

TESTING FOR CONCORDANCE ORDERING

BY

ANA C. CEBRIÁN, MICHEL DENUIT AND OLIVIER SCAILLET

ABSTRACT

We propose inference tools to analyse the concordance (or correlation) order of random vectors. The analysis in the bivariate case relies on tests for upper and lower quadrant dominance of the true distribution by a parametric or semiparametric model, i.e. for a parametric or semiparametric model to give a probability that two variables are simultaneously small or large at least as great as it would be were they left unspecified. Tests for its generalisation in higher dimensions, namely joint lower and upper orthant dominance, are also analysed. The parametric and semiparametric settings are based on the copula representation for multivariate distribution, which allows for disentangling behaviour of margins and dependence structure. 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 insurance claim data.

KEYWORDS

Nonparametric, Concordance (Correlation) Order, Quadrant Dominance, Orthant Dominance, Copula, Inequality Constraint Tests, Risk Management, Loss Severity Distribution.

INTRODUCTION

Random variables are concordant if they tend to be all large together or small together. Concordance of random variables conveys the idea of clustering of large and small events. An ordering of concordance was initially considered for two random variables by YANAGIMOTO and OKAMOTO (1960), CAMBANIS, SIMONS and STOUT (1976) and TCHEN (1980), and then extended by JOE (1990) to the multivariate case. This ordering corresponds to a natural notion of stochastic dominance between two distribution functions with fixed marginals. Large and small values will tend to be more often associated under the distribution which dominates the other one. This tool appears to be of particularly relevance in actuarial science, as demonstrated by DHAENE and GOOVAERTS (1996) (see also the review of applications of the concordance order in CEBRIÁN, DENUIT and SCAILLET (2002)).

In fact detection of concordant behaviour is especially important in risk management of large portfolios of insurance contracts or financial assets. In these portfolios the main risk is the occurrence of many joint default events or simultaneous downside evolution of prices. An accurate knowledge of concordance between claims or financial asset prices will help to assess this risk of loss clustering and there from allows for taking appropriate action to ensure that the risk incurred by the financial institution remains within its stated risk appetite. Clearly the presence of concordance affects risk measures and asset allocations resulting from optimal portfolio selection. Analysis of concordance cannot be neglected and reveals much of the danger associated to a given position.

Modelling of concordance can be fully parametric or semiparametric. In the first case specific parametric forms are selected for the dependence structure and the margins, while in the second case margins are left unspecified. The dependence structure is expressed by means of a parametric copula function. Once these models estimated a natural concern of a risk manager ought to be: does the adopted modelling reflect the dependence structure present in the data safely enough? The aim of this paper is to provide inference tools to answer this question. These tools will be tests for the ordering of concordance for random variables. In fact we propose testing procedures for concordance order between the chosen model and the empirical distribution. This allows for checking whether the estimated parametric or semiparametric model gives a safe picture of the association between small and large observed losses.

It is worth emphasising the difference between the actuarial approach and the statistical one. For a statistician, the aforementioned problem is not really relevant. Instead of addressing the issue whether the adopted modelling reflects the dependence structure present in the data safely enough, the statistical primary concern is that the model adopted reflects the dependence structure as accurately as possible, i.e. that it be realistic rather than conservative. For the actuary, the model should also be realistic but primarily conservative. Indeed, the adopted model will be used to price insurance products and any underestimation of the associated premium may result in severe financial losses for the company. Here we take the point of view of a seller, for example an insurance company whose core business is selling contracts with premiums high enough to cover potential losses. The buyer point of view will be different, but insurance premiums may often be considered to have only a marginal impact on business revenues, as well as to be difficult to negotiate. Besides regulators may also favour a conservative behaviour to guarantee the stability of the insurance system.

The paper is organized as follows. In Section 2, we formally define concordance and its ordering. We also recall the definition of copula functions, as well as the classical Sklar's representation theorem for multivariate distributions. Sections 3 and 4 are devoted to inference. The first one deals with the parametric setting, and the second one with the semiparametric setting. Since concordance is a copula-based concept, the natural way to proceed is to compare the empirical copula with the estimated copula selected from a given parametric family. In Section 5 we develop the testing procedures, and describe the

null and alternative hypotheses we are interested in. These procedures are closely related to inference tools for traditional first order and second order stochastic dominance, or for positive quadrant dependence. These tools also rely on distance and intersection-union tests for inequality constraints (see DAVIDSON and DUCLOS (2000), DENUIT and SCAILLET (2002) and the references therein). An empirical illustration on US insurance claim data is proposed in Section 6. Section 7 contains some concluding remarks. Proofs of all propositions are gathered in an appendix.

2. CONCORDANCE ORDER

Let F and G denote a n -dimensional cdf, and \bar{F} and \bar{G} their corresponding survival function. Then G is more concordant than F , written $F \leq_c G$, if

$$F(\mathbf{y}) \leq G(\mathbf{y}) \quad \text{and} \quad \bar{F}(\mathbf{y}) \leq \bar{G}(\mathbf{y}), \quad \forall \mathbf{y} \in \mathbb{R}^n. \quad (2.1)$$

The first inequality $F(\mathbf{y}) \leq G(\mathbf{y})$ corresponds to lower orthant dominance, while the second one $\bar{F}(\mathbf{y}) \leq \bar{G}(\mathbf{y})$ corresponds to upper orthant dominance. If both inequalities hold, large and small values will tend to be more often associated under G than F . Condition (2.1) implies that F and G have the same j th univariate marginal distribution ($j = 1, \dots, n$), and that all bivariate and higher dimensional marginals of G are more concordant than the corresponding ones for F . Hence we see that the ordering of concordance of variables is derived from comparisons of pairs of distributions with identical marginals (YANAGIMOTO and OKAMOTO (1960), CAMBANIS, SIMONS and STOUT (1976), TCHEN (1980), JOE (1990)). Besides the concordance order is equivalent to a partial order among the parameters for elliptically contoured distributions, such as the multivariate normal and student distributions. Henceforth we will freely use \leq_c between cdf's or random vectors to indicate that (2.1) holds.

At this stage, it is interesting to stress the very particular nature of the bivariate case (compared to dimension ≥ 3). Indeed, consider F and G with identical marginals. Then, it is easily seen that one inequality in (2.1) implies the other. The bivariate case is really particular, mainly because the concordance order coincides with the supermodular order for random couples (this is not the case for random vectors of larger dimension).

The concordance order can also be characterised in terms of copulas. The marginal pdf and cdf of each element Y_j of \mathbf{Y} at point y_j , $j = 1, \dots, n$, will be written $f_j(y_j)$, and $F_j(y_j)$, respectively. 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 \mathbf{Y} , copulas allow for separating dependence and marginal behaviour of the elements constituting $\mathbf{Y} = (Y_1, \dots, Y_n)'$. 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 2.1. (Sklar's Theorem)

Let F be an n -dimensional cdf with margins F_1, \dots, F_n . Then there exists an n -copula C such that for all \mathbf{y} in \mathbb{R}^n ,

$$F(\mathbf{y}) = C(F_1(y_1), \dots, F_n(y_n)). \quad (2.2)$$

If F_1, \dots, F_n are all continuous, then C is uniquely defined. Otherwise, C is uniquely determined on range $F_1 \times \dots \times \text{range } F_n$. Conversely, if C is an n -copula and F_1, \dots, F_n are cdf's, then the function F defined by (2.2) is an n -dimensional cdf with margins F_1, \dots, 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 (see EMBRECHTS, MCNEIL and STRAUMANN (2000) for implications on risk management). 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, the representation

$$C(\mathbf{u}) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \quad (2.3)$$

holds true for any $\mathbf{u} \in [0, 1]^n$ provided all margins F_1, \dots, F_n are continuous. Expression (2.3) shows that the dependence structure embodied by the copula can be recovered from the knowledge of the joint cdf F and its margins F_j , $j = 1, \dots, n$.

