doi:10.1111/j.1467-9469.2006.00470.x

Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation

CHRISTIAN GENEST

Département de mathématiques et de statistique, Université Laval

JEAN-FRANÇOIS QUESSY

Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières

BRUNO RÉMILLARD

GERAD and Service de l'enseignement des méthodes quantitatives de gestion, HEC Montréal

ABSTRACT. Wang & Wells [*J. Amer. Statist. Assoc.* 95 (2000) 62] describe a non-parametric approach for checking whether the dependence structure of a random sample of censored bivariate data is appropriately modelled by a given family of Archimedean copulas. Their procedure is based on a truncated version of the Kendall process introduced by Genest & Rivest [*J. Amer. Statist. Assoc.* 88 (1993) 1034] and later studied by Barbe *et al.* [*J. Multivariate Anal.* 58 (1996) 197]. Although Wang & Wells (2000) determine the asymptotic behaviour of their truncated process, their model selection method is based exclusively on the observed value of its L^2-norm. This paper shows how to compute asymptotic *p*-values for various goodness-of-fit test statistics based on a non-truncated version of Kendall's process. Conditions for weak convergence are met in the most common copula models, whether Archimedean or not. The empirical behaviour of the proposed goodness-of-fit tests is studied by simulation, and power comparisons are made with a test proposed by Shih [*Biometrika* 85 (1998) 189] for the gamma frailty family.

Key words: empirical process, Kendall's tau, probability integral transformation, pseudo-observation

1. Introduction

Due in part to their connection with frailty models in survival analysis (Oakes, 1989, 2001), Archimedean copulas have become quite popular as a tool for describing the dependence between two random variables X and Y with continuous marginal distributions F and G, respectively. Given a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ with joint cumulative distribution function

$$H(x, y) = C\{F(x), G(y)\}, \quad x, y \in \mathbb{R}$$

there is thus considerable interest in testing whether the unique underlying copula C belongs to a parametric class $\mathcal{C} = (C_{\phi_\theta})$ of Archimedean copulas

$$C_{\phi_\theta}(u, v) \equiv \phi_\theta^{-1}\{\phi_\theta(u) + \phi_\theta(v)\},$$

where $\phi_\theta : (0, 1] \mapsto [0, \infty)$ is a mapping, indexed by a parameter $\theta \in \mathbb{R}$, which satisfies the following conditions:

$$\phi_\theta(1) = 0, \quad (-1)^i \frac{d^i}{\mathrm{d}t^i}\, \phi_\theta^{-1}(t) > 0, \quad i \in \{1, 2\}.$$

Using the fact that the distribution function K of the probability integral transformation $V = H(X, Y)$ is of the form

$$K(\theta,t)=t-\frac{\phi_\theta(t)}{\phi'_\theta(t)},\quad t\in(0,1]$$

whenever $C\in\mathcal{C}$, Genest & Rivest (1993) proposed a graphical procedure for selecting an Archimedean model through a visual comparison of a non-parametric estimate K_n of K with a parametric estimate $K(\theta_n,\cdot)$ obtained under the composite null hypothesis $H_0: C\in\mathcal{C}$.

In their work, Genest & Rivest (1993) simply defined K_n as an empirical cumulative distribution function allocating a weight of $1/n$ to each pseudo-observation

$$\mathcal{V}_i=\frac{1}{n-1}\,\#\{j: X_j<X_i, Y_j<Y_i\}.$$

As for $K(\theta_n,\cdot)$, it was obtained by finding the value θ_n of θ such that, under H_0, the population value $\tau(\theta)=4E(V)-1$ of Kendall's tau matches its standard empirical version, given by $\tau_n=4\bar{\mathcal{V}}-1$, where $\bar{\mathcal{V}}=(\mathcal{V}_1+\cdots+\mathcal{V}_n)/n$. By identifying the pointwise limit of Kendall's process $\sqrt{n}\{K_n(\cdot)-K(\theta,\cdot)\}$, Genest & Rivest (1993) were able to construct confidence bands to help with the choice of a proper family $\mathcal{C}$. The limit of the process as such was later identified for arbitrary d-dimensional copulas by Barbe *et al.* (1996).

Restricting themselves to the bivariate Archimedean case, but allowing for censorship, Wang & Wells (2000) furthered this work by proposing a goodness-of-fit statistic

$$S_{\xi n}=\int_\xi^1\{\mathbb{K}_n(t)\}^2\,\mathrm{d}t,$$

which is a continuous functional of the process

$$\mathbb{K}_n(t)=\sqrt{n}\{K_n(t)-K(\theta_n,t)\}. \tag{1}$$

In order to avoid technical difficulties related to censorship and unboundedness of the density $k(\theta,\cdot)$ of $K(\theta,\cdot)$ at the origin, which is common in practice, their statistic involves an arbitrary cut-off point $\xi>0$. Mimicking the approach of Barbe *et al.* (1996), they were able to identify the limit of $\mathbb{K}_n$, and hence that of $S_{\xi n}$, even under the presence of censoring. However, because of an observed bias in a parametric bootstrap procedure they describe for approximating the variance of $S_{\xi n}$, Wang & Wells (2000) ended up recommending that the selection of a model from a set of Archimedean copula families be based on a comparison of the raw values of $S_{\xi n}$.

This paper extends the work of Wang & Wells (2000) in a number of ways. Expressed in the simplest of terms, what is proposed here are alternatives to $S_{\xi n}$ given by

$$S_n=\int_0^1|\mathbb{K}_n(t)|^2k(\theta_n,t)\,\mathrm{d}t\quad\text{and}\quad T_n=\sup_{0\le t\le1}|\mathbb{K}_n(t)|.$$

It will be seen that the use of these statistics has several advantages. Specifically:

(a) simple formulas are available for S_n and T_n in terms of the ranks of the observations, which is not the case for $S_{\xi n}$;
(b) the procedures are free of any extraneous constant ξ, whose selection and influence on the limiting distribution of $S_{\xi n}$ were not addressed by Wang & Wells (2000);
(c) the large-sample distribution of S_n and T_n can be found not only for bivariate Archimedean copulas, but also in arbitrary dimension $d\ge2$ and for general copulas satisfying weak regularity conditions;
(d) although the limits are not explicit, a parametric bootstrap procedure which is demonstrably valid can be used to approximate p-values associated with any continuous functional of $\mathbb{K}_n$, and in particular with S_n and T_n.

In the course of these developments, an explanation will be given for the bias observed by Wang & Wells (2000) in their own parametric bootstrap, and a correction will be proposed. Furthermore, numerical examples will illustrate how a selection procedure based only on the comparison of raw values of $S_{\xi n}$ may sometimes lead to models that should be rejected on the basis of their p-value. Of course, comparing raw values of S_n and T_n could be just as misleading, whence the importance of the valid parametric bootstrap procedure proposed herein.

Conditions which ensure the weak convergence of $\mathbb{K}_n$ in arbitrary dimension $d \geq 2$ are given in section 2 and verified in section 3 for a number of common copula families, including non-Archimedean models. In section 4, simple formulas are presented for the goodness-of-fit statistics S_n and T_n, and the implementation of the parametric bootstrap is discussed. In section 5, simulations are then used to assess the power of goodness-of-fit tests based on S_n and T_n. Comparisons are also made there with a statistic proposed by Shih (1998) for testing the adequacy of the Clayton family of copulas, also known as the gamma frailty model. Two concrete examples of application of the new procedures are discussed in section 6, and concluding remarks are made in the final section.

While the work of Wang & Wells (2000) was motivated by biostatistical applications in which data are often censored, the present paper does not address the issue of censorship, as it arose from modelling issues in actuarial science and finance, where this problem is much less frequent. For illustrations of copula modelling in the latter fields, see for instance Frees & Valdez (1998), Klugman & Parsa (1999), Li (2000), Belguise & Lévi (2001, 2002), Cherubini & Luciano (2002), Embrechts *et al.* (2002), Hennessy & Lapan (2002), Lauprete *et al.* (2002), Dakhli (2004) and van den Goorbergh *et al.* (2005).

2. Distributional results

Let $(X_{11}, \ldots, X_{d1}), \ldots, (X_{1n}, \ldots, X_{dn})$ be $n \geq 2$ independent copies of a vector $\mathbf{X} = (X_1, \ldots, X_d)$ from some continuous d-variate copula model $\mathcal{C} = (C_\theta)$ with unknown continuous margins $F_1, \ldots, F_d$. In other words, suppose that the cumulative distribution function H of $\mathbf{X}$ is of the form

$$H(x_1, \ldots, x_d) = C\{F_1(x_1), \ldots, F_d(x_d)\},$$

for some copula $C = C_\theta \in \mathcal{C}$, whose parameter θ takes its value in an open set $\mathcal{O} \subset \mathbb{R}^m$. It is not assumed that C_θ is Archimedean in the sequel.

Let $K(\theta, t) = P\{H(\mathbf{X}) \leq t\}$, and define its empirical version as

$$K_n(t) = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}(V_{jn} \leq t), \quad t \in [0, 1]$$

where the V_{jn} are pseudo-observations defined by

$$V_{jn} = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}\big(X_{1k} \leq X_{1j}, \ldots, X_{dk} \leq X_{dj}\big) = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}\big(R_{1k} \leq R_{1j}, \ldots, R_{dk} \leq R_{dj}\big),$$

with R_{ij} standing for the rank of X_{ij} among $X_{i1}, \ldots, X_{in}$.

Proposition 1 below identifies the weak limit $\mathbb{K}$ of the process $\mathbb{K}_n$ defined in (1). The conditions are sufficient to ensure that the statistics S_n and T_n are continuous functionals of $\mathbb{K}_n$, whose limits are thus given, respectively, by

$$S = \int_0^1 |\mathbb{K}(t)|^2 k(\theta, t)\, \mathrm{d}t \quad \text{and} \quad T = \sup_{t\in[0,1]} |\mathbb{K}(t)|. \tag{2}$$

Hypothesis I

For all $\theta \in \mathcal{O}$, the distribution function $K(\theta, t)$ of $H(\mathbf{X})$ admits a density $k(\theta, t)$ which is continuous on $\mathcal{O} \times (0, 1]$ and such that

$$k(\theta, t) = o\left\{ t^{-1/2} \log^{-1/2-\epsilon}\left(\frac{1}{t}\right)\right\}$$

for some $\epsilon > 0$ as $t \to 0$.

Hypothesis II

For all $\theta \in \mathcal{O}$, there exists a version of the conditional distribution of the vector $\mathbf{X} = (X_1, \ldots, X_d)$ given $H(\mathbf{X}) = t$ such that, for any continuous real-valued function f on $[0, 1]^d$, the mapping

$$t \mapsto \mu(t, f) = k(\theta, t) E\{f(X_1, \ldots, X_d) \mid H(\mathbf{X}) = t\}$$

is continuous on $(0, 1]$ with $\mu(1, f) = k(\theta, 1)\, f(1, \ldots, 1)$.

As noted by Barbe *et al.* (1996) and Ghoudi & Rémillard (1998), these two hypotheses are sufficient already to imply the weak convergence of Kendall's process, namely

$$\mathbb{K}_{n,\theta}(t) = \sqrt{n}\,\{K_n(t) - K(\theta, t)\},$$

to a continuous, centred Gaussian process $\mathbb{K}_\theta$ whose asymptotic covariance function $\Gamma_\theta(s, t)$ is identified in theorem 1 of Barbe *et al.* (1996). To guarantee the convergence of the process $\mathbb{K}_n$, however, it is also necessary to restrict the large-sample behaviour of $\Theta_n = \sqrt{n}(\theta_n - \theta)$ in such a way that θ_n is a 'good' estimator of the parameter θ.

