréal.

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

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

[11] 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.

[12] Denuit, M., J. Dhaene and C. Ribas (2001): "Does positive dependence between individual risks increase stop-loss premiums?", *Insurance: Mathematics and Economics*, to appear.

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

[14] Embrechts, P., C. Kluppelberg and T. Mikosch (1997): *Modelling Extremal Events for Insurance and Finance*, Springer Verlag, Berlin.

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

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

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

[18] Gerber, H. (1995): *Life Insurance Mathematics*, Springer, Berlin.

[19] 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.

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

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

[22] Kaufmann, R., A. Gadmer and R. Klett (1999): "Introduction to Dynamic Financial Analysis", ETH Technical Report, Zürich.

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

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

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

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

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

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

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

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

[31] McNeil, A. (1997): "Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory", *Astin Bulletin*, **27**, 117-137.

[32] McFadden, D. (1989): "Testing for Stochastic Dominance", in *Studies in the Economics of Uncertainty*, eds Fomby T. and Seo T., Springer-Verlag, New York.

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

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

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

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

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

[38] Pinquet, J. (1998): "Designing Optimal Bonus-Malus Systems from Different Types of Claims," *ASTIN Bulletin* 28, No. 2, 205-220.

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

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

[41] Scaillet, O. (2000): "Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall", DP IRES.

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

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

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

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

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

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

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

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

[50] 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.