Let us now state the definition (2.1) of concordance in terms of copulas. We denote by \bar{C} the survival copula associated with C , and use subscripts F and G to reflect their associated distribution. If F and G share the same univariate margins, G is more concordant than F if

$$C_F(\mathbf{u}) \leq C_G(\mathbf{u}) \quad \text{and} \quad \bar{C}_F(\mathbf{u}) \leq \bar{C}_G(\mathbf{u}), \quad \forall \mathbf{u} \in [0, 1]^n. \quad (2.4)$$

Note that the density c associated with the copula C is given by:

$$c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}.$$

The density f of F can be expressed in terms of the copula density c and the product of the univariate marginal densities f_j :

$$f(y_1, \dots, y_n) = c(F_1(y_1), \dots, F_n(y_n)) \prod_{j=1}^n f_j(y_j).$$

Obviously the latter equality does not depend on a parametric assumption for the multivariate distribution. In the inference part of this paper we will consider

two cases. The first case will put some parametric assumption $f_j(y_j; \beta_j)$ on the margins, while the second will not. The copula will be parameterised according to $c(u_1, \dots, u_n; \theta)$ in both cases.

Let us remark that independence between random variables can be characterised through copulas. Indeed, n random variables are independent if, and only if, their copula is $C^\perp(\mathbf{u}) = \prod_{j=1}^n u_j$, for all $\mathbf{u} \in [0, 1]^n$. C^\perp is further referred to as the independence copula.

3. INFERENCE UNDER PARAMETRIC SPECIFICATION

Now that the relevant theoretical concepts have been presented, we may turn our attention to inference. We consider a setting made of i.i.d. observations $\{\mathbf{Y}_t; t = 1, \dots, T\}$ of a random vector \mathbf{Y} taking values in \mathbb{R}^n . These data may correspond to either observed individual losses on n insurance contracts, amounts of claims reported by a given policyholder on n different guarantees in a multiline product or observed returns of n financial assets.

We consider a fully parametric specification in this section, namely the parametric family:

$$\{F(\mathbf{y}; \nu) = C(F_1(y_1; \beta_1), \dots, F_n(y_n; \beta_n); \theta), \nu = (\beta', \theta')' \in \Psi \subset \mathbb{R}^{q+p}\}.$$

This parametric family is specified in terms of a parametric copula $C(\mathbf{u}; \theta)$ and parametric margins $F_j(y_j; \beta_j)$, $j = 1, \dots, n$. The q -dimensional vector $\beta = (\beta'_1, \dots, \beta'_n)'$ and the p -dimensional vector θ forming ν are jointly estimated by pseudo maximum likelihood. The estimator $\hat{\nu}$ is derived from

$$\max_{\beta, \theta} \frac{1}{T} \sum_{t=1}^T \ln c(F_1(Y_{1t}; \beta_1), \dots, F_n(Y_{nt}; \beta_n); \theta) \sum_{j=1}^n \ln f_j(Y_{jt}; \beta_j),$$

and its limit, i.e. the pseudo true value, is denoted by $\nu_0 = (\beta'_0, \theta'_0)'$. Standard regularity conditions ensure consistency and asymptotic normality of the estimator.

Let us denote as F_0 , resp. F_{0j}, f_{0j} , the true underlying joint distribution, resp. j th marginal cdf, j th marginal pdf, of the \mathbf{Y}_t 's. Here we assume that the marginals are correctly specified, i.e. misspecification if any is in the copula only. This is because a comparison in the \leq_c -sense implies that the distributions have the same univariate marginals. Now, $F(\cdot; \nu_0)$ is the “best” approximation to F_0 in a parametric family $\{F(\cdot; \nu)\}$ in the Kullback-Leibler divergence. The question is whether $F(\cdot; \nu_0)$ dominates the true F_0 in the \leq_c -sense, which is equivalent to wonder whether the same \leq_c ranking holds for the underlying copulas.

Let us take d points $\mathbf{u}_i = (u_{i1}, \dots, u_{in})'$, with $u_{ij} \in (0, 1)$, $i = 1, \dots, d$, $j = 1, \dots, n$. We assume hereafter that the cdf F_{j0} is such that the equation $F_{j0}(y) = u_{ij}$ admits a unique solution denoted ζ_{ij} , $i = 1, \dots, d$, $j = 1, \dots, n$, while $f_{j0}(\zeta_{ij}) > 0$ at each quantile ζ_{ij} . We denote the stack of the univariate quantiles ζ_{ij} by ζ_i .

We may then define $D_1^i = C(\mathbf{u}_i; \theta_0) - F_0(\zeta_i)$, and $\mathbf{D}_1 = (D_1^1, \dots, D_1^d)'$. The survival quantities will be $\bar{D}_1^i = \bar{C}(\mathbf{u}_i; \theta_0) - \bar{F}_0(\zeta_i)$, and $\bar{\mathbf{D}}_1 = (\bar{D}_1^1, \dots, \bar{D}_1^d)'$. The empirical

counterparts are then $\hat{D}_1^i = C(\mathbf{u}_i; \hat{\theta}) - \hat{F}(\hat{\xi}_i)$, and $\hat{\bar{D}}_1^i = \bar{C}(\mathbf{u}_i; \hat{\theta}) - \hat{\bar{F}}(\hat{\xi}_i)$, where $\hat{\xi}_i = (\hat{\xi}_{i1}, \dots, \hat{\xi}_{in})'$ is made of the empirical univariate quantiles $\hat{\xi}_{ij}$.

Proposition 3.1. *The random vector $\sqrt{T}(\hat{\mathbf{D}}_1 - \mathbf{D}_1)$, resp. $\sqrt{T}(\hat{\bar{\mathbf{D}}}_1 - \bar{\mathbf{D}}_1)$, converges in distribution to a d -dimensional normal random variable with mean zero and covariance matrix \mathbf{V}_1 , resp. $\bar{\mathbf{V}}_1$, whose elements are*

$$v_{1,kl} = \lim_{T \rightarrow \infty} T \text{Cov}_0[\hat{D}_1^k, \hat{D}_1^l] = B_{\theta_0}^{k'} \text{Cov}_0[S_{\theta_0}^k, S_{\theta_0}^l] B_{\theta_0}^l, \quad k, l = 1, \dots, d,$$

with

$$B_{\theta_0}^i = \left(\nabla_{\theta_0}^i C' J_{\theta_0}^{-1}, -1, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_1}}{f_{10}(\zeta_{i1})}, \dots, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_n}}{f_{n0}(\zeta_{in})} \right)',$$

$$S_{\theta_0}^i = \left(\frac{\partial}{\partial \theta'} \log f(\mathbf{Y}; v_0), \mathbb{I}[\mathbf{Y} \leq \zeta_i], \mathbb{I}[Y_1 \leq \zeta_{i1}], \dots, \mathbb{I}[Y_n \leq \zeta_{in}] \right)',$$

and

$$\bar{v}_{1,kl} = \lim_{T \rightarrow \infty} T \text{Cov}_0[\hat{\bar{D}}_1^k, \hat{\bar{D}}_1^l] = \bar{B}_{\theta_0}^{k'} \text{Cov}_0[\bar{S}_{\theta_0}^k, \bar{S}_{\theta_0}^l] \bar{B}_{\theta_0}^l, \quad k, l = 1, \dots, d,$$

with

$$\bar{B}_{\theta_0}^i = \left(\nabla_{\theta_0}^i \bar{C}' J_{\theta_0}^{-1}, -1, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_1}}{f_{10}(\zeta_{i1})}, \dots, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_n}}{f_{n0}(\zeta_{in})} \right)',$$

$$\bar{S}_{\theta_0}^i = \left(\frac{\partial}{\partial \theta'} \log f(\mathbf{Y}; v_0), \mathbb{I}[\mathbf{Y} > \zeta_i], \mathbb{I}[Y_1 \leq \zeta_{i1}], \dots, \mathbb{I}[Y_n \leq \zeta_{in}] \right)',$$

where

$$\nabla_{\theta_0}^i C = \frac{\partial}{\partial \theta} C(\mathbf{u}_i; \theta_0), \quad J_{\theta_0} = E_0 \left[-\frac{\partial^2}{\partial \theta \partial \theta'} \log f(\mathbf{Y}; v_0) \right],$$

while the elements of the cross covariance matrix $\mathbf{C}\mathbf{V}_1$ are

$$cv_{1,kl} = \lim_{T \rightarrow \infty} T \text{Cov}_0[\hat{D}_1^k, \hat{\bar{D}}_1^l] = B_{\theta_0}^{k'} \text{Cov}_0[S_{\theta_0}^k, \bar{S}_{\theta_0}^l] \bar{B}_{\theta_0}^l, \quad k, l = 1, \dots, d.$$

Some of the asymptotic covariances involve derivatives of F_0 and the univariate densities f_{j0} . These quantities may be estimated by standard kernel methods (see e.g. SCOTT (1992)) in order to deliver a consistent covariance estimate. For example we may take a Gaussian kernel and different bandwidth values h_j in each dimension, which leads to:

$$\frac{\partial \hat{F}(\xi_i)}{\partial x_j} = (Th_j)^{-1} \sum_{t=1}^T \varphi \left(\frac{Y_{jt} - \hat{\xi}_{ij}}{h_j} \right) \prod_{l \neq j}^n \Phi \left(\frac{Y_{lt} - \hat{\xi}_{il}}{h_l} \right),$$

$$\hat{f}_j(\hat{\xi}_{ij}) = (Th_j)^{-1} \sum_{t=1}^T \varphi \left(\frac{Y_{jt} - \hat{\xi}_{ij}}{h_j} \right),$$

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 . Note that since the marginals are assumed to have a correct parametric specification, the univariate densities f_{j0} can also be estimated by $f_j(\hat{\xi}_{ij}; \hat{\beta}_j)$.