Hypothesis III

One has $(\mathbb{K}_{n,\theta}, \Theta_n) \rightsquigarrow (\mathbb{K}_\theta, \Theta)$ in $\mathcal{D}[0, 1] \times \mathbb{R}^m$, and the limit is Gaussian and centred. In the sequel, $\Sigma = var(\Theta)$ and

$$\gamma(t) = (\gamma_1(t), \ldots, \gamma_m(t))^\top = (\mathrm{cov}\{\mathbb{K}_\theta(t), \Theta_1\}, \ldots, \mathrm{cov}\{\mathbb{K}_\theta(t), \Theta_m\})^\top, \quad t \in [0, 1].$$

The last hypothesis is a technical condition concerning the existence and smoothness of the gradient of $K(\theta, t)$ with respect to θ, defined as

$$\dot{K}(\theta, t) = \nabla_\theta K(\theta, t) = \left(\frac{\partial}{\partial\theta_1} K(\theta, t), \ldots, \frac{\partial}{\partial\theta_m} K(\theta, t)\right)^\top.$$

Hypothesis IV

For every given $\theta \in \mathcal{O}$, $\dot{K}(\theta, t)$ exists and is continuous for all $t \in [0, 1]$. Moreover,

$$\sup_{\|\theta^\star - \theta\| < \varepsilon} \sup_{t\in[0,1]} \left|\dot{K}(\theta^\star, t) - \dot{K}(\theta, t)\right| \longrightarrow 0 \quad as \quad \varepsilon \to 0. \tag{3}$$

This condition is needed in particular to show that S_n is indeed a continuous functional of $\mathbb{K}_n$. For, whatever $t_0 \in (0, 1)$, one has

$$\int_0^1 |k(\theta_n,t)-k(\theta,t)|\,\mathrm{d}t \le \int_{t_0}^1 |k(\theta_n,t)-k(\theta,t)|\,\mathrm{d}t + |K(\theta_n,t_0)-K(\theta,t_0)| + 2K(\theta,t_0).$$

Therefore, hypotheses I, III and IV together imply that the left-hand side converges in probability to zero. Observe also in passing that if $\dot{K}(\theta,t)=\nabla_\theta K(\theta,t)$, then condition (3) is verified whenever there exists $\epsilon>0$ such that

$$\sup_{\|\theta^\star-\theta\|<\varepsilon}\sup_{t\in[0,1]}\left\|\ddot{K}(\theta^\star,t)\right\|<\infty.$$

Proposition 1

Under hypotheses I–IV, $\mathbb{K}_n \rightsquigarrow \mathbb{K}$ in $\mathcal{D}[0,1]$, where the weak limit $\mathbb{K}$ is a continuous, centred Gaussian process having representation

$$\mathbb{K}(t)=\mathbb{K}_\theta(t)-\dot{K}(\theta,t)^\top\Theta,\quad t\in[0,1]$$

and covariance function

$$\Gamma(s,t)=\Gamma_\theta(s,t)+\dot{K}(\theta,s)^\top\Sigma\,\dot{K}(\theta,t)-\dot{K}(\theta,s)^\top\gamma(t)-\dot{K}(\theta,t)^\top\gamma(s),\quad s,t\in[0,1].$$

Proof. Write $\mathbb{K}_n(t)=\mathbb{K}_{n,\theta}(t)-B_n(t)$ with $B_n(t)=\sqrt{n}\{K(\theta_n,t)-K(\theta,t)\}$ for all $t\in[0,1]$. As mentioned already, it follows from hypotheses I and II that $\mathbb{K}_{n,\theta}\rightsquigarrow\mathbb{K}_\theta$ in $\mathcal{D}[0,1]$, where $\mathbb{K}_\theta$ is the continuous, centred Gaussian process introduced earlier. Furthermore, it is shown in appendix A that

$$\sup_{t\in[0,1]}\left|B_n(t)-\dot{K}(\theta,t)^\top\Theta_n\right|\xrightarrow{P}0$$

under hypotheses III and IV. Finally, making use of hypothesis III, one has

$$\begin{aligned}\Gamma(s,t)&=\mathrm{cov}\{\mathbb{K}(s),\mathbb{K}(t)\}\\&=\mathrm{cov}\{\mathbb{K}_\theta(s),\mathbb{K}_\theta(t)\}+\mathrm{cov}\left\{\dot{K}(\theta,s)^\top\Theta,\dot{K}(\theta,t)^\top\Theta\right\}\\&\quad-\mathrm{cov}\left\{\dot{K}(\theta,s)^\top\Theta,\mathbb{K}_\theta(t)\right\}-\mathrm{cov}\left\{\dot{K}(\theta,t)^\top\Theta,\mathbb{K}_\theta(s)\right\}\\&=\Gamma_\theta(s,t)+\dot{K}(\theta,s)^\top\Sigma\dot{K}(\theta,t)-\dot{K}(\theta,s)^\top\gamma(t)-\dot{K}(\theta,t)^\top\gamma(s).\end{aligned}$$

Thus the proof is complete.

Remark. From the work of Barbe *et al.* (1996), it is known that

$$\begin{aligned}\Gamma_\theta(s,t)&=K(\theta,s\wedge t)-K(\theta,s)K(\theta,t)\\&\quad+k(\theta,s)k(\theta,t)R_\theta(s,t)-k(\theta,t)Q_\theta(s,t)-k(\theta,s)Q_\theta(t,s),\end{aligned}$$

where $s\wedge t=\min(s,t)$ and for all $s,t\in[0,1]$,

$$R_\theta(s,t)=P\{\mathbf{X}_1\le\mathbf{X}_2\wedge\mathbf{X}_3\,|\,H(\mathbf{X}_2)=s,H(\mathbf{X}_3)=t\}-st$$

and

$$Q_\theta(s,t)=P\{H(\mathbf{X}_1)\le s,\mathbf{X}_1\le\mathbf{X}_2\,|\,H(\mathbf{X}_2)=t\}-tK(\theta,s)$$

are defined in terms of mutually independent copies $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$ of $\mathbf{X}$.

A potential candidate for θ_n is the omnibus rank-based estimator of Genest *et al.* (1995) or Shih & Louis (1995), obtained through a maximization of the pseudo-likelihood

$$\sum_{j=1}^{n} \log c_\theta\left(\frac{R_{1j}}{n+1}, \ldots, \frac{R_{dj}}{n+1}\right),$$

where c_θ is the density associated with C_θ. It follows from example 3.2.1 of Ghoudi & Rémillard (2004) that provided that c_θ is smooth enough hypothesis III is automatically verified when hypotheses I and II hold, so that proposition 1 then holds under hypotheses I, II and IV only.

In the special case where θ is real, another common procedure considered by Wang & Wells (2000) among others, consists of estimating θ by $\theta_n = \tau^{-1}(\tau_n)$, where $\tau(\theta)$ is the multivariate extension of Kendall's tau defined by

$$\tau = \left(\frac{2^d - 1}{2^{d-1} - 1}\right) - \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 K(\theta, t)\, \mathrm{d}t, \tag{4}$$

as in Barbe *et al.* (1996) or Jouini & Clemen (1996). Assume that the mapping $\theta \mapsto \tau(\theta)$ has a continuous non-vanishing derivative

$$\dot{\tau}(\theta) = -\left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \dot{K}(\theta, t)\, \mathrm{d}t$$

in $\mathcal{O}$. Redefining

$$\tau_n = \left(\frac{2^d - 1}{2^{d-1} - 1}\right) - \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 K_n(t)\, \mathrm{d}t,$$

one can see that $\sqrt{n}(\tau_n - \tau)$ is related to Kendall's process $\mathbb{K}_{n,\theta}$ through the linear functional

$$\sqrt{n}(\tau_n - \tau) = -\left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \mathbb{K}_{n,\theta}(t)\, \mathrm{d}t.$$

An application of Slutsky's theorem then implies that, under hypotheses I and II,

$$\Theta_n = \sqrt{n}\{\tau^{-1}(\tau_n) - \theta\} = -\frac{1}{\dot{\tau}(\theta)} \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \mathbb{K}_{n,\theta}(t)\, \mathrm{d}t + o_P(1)$$

converges in law to

$$\Theta = -\frac{1}{\dot{\tau}(\theta)} \left(\frac{2^d}{2^{d-1} - 1}\right) \int_0^1 \mathbb{K}_\theta(t)\, \mathrm{d}t = \kappa_\theta \int_0^1 \mathbb{K}_\theta(t)\, \mathrm{d}t,$$

where

$$\frac{1}{\kappa_\theta} = \int_0^1 \dot{K}(\theta, t)\, \mathrm{d}t$$

is assumed to be non-zero. The weak convergence $(\mathbb{K}_{n,\theta}, \Theta_n) \rightsquigarrow (\mathbb{K}_\theta, \Theta)$ required in hypothesis III is thus immediate, and

$$\gamma(t) = \mathrm{cov}\{\mathbb{K}_\theta(t), \Theta\} = \kappa_\theta \int_0^1 \Gamma_\theta(u, t)\, \mathrm{d}u$$

while

$$\mathrm{var}(\Theta) = \kappa_\theta^2 \int_0^1 \int_0^1 \Gamma_\theta(u, v)\, \mathrm{d}u\, \mathrm{d}v.$$

This suggests the following consequence of the main result.

Proposition 2
If $\theta \in \mathcal{O} \subseteq \mathbb{R}$ is estimated by $\theta_n = \tau^{-1}(\tau_n)$ and κ_θ is finite, then under hypotheses I, II and IV, one has $\mathbb{K}_n \rightsquigarrow \mathbb{K}$, where the weak limit $\mathbb{K}$ is a centred Gaussian process having representation

$$\mathbb{K}(t) = \mathbb{K}_\theta(t) - \kappa_\theta \dot{K}(\theta, t) \int_0^1 \mathbb{K}_\theta(v)\, \mathrm{d}v, \quad t \in [0, 1]$$

and limiting covariance function

$$\begin{aligned}\Gamma(s, t) = {} & \Gamma_\theta(s, t) + \kappa_\theta^2 \dot{K}(\theta, s)\dot{K}(\theta, t) \int_0^1 \int_0^1 \Gamma_\theta(u, v)\, \mathrm{d}u\, \mathrm{d}v \\ & - \kappa_\theta \dot{K}(\theta, s) \int_0^1 \Gamma_\theta(u, t)\, \mathrm{d}u - \kappa_\theta \dot{K}(\theta, t) \int_0^1 \Gamma_\theta(u, s)\, \mathrm{d}u, \quad s, t \in [0, 1].\end{aligned}$$

3. Examples

This section presents a few popular classes of multivariate copulas that satisfy hypotheses I–IV stated above. The list is by no means exhaustive, of course.

3.1. Archimedean copulas

Copulas are called Archimedean when they may be expressed in the form

$$C(u_1, \ldots, u_d) = \phi^{-1}\{\phi(u_1) + \cdots + \phi(u_d)\}$$

in terms of a bijection $\phi : (0, 1] \to [0, \infty)$ such that $\phi(1) = 0$ and

$$\frac{(-1)^i d^i}{\mathrm{d}x^i} \phi^{-1}(x) > 0, \quad i \in \{1, \ldots, d\}. \tag{5}$$

As shown by Genest & Rivest (1993) in the case $d = 2$, the generator ϕ can be recovered from K, since $K(t) = t - \phi(t)/\phi'(t)$, $t \in (0, 1]$. Among the multivariate copula models that fall into this category (see Nelsen, 1999, Chapter 4), Table 1 presents summary information for those of Ali *et al.* (1978), Clayton (1978), Gumbel (1960) and Frank (1979). Note that in this table, the parameter space $\mathcal{O}$ is limited to positive degrees of association, as those are the only values that can be achieved in all dimensions for Archimedean copulas in general (Marshall & Olkin, 1988), and for these four models in particular.

One key characteristic of Archimedean copulas is the fact that all the information about the d-dimensional dependence structure is contained in a univariate generator, ϕ_θ. From Barbe *et al.* (1996),

$$K(\theta, t) = t + \sum_{i=1}^{d-1} \frac{(-1)^i}{i!} \{\phi_\theta(t)\}^i f_i(\theta, t), \tag{6}$$

where

Table 1. *Families of multivariate Archimedean copulas*

Model	$\phi_\theta(t)$	$K(\theta, t)$ for $d = 2$	$\tau = g(\theta)$	$\mathcal{O}$
Ali–Mikhail–Haq	$\dfrac{\log((1-\theta)/t+\theta)}{1-\theta}$	$t+\dfrac{t^2}{1-\theta}\left(\dfrac{1-\theta}{t}+\theta\right)$ $\times\log\left(\dfrac{1-\theta}{t}+\theta\right)$	$\dfrac{3\theta-2}{3\theta}$ $-\dfrac{2(1-\theta)^2\log(1-\theta)}{3\theta^2}$	(0, 1)
Clayton	$\dfrac{t^{-\theta}-1}{\theta}$	$t+\dfrac{t(1-t^\theta)}{\theta}$	$\dfrac{\theta}{\theta+2}$	$(0, \infty)$
Frank	$\log\left(\frac{1-\mathrm{e}^{-\theta}}{1-\mathrm{e}^{-\theta t}}\right)$	$t-\dfrac{(1-\mathrm{e}^{\theta t})}{\theta}\log\left(\dfrac{1-\mathrm{e}^{-\theta}}{1-\mathrm{e}^{-\theta t}}\right)$	$1-\dfrac{4}{\theta}+\dfrac{4D_1(\theta)}{\theta}$	$(0, \infty)$
Gumbel–Hougaard	$\lvert\log t\rvert^{1/(1-\theta)}$	$t-(1-\theta)t\log t$	θ	(0, 1)

$D_1(\theta)=\theta^{-1}\int_0^\theta \frac{x}{\mathrm{e}^x-1}\,\mathrm{d}x$ stands for the first Debye function.

$$f_i(\theta,t)=\frac{d^i}{\mathrm{d}x^i}\phi_\theta^{-1}(x)\bigg|_{x=\phi_\theta(t)},$$

provided that $\{\phi_\theta(t)\}^i f_i(\theta, t) \to 0$ as $t \to 0$ for all $i \in \{1, \ldots, d-1\}$. Note in passing that

$$f_{i+1}(\theta,t)=f_1(\theta,t)\frac{\partial}{\partial t}f_i(\theta,t), \quad i\in\{1,\ldots,d-1\}. \tag{7}$$

As a consequence, the moments of C_θ are also functions of ϕ_θ only. In addition, the multivariate version of Kendall's measure of association may be computed, in view of (4), through the formula

$$\tau=1-\left(\frac{2^d}{2^{d-1}-1}\right)\sum_{i=1}^{d-1}\frac{(-1)^i}{i!}\int_0^1\{\phi_\theta(t)\}^i f_i(\theta,t)\,\mathrm{d}t,$$

which for $d = 2$ reduces to the well-known expression (Genest & MacKay, 1986; Nelsen, 1999, section 5.1)

$$\tau=1+4\int_0^1\frac{\phi_\theta(t)}{\phi'_\theta(t)}\,\mathrm{d}t.$$

Barbe *et al.* (1996) prove that the four families of copulas listed in Table 1 meet hypotheses I and II. It is shown in appendices B1–B4 that they also satisfy hypothesis IV for all values of $\theta \in \mathcal{O}$. Explicit expressions for K may also be found there.