4. INFERENCE UNDER SEMIPARAMETRIC SPECIFICATION

The previous section was devoted to the fully parametric specification. If we wish to be less restrictive a priori on the univariate margins, we may leave them unspecified, and use the family

$$\{F(\mathbf{y}; \theta) = C(F_1(y_1), \dots, F_n(y_n); \theta), \theta \in \Theta \subset \mathbb{R}^p\}.$$

Hence we get a semiparametric setting only parameterised through $C(\mathbf{u}; \theta)$. The estimator $\hat{\theta}$ of θ is obtained by

$$\max_{\theta} \frac{1}{T} \sum_{t=1}^T \ln c(\hat{F}_1(Y_{1t}), \dots, \hat{F}_n(Y_{nt}); \theta),$$

where

$$\hat{F}_j(y) = \frac{1}{T} \sum_{t=1}^T \mathbb{I}[Y_{jt} \leq y], \quad j = 1, \dots, n. \quad (4.1)$$

Its limit is denoted by θ_0^* , and will correspond to θ_0 (the true value) if both copula and margins are well specified in the parametric case. The asymptotic distribution of $\hat{\theta}$ under correct specification is given in GENEST, GHOUDI and RIVEST (1995) and SHIH and LOUIS (1995) (its distribution in the misspecified case is established in Appendix A.2). Again we wish to check whether $F(\cdot; \theta_0^*)$ is more concordant than the true distribution function $F_0(\cdot)$, namely

$$F_0(\mathbf{y}) \leq F(\mathbf{y}; \theta_0^*) \quad \text{and} \quad \bar{F}_0(\mathbf{y}) \leq \bar{F}(\mathbf{y}; \theta_0^*), \quad \forall \mathbf{y} \in \mathbb{R}^n. \quad (4.2)$$

We use

$$\hat{D}_2^i = C(\mathbf{u}_i; \hat{\theta}) - \hat{F}(\hat{\xi}_i),$$

and

$$\hat{D}_2^i = \bar{C}(\mathbf{u}_i; \tilde{\theta}) - \hat{F}(\hat{\xi}_i).$$

This leads to the following proposition which is equivalent to Proposition 3.1 of the parametric setting.

Proposition 4.1. *The random vector $\sqrt{T}(\hat{\mathbf{D}}_2 - \mathbf{D}_2)$, resp. $\sqrt{T}(\hat{\hat{\mathbf{D}}}_2 - \bar{\mathbf{D}}_2)$, converges in distribution to a d -dimensional normal random variable with mean zero and covariance matrix \mathbf{V}_2 , resp. $\bar{\mathbf{V}}_2$, whose elements are*

$$v_{2,kl} = \lim_{T \rightarrow \infty} T \text{Cov}_0[\hat{D}_2^k, \hat{D}_2^l] = B_{\theta_0^*}^{k'} \text{Cov}_0[S_{\theta_0^*}^k, S_{\theta_0^*}^l] B_{\theta_0^*}^l, \quad k, l = 1, \dots, d,$$

with

$$B_{\theta_0^*}^i = \left(\nabla_{\theta_0^*}^i C' J_{\theta_0^*}^{-1}, -1, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_1}}{f_{10}(\zeta_{i1})}, \dots, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_n}}{f_{n0}(\zeta_{in})} \right)',$$

$$S_{\theta_0^*}^i = \left(U_{\theta_0^*}'^i, \mathbb{I}[Y \leq \zeta_i], \mathbb{I}[Y_1 \leq \zeta_{i1}], \dots, \mathbb{I}[Y_n \leq \zeta_{in}] \right)',$$

and

$$\bar{v}_{2,kl} = \lim_{T \rightarrow \infty} T \text{Cov}_0[\bar{D}_2^k, \bar{D}_2^l] = \bar{B}_{\theta_0^*}^{k'} \text{Cov}_0[\bar{S}_{\theta_0^*}^k, \bar{S}_{\theta_0^*}^l] \bar{B}_{\theta_0^*}^l, \quad k, l = 1, \dots, d,$$

with

$$\bar{B}_{\theta_0^*}^i = \left(\nabla_{\theta_0^*}^i \bar{C}' J_{\theta_0^*}^{-1}, -1, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_1}}{f_{10}(\zeta_{i1})}, \dots, \frac{\frac{\partial F_0(\zeta_i)}{\partial x_n}}{f_{n0}(\zeta_{in})} \right)',$$

$$\bar{S}_{\theta_0^*}^i = \left(U_{\theta_0^*}'^i, \mathbb{I}[Y > \zeta_i], \mathbb{I}[Y_1 \leq \zeta_{i1}], \dots, \mathbb{I}[Y_n \leq \zeta_{in}] \right)',$$

where

$$J_{\theta_0^*} = E_0 \left[-\frac{\partial^2}{\partial \theta \partial \theta'} \log c(F_{10}(Y_1), \dots, F_{n0}(Y_n); \theta_0^*) \right]$$

and

$$U_{\theta_0^*} = \frac{\partial}{\partial \theta} \log c(F_{10}(Y_1), \dots, F_{n0}(Y_n); \theta_0^*)$$

$$+ \sum_{j=1}^n \int_{\mathbb{R}^n} \mathbb{I}[Y_j \leq z_j] \frac{\partial^2}{\partial \theta \partial u_j} \log c(F_{10}(z_1), \dots, F_{n0}(z_n); \theta_0^*) dF_0(z_1, \dots, z_n)$$

while

$$\lim_{T \rightarrow \infty} T \text{Cov}_0[\hat{D}_2^k, \hat{D}_2^l] = B_{\theta_0^*}^{k'} \text{Cov}_0[S_{\theta_0^*}^k, \tilde{S}_{\theta_0^*}^l] \tilde{B}_{\theta_0^*}^l, \quad k, l = 1, \dots, d.$$

5. TESTING PROCEDURES

The distributional results of Propositions 3.1 and 4.1 are the building blocks of the testing procedures. Let \mathbf{Z}_k , resp. $\hat{\mathbf{Z}}_k$, be the stack of \mathbf{D}_k , resp. $\hat{\mathbf{D}}_k$, and $\tilde{\mathbf{D}}_k$, resp. $\tilde{\mathbf{D}}_k$, $k = 1, 2$.

The null hypothesis of a test for concordance may be written as

$$H_k^0 = \{\mathbf{Z}_k : \mathbf{Z}_k \geq 0\},$$

with alternative hypothesis:

$$H_1^k = \{\mathbf{Z}_k : \mathbf{Z}_k \text{ unrestricted}\}.$$

The vector inequality in the definition of the null hypothesis means that each component of the vector should be positive. To examine such an hypothesis, 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)). They are relevant when one or several components of $\hat{\mathbf{Z}}_k$ are found to be negative (in such a case one wants to know whether this invalidates concordance).

Let $\tilde{\mathbf{Z}}_k$, be solution of the constrained quadratic minimisation problem:

$$\inf_{\mathbf{Z}} T(\mathbf{Z} - \hat{\mathbf{Z}}_k)' \hat{\Sigma}_k^{-1} (\mathbf{Z} - \hat{\mathbf{Z}}_k) \quad \text{s.t.} \quad \mathbf{Z} \geq 0, \quad (5.1)$$

where $\hat{\Sigma}_k$ is a consistent estimate of the asymptotic covariance matrix Σ_k of $\sqrt{T} \hat{\mathbf{Z}}_k$, and put

$$\hat{\xi}_k = T(\tilde{\mathbf{Z}}_k - \hat{\mathbf{Z}}_k)' \hat{\Sigma}_k^{-1} (\tilde{\mathbf{Z}}_k - \hat{\mathbf{Z}}_k).$$

Roughly speaking, $\tilde{\mathbf{Z}}_k$ is the closest point to $\hat{\mathbf{Z}}_k$ under the null hypothesis in the distance measured via the metric of $\hat{\Sigma}_k$, and the test statistic $\hat{\xi}_k$ is the distance between $\tilde{\mathbf{Z}}_k$ and $\hat{\mathbf{Z}}_k$. The idea is to reject H_0^k when this distance becomes too large.

The asymptotic distribution of $\hat{\xi}_k$ under the null hypothesis (see e.g. GOURIEROUX, HOLLY and MONFORT (1982), KODDE and PALM (1986), WOLAK (1989a, b)) is such that for any positive x :

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

where the weight $w(2d, 2d-i, \hat{\Sigma}_k)$ is the probability that $\tilde{\mathbf{Z}}_k$ has exactly $2d-i$ positive elements.

Computation of the solution $\tilde{\mathbf{Z}}_k$ can be performed by a numerical optimisation routine for constrained quadratic programming problems available in most statistical softwares. Closed form solutions for the weights are available for $2d \leq 4$ (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 $\hat{\Sigma}_k$. Then use these realisations as $\hat{\mathbf{Z}}_k$ in the above minimisation problem (5.1), compute $\tilde{\mathbf{Z}}_k$, and count the number of elements of the vector greater than zero. The proportion of draws such that $\tilde{\mathbf{Z}}_k$ has exactly $2d-i$ elements greater than zero gives a Monte Carlo estimate of $w(2d, 2d-i, \hat{\Sigma}_k)$. 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 for non-concordance. It is based on the null hypothesis:

$$\tilde{H}_k^0 = \{\mathbf{Z}_k : Z_k^l \leq 0 \text{ for some } l, l = 1, \dots, 2d\},$$

and the alternative hypothesis:

$$\tilde{H}_k^1 = \{\mathbf{Z}_k : Z_k^l > 0 \text{ for all } l\}.$$

These hypotheses will be tested through intersection-union tests based on the minimum of a t -statistic. They are used when all components of $\hat{\mathbf{Z}}_k$ are found to be positive. The question is then whether this suffices to ensure concordance.

Let $\hat{\gamma}_k^l = \sqrt{T} \hat{\mathbf{Z}}_k^l / \sqrt{\hat{\sigma}_{k,l}}$, where $\hat{\sigma}_{k,l}$ is a consistent estimate of the asymptotic standard deviation of $\sqrt{T} \hat{\mathbf{Z}}_k^l$, $l = 1, \dots, 2d$. Then under \tilde{H}_k^0 , the limit of $P[\inf \hat{\gamma}_k^l > z_{1-\alpha}]$ will be less or equal to α , and exactly equal to α if $Z_k^l = 0$ for a given l and $Z_k^s > 0$ for $s \neq l$, while its limit is one under \tilde{H}_k^1 . Hence the test consisting of rejecting \tilde{H}_k^0 when $\inf \hat{\gamma}_k^l$ 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 non-dominance tests in DARDANONI and FORCINA (1999) (see also the comments in DAVIDSON and DUCLOS (2000)). They carry over to our case. In principle the second testing procedure should have better properties since its definition of the null and alternative hypotheses is a better indication of the “truth” of the hypothesis of concordance. However two points are worth being mentioned. 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 thorough 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 hypothesis of dominance, here concordance, by distance tests occurs along with the nonrejection of the null hypothesis of nondominance, here

non-concordance, by intersection-union tests. This is due to the highly conservative nature of the latter, and will typically occur in our setting if \hat{Z}_k is close enough to zero for a number of coordinates. This empirical feature has already been observed on tests for positive quadrant dependence (PQD) in DENUIT and SCAILLET (2002).

6. AN EMPIRICAL ILLUSTRATION: US LOSSES AND ALAE'S

6.1. Presentation of the data

Often insurance processes involve correlated pairs of variables. A fine example is the LOSS and allocated adjustment expenses (ALAE, in short) on a single claim. ALAE's 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 is examined by FREES and VALDEZ (1998) who choose the Pareto distribution to model the margins and select Gumbel and Frank's copulas. Both models express PQD by their estimated parameter values. KLUGMAN and PARSA (1999) opt for the Inverse Paralogistic for LOSS and for the Inverse Burr for ALAE's and use Frank's copula to model the dependence between them. DENUIT and SCAILLET (2002) test the existence of PQD for LOSS and ALAE using a nonparametric approach and find that, as both previous models suggest, significant positive quadrant dependence exists.

The database we have considered consist in $T = 1,466$ uncensored observed values of the random vector $(LOSS, ALAE)$. The estimated values for Pearson's r , Kendall's τ and Spearman's ρ are 0.381, 0.307 and 0.444, respectively; all of them are significantly positive at a 1% level. Summary statistics for $(LOSS, ALAE)$ are provided in Table 6.1.

TABLE 6.1.
SUMMARY STATISTICS FOR VARIABLES LOSS AND ALAE.

	<i>LOSS</i>	<i>ALAE</i>
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

Because some very high values of the variables are contained in the data set, we will work on a logarithmic scale to represent the data. This will not affect

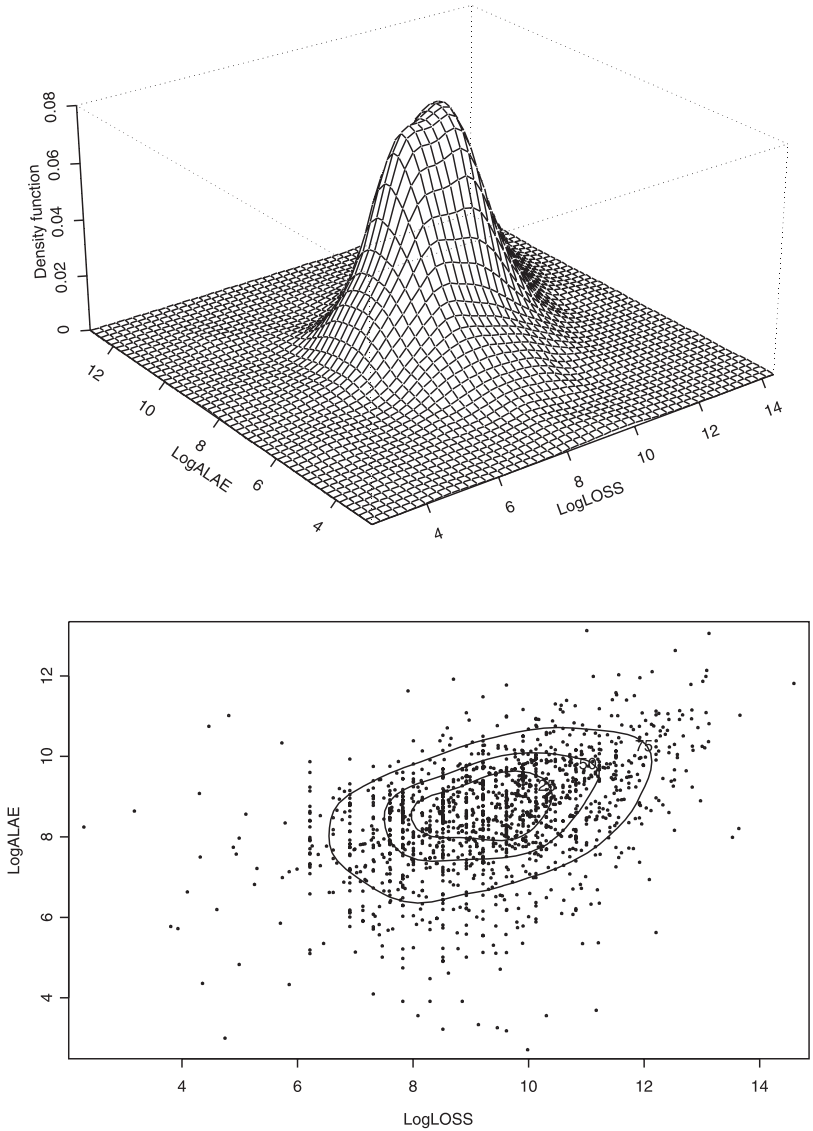


FIGURE 6.1: Kernel estimation of the bivariate pdf for $(\log(\text{LOSS}), \log(\text{ALAE}))$.

testing for concordance ordering since this order enjoys a functional invariance property (cf. Section 3.2). Figure 6.1 shows the kernel estimator of the bivariate pdf of the couple $(\log(\text{LOSS}), \log(\text{ALAE}))$, 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.