3.2. Bivariate extreme-value copulas

It has been known since the work of Pickands (1981) that bivariate extreme-value distributions have underlying copulas of the form

$$C_A(u,v)=\exp\left[\log(uv)A\left\{\frac{\log(u)}{\log(uv)}\right\}\right],$$

where the dependence function A, defined on $[0, 1]$, is convex and such that $\max(t, 1-t) \le A(t) \le 1$ for all $t \in [0, 1]$. The most common parametric models of bivariate extreme-value copulas are presented in Table 2. For additional details see, for instance, Tawn (1988), Capéraà *et al.* (1997, 2000) or Capéraà & Fougères (2001).

Ghoudi *et al.* (1998) note that if (U_1, U_2) is distributed as C_A, then

$$K_A(t)=P\{C_A(U_1,U_2)\le t\}=t-(1-\tau)t\log t, \quad t\in(0,1]$$

Table 2. *Families of bivariate extreme-value copulas*

Model	$A_\theta(t)$	$C_{A_\theta}(u, v)$	$\mathcal{O}$
Gumbel	$\theta t^2 - \theta t + 1$	$uv \exp\left(-\theta \frac{\log u \log v}{\log uv}\right)$	$(0, 1)$
Gumbel–Hougaard	$\left\{t^{\frac{1}{1-\theta}} + (1-t)^{\frac{1}{1-\theta}}\right\}^{1-\theta}$	$\exp\left[-\left\{\lvert\log u\rvert^{\frac{1}{1-\theta}} + \lvert\log v\rvert^{\frac{1}{1-\theta}}\right\}^{1-\theta}\right]$	$(0, 1)$
Galambos	$1 - \{t^{-\theta} + (1-t)^{-\theta}\}^{-1/\theta}$	$uv \exp\left[(\lvert\log u\rvert^{-\theta} + \lvert\log v\rvert^{-\theta})^{-1/\theta}\right]$	$(0, \infty)$
Generalized Marshall–Olkin	$\max\{1 - \theta_1 t, 1 - \theta_2(1-t)\}$	$u^{1-\theta_1} v^{1-\theta_2} \min(u^{\theta_1}, v^{\theta_2})$	$(0, 1)^2$

depends only on the population version of Kendall's measure of association computed as a function of A through the identity

$$\tau = \tau(A) = \int_0^1 \frac{t(1-t)}{A(t)} \, \mathrm{d}A'(t).$$

Thus if two bivariate extreme-value copulas with generators A and A^* are such that $\tau(A) = \tau(A^*)$, then K_A is the same as K_{A^*} and they could not possibly be distinguished by goodness-of-fit procedures based on the process $\mathbb{K}_n$. To avoid the ensuing identifiability issues, suppose that θ is real. There is then no loss of generality in taking $\theta = \tau$ and $\theta_n = \tau_n$. In this case, it can be checked easily that the conditions of the main result are satisfied. To this end, first note that $\dot{K}(\theta, t) = t \log t$, so that hypothesis IV trivially holds. As hypothesis I is also readily verified, an application of proposition 2 implies that

$$\mathbb{K}_n(t) = \sqrt{n}\{K_n(t) - K(\theta_n, t)\} \rightsquigarrow \mathbb{K}(t) = \mathbb{K}_\theta(t) + 4t \log t \int_0^1 \mathbb{K}_\theta(v) \, \mathrm{d}v.$$

Furthermore, the limiting covariance function, for which no explicit representation seems possible, is given by

$$\begin{aligned}\Gamma(s, t) = \Gamma_\theta(s, t) &+ 16st \log s \log t \int_0^1 \int_0^1 \Gamma_\theta(u, v) \, \mathrm{d}u \, \mathrm{d}v \\ &+ 4s \log s \int_0^1 \Gamma_\theta(u, t) \, \mathrm{d}u + 4t \log t \int_0^1 \Gamma_\theta(s, v) \, \mathrm{d}v.\end{aligned}$$

3.3. *Fréchet copulas*

These bivariate copulas are mixtures of the independence copula $C_\Pi(u, v) = uv$ and of the upper Fréchet–Hoeffding bound $C_M(u, v) = \min(u, v)$, that is,

$$C_\theta(u, v) = (1-\theta)uv + \theta \min(u, v), \quad \theta \in [0, 1].$$

Letting $\zeta(\theta, t) = 4t/\{I(\theta, t) + \theta\}^2$ and $I(\theta, t) = \{\theta^2 + 4t(1-\theta)\}^{1/2}$, one can show that

$$K(\theta, t) = t - t \log t + t \log\{\zeta(\theta, t)\}, \quad t \in [0, 1].$$

See Genest & Rivest (2001) for a proof of this result. Note that ζ is continuous on $[0, 1]^2$ and bounded above by 1.

For this model, it is known (see e.g. Nelsen 1999, p. 130) that $\tau = \theta(\theta + 2)/3$, and hence $\theta = \sqrt{3\tau + 1} - 1$. Furthermore, a simple calculation shows that $\zeta'(\theta, t)/\zeta(\theta, t) = \theta/\{tI(\theta, t)\}$, whence the density associated with $K(\theta, t)$ is given by $k(\theta, t) = -\log t + \log\{\zeta(\theta,t)\} + \theta/I(\theta,t)$. The latter function is continuous and since $\theta \leq I(\theta, t)$, one

can see at once that it has the appropriate behaviour as $t \to 0$ for hypothesis I to hold. The verification of hypothesis IV is deferred to appendix B5.

3.4. Bivariate Farlie–Gumbel–Morgenstern copulas

It is not always necessary to be able to compute $K(\theta, t)$ explicitly to obtain the weak convergence of $\mathbb{K}_{n,\theta}$. Such is the case for the Farlie–Gumbel–Morgenstern system of distributions, whose members have the form

$$C_\theta(u, v) = uv + \theta uv(1-u)(1-v), \quad \theta \in [-1, 1].$$

Barbe *et al.* (1996) show that

$$k(\theta, t) = \int_t^1 h(\theta, x, t)\, \mathrm{d}x,$$

where

$$h(\theta, x, t) = \left\{ \frac{1}{(1-x)r(\theta, x, t)} + \frac{1}{x} - \frac{1}{1-x} \right\} \mathbf{1}(t < x \leq 1)$$

and

$$r(\theta, x, t) = \left[\{1 - \theta(1-x)\}^2 + 4\theta(1-x)\left(1 - \frac{t}{x}\right) \right]^{1/2}.$$

Although in this model $\tau = 2\theta/9$, there is no explicit expression for

$$K(\theta, t) = \int_0^t k(\theta, s)\, \mathrm{d}s = \int_0^t \int_s^1 h(\theta, x, s)\, \mathrm{d}x\, \mathrm{d}s.$$

Nevertheless, Barbe *et al.* (1996) prove that $k(\theta, t)$ satisfies hypothesis I. They further mention that hypothesis II is verified as well. The proof that hypothesis IV also holds may be found in appendix B6.

4. Implementation of the goodness-of-fit tests

Straightforward calculations show that

$$S_n = \frac{n}{3} + n \sum_{j=1}^{n-1} K_n^2\left(\frac{j}{n}\right) \left\{ K\left(\theta_n, \frac{j+1}{n}\right) - K\left(\theta_n, \frac{j}{n}\right) \right\} - n \sum_{j=1}^{n-1} K_n\left(\frac{j}{n}\right) \left\{ K^2\left(\theta_n, \frac{j+1}{n}\right) - K^2\left(\theta_n, \frac{j}{n}\right) \right\}$$

and

$$T_n = \sqrt{n} \max_{i=0,1;\ 0 \leq j \leq n-1} \left\{ \left| K_n\left(\frac{j}{n}\right) - K\left(\theta_n, \frac{j+i}{n}\right) \right| \right\}.$$

Formal testing procedures based on these statistics would consist of rejecting $H_0 : C \in \mathcal{C}$ when the observed value of S_n or T_n is greater than the $100(1-\alpha)$th percentile of its distribution under the null hypothesis. As implied by formula (2), however, this distribution depends on the unknown association parameter θ, even in the limit.

Table 3. *Estimation based on 1000 replicates of the 95th percentile of the distribution of the Cramér–von Mises statistic S_n and the Kolmogorov–Smirnov statistic T_n*

		S_n			T_n		
Model	τ	$n = 100$	$n = 250$	$n = 1000$	$n = 100$	$n = 250$	$n = 1000$
Clayton	0.20	0.1872	0.1589	0.1567	1.0402	0.9725	0.9824
	0.40	0.1410	0.1336	0.1278	0.9244	0.9015	0.8805
	0.60	0.0992	0.0902	0.0977	0.8563	0.8071	0.7863
	0.80	0.0573	0.0518	0.0511	0.6775	0.6578	0.6650
Frank	0.20	0.1515	0.1328	0.1216	0.9743	0.9294	0.9210
	0.40	0.1254	0.1186	0.1150	0.8883	0.8614	0.8439
	0.60	0.1017	0.0979	0.0975	0.7945	0.7982	0.7875
	0.80	0.0591	0.0536	0.0492	0.6410	0.6472	0.6496
Gumbel–Hougaard	0.20	0.1516	0.1266	0.1265	0.9740	0.9431	0.9325
	0.40	0.1255	0.1155	0.1117	0.9284	0.8812	0.8737
	0.60	0.0946	0.0961	0.0833	0.8033	0.7839	0.8000
	0.80	0.0567	0.0550	0.0508	0.6381	0.6573	0.6469

This fact is illustrated in Table 3, where the 95th percentiles of the distributions of S_n and T_n are evaluated for Archimedean copulas of Table 1 for some values of τ. Of course, the statistic $S_{\xi n}(\theta_n)$ of Wang & Wells (2000) suffers from the same limitation, let alone its dependence on the arbitrary cut-off point ξ.

To circumvent these methodological issues and obtain an approximate p-value for either S_n or T_n, one may call on a parametric bootstrap or Monte Carlo testing approach based on C_{θ_n}. This is described is section 4.2. As the tests and their distributions only involve C_{θ_n} through $K(\theta_n, \cdot)$, it may be tempting to base the bootstrap exclusively on the latter, as done by Wang & Wells (2000). Section 4.1 explains why this shortcut is inappropriate.

4.1. The parametric bootstrap method in Wang & Wells (2000)

To find an estimate of the variance of their statistic $S_{\xi n}$, Wang & Wells (2000) propose to generate a sample $V_{1,n}^*, \ldots, V_{n,n}^*$, where $V_{j,n}^* \sim K(\theta_n, \cdot)$ for $j \in \{1, \ldots, n\}$, and then to calculate

$$K_n^*(v) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}\left(V_{j,n}^* \le v\right), \quad \tau_n^* = -1 + \frac{4}{n}\sum_{j=1}^{n} V_{j,n}^*, \quad \theta_n^* = \tau^{-1}\left(\tau_n^*\right)$$

and

$$S_{\xi n}^* = n\int_{\xi}^{1} \left\{K_n^*(v) - K\left(\theta_n^*, v\right)\right\}^2 \mathrm{d}v.$$

By repeating the procedure N times, one ends up with values $S_{\xi n,1}^*, \ldots, S_{\xi n,N}^*$. Wang & Wells (2000) thus suggest that the variance of $S_{\xi n}$ could be estimated by the sample variance of $S_{\xi n,1}^*, \ldots, S_{\xi n,N}^*$.

Unfortunately, this algorithm is invalid. As shown in appendix C, the empirical bootstrap process $\sqrt{n}\{\mathbb{K}_n^* - K(\theta_n^*, \cdot)\}$ converges in $\mathcal{D}[0, 1]$ to a limit $\mathbb{K}^*$ which is independent of $\mathbb{K}$ but generally different in law. Consequently, this bootstrap does not yield a valid

estimate of the variance of $S_{\xi n}$, nor reliable p-values of any goodness-of-fit test based thereon.

4.2. The parametric bootstrap method based on C_{θ_n}

In order to compute p-values for any test statistic based on the empirical process $\mathbb{K}_n$, one requires generating a large number, N, of independent samples of size n from C_{θ_n} and computing the corresponding values of the selected statistic, such as S_n or T_n. In the former case, for example, the procedure would work as follows:

Step 1: Estimate θ by a consistent estimator θ_n.

Step 2: Generate N random samples of size n from C_{θ_n} and, for each of these samples, estimate θ by the same method as before and determine the value of the test statistic.