6.2. Inference under parametric specification

First, the parametric framework suggested by FREES and VALDEZ (1998) is studied. It relies on a Gumbel copula

$$C(u_1, u_2; \theta) = \exp \left\{ - \left[(-\ln(u_1))^\theta + (-\ln(u_2))^\theta \right]^{1/\theta} \right\},$$

for the dependence structure and Pareto distributions

$$F_i(x) = 1 - \left(1 + \xi_i \frac{x}{\gamma_i} \right)^{-1/\xi_i}, \quad i = 1, 2,$$

for the marginal behaviours.

Estimated values for the parameter $\nu = (\xi_1, \gamma_1, \xi_2, \gamma_2, \theta)'$ are shown in Table 6.2.

TABLE 6.2.

ESTIMATED PARAMETER VALUES OF THE BIVARIATE DISTRIBUTION OF (LOSS, ALAE).

LOSS	Pareto	$\hat{\xi}_1 = 0.760, \hat{\beta}_1 = 12,816.9$
ALAE	Pareto	$\hat{\xi}_2 = 0.425, \hat{\beta}_2 = 6,756.5$
Copula	Gumbel	$\hat{\theta} = 1.425$

In the testing procedures we use 81 points built on the grid of probability levels $\{0.1, 0.2, \dots, 0.9\} \times \{0.1, 0.2, \dots, 0.9\}$. Since 105 of the 162 components of the vector $\hat{\mathbf{Z}}_1$ are negative, with 30 among them less than -0.1 , a concordance test is applied. See Figures 6.2 and 6.3 for a representation of the Gumbel and the empirical estimation of the copula function and the difference between both estimations to make comparisons easier.

To compute the solution $\tilde{\mathbf{Z}}_1$ of the quadratic minimisation problem (5.1), a local minimiser for nonlinear functions subject to boundary constraints is used (specifically, `nlminb` in S-plus). Since $\tilde{\mathbf{Z}}_1$ represents in a way the closest

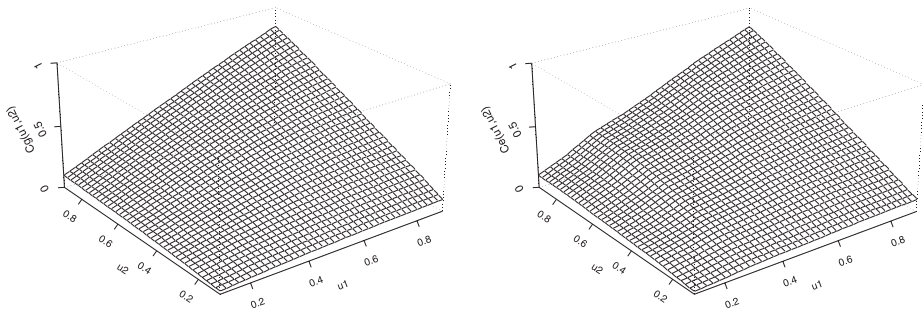


FIGURE 6.2: Plot of the estimation of the copula function $C(u_1, u_2)$: Gumbel estimation C_g (left), empirical estimation C_e (right).

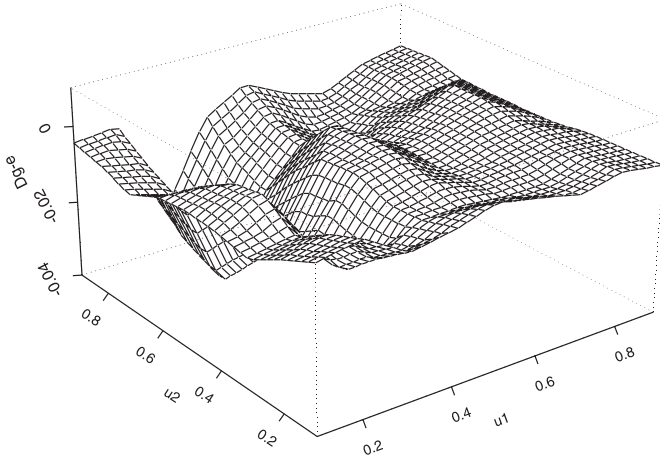


FIGURE 6.3: Plot of the difference between the Gumbel and the empirical estimations of the copula function, $D_{g-e} = C_g(u_1, u_2) - C_e(u_1, u_2)$.

point to $\hat{\mathbf{Z}}_1$ under the null hypothesis, we take the vector $\max(0, \hat{\mathbf{Z}}_1)$ as starting point for the numerical optimisation routine. This initial value satisfies the boundary restrictions. The minimum value of the function is then found to be $\xi_1 = 2.096$.

According to the bounds given in KODDE and PALM (1986), the null hypothesis of a greater concordance of the fitted distribution cannot be rejected at any level lower than 5%. This indicates that the amount of positive dependence expressed by the parametric framework is at least as large as that suggested by the data. This is particularly appealing to actuaries since it ensures that most actuarial quantities computed in the Gumbel-Pareto model will not be underestimated.

It can also be of interest to test the concordance behaviour only in the upper tails. We consider a 81 grid formed by the percentiles in $\{0.91, 0.92, \dots, 0.99\} \times \{0.91, 0.92, \dots, 0.99\}$. See Figures 6.4 and 6.5 for a representation of

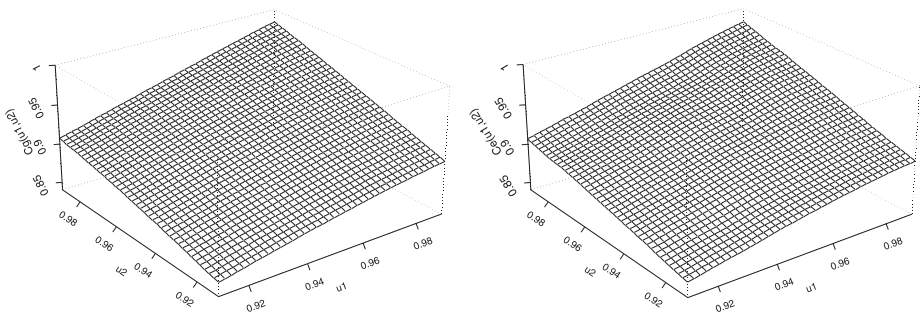


FIGURE 6.4: Plot of the estimation of the copula function $C(u_1, u_2)$ in the upper tail: Gumbel estimation C_g (left), empirical estimation C_e (right).

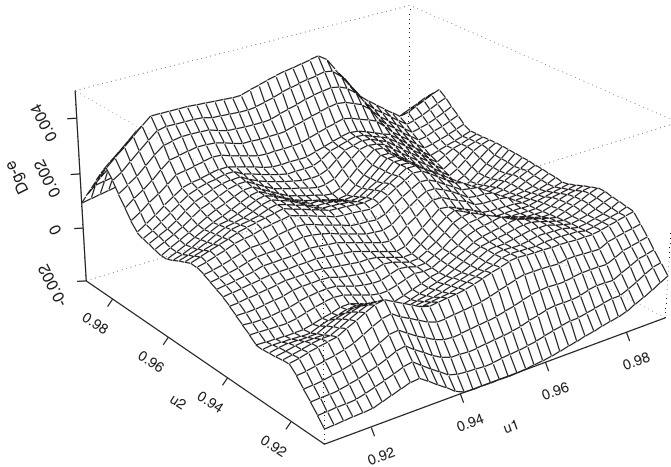


FIGURE 6.5: Plot of the difference between the Gumbel and the empirical estimations of the copula function, $D_{g-e} = C_g(u_1, u_2) - C_e(u_1, u_2)$, in the upper tail.

the estimations of the copula function. In this case only 10 components of \hat{Z}_1 are found to be negative and the minimum value of the function is $\hat{\xi}_1 = 0.000035$. Thus, again the null hypothesis cannot be rejected.