Step 3: If $S_{1:N}^* \leq \cdots \leq S_{N:N}^*$ denote the ordered values of the test statistics calculated in step 2, an estimate of the critical value of the test at level α based on S_n is given by

$$S_{\lfloor (1-\alpha)N \rfloor :N}^* \quad \text{and} \quad \frac{1}{N} \# \left\{ j : S_j^* \geq S_n \right\}$$

yields an estimate of the p-value associated with the observed value S_n of the statistic. Here, $\lfloor x \rfloor$ denotes the integer part of x.

The validity of this approach is established in a companion paper by Genest & Rémillard (2005). The assumptions needed for the method to work are stated in appendix D.

5. Numerical studies

Simulation studies were conducted to assess the finite-sample properties of the proposed goodness-of-fit tests for various classes of copula models under the null hypothesis and under the alternative. Three copula families were used under H_0, namely those of Clayton, Frank and Gumbel–Hougaard. All of them are Archimedean and complete, in the sense that they cover all possible degrees of positive dependence, as measured by Kendall's tau. Three complete systems of non-Archimedean copulas were also used as alternatives, namely the Fréchet, Gaussian and Plackett families. In all models considered, the validity conditions for the parametric bootstrap are verified.

In each case, 10,000 pseudo-random samples of size $n = 250$ were generated from the selected model with a specified value of Kendall's tau, chosen in the set $\{0.2, 0.4, 0.6, 0.8\}$. For each of these 10,000 samples, the dependence parameter of the copula model under the null hypothesis was estimated by inversion of Kendall's tau, that is, by setting $\theta_n = \tau^{-1}(\tau_n)$. Statistics S_n, T_n and S_{0n} were then calculated, and their respective p-value was estimated by generating $N = 1000$ bootstrap samples from C_{θ_n} from the copula model under the null hypothesis, as detailed in steps 1–3 above. The proportion of such p-values that were inferior to 5% was then determined. As the bootstrap procedure is consistent, this proportion yields an estimate of the size and power of the test, depending on whether or not the original data are from the assumed copula family under the null hypothesis.

For the Clayton and Gumbel–Hougaard copulas, the rank-based estimators derived by the formula $\theta_n = \tau^{-1}(\tau_n)$ are simply given by

Table 4. *Percent relative bias of the estimator* $\theta_n = \tau^{-1}(\tau_n)$ *based on 10,000 samples of size* $n = 100$ *and* $n = 250$ *from the Clayton, Frank and Gumbel–Hougaard models with various degrees of dependence*

Model	τ	$n = 100$	$n = 250$
Clayton	0.20	3.22	1.98
	0.40	3.25	0.95
	0.60	2.39	0.93
	0.80	2.29	0.98
Frank	0.20	1.57	0.56
	0.40	1.26	0.23
	0.60	1.10	0.44
	0.80	1.28	0.28
Gumbel–Hougaard	0.20	0.12	0.18
	0.40	0.07	0.01
	0.60	−0.05	0.05
	0.80	0.01	0.01

$$\theta_n = \frac{2\tau_n}{1-\tau_n} \quad \text{and} \quad \theta_n = \tau_n,$$

respectively, but the inversion needs to be carried out numerically for Frank's model. Table 4 shows that the resulting rank-based estimators have reasonably small percent relative bias for samples of size $n = 100$ and 250.

5.1. Comparison between S_n, T_n and $S_{\xi n}$

Table 5 shows the power and size of S_n and T_n as goodness-of-fit test statistics of the Clayton, Frank and Gumbel–Hougaard null hypothesis under 24 choices of copula for the true distribution. Results for the $S_{\xi n}$ statistic of Wang & Wells (2000) with $\xi = 0$ were also added for comparison purposes. In interpreting these results, it must be kept in mind that the error associated with the power estimates is larger than might usually be expected under 10,000 replications. This is because the distribution of the observations under the null hypothesis involves a parameter that must be estimated. Additional variation thus arises from the use of C_{θ_n} rather than C_θ in the calculation of the p-values associated with the tests.

Table 5 shows that when the null hypothesis holds true, all three statistics are at the right level, up to sampling error. Detailed inspection of the results leads to the following additional insights:

(a) As a general rule, the tests S_n and S_{0n} based on Cramér–von Mises functionals outperform that which is founded on the Kolmogorov–Smirnov statistic T_n.
(b) All tests appear to distinguish rather easily between the Clayton model and the various alternatives considered; the best performance is achieved by S_{0n}.
(c) Testing for the Frank or the Gumbel–Hougaard families seems to be more difficult, at least given the alternatives considered; each of S_n and S_{0n} delivers the best power in roughly half the cases.

There is also a hint in Table 5 that as the value of τ goes from 0 to 1, the power of the three tests increases, levels off, and ultimately starts decreasing again. This was only to be expected, as all families considered have independence and the Fréchet–Hoeffding upper bound $M(u, v) = u \wedge v$ as their limiting copulas at $\tau = 0$ and $\tau = 1$, respectively.

Table 5. *Percentage of rejection of three different null hypotheses using statistics S_n, T_n or S_{0n} at the 5% level when $n = 250$ for various copula alternatives, based on 10,000 replicates*

Alternative		Model under the null hypothesis								
		Clayton			Frank			Gumbel–Hougaard		
Family	τ	S_n	T_n	S_{0n}	S_n	T_n	S_{0n}	S_n	T_n	S_{0n}
Clayton	0.2	4.6	5.5	4.9	82.2	78.9	73.6	95.3	88.4	94.4
	0.4	4.2	4.7	3.6	99.8	99.3	99.6	100.0	99.9	100.0
	0.6	4.6	4.2	5.0	100.0	99.9	99.9	100.0	100.0	100.0
	0.8	4.7	4.4	5.6	100.0	99.8	100.0	100.0	100.0	100.0
Frank	0.2	47.1	38.9	51.1	4.7	4.3	5.7	24.5	14.7	30.9
	0.4	97.1	88.4	97.4	3.7	5.0	5.9	55.8	37.4	65.6
	0.6	99.9	98.2	100.0	5.3	4.9	5.7	75.2	54.2	85.7
	0.8	100.0	96.2	100.0	4.4	5.4	3.8	89.1	62.3	88.7
Gumbel–Hougaard	0.2	68.4	57.6	80.6	10.6	9.3	25.7	5.1	4.3	5.2
	0.4	99.8	97.9	100.0	36.8	24.3	58.4	6.7	5.8	4.3
	0.6	100.0	100.0	100.0	69.5	49.5	82.9	4.3	5.4	4.5
	0.8	100.0	100.0	100.0	91.2	59.6	89.9	4.4	5.1	4.0
Fréchet	0.2	35.0	21.9	41.8	25.3	22.5	20.1	33.1	26.0	22.3
	0.4	81.7	60.2	89.2	52.1	38.4	52.2	54.6	40.9	38.1
	0.6	98.0	83.1	98.9	77.2	45.5	76.9	51.7	31.2	45.6
	0.8	98.3	86.3	99.4	80.1	30.2	78.5	33.9	8.5	28.6
Gaussian	0.2	33.3	23.7	38.7	9.9	8.9	9.1	29.2	17.9	24.2
	0.4	86.9	71.6	90.7	23.6	17.4	23.5	52.4	34.5	47.5
	0.6	99.1	92.1	99.8	50.6	37.7	50.4	53.1	36.5	54.6
	0.8	100.0	98.0	100.0	76.8	42.6	75.0	41.3	23.1	40.6
Plackett	0.2	49.5	35.1	49.7	5.2	6.4	6.2	20.3	15.7	28.5
	0.4	94.0	83.0	95.3	6.3	5.4	7.6	46.5	28.1	48.9
	0.6	99.5	94.8	99.8	15.5	10.2	15.7	48.5	31.6	57.5
	0.8	99.9	93.5	99.7	36.1	14.6	27.9	40.2	22.9	42.9

5.2. *Comparison with the test of Shih for the Clayton family*

Clayton's copula is often referred to as the gamma frailty model in the survival analysis literature. For this specific choice of null hypothesis, two goodness-of-fit tests are already available, which were developed by Shih (1998) in the bivariate case and by Glidden (1999) for arbitrary dimension $d \geq 2$. It may thus be of interest to compare the power of these specific test statistics to those of the omnibus procedures based on S_n and T_n.

In attempting to make such comparisons, difficulties were encountered with the implementation of both Shih's and Glidden's procedures. Specifically:

(a) The limiting variance of the test statistic given on p. 198 of Shih (1998) is erroneous; in fact, her expression tends to $-\infty$ as $\theta = 1/\eta \to 0$, while the correct result is 7/9 in the limiting case of independence. As shown by Genest *et al.* (2006), the correct expression for the asymptotic variance should be

$$\frac{18\eta^7 + 240\eta^6 + 3001\eta^5 + 8281\eta^4 + 9449\eta^3 + 5171\eta^2 + 1352\eta + 136}{3\eta^2(\eta+1)^2(3\eta+1)} + \frac{8(2\eta+1)^4}{(\eta+1)^2}L(\eta) - (\eta+1)^4\left\{\Psi'\left(1+\frac{\eta}{2}\right) - \Psi'\left(\frac{1}{2}+\frac{\eta}{2}\right)\right\} - 8(\eta+1)(2\eta+1)^2 J(\eta),$$

where

$$L(\eta) = \frac{1}{4\eta^2}\,\text{hypergeom}([1,1,\eta],[2\eta+1,2\eta+1],1),$$

$$J(\eta) = \frac{1}{2\eta^2}\,\text{hypergeom}([1,1,\eta],[\eta+1,2\eta+1],1),$$

and Ψ' denotes the trigamma function.

(b) Glidden's test is overly conservative, especially in cases of weak dependence. This phenomenon, documented by Glidden (1999) himself, hints to the fact that the asymptotic distribution may be incorrectly approximated by his proposed computational procedure. In an attempt to reproduce his calculations, it was further discovered that in Glidden's paper, many expressions leading to the identification of the limit were themselves incorrect. In the formulas for ϵ_i and π_k on pp. 385–386, for example, one should replace every instance of $X_{ik\ell}$ by $\tau \wedge X_{ik\ell}$. Furthermore, the expression given for $V(\theta)$ on top of p. 386 should be the limit of $-(1/n)\partial^2\hat{\ell}_n(\theta)/\partial\theta^2$. Also, a factor of $\exp\{\hat{\theta}\hat{\Lambda}_k(\tau \wedge X_{ik\ell})\}$ appears to be missing in the definition of $\hat{\epsilon}_i$, on p. 392.

In view of the numerous difficulties encountered in trying to implement Glidden's test, it was ultimately decided to restrict attention to the corrected version of Shih's procedure. The power of the latter test is compared in Table 6 with those of the tests based on S_n, T_n and S_{0n}. As might have been expected, differences in power between the four procedures are tenuous in

Table 6. *Percentage of rejection of the null hypothesis of Clayton's copula for tests based on S_n, T_n, S_{0n} and Shih's statistic at the 5% level when $n = 250$ for various copula alternatives, based on 10,000 replicates*

Alternative		Estimated power			
Family	τ	S_n	T_n	S_{0n}	Shih
Clayton	0.20	4.6	5.5	4.9	4.2
	0.40	4.2	4.7	3.6	4.2
	0.60	4.6	4.2	5.0	4.9
	0.80	4.7	4.4	5.6	5.9
Frank	0.20	47.1	38.9	51.1	73.6
	0.40	97.1	88.4	97.4	99.9
	0.60	99.9	98.2	100.0	100.0
	0.80	100.0	96.2	100.0	100.0
Gumbel–Hougaard	0.20	68.4	57.6	80.6	8.8
	0.40	99.8	97.9	100.0	37.1
	0.60	100.0	100.0	100.0	78.8
	0.80	100.0	100.0	100.0	97.3
Fréchet	0.20	35.0	21.9	41.8	36.1
	0.40	81.7	60.2	89.2	85.1
	0.60	98.0	83.1	98.9	97.2
	0.80	98.3	86.3	99.4	98.3
Gaussian	0.20	33.3	23.7	38.7	52.3
	0.40	86.9	71.6	90.7	97.3
	0.60	99.1	92.1	99.8	100.0
	0.80	100.0	98.0	100.0	100.0
Plackett	0.20	49.5	35.1	49.7	70.5
	0.40	94.0	83.0	95.3	99.8
	0.60	99.5	94.8	99.8	100.0
	0.80	99.9	93.5	99.7	100.0

cases of strong association. When the dependence is weak, however, the (corrected) Shih statistic turns out to be significantly more powerful than the other three, except when the alternative is Gumbel–Hougaard.

6. Illustrations

This section presents two applications of the proposed methodology to data sets originally considered by Frees & Valdez (1998) and by Cook & Johnson (1981, 1986).

6.1. Insurance data

Figure 1 displays the relation between the natural logarithms of an indemnity payment X_1 and an allocated loss adjustment expense X_2 (comprising lawyers' fees and claim investigation expenses, among others) for 1500 general liability claims. These data were used by Frees & Valdez (1998), Klugman & Parsa (1999) and Chen & Fan (2005), among others, to illustrate copula-model selection and fitting in an insurance context. In their analysis, Frees & Valdez (1998) ignored the censoring present in 34 claims in their visual procedure for the selection of an appropriate copula model, although they used the full sample in their formal estimation of the dependence parameter in the Clayton, Frank and Gumbel–Hougaard copulas. (Although this is irrelevant here, they used generalized Pareto distributions for the margins.)

For simplicity, the analysis presented in the sequel is limited to the 1466 uncensored claims. This restriction has little effect on the estimation of the dependence parameters, as evidenced by a comparison of the numerical estimates obtained by Frees & Valdez (1998) and Genest *et al.* (1998) with and without censoring, respectively. For the uncensored sample, the observed value of Kendall's tau is 0.3195, which is also the estimate of the dependence parameter θ in the Gumbel–Hougaard model. Frees & Valdez (1998) identified this model as providing the best fit of the three. Their judgement was based on a visual comparison of $K_n(t)$ and the parametric distribution functions $K(\theta_n, t)$ corresponding to the Clayton, Frank and Gumbel–Hougaard dependence structures. The same conclusion was reached by Genest *et al.* (1998) and by Chen & Fan (2005) using more formal, pseudo-likelihood ratio based, procedures.

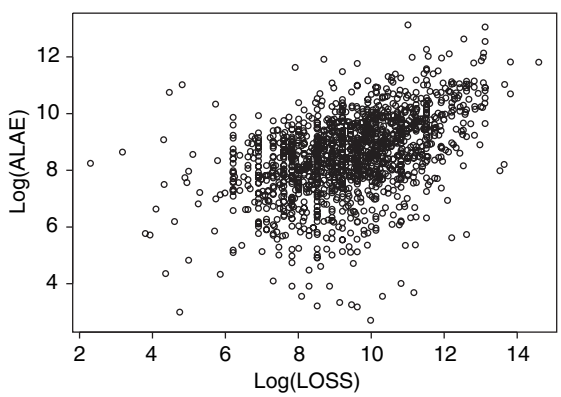

Fig. 1. Scatter plot of the natural logarithms of the indemnity payment (LOSS) and the allocated loss adjustment expense (ALAE) for 1500 general liability claims.

The non-parametric estimator $K_n(t)$ of $K(t)$ is shown in Fig. 2, along with the parametric curves $K(\theta_n, t)$ corresponding to the Clayton, Frank, and Gumbel–Hougaard models, with θ_n estimated in each case through inversion of Kendall's tau. The graph suggests that the Gumbel–Hougaard copula is preferable. This conclusion is confirmed by formal tests based on S_n, T_n and Wang and Wells' statistic $S_{\xi n}$ with $\xi = 0$. The critical points and p-values reported in Table 7 were derived using $N = 10{,}000$ repetitions of the parametric bootstrap procedure described in section 4.2, which is based on C_{θ_n}. While the p-values of S_n, T_n and S_{0n} lead to rejection of the Clayton and Frank dependence structures at the 5% level, they exceed 80% for the Gumbel–Hougaard model.

Additional evidence in favour of the Gumbel–Hougaard extreme-value structure is supplied in Fig. 3, in which the non-parametric estimator $K_n(t)$ is displayed, along with a global 95% confidence band for each of the three Archimedean models considered. Its limits are of the form

$$K(\theta_n, t) \pm \frac{1}{\sqrt{n}} c_{2n}(0.95),$$

where $c_{2n}(0.95)$ is the 95% quantile of T_n under the null hypothesis, as reported in Table 7.

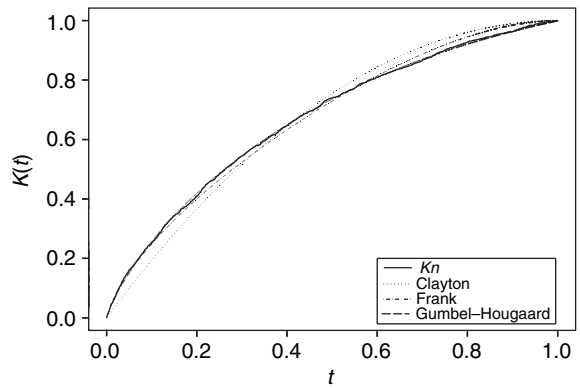

Fig. 2. Non-parametric estimator $K_n(t)$ for the LOSS and ALAE data, along with three parametric estimators $K(\theta_n, t)$ corresponding to the Clayton, Frank and Gumbel–Hougaard copula models, with θ_n estimated by inversion of the empirical version τ_n of Kendall's tau.

Table 7. *Results of the goodness-of-fit tests based on the statistics S_n, T_n and $S_{\xi n}$ with $\xi = 0$ for the data of LOSS and ALAE insurance data*

Model	θ_n	S_n T_n S_{0n}	Critical value c_{2n} (0.95)	p-value (in %)
Clayton	0.939	2.330	0.135	0.0
		2.517	0.910	0.0
		1.892	0.126	0.0
Frank	3.143	0.244	0.123	0.0
		0.903	0.873	3.6
		0.330	0.128	0.0
Gumbel–Hougaard	0.319	0.027	0.117	88.8
		0.483	0.902	84.0
		0.051	0.127	90.2

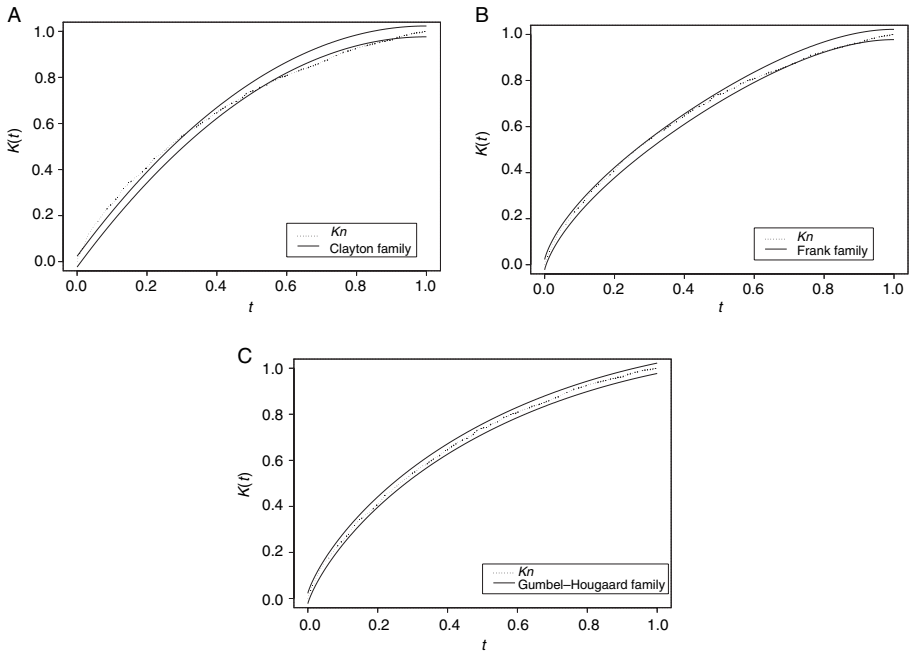

Fig. 3. Non-parametric estimator $K_n(t)$ for the LOSS and ALAE data, along with global 95% confidence bands based on the Kolmogorov–Smirnov statistic T_n for each of the three Archimedean models considered: the Clayton (top left panel), the Frank (top right panel), and Gumbel–Hougaard (bottom panel).

6.2. *Uranium exploration data*

As a second illustration, the analysis of the uranium exploration data set originally considered by Cook & Johnson (1981, 1986) was revisited. These data consist of 655 chemical analyses from water samples collected from the Montrose quadrangle of western Colorado (USA). Concentrations were measured for the following elements: uranium (U), lithium (Li), cobalt (Co), potassium (K), caesium (Cs), scandium (Sc) and titanium (Ti).

Table 8 shows the values of the test statistics S_n, T_n and S_{0n} for selected pairs of variables, along with the corresponding p-values. The latter are based on $N = 10{,}000$ repetitions of the parametric bootstrap procedure.

The following observations can be drawn from Table 8:

(a) In the authors' experience, the three tests are generally in agreement, as for the pairs (U, Li) and (Co, Ti). Occasionally, they lead to different choices of models, as for the pairs (U, Co), (Li, Sc) and (Cs, Sc).
(b) The p-values associated with the various tests can sometimes differ markedly. This is often inconsequential, as in the case of the Gumbel–Hougaard copula for the pair (U, Li). In other occasions, however, the choice of statistic could make a difference between acceptance and rejection at a given level. At the 5% level, for example, Frank's model is acceptable for the pair (Co, Ti), both according to S_n and S_{0n}, but not under T_n. A similar phenomenon can be observed in the pair (Li, Ti), for which the four models considered would be accepted at the 15% level if T_n were used, but rejected by the other two statistics.
(c) Although the statistics S_{0n} computed for different models have different distributions, it can be observed empirically that the model for which the statistic is smallest generally has

Table 8. *Values taken by the goodness-of-fit statistics S_n, T_n and S_{0n} and associated* p*-values, for selected pairs of variables in the uranium exploration data*

Pair	Model	S_n	p-value (in %)	T_n	p-value (in %)	S_{0n}	p-value (in %)
(U, Li)	Ali–Mikhail–Haq	0.0880	22.4	0.6723	43.0	0.0690	33.0
	Clayton	0.3328	0.2	1.2329	0.5	0.2175	0.4
	Frank	**0.0538**	**52.5**	**0.5742**	**67.0**	**0.0448**	**74.2**
	Gumbel–Hougaard	0.1033	14.1	0.6727	42.5	0.1190	5.0
(U, Co)	Ali–Mikhail–Haq	**0.0836**	**24.7**	0.8328	13.7	**0.0679**	**33.4**
	Clayton	0.1057	20.6	**0.7730**	**28.3**	0.0793	24.7
	Frank	0.1099	10.7	0.8614	9.5	0.0862	12.6
	Gumbel–Hougaard	0.1448	6.6	0.9307	6.1	0.1351	4.2
(U, Sc)	Ali–Mikhail–Haq	0.2344	1.0	**1.2362**	0.2	**0.2077**	0.3
	Clayton	0.4042	0.2	1.3436	0.2	0.2934	0.1
	Frank	**0.2285**	**1.2**	1.2402	**0.3**	0.2140	**0.4**
	Gumbel–Hougaard	0.3203	0.1	1.4487	0.1	0.3470	0.0
(Li, Sc)	Ali–Mikhail–Haq	0.1347	6.4	0.9182	5.6	**0.0716**	**26.0**
	Clayton	**0.1120**	**16.6**	**0.8083**	**19.5**	0.0891	17.2
	Frank	0.1634	3.1	1.0443	1.6	0.0817	16.3
	Gumbel–Hougaard	0.2187	0.8	1.1089	0.9	0.1426	1.9
(Li, Ti)	Ali–Mikhail–Haq	0.1493	6.0	0.7456	28.1	0.1553	1.8
	Clayton	0.1382	10.8	0.7724	29.6	0.1578	3.4
	Frank	**0.1256**	12.0	**0.6463**	**53.0**	0.1401	**5.2**
	Gumbel–Hougaard	0.1327	**12.1**	0.6942	41.8	**0.1330**	4.9
(Co, Cs)	Ali–Mikhail–Haq	0.0926	18.4	**0.8172**	13.5	0.0677	33.8
	Clayton	**0.0875**	**27.9**	0.8642	**14.1**	**0.0611**	**43.2**
	Frank	0.2372	0.4	0.9941	0.2	0.1657	0.8
	Gumbel–Hougaard	0.4300	0.0	1.2421	0.0	0.3516	0.0
(Co, Ti)	Ali–Mikhail–Haq	0.6294	0.0	1.5078	0.1	0.5916	0.0
	Clayton	0.6916	0.0	1.5009	0.0	0.5437	0.0
	Frank	**0.0731**	**23.0**	**0.9230**	**2.9**	**0.0539**	**48.5**
	Gumbel–Hougaard	0.2252	0.0	0.9899	1.4	0.2687	0.0
(Cs, Sc)	Ali–Mikhail–Haq	**0.1087**	**16.3**	0.8675	12.5	**0.0851**	**21.3**
	Clayton	0.1657	3.0	**0.8429**	**13.0**	0.1395	2.1
	Frank	0.2382	0.0	1.1769	0.0	0.1320	2.6
	Gumbel–Hougaard	0.3732	0.0	1.1256	1.0	0.2657	0.1

the highest p-value. The pairs (U, Sc) and (Li, Ti) provide counterexamples which might or might not be due to sampling error. The same phenomenon could be observed for the statistics S_n and T_n; see, for example, pairs (U, Sc), (Li, Ti) and (Co, Cs).