6.3. Inference under semiparametric specification

In this section we wish to test the same type of hypothesis than in the previous subsection but using the semiparametric approach. We thus drop the Pareto modelling of the marginals and leave them unspecified. Note, in Figures 6.6 and 6.7, the similarity between the parametric and empirical estimations of the cumulative distribution function (CDF) of both margins, LOSS and ALAE. This explains why the estimated value $\hat{\theta} = 1.415$ of the Gumbel copula is not much affected.

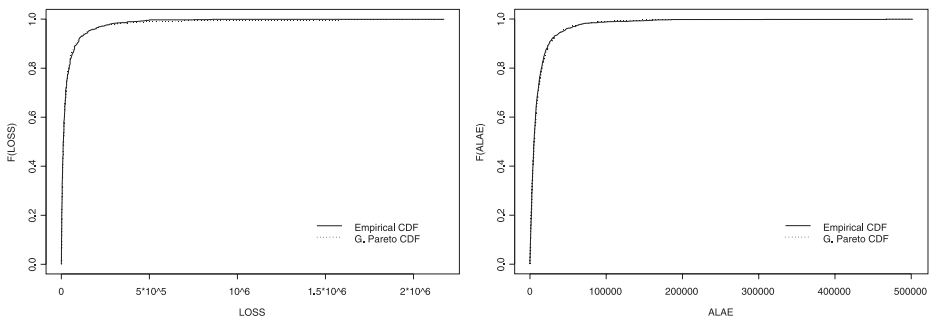


FIGURE 6.6: Parametric (Generalized Pareto) and empirical estimations of the CDF of LOSS (left) and ALAE (right).

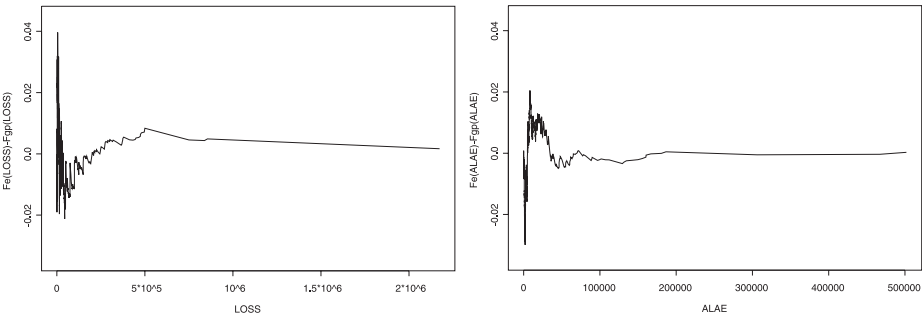


FIGURE 6.7: Difference between the Generalized Pareto and the empirical estimations of the CDF of LOSS (left) and ALAE (right).

As before we perform inference based on the grid $\{0.1, 0.2, \dots, 0.9\} \times \{0.1, 0.2, \dots, 0.9\}$. 6.8 displays the differences between the semiparametric and empirical estimations and the semiparametric and parametric estimations of the copula on our data. Note again the small difference between the semiparametric and parametric estimations. We thus expect to get the same conclusion under the semiparametric framework as under the parametric one.

The minimum value of the function is now found to be $\hat{\xi}_2 = 14.161$. Since this value does not allow us to get a conclusion about its significance using the bounds of KODDE and PALM (1986), we need to rely on the simple Monte Carlo technique described in Section 6. A p -value equal to 0.98 has been obtained which clearly yields to not reject the null hypothesis of concordance.

Besides the results about the concordance behaviour in the upper tails are equivalent to the ones from the parametric approach using the same grid. Differences between the empirical and the semiparametric estimations and between the semiparametric and the parametric estimations of the copula function are shown in Figure 6.9. Only 12 components of $\hat{\mathbf{Z}}_2$ are negative and the minimum value of the function is $\hat{\xi}_2 = 0.000036$, which does not allow to reject the null hypothesis.

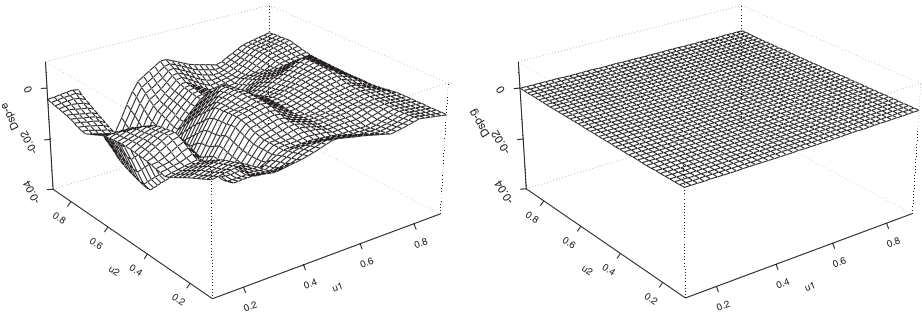


FIGURE 6.8: Plot of the difference between the semiparametric (Gumbel) and the empirical estimations of the copula function, $D_{sp-e} = C_{sp}(u_1, u_2) - C_e(u_1, u_2)$ (left) and between the semiparametric (Gumbel) and the parametric (Gumbel) estimations of the copula function, $D_{sp-g} = C_{sp}(u_1, u_2) - C_g(u_1, u_2)$ (right).

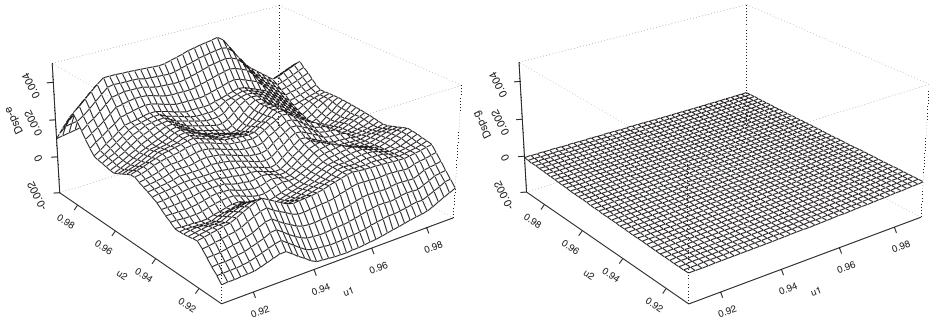


FIGURE 6.9: Plot of the difference in the upper tail between the semiparametric (Gumbel) and the empirical estimations of the copula function, $D_{sp-e} = C_{sp} - C_e$ (left) and between the semiparametric (Gumbel) and the parametric (Gumbel) estimations of the copula function, $D_{sp-g} = C_{sp} - C_g$ (right).

7. CONCLUDING REMARKS

In this paper we have analysed simple distributional free inference for concordance ordering. The testing procedures have proven to be empirically relevant to the analysis of dependencies among US insurance claim data. In particular they suggest that the Gumbel copula reflects the dependence structure in the data safely enough.

APPENDIX

A. ASYMPTOTIC DISTRIBUTIONS

We first derive the asymptotic distribution of the parametric estimator $\hat{v} = (\hat{\beta}', \hat{\theta}')'$ and the semiparametric estimator $\tilde{\theta}$ in a misspecified framework. For the well specified case the results can be found in GENEST, GHOUDI and RIVEST (1995) and SHIH and LOUIS (1995). Then we proceed with the asymptotic distribution of the various difference vectors \hat{D}_k and \tilde{D}_k , $k = 1, 2$.

A.1. Asymptotic distribution of the parametric estimator

The asymptotic distribution of \hat{v} immediately results from usual pseudo maximum likelihood theory (see e.g. WHITE (1982), GOURIÉROUX, MONFORT and TROGNON (1984)). Indeed from a standard Taylor expansion of the first order condition of the maximum likelihood criterion and the law of large numbers, we get:

$$\sqrt{T}(\hat{v} - v_0) = J_{v_0}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial}{\partial v} \log f(Y_t; v_0) + o_p(1),$$

with

$$J_{v_0} = E_0 \left[- \frac{\partial^2}{\partial v \partial v'} \log f(Y; v_0) \right],$$

where E_0 denotes expectation w.r.t. the true distribution F_0 , and by application of the central limit theorem

$$\sqrt{T}(\hat{v} - v_0) \Rightarrow N(0, J_{v_0}^{-1} I_{v_0} J_{v_0}^{-1}),$$

where

$$I_{v_0} = E_0 \left[\frac{\partial}{\partial v} \log f(Y; v_0) \frac{\partial}{\partial v'} \log f(Y; v_0) \right].$$

When the parametric model is well specified, i.e. $F(\cdot, v_0) = F_0(\cdot)$, we have $J_{v_0} = I_{v_0}$.

A.2. Asymptotic distribution of the semiparametric estimator

From a Taylor expansion of the first order condition of the maximum likelihood criterion and the law of large numbers, we get:

$$\sqrt{T}(\tilde{\theta} - \theta_0^*) = J_{\theta_0^*}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial}{\partial \theta} \log c(\hat{F}_1(Y_{1t}), \dots, \hat{F}_n(Y_{nt}); \theta_0^*) + o_p(1), \quad (\text{A.1})$$

where

$$J_{\theta_0^*} = E_0 \left[-\frac{\partial^2}{\partial \theta \partial \theta'} \log c(F_{10}(Y_1), \dots, F_{n0}(Y_n); \theta_0^*) \right].$$

The random part of the right hand side in (A.1) can be written:

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial}{\partial \theta} \log c(\hat{F}_1(Y_{1t}), \dots, \hat{F}_n(Y_{nt}); \theta_0^*) \\ &= \sqrt{T} \int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} \log c(\hat{F}_1(z_1), \dots, \hat{F}_n(z_n); \theta_0^*) d\hat{F}(z_1, \dots, z_n), \end{aligned}$$

and decomposed into three terms:

$$\begin{aligned} & \sqrt{T} \int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} \log c(\hat{F}_1(z_1), \dots, \hat{F}_n(z_n); \theta_0^*) d\hat{F}(z_1, \dots, z_n) \\ &= \sqrt{T} \int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} \log c(\hat{F}_1(z_1), \dots, \hat{F}_n(z_n); \theta_0^*) dF_0(z_1, \dots, z_n) \\ &+ \sqrt{T} \int_{\mathbb{R}^n} \left[\frac{\partial}{\partial \theta} \log c(\hat{F}_1(z_1), \dots, \hat{F}_n(z_n); \theta_0^*) - \frac{\partial}{\partial \theta} \log c(F_{10}(z_1), \dots, F_{n0}(z_n); \theta_0^*) \right] \\ &\quad \times d[\hat{F}(z_1, \dots, z_n) - F_0(z_1, \dots, z_n)] \\ &+ \sqrt{T} \int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} \log c(F_{10}(z_1), \dots, F_{n0}(z_n); \theta_0^*) d[\hat{F}(z_1, \dots, z_n) - F_0(z_1, \dots, z_n)]. \end{aligned}$$

The second term converges to zero, and by the central limit theorem the third term converges to $N(0, I_{\theta_0^*})$, where

$$I_{\theta_0^*} = E_0 \left[\frac{\partial}{\partial \theta} \log c(F_1(Y_1), \dots, F_n(Y_n); \theta_0^*) \frac{\partial}{\partial \theta'} \log c(F_1(Y_1), \dots, F_n(Y_n); \theta_0^*) \right].$$

Now a Taylor expansion of the first term (see SERFLING (1980) for expansion of von Mises differentiable statistical functions) leads to

$$\begin{aligned} & \sqrt{T} \int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} \log c(\hat{F}_1(z_1), \dots, \hat{F}_n(z_n); \theta_0^*) dF_0(z_1, \dots, z_n) \\ &= \sqrt{T} \int_{\mathbb{R}^n} \frac{\partial}{\partial \theta} \log c(F_{10}(z_1), \dots, F_{n0}(z_n); \theta_0^*) dF_0(z_1, \dots, z_n) \\ &+ \sqrt{T} \sum_{j=1}^n \int_{\mathbb{R}^n} \frac{\partial^2}{\partial \theta \partial u_j} \log c(F_{10}(z_1), \dots, F_{n0}(z_n); \theta_0^*) dF_0(z_1, \dots, z_n) \\ &\quad \times (\hat{F}_j(z_j) - F_{j0}(z_j)) + o_p(1) \\ &= 0 + \sqrt{T} \sum_{j=1}^n \int_{\mathbb{R}} \left[\int_{\mathbb{R}^n} \mathbb{I}[w_j \leq z_j] \frac{\partial^2}{\partial \theta \partial u_j} \log c(F_{10}(z_1), \dots, F_{n0}(z_n); \theta_0^*) dF_0(z_1, \dots, z_n) \right] \\ &\quad \times d[\hat{F}_j(w_j) - F_{j0}(w_j)] + o_p(1). \end{aligned}$$

Hence by the central limit theorem the first term converges to $N(0, M_{\theta_0^*})$, where

$$M_{\theta_0^*} = \mathbb{V}\text{ar}_0 \left[\sum_{j=1}^n \int_{\mathbb{R}^q} \mathbb{I}[Y_j \leq z_j] \frac{\partial^2}{\partial \theta \partial u_j} \log c(F_{10}(z_1), \dots, F_{n0}(z_n); \theta_0^*) dF_0(z_1, \dots, z_n) \right].$$

Since the conditional expectation of $\frac{\partial}{\partial \theta} \log c(F_{10}(Y_1), \dots, F_{n0}(Y_n); \theta_0^*)$ w.r.t. any Y_j is null, the first term and the third term are uncorrelated, and we finally get:

$$\sqrt{T}(\tilde{\theta} - \theta_0^*) \implies N\left(0, J_{\theta_0^*}^{-1}(I_{\theta_0^*} + M_{\theta_0^*})J_{\theta_0^*}^{-1}\right).$$

When the parametric model is well specified, i.e. $F(\cdot; \theta_0^*) = F_0(\cdot)$, we have $J_{\theta_0^*} = I_{\theta_0^*}$.

A.3. Asymptotic distribution of the difference vectors

A.3.1. Asymptotic distribution of \hat{D}_1 and $\hat{\bar{D}}_1$

A Taylor expansion of $C(\mathbf{u}_i; \hat{\theta})$ around θ_0 gives:

$$C(\mathbf{u}_i; \hat{\theta}) = C(\mathbf{u}_i; \theta_0) + \frac{\partial}{\partial \theta} C(\mathbf{u}_i; \bar{\theta}_0)'(\hat{\theta} - \theta_0) + o_p(T^{-1/2}),$$

where $\bar{\theta}_0$ lies between $\hat{\theta}$ and θ_0 .

We get using Subsection A.1:

$$\sqrt{T}(C(\mathbf{u}_i; \hat{\theta}) - C(\mathbf{u}_i; \theta_0)) = \nabla_{\theta_0}^i C' J_{\theta_0}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial}{\partial \theta} \log f(Y_t; v_0) + o_p(1).$$

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

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

is stochastically differentiable at ζ_i with random derivative $(n \times 1)$ -vector $D\hat{F}(\zeta_i)$ (see e.g. POLLARD (1985), ANDREWS (1994, 1999) for definition, use and check of stochastic differentiability). It means that we have the approximation:

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

where $\bar{\zeta}_i$ is a mean value located between $\hat{\zeta}_i$ and ζ_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 $F_{j0}(\zeta_{ij}) = u_{ij} = \hat{F}_j(\hat{\zeta}_{ij})$ leads to

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

where S_i is the stack of $(\hat{F}_j(\zeta_{ij}) - u_{ij}) / D\hat{F}_j(\bar{\zeta}_{ij})$, $j = 1, \dots, n$. Hence we get:

$$\begin{aligned} & \sqrt{T}(\hat{D}_1^i - D_1^i) \\ &= \nabla_{\theta_0}^i C' J_{\theta_0}^{-1} \frac{1}{\sqrt{T}} \sum_{i=1}^T \frac{\partial}{\partial \theta} \log f(Y_i; v_0) - \sqrt{T}(\hat{F}(\zeta_i) - F_0(\zeta_i)) + D\hat{F}(\bar{\zeta}_i)' \sqrt{T} S_i + o_p(1). \end{aligned}$$