7. Discussion

The goodness-of-fit procedures proposed herein will be consistent so long as $K(t) = P\{H(\mathbf{X}) \leq t\}$ assumes different functional forms under the hypothesized copula model and the true one. As already argued by Wang & Wells (2000) in the case of $S_{\xi n}$, such is the case for bivariate Archimedean copulas, a result which stems from the fact established by Genest & Rivest (1993) that the Archimedean generator ϕ is completely determined by K. This argument extends readily to S_n and T_n, or any other continuous functional of the process $\mathbb{K}_n$. It may be conjectured that this characterization of Archimedean generators extends to arbitrary dimension $d \geq 2$. For, in the light of Equation (6) from Barbe *et al.* (1996), one can check easily that $y = \phi^{-1}$ is a solution of the differential equation

$$\sum_{i=0}^{d-1} \frac{(-1)^i}{i!} t^i y^{(i)}(t) - K\{\theta, y(t)\} = 0, \tag{8}$$

of order $d-1$ with boundary conditions $y(0) = 1$, and $t^i y^{(i)}(t) \to 0$ as $t \to \infty$ for all $i \in \{1, \ldots, d-1\}$. It follows that if, given $K(\theta, \cdot)$, the solution of (8) is unique up to a scaling parameter, then the copula C_θ associated with $K(\theta, \cdot)$ is unique. In other words, the copula is uniquely determined by K whenever any two solutions y_1 and y_2 of the above equation satisfy $y_2(t) = y_1(\alpha t)$, for some $\alpha > 0$.

Nevertheless, there are circumstances when goodness-of-fit tests based on the process $\mathbb{K}_n$ will not be consistent. Suppose, for example, that C_1 and C_2 are two extreme-value copulas with the same value, τ, of Kendall's tau. In such a case, one would have $K(t) = t - (1-\tau)t \log(t)$ for both models. Although the process $\mathbb{K}_n$ may not have the same asymptotic distribution accordingly as the data arise from C_1 or C_2, the limit would be a centred Gaussian process in both cases. Accordingly, the power of test statistics such as S_n and T_n, or indeed any other continuous functional of $\mathbb{K}_n$, could not approach 1 as $n \to \infty$.

One important advantage of the procedures proposed herein is that they are applicable to situations involving more than two variables. As a brief illustration, the goodness-of-fit of trivariate Ali–Mikhail–Haq, Clayton, Frank and Gumbel–Hougaard copula models was checked on the triplet (Li, K, Ti).

Table 9 summarizes the results of the tests based on statistics S_n and T_n with $d = 3$. The p-values associated with the Frank and Gumbel–Hougaard dependence structures clearly lead to the rejection of those models. The Ali–Mikhail–Haq copula is also rejected at the 5% level by the Kolmogorov–Smirnov test. And indeed, for these three models, it may also be checked graphically (figures not provided) that the non-parametric estimator K_n lies in part outside the global 95% confidence band, while K_n lies inside the global 95% confidence band for the Clayton copula (Fig. 4). There is thus no evidence to conclude that formula

$$C_\theta(u, v, w) = \left(u^{-\theta} + v^{-\theta} + w^{-\theta} - 2\right)^{-1/\theta}, \quad \theta > 0 \tag{9}$$

should be discarded as a potential model for these data.

In subsequent work, it would be of interest to extend the set of copulas for which hypotheses I–IV are met. Because of their popularity in survival data analysis, where their mixture representation allows them to be viewed as a natural extension of Cox's proportional hazards model (Oakes, 2001), multiparameter Archimedean models such as those considered by Joe (1997) and Genest *et al.* (1998) should probably be considered first. For the (Li, K, Ti) data considered just above, for example, it may well be that a multiparameter

Table 9. *Results of the goodness-of-fit tests based on the Cramér–von Mises and Kolmogorov–Smirnov statistics S_n and T_n for the trivariate data involving concentrations of lithium, potassium and titanium*

Model	θ_n		Critical value	p-value (in %)
		S_n		
		T_n		
Ali–Mikhail–Haq	0.242	1.106	1.370	14.8
		2.184	2.154	4.3
Clayton	0.122	0.264	0.837	50.8
		1.225	1.911	47.2
Frank	0.548	1.140	0.724	0.6
		2.193	1.702	0.5
Gumbel–Hougaard	0.055	1.347	0.656	0.0
		2.235	1.645	0.2

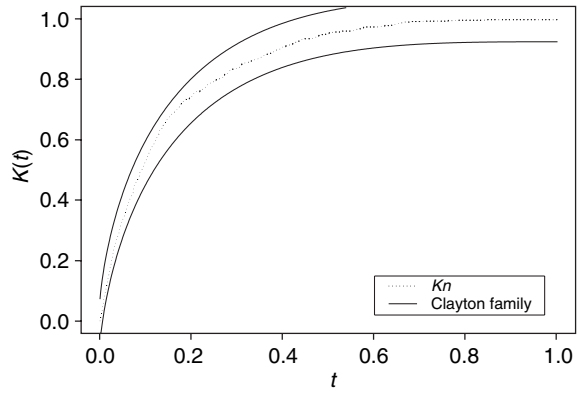

Fig. 4. Non-parametric estimator $K_n(t)$ for the lithium, potassium and titanium concentration data, along with global 95% confidence bands based on the Kolmogorov–Smirnov statistic T_n for the trivariate Clayton model.

copula model with different Clayton margins for different pairs of variables (see e.g. Bandeen-Roche & Liang, 1996) might be more appropriate than the somewhat restrictive dependence structure (9).

Acknowledgements

Partial funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada, the Fonds québécois de la recherche sur la nature et les technologies, as well as the Institut de finance mathématique de Montréal. The authors thank the Editor, the Associate Editor and a referee for their useful comments on an earlier version of this paper.

References

Ali, M. M., Mikhail, N. N. & Haq, M. S. (1978). A class of bivariate distributions including the bivariate logistic. *J. Multivariate Anal.* **8**, 405–412.

Bandeen-Roche, K. J. & Liang, K.-Y. (1996). Modelling failure-time associations in data with multiple levels of clustering. *Biometrika* **83**, 29–39.

Barbe, P., Genest, C., Ghoudi, K. & Rémillard, B. (1996). On Kendall's process. *J. Multivariate Anal.* **58**, 197–229.

Belguise, O. & Lévi, C. (2001–2002). Tempêtes: Étude des dépendances entre les branches automobile et incendie à l'aide de la théorie des copulas. *Bulletin français d'actuariat* **5**, 135–174.

Capéraà, P. & Fougères, A.-L. (2001). Estimation of a bivariate extreme value distribution. *Extremes* **3**, 311–329.

Capéraà, P., Fougères, A.-L. & Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika* **84**, 567–577.

Capéraà, P., Fougères, A.-L. & Genest, C. (2000). Bivariate distributions with given extreme value attractor. *J. Multivariate Anal.* **72**, 30–49.

Chen, X. & Fan, Y. (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. *Can. J. Statist.* **33**, 389–414.

Cherubini, U. & Luciano, E. (2002). Bivariate option pricing with copulas. *Appl. Math. Finance* **9**, 69–85.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.

Cook, R. D. & Johnson, M. E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *J. Roy. Statist. Soc. Ser. B* **43**, 210–218.

Cook, R. D. & Johnson, M. E. (1986). Generalized Burr–Pareto–logistic distributions with applications to a uranium exploration data set. *Technometrics* **28**, 123–131.

Dakhli, T. (2004). *Analyse de la dépendance de défaut et évaluation des dérivés de crédit sur portefeuille*. Master's thesis, HEC Montréal, Canada.

Embrechts, P., McNeil, A. J. & Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk management: value at risk and beyond* (ed. M. A. H. Dempster), 176–223. Cambridge University Press, Cambridge.

Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Math.* **19**, 194–226.

Frees, E. W. & Valdez, E. A. (1998). Understanding relationships using copulas. *North Am. Act. J.* **2**, 1–25.

Genest, C. & MacKay, R. J. (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *Can. J. Statist.* **14**, 145–159.

Genest, C. & Rémillard, B. (2005). *Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models*. Les Cahiers du GERAD, no. G–2005–51, HEC Montréal, Canada.

Genest, C. & Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *J. Amer. Statist. Assoc.* **88**, 1034–1043.

Genest, C. & Rivest, L.-P. (2001). On the multivariate probability integral transformation. *Statist. Probab. Lett.* **53**, 391–399.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1998). Comment on a paper by E. W. Frees and E. A. Valdez entitled "Understanding relationships using copulas". *North. Am. Act. J.* **2**, 143–149.

Genest, C., Quessy, J.-F. & Rémillard, B. (2006). On the joint asymptotic behavior of two rank-based estimators of the association parameter in the gamma frailty model. *Statist. Probab. Lett.* **76**, 10–18.

Ghoudi, K. & Rémillard, B. (1998). Empirical processes based on pseudo-observations. In *Asymptotic methods in probability and statistics: a volume in honour of Miklós Csörgő* (ed. B. Szyskowitz), 171–197. Elsevier, Amsterdam.

Ghoudi, K. & Rémillard, B. (2004). Empirical processes based on pseudo-observations II: the multivariate case. In *Asymptotic methods in stochastics: festschrift for Miklós Csörgő* (eds L. Horváth & B. Szyskowitz), 381–406. The Fields Institute Communications Series, vol. 44, American Mathematical Society, Providence, RI.

Ghoudi, K., Khoudraji, A. & Rivest, L.-P. (1998). Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. *Can. J. Statist.* **26**, 187–197.

Glidden, D. V. (1999). Checking the adequacy of the gamma frailty model for multivariate failure times. *Biometrika* **86**, 381–393.

van den Goorbergh, R. W. J., Genest, C. & Werker, B. J. M. (2005). Bivariate option pricing using dynamic copula models. *Insur. Math. Econ.* **37**, 101–114.

Gumbel, E. J. (1960). Distribution des valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris* **9**, 171–173.

Hennessy, D. A. & Lapan, H. E. (2002). The use of Archimedean copulas to model portfolio allocations. *Math. Finance* **12**, 143–154.

Joe, H. (1997). *Multivariate models and dependence concepts*. Chapman & Hall, London.

Jouini, M. N. & Clemen, R. T. (1996). Copula models for aggregating expert opinions. *Oper. Res.* **44**, 444–457.

Klugman, S. A. & Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insur. Math. Econ.* **24**, 139–148.

Lauprete, G. J., Samarov, A. M. & Welsch, R. E. (2002). Robut portfolio optimization. *Metrika* **55**, 139–149.

Li, D. X. (2000). On default correlation: a copula function approach. *J. Fixed Income* **9**, 43–54.

Marshall, A. W. & Olkin, I. (1988). Families of multivariate distributions. *J. Amer. Statist. Assoc.* **83**, 834–841.

Nelsen, R. B. (1999). *An introduction to copulas*. Lecture Notes in Statistics no. 139. Springer, New York.

Oakes, D. (1989). Bivariate survival models induced by frailties. *J. Amer. Statist. Assoc.* **84**, 487–493.

Oakes, D. (2001). *Biometrika* centenary: survival analysis. *Biometrika* **88**, 99–142.

Pickands, J. (1981). Multivariate extreme value distributions. *Bull. Int. Statist. Inst.*, 859–878.

Shih, J. H. (1998). A goodness-of-fit test for association in a bivariate survival model. *Biometrika* **85**, 189–200.

Shih, J. H. & Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384–1399.

Tawn, J. A. (1988). Bivariate extreme value theory: models and estimation. *Biometrika* **75**, 397–415.
Wang, W. & Wells, M. T. (2000). Model selection and semiparametric inference for bivariate failure-time data. *J. Amer. Statist. Assoc.* **95**, 62–72.

Received September 2003, in final form June 2005

Christian Genest, Département de mathématiques et de statistique, Université Laval, Québec, Canada G1K 7P4.
E-mail: christian.genest@mat.ulaval.ca

Appendix A: A convergence result

This appendix offers a proof of the fact that under hypotheses III and IV given in section 2, the process $B_n(t) = \sqrt{n}\{K(\theta_n, t) - K(\theta, t)\}$ is such that

$$\sup_{t\in[0,1]} \left|B_n(t) - \dot{K}(\theta, t)^\top \Theta_n\right| \xrightarrow{P} 0.$$

To see this, let $\lambda > 0$ be arbitrary. By hypothesis III, the sequence (Θ_n) is tight as it converges in law to Θ. Hence for any given $\delta > 0$, there exist $M = M_\delta \in \mathbb{R}^+$ and N_0 such that $P(||\Theta_n|| > M) < \delta$ for all $n \geq N_0$. For any such n, one has

$$\begin{aligned}
&P\left\{\sup_{t\in[0,1]} \left|B_n(t) - \dot{K}(\theta, t)^\top \Theta_n\right| > \lambda\right\} \\
&\leq P\left\{\sup_{t\in[0,1]} \left|B_n(t) - \dot{K}(\theta, t)^\top \Theta_n\right| > \lambda, ||\Theta_n|| \leq M\right\} + P(||\Theta_n|| > M) \\
&< P\left\{\sup_{t\in[0,1]} \left|B_n(t) - \dot{K}(\theta, t)^\top \Theta_n\right| > \lambda, ||\Theta_n|| \leq M\right\} + \delta.
\end{aligned}$$

Next, the mean-value theorem implies that for any realization of Θ_n, there exists θ_n^* with $|\theta_n^* - \theta| \leq |\Theta_n|/\sqrt{n}$ such that $B_n(t) = \dot{K}(\theta_n^*, t)^\top \Theta_n$. Hence, using hypothesis IV, one obtains

$$\begin{aligned}
&\lim_{n\to\infty} P\left\{\sup_{t\in[0,1]} \left|B_n(t) - \dot{K}(\theta, t)^\top \Theta_n\right| > \lambda, ||\Theta_n|| \leq M\right\} \\
&\leq \lim_{n\to\infty} P\left\{||\Theta_n|| \sup_{t\in[0,1]} ||\dot{K}(\theta_n^*, t) - \dot{K}(\theta, t)|| > \lambda, ||\Theta_n|| \leq M\right\} \\
&\leq \lim_{n\to\infty} P\left\{\sup_{||\theta^*-\theta||\leq M/\sqrt{n}} \sup_{t\in[0,1]} ||\dot{K}(\theta^*, t) - \dot{K}(\theta, t)|| > \frac{\lambda}{M}\right\} = 0.
\end{aligned}$$

As δ can be chosen arbitrarily small, the result follows.