An application of the central limit theorem delivers:

$$\sqrt{T}(\hat{D}_1^i - D_1^i) \Rightarrow N\left(0, B_{\theta_0}^{i'} \text{Cov}_0[S_{\theta_0}^i, S_{\theta_0}^i] B_{\theta_0}^i\right).$$

We also get:

$$\sqrt{T}(\hat{D}_1^i - \bar{D}_1^i) \Rightarrow N\left(0, \bar{B}_{\theta_0}^{i'} \text{Cov}_0[\bar{S}_{\theta_0}^i, \bar{S}_{\theta_0}^i] \bar{B}_{\theta_0}^i\right).$$

A.3.2. Asymptotic distribution of \hat{D}_2 and $\hat{\bar{D}}_2$

The only difference between \hat{D}_1^i and \hat{D}_2^i lies in the replacement of the parametric estimator $\hat{\theta}$ by the semiparametric estimator $\hat{\theta}$. Hence the asymptotic distribution of \hat{D}_2^i is obtained after substituting $\nabla_{\theta_0}^i C' J_{\theta_0}^{-1}$ for $\nabla_{\theta_0}^i C' J_{\theta_0}^{-1}$, and $U_{\theta_0}^i$ for $\frac{\partial}{\partial \theta} \log f(Y; \theta_0)$ in the asymptotic results for \hat{D}_1^i .

Similarly in order to derive the asymptotic normality of $\hat{\bar{D}}_2^i$, we only have to replace $\nabla_{\theta_0}^i \bar{C}' J_{\theta_0}^{-1}$ with $\nabla_{\theta_0}^i \bar{C}' J_{\theta_0}^{-1}$, and $\frac{\partial}{\partial \theta} \log f(Y; \theta_0)$ with $U_{\theta_0}^i$ in the asymptotic results for $\hat{\bar{D}}_1^i$.

ACKNOWLEDGEMENTS

The authors would like to thank Paul Embrechts and the two referees for constructive criticism, as well as Professors Frees and Valdez for kindly providing the Loss-ALAE data, which were collected by the US Insurance Services Office (ISO). We are also grateful to participants at the copula workshop in Kaiserslautern and the common features conference in Rio for helpful comments. Ana Cebrián thanks the Université catholique de Louvain, Louvain-la-Neuve, Belgium, for financial support through an FSR research grant. Michel Denuit acknowledges financial support of the Belgian Government under “Projet d’Actions de Recherche Concertées” (No. 98/03-217). Olivier Scaillet receives support by the Swiss National Science Foundation through the National Center of Competence: Financial Valuation and Risk Management (NCCR FINRISK). Part of this research was done when he was visiting THEMA and IRES.

Downloadable at

http://www.hec.unige.ch/professeurs/SCAILLET_Olivier/pages_web/Home_Page_of_Olivier_Scaillet.htm

REFERENCES

- ANDREWS, D. (1994) "Empirical Process Methods in Econometrics", in *Handbook of Econometrics*, Vol. IV, eds R. Engle and D. McFadden, 2247-2294.
- ANDREWS, D. (1999) "Estimation when a Parameter is on a Boundary", *Econometrica*, **67**, 1341-1384.
- BARTHOLOMEW, D. (1959a) "A Test of Homogeneity for Ordered Alternatives, I", *Biometrika*, **46**, 36-48.
- BARTHOLOMEW, D. (1959b) "A Test of Homogeneity for Ordered Alternatives, II", *Biometrika*, **46**, 328-335.
- CAMBANIS, S., SIMONS, G. and STOUT, W. (1976) "Inequalities for $E_k(X, Y)$ when the Marginals are Fixed", *Z. Wahrsch. Verw. Gebiete*, **36**, 285-294.
- CEBRIAN, A., DENUIT, M. and SCAILLET, O. (2002) "Testing for Concordance Ordering", FAME DP 41.
- DARDANONI, V. and FORCINA, A. (1999) "Inference for Lorenz Curve Orderings", *Econometrics Journal*, **2**, 48-74.
- DAVIDSON, R. and DUCLOS, J.-Y. (2000) "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality", *Econometrica*, **68**, 1435-1464.
- DENUIT, M. and SCAILLET, O. (2002) "Nonparametric Tests for Positive Quadrant Dependence", FAME DP 44.
- DHAENE, J. and GOOVAERTS, M. (1996) "Dependency of Risks and Stop-Loss Order", *ASTIN Bulletin*, **26**, 201-212.
- EMBRECHTS, P., MCNEIL, A. and STRAUMANN, D. (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.
- FREES, E. and VALDEZ, E. (1998) "Understanding Relationships Using Copulae", *North American Actuarial Journal*, **2**, 1-25.
- GENEST, C., GHOUDI, K. and RIVEST, L.-P. (1998) "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions", *Biometrika*, **82**, 543-552.
- GOURIÉROUX, C., HOLLY, A. and MONFORT, A. (1982) "Likelihood Ratio, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters", *Econometrica*, **50**, 63-80.
- GOURIÉROUX, C., MONFORT, A. and TROGNON, A. (1984) "Pseudo Maximum Likelihood Theory", *Econometrica*, **52**, 681-700.
- HOWES, S. (1993) "Inferring Population Rankings from Sample Data", STICERD discussion paper.
- JOE, H. (1990) "Multivariate concordance". *Journal of Multivariate Analysis*, **35**, 12-30.
- JOE, H. (1997) *Multivariate Models and Dependence Concepts*, Chapman & Hall, London.
- KAUR, A., PRAKASA RAO, B. and SINGH, H. (1994) "Testing for Second-Order Dominance of Two Distributions", *Econometric Theory*, **10**, 849-866.
- KLUGMAN, S. and PARSIA, R. (1999) "Fitting Bivariate Loss Distributions with Copulas", *Insurance Mathematics and Economics*, **24**, 139-148.
- KODDE, D. and PALM, F. (1986) "Wald Criteria for Jointly Testing Equality and Inequality Restrictions", *Econometrica*, **54**, 1243-1248.
- KUDO, A. (1963) "A Multivariate Analogue for the One-Sided Test", *Biometrika*, **50**, 403-418.
- NELSEN, R. (1999) *An Introduction to Copulas*, Lecture Notes in Statistics, Springer-Verlag, New-York.
- POLLARD, D. (1985) "New Ways to Prove Central Limit Theorems", *Econometric Theory*, **1**, 295-314.
- SERFLING, R. (1980) *Approximation Theorems of Mathematical Statistics*, Wiley, New-York.
- SHIH, J. and LOUIS, T. (1995) "Inferences on the Association Parameter in Copula Models for Bivariate Survival Data", *Biometrics*, **51**, 1384-1399.
- SCOTT, D. (1992) "Multivariate Density Estimation: Theory, Practice and Visualisation", John Wiley & Sons, New-York.
- SKLAR, A. (1959) "Fonctions de Répartition à n Dimensions et leurs Marges", *Publ. Inst. Stat. Univ. Paris*, **8**, 229-231.

- TCHEN, A. (1980) "Inequalities for Distributions with Given Marginals", *The Annals of Probability*, **8**, 814-827.
- WHITE, H. (1982) "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, **50**, 1-26.
- WOLAK, F. (1989a) "Testing Inequality Constraints in Linear Econometric Models", *Journal of Econometrics*, **41**, 205-235.
- WOLAK, F. (1989b) "Local and Global of Linear and Nonlinear Inequality Constraints in Non-linear Econometric Models", *Econometric Theory*, **5**, 1-35.
- YANAGIMOTO, Y. and OKAMOTO, M. (1969) "Partial orderings of permutations and monotonicity of a rank correlation statistic", *Annals of the Institute of Statistical Mathematics*, **21**, 489-506.

ANA C. CEBRIÁN
Dpto. Métodos Estadísticos. Ed. Matemáticas
Facultad de Ciencias
Universidad de Zaragoza
CP. Cerbuna, 12
S-Zaragoza 50009, Spain
E-mail: acebrian@unizar.es

MICHEL DENUIT
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve, Belgium
E-mail: denuit@stat.ucl.ac.be

OLIVIER SCAILLET
HEC Genève and FAME
Université de Genève
Bd Carl Vogt, 102
CH-1211 Genève 4, Suisse
E-mail: olivier.scaillet@hec.unige.ch