Appendix B: Verification of hypothesis IV for various copula models

B1. Ali–Mikhail–Haq copulas

In view of relation (7), $f_i(\theta, t)$ is a polynomial in t with coefficients that are non-negative whenever $\theta \in (0, 1)$, so that

$$\phi_\theta^{-1}(x) = \frac{1-\theta}{\mathrm{e}^{-x(1-\theta)} - \theta}, \quad x > 0$$

satisfies condition (5) for every integer $d \geq 1$, and hence is completely monotone. It is easy to see that

$$K(\theta, t) = t + t \sum_{i=1}^{d-1} \frac{p_i(\theta, t)}{(1-\theta)^i} \left\{ \log\left(\frac{1-\theta}{t} + \theta \right) \right\}^i,$$

where $p_i(\theta, t) = f_i(\theta, t)/t$ is also a polynomial in both θ and t. Then

$$\dot{K}(\theta, t) = t \sum_{i=1}^{d-1} \frac{(1-\theta)\dot{p}_i(\theta, t) + i p_i(\theta, t)}{(1-\theta)^{i+1}} \left\{ \log\left(\frac{1-\theta}{t} + \theta \right) \right\}^i - \frac{t(1-t)}{1-\theta+\theta t} \sum_{i=1}^{d-1} \frac{i p_i(\theta, t)}{(1-\theta)^i} \left\{ \log\left(\frac{1-\theta}{t} + \theta \right) \right\}^{i-1}$$

is continuous on $(-1, 1) \times [0, 1]$. Hence (3) holds for all $\theta \in (-1, 1)$.

Note, however, that hypothesis IV does not generally hold at $\theta = 1$. In the case $d = 2$, for example,

$$0 = \lim_{\theta \to 1} \lim_{t \to 0} \dot{K}(\theta, t) \neq \lim_{t \to 0} \lim_{\theta \to 1} \dot{K}(\theta, t) = \frac{1}{2}.$$

B2. Clayton copulas

First, it is easily seen that

$$f_{i,\theta}(t) = \frac{d^i}{\mathrm{d}s^i} \phi_\theta^{-1}(s) \bigg|_{s=\phi_\theta(t)} = (-1)^i q(\theta, i, 1) t^{1+i\theta},$$

where $q(\theta, i, m) = \prod_{j=0}^{i-1}(m + j\theta)$ is a polynomial of degree $i - 1$ in θ. It then follows that

$$K(\theta, t) = t + t \sum_{i=1}^{d-1} \left(\frac{1-t^\theta}{\theta} \right)^i \frac{q(\theta, i, 1)}{i!} \quad \text{and} \quad k(\theta, t) = \left(\frac{1-t^\theta}{\theta} \right)^{d-1} \frac{q(\theta, d, 1)}{(d-1)!}.$$

Now since

$$\int_0^1 t\left(1 - t^\theta\right)^i \mathrm{d}t = \frac{\Gamma(2/\theta) i!}{\theta \Gamma(2/\theta + i + 1)} = \frac{\theta^{i+1} i!}{\theta \prod_{j=0}^{i}(2 + j\theta)} = \frac{\theta^i i!}{q(\theta, i+1, 2)},$$

one has, according to formula (4),

$$\tau(\theta) = 1 - \left(\frac{2^d}{2^{d-1} - 1} \right) \sum_{i=1}^{d-1} \frac{q(\theta, i, 1)}{q(\theta, i+1, 2)} = \left(\frac{2^d}{2^{d-1} - 1} \right) \frac{q(\theta, d, 1)}{q(\theta, d, 2)} - \frac{1}{2^{d-1} - 1}.$$

Writing $\log q(\theta, i, m) = \sum_{j=0}^{i-1} \log(m + j\theta)$, it follows that

$$q'(\theta,i,m)=q(\theta,i,m)\sum_{j=1}^{i-1}\left(\frac{j}{m+j\theta}\right)$$

and

$$\dot{k}(\theta,t)=k(\theta,t)\left(\sum_{j=1}^{d-1}\frac{j}{1+j\theta}\right)-\left(\frac{1-t^{\theta}}{\theta}\right)^{d-2}\frac{q(\theta,d,1)}{(d-2)!}\left(\frac{t^{\theta}}{\theta}\log t+\frac{1-t^{\theta}}{\theta^2}\right),$$

which is clearly continuous for $(0,\infty)\times[0,1]$.

Note that in this case, hypothesis IV also holds true at the boundary value $\theta=0$. For,

$$\lim_{\theta\to 0^{+}}\dot{K}(\theta,t)=\frac{-t(-\log t)^{d}}{2(d-2)!},$$

so that

$$\dot{K}(\varepsilon,t)+\frac{t(-\log t)^{d}}{2(d-2)!}=\frac{F(\varepsilon,t)}{\varepsilon^{d}},$$

where

$$F(\varepsilon,t)=t\sum_{i=1}^{d-1}\frac{\varepsilon^{d-i-1}(1-t^{\varepsilon})^{i-1}}{(i-1)!}\left\{t^{\varepsilon}-\varepsilon t^{\varepsilon}\log t-1+\frac{\varepsilon(1-t^{\varepsilon})}{i}\sum_{j=1}^{i-1}\frac{j}{1+j\varepsilon}\right\}+\frac{t(-\varepsilon\log t)^{d}}{2(d-2)!}.$$

Next, in view of the general fact that

$$\frac{d^{p}}{\mathrm{d}\varepsilon^{p}}f(\varepsilon)g(\varepsilon)h(\varepsilon)=\sum_{k=0}^{p}\sum_{\ell=0}^{k}\binom{p}{k}\binom{k}{\ell}f^{(\ell)}(\varepsilon)g^{(k-\ell)}(\varepsilon)h^{(p+1-k)}(\varepsilon),$$

and since

$$\frac{\partial^{d}}{\partial\varepsilon^{d}}(1-t^{\varepsilon})^{i}=(\log t)^{d}\sum_{k=1}^{i}\binom{i}{k}(-t^{\varepsilon})^{k}k^{d}$$

is continuous and bounded for all $(\varepsilon,t)\in[0,1]^2$, one finds that

$$F^{(d+1)}(\varepsilon,t)=\frac{\partial^{d+1}}{\partial\varepsilon^{d+1}}F(\varepsilon,t)$$

is bounded by a constant $M>0$ on the unit square. Hence

$$\sup_{(\varepsilon,t)\in[0,1]^2}|F(\varepsilon,t)|=\left|\int_0^{\varepsilon}\int_0^{u_{d+1}}\cdots\int_0^{u_2}F^{(d+1)}(u_1,t)\,\mathrm{d}u_1\cdots\mathrm{d}u_{d+1}\right|\le\frac{\varepsilon^{d+1}M}{(d+1)!}.$$

Therefore

$$\lim_{\varepsilon\to 0}\sup_{t\in[0,1]}\left|\frac{F(\varepsilon,t)}{\varepsilon^{d}}\right|\le\lim_{\varepsilon\to 0}\frac{\varepsilon M}{(d+1)!}=0.$$

Consequently, hypothesis IV is satisfied for all $\theta\in(0,\infty)$, as well as at the boundary value $\theta=0$.

B3. Frank copulas

In the light of an observation made in section 2, and from the fact that

$$\left|\ddot{K}(\theta,t)\right| \le \int_0^t \left|\ddot{k}(\theta,s)\right| \mathrm{d}s \le \int_0^1 \left|\ddot{k}(\theta,s)\right| \mathrm{d}s,$$

it suffices to show that $\ddot{k}(\theta,s)$ is uniformly bounded for all $\theta \in \mathcal{O}$ and $s \in [0, 1]$ in order to verify hypothesis IV. For that purpose, introduce the continuous and bounded function

$$\psi(x) = \log\left(\frac{1-\mathrm{e}^{-x}}{x}\right), \quad x > 0$$

so that $\phi_\theta(t) = \psi(\theta) - \psi(\theta t) - \log t$. Next, let $p_0(x) = x - 1$ and $p_i(x) = x(1-x)p'_{i-1}(x)$ for arbitrary integer $i \ge 1$. With this notation, the formula already derived by Barbe *et al.* (1996) can be written as

$$k(\theta,s) = \mathrm{e}^{\theta t} p'_{d-2}(\mathrm{e}^{\theta t}) \frac{\{\log t + \psi(\theta t) - \psi(\theta)\}^{d-1}}{(d-1)!}, \quad \theta \ge 0.$$

One can show that

$$\begin{aligned}
&\ddot{k}(\theta,s) \\
&= s\dot{k}(\theta,s) + \frac{s^2\,\mathrm{e}^{2\theta s}}{(d-1)!}\{\log s + \psi(\theta s) - \psi(\theta)\}^{d-1}\{2p''_{d-2}(\mathrm{e}^{\theta s}) + \mathrm{e}^{\theta s}p'''_{d-2}(\mathrm{e}^{\theta s})\} \\
&\quad + \frac{s\,\mathrm{e}^{\theta s}}{(d-2)!}\{\log s + \psi(\theta s) - \psi(\theta)\}^{d-2}\{s\psi'(\theta s) - \psi'(\theta)\}\{2\,\mathrm{e}^{\theta s}p''_{d-2}(\mathrm{e}^{\theta s}) + p'_{d-2}(\mathrm{e}^{\theta s})\} \\
&\quad + \frac{\mathrm{e}^{\theta s}p'_{d-2}(\mathrm{e}^{\theta s})}{(d-3)!}\{\log s + \psi(\theta s) - \psi(\theta)\}^{d-3}\{s\psi'(\theta s) - \psi'(\theta)\}^2 \\
&\quad + \frac{\mathrm{e}^{\theta s}p'_{d-2}(\mathrm{e}^{\theta s})}{(d-2)!}\{\log s + \psi(\theta s) - \psi(\theta)\}^{d-2}\{s^2\psi''(\theta s) - \psi''(\theta)\}
\end{aligned}$$

is bounded for all $(\theta, s) \in (-\infty, \infty) \times [0, 1]$ since both ψ' and ψ'' are bounded and

$$\begin{aligned}
\dot{k}(\theta,s) = sk(\theta,s) &+ \frac{s\,\mathrm{e}^{2\theta s}p''_{d-2}(\mathrm{e}^{\theta s})}{(d-1)!}\{\log s + \psi(\theta s) - \psi(\theta)\}^{d-1} \\
&+ \frac{\mathrm{e}^{\theta s}p'_{d-2}(\mathrm{e}^{\theta s})}{(d-2)!}\{\log s + \psi(\theta s) - \psi(\theta)\}^{d-2}\{s\psi'(\theta s) - \psi'(\theta)\}.
\end{aligned}$$

B4. Gumbel–Hougaard copulas

This copula also belongs to the family of extreme-value copulas, further discussed in section 3.2. From Barbe *et al.* (1996),

$$k(\theta,t) = \frac{p_{d-1}(-\log t)}{(d-1)!},$$

where $p_0(x) \equiv 1$ and for integer $i \ge 1$,

$$p_i(x) = (1-\theta)x\{p_{i-1}(x) - p'_{i-1}(x)\} + (\theta + i - 1)p_{i-1}(x)$$

is a polynomial of degree i in x and in θ. This corrects a typographical error on p. 207 of Barbe *et al.* (1996), where one should have read

$$p_i(x) = \theta x\{p_{i-1}(x) - p'_{i-1}(x)\} + (i - \theta)p_{i-1}(x)$$

in the parametrization used there.

Writing $p_{d-1}(x) = \sum_{k=0}^{d-1} r_k(\theta)x^k$, one may thus conclude that

$$K(\theta,t) = \int_0^t k(\theta,s)\,\mathrm{d}s = \sum_{k=0}^{d-1} \frac{r_k(\theta)}{(d-1)!} \int_0^t (-\log s)^k\,\mathrm{d}s = \frac{t}{(d-1)!} \sum_{k=0}^{d-1} k\; r_k(\theta) \sum_{i=0}^{k} \frac{(-\log t)^i}{i!},$$

so that

$$\dot{K}(\theta,t) = \frac{t}{(d-1)!} \sum_{k=0}^{d-1} k! r_k'(\theta) \sum_{i=0}^{k} \frac{(-\log t)^i}{i!}$$

is clearly continuous on $[0,\,1]^2$.

B5. Fréchet copulas

Note that for copulas in this class, one has

$$\dot{K}(\theta,t) = t\frac{\dot{\zeta}(\theta,t)}{\zeta(\theta,t)} = -2\frac{t}{I(\theta,t)} + 4\frac{t^2}{I(\theta,t)\{I(\theta,t)+\theta\}} = -2\frac{t}{I(\theta,t)} + \frac{t\zeta(\theta,t)\{I(\theta,t)+\theta\}}{I(\theta,t)}.$$

It is easy to check that $t/I\,(\theta,\,t)$ is continuous on $[0,\,1]^2$, whence hypothesis IV holds true on $\mathcal{O} = (0,\,1)$. To see that the latter is also verified at $\theta = 1$, note that $\dot{K}(\theta,t) \to -t$ as $\theta \to 1$. Putting $\delta = 1 - \varepsilon$, one gets $\dot{K}(\varepsilon,t) + t = F(\delta,t)/\delta$, where

$$F(\delta,t) = t\left\{\delta + 1 - \frac{\delta+1}{I(1-\delta,t)}\right\}.$$

It can easily be shown that

$$\sup_{(\delta,t)\in[0,1)\times[0,1]} \left|\ddot{K}(\delta,t)\right| = \sup_{(\delta,t)\in[0,1)\times[0,1]} \left|\frac{\left\{4t(1-3t) + 4\delta t(\delta+t-3) - 4(\delta-1)^2\right\}t}{\left(I_{1-\delta,t}\right)^{5/2}}\right|$$

is bounded above by some constant M. Thus,

$$\sup_{(\delta,t)\in[0,1)\times[0,1]} |F(\delta,t)| = \sup_{(\delta,t)\in[0,1)\times[0,1]} \left|\int_0^\delta \int_0^v F(u,t)\,\mathrm{d}u\,\mathrm{d}v\right| \le \frac{\delta^2 M}{2},$$

which implies that $\sup_{(\delta,t)\,\in\,[0,1)\times[0,1]} |F(\delta,\,t)/\delta| \le \delta M/2 \to 0$ as $\varepsilon = 1 - \delta \to 1$.

B6. Farlie–Gumbel–Morgenstern copulas

Here, one has

$$\begin{aligned}\dot{K}(\theta,t) &= \int_0^t \int_s^1 \dot{h}(\theta,x,s)\,\mathrm{d}x\,\mathrm{d}s = -\int_0^t \int_s^1 \frac{\dot{r}(\theta,x,t)}{(1-x)\{r(\theta,x,t)\}^2}\,\mathrm{d}x\,\mathrm{d}s \\ &= \int_0^t \int_s^1 \left[\frac{\{1-\theta(1-x)\} - 2(1-s/x)}{\{r(\theta,x,t)\}^3}\right]\mathrm{d}x\,\mathrm{d}s.\end{aligned}$$

Now for arbitrary $\theta_1,\,\theta_2 \in (-1+\delta,\,1-\delta)$ for fixed $0 < \delta < 1$, one finds

$$
\begin{aligned}
\left|\dot{K}(\theta_2,t)-\dot{K}(\theta_1,t)\right| &= \left|\int_0^t\int_s^1 \left\{2\left(1-\frac{s}{x}\right)-1\right\}\left[\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right]\mathrm{d}x\,\mathrm{d}s\right. \\
&\quad \left. +\int_0^t\int_s^1 (1-x)\left[\frac{\theta_2}{\{r(\theta_2,x,s)\}^3}-\frac{\theta_1}{\{r(\theta_1,x,s)\}^3}\right]\mathrm{d}x\,\mathrm{d}s\right| \\
&\le 3\int_0^1\int_s^1\left|\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right|\mathrm{d}x\,\mathrm{d}s \\
&\quad +\int_0^1\int_s^1\left|\frac{\theta_2}{\{r(\theta_2,x,s)\}^3}-\frac{\theta_1}{\{r(\theta_1,x,s)\}^3}\right|\mathrm{d}x\,\mathrm{d}s \\
&\le 4\int_0^1\int_s^1\left|\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right|\mathrm{d}x\,\mathrm{d}s \\
&\quad +|\theta_2-\theta_1|\int_0^1\int_s^1\frac{1}{\{r(\theta_2,x,s)\}^3}\,\mathrm{d}x\,\mathrm{d}s.
\end{aligned}
$$

As $r(\theta, x, s) \ge 1 - |\theta| > \delta$, the second summand is bounded above by $|\theta_2 - \theta_1|/\delta^3$. To handle the first summand, note that $r(\theta,x,s) \le \sqrt{8}$ and $|\dot{r}(\theta,x,s)| \le 4/\delta$. It follows that

$$
\left|\{r(\theta_2,x,s)\}^3-\{r(\theta_2,x,s)\}^3\right| = 3\left|\int_{\theta_1}^{\theta_2}\{r(\theta,x,s)\}^2\dot{r}(\theta,x,s)\,\mathrm{d}\theta\right| \le \frac{96}{\delta}|\theta_2-\theta_1|.
$$

Therefore,

$$
\left|\frac{1}{\{r(\theta_2,x,s)\}^3}-\frac{1}{\{r(\theta_1,x,s)\}^3}\right| \le \frac{96}{\delta^7}|\theta_2-\theta_1|,
$$

so the first summand is bounded by $384|\theta_2 - \theta_1|/\delta^7$. Hence hypothesis IV is satisfied for all $\theta \in \mathcal{O} = (0, 1)$.

Appendix C: Asymptotic behaviour of the parametric bootstrap method of Wang & Wells (2000)

Without loss of generality, one may assume that $\theta = \tau$. To show that the suggested methodology is incorrect, even in the uncensored case, let $U_1, \ldots, U_n$ be independent uniformly distributed random variables in (0, 1) and for a given value τ_n, set $V_{j,n}^* = K_{\tau_n}^{-1}(U_j)$. Then $V_{j,n}^*$ has distribution $K(\tau_n, \cdot)$.

For simplicity, set $\dot{K}(u) = \partial K(\tau,u)/\partial\tau$, $K(\cdot) = K(\tau, \cdot)$ and $k(\cdot) = k(\tau, \cdot)$. Then, one can show that

$$
\dot{Q}(u) = \frac{\partial}{\partial\tau}K_\tau^{-1}(u) = -\frac{\dot{K}\{K^{-1}(u)\}}{k\{K^{-1}(u)\}},
$$

and under appropriate regularity conditions on K, one gets

$$
\begin{aligned}
E\{\dot{Q}(U_j)\} &= -\int_0^1\frac{\dot{K}\{K^{-1}(u)\}}{k\{K^{-1}(u)\}}\,\mathrm{d}u = -\int_0^1\dot{K}(t)\,\mathrm{d}t \\
&= \frac{\partial}{\partial\tau}\int_0^1\{1-K_\tau(t)\}\,\mathrm{d}t = \frac{\partial}{\partial\tau}\left(\frac{\tau+1}{4}\right) = \frac{1}{4}.
\end{aligned}
$$

Next,

$$K_n^*(t) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}\{U_j \le K(\tau_n, t)\} = \frac{1}{\sqrt{n}}\beta_n^* \circ K(\tau_n, t) + K(\tau_n, t),$$

where

$$\beta_n^*(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\mathbf{1}(U_j \le t) - t\}$$

converges in law to a Brownian bridge β^*, independent of $\mathbb{K}_\theta$. Hence,

$$\begin{aligned}\sqrt{n}\{K_n^*(t) - K(\tau_n^*, t)\} &= \beta_n^* \circ K(\tau_n, t) + \sqrt{n}\{K(\tau_n, t) - K(\tau_n^*, t)\} \\ &= \beta_n^* \circ K(\tau_n, t) - \sqrt{n}(\tau_n^* - \tau_n)\dot{K}(t) + o_P(1).\end{aligned}$$

Moreover, denoting by E_U the expectation with respect to U_j, and setting

$$\bar{\tau}_n = -1 + \frac{4}{n}\sum_{j=1}^{n} K^{-1}(U_j),$$

one obtains

$$\begin{aligned}\sqrt{n}(\tau_n^* - \tau_n) &= \sqrt{n}\left[\frac{4}{n}\sum_{j=1}^{n} K_{\tau_n}^{-1}(U_j) - 4E_U\left\{K_{\tau_n}^{-1}(U_j)\right\}\right] \\ &= \frac{4}{\sqrt{n}}\sum_{j=1}^{n}\left\{K_{\tau_n}^{-1}(U_j) - K^{-1}(U_j)\right\} + \sqrt{n}(\bar{\tau}_n - \tau) \\ &\qquad - 4\sqrt{n}\,E_U\left\{K_{\tau_n}^{-1}(U_1) - K^{-1}(U_1)\right\} \\ &= 4\sqrt{n}(\tau_n - \tau)\left\{\frac{1}{n}\sum_{j=1}^{n}\dot{Q}(U_j)\right\} + \sqrt{n}(\bar{\tau}_n - \tau) \\ &\qquad - 4\sqrt{n}(\tau_n - \tau)\,E_U\{\dot{Q}(U_1)\} + o_P(1) \\ &= \sqrt{n}(\bar{\tau}_n - \tau) + o_P(1).\end{aligned}$$

It follows that $\sqrt{n}\{K_n^*(t) - K(\tau_n^*, t)\}$ converges in distribution to $\beta^* \circ K(t) - Z^*\dot{K}(t)$, where

$$Z^* = -4\int_0^1 \beta^* \circ K(t)\,\mathrm{d}t$$

is the limit in distribution of $\sqrt{n}(\tau_n^* - \tau_n)$, since

$$\sqrt{n}(\bar{\tau}_n - \tau) = -4\int_0^1 \beta_n^* \circ K(t)\,\mathrm{d}t.$$

Note also that Z^* and the limit Θ of $\sqrt{n}(\tau_n - \tau)$ are independent random variables, and that their distributions are generally different. In fact, $\Theta = 2Z - 4\mathcal{X}_1 - 4\mathcal{X}_2$, where

$$Z = -4\int_0^1 \beta \circ K(t)\,\mathrm{d}t$$

is an independent copy of Z, and

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\{F_i(X_{ij})-\frac{1}{2}\right\} \rightsquigarrow \mathcal{X}_i, \quad i=1,2.$$

Thus, while the limiting process $\mathbb{K}^* = \beta^* \circ K - Z^*\dot{K}$ is independent of the limit $\mathbb{K} = \mathbb{K}_\tau - \Theta\dot{K}$ appearing in proposition 1, their distributions are clearly not identical, as follows from Barbe *et al.* (1996). Therefore, the procedure cannot be used to estimate any functional of $\mathbb{K}_n(\cdot) - K(\tau_n, \cdot)$; in particular it cannot be used to estimate the variance of $S_{\xi n}$ or any p-value.

Appendix D: Assumptions for the parametric bootstrap based on C_{θ_n}

Here are the conditions under which Genest & Rémillard (2005) prove that the parametric bootstrap procedure described in section 4 is valid:

(R1) For any $\theta \in \mathcal{O}$, the densities c_θ of C_θ exists, are strictly positive on $(0, 1)^d$, and are twice differentiable with respect to θ. Moreover, for any $\theta_0 \in \mathcal{O}$,
- $\theta \mapsto \dot{c}_\theta(u)/c_\theta(u)$ and $\theta \mapsto \ddot{c}_\theta(x)/c_\theta(x)$ are continuous at θ_0, for almost every $u \in (0, 1)^d$;
- there is a neighbourhood $\mathcal{N} = \mathcal{N}(\theta_0)$ of θ_0 such that for all $u \in (0, 1)^d$,

$$\sup_{\theta\in\mathcal{N}}\left\|\frac{\dot{c}_\theta(u)}{c_\theta(u)}\right\| \le h_1(u), \qquad \sup_{\theta\in\mathcal{N}}\left\|\frac{\ddot{c}_\theta(u)}{c_\theta(u)}\right\| \le h_2(u),$$

where h_1^2 and h_2 are integrable with respect to C_{θ_0}.

(R2) For any fixed $\theta_0 \in \mathcal{O}$, $\theta \mapsto \dot{k}_\theta(t)$ is continuous at θ_0, for almost all $t \in (0, 1)$, and there is a neighbourhood $\mathcal{N}$ of θ_0 such that $\sup_{\theta\in\mathcal{N}} \|\dot{k}_\theta(t)\| \le h_3(t)$, with h_3 integrable over $(0, 1)$.

(R3) For any fixed $\theta_0 \in \mathcal{O}$, there exists a square integrable function J_{θ_0}, with respect to C_{θ_0}, such that

$$\theta_n = \frac{1}{n}\sum_{i=1}^{n} J_{\theta_0}\{F_1(X_{1i}), \ldots, F_d(X_{di})\} + o_P\big(1/\sqrt{n}\big),$$

where

$$\int_{(0,1)^d} c_{\theta_0}(u_1, \ldots, u_d) J_{\theta_0}(u_1, \ldots, u_d)\, \mathrm{d}u_1 \cdots \mathrm{d}u_d = \theta_0,$$

and

$$\int_{(0,1)^d} \dot{c}_{\theta_0}(u_1, \ldots, u_d) J_{\theta_0}(u_1, \ldots, u_d)\, \mathrm{d}u_1 \cdots \mathrm{d}u_d = I,$$

where I is the $m \times m$ identity matrix.

For example, if the pseudo-maximum likelihood of Genest *et al.* (1995) exists, then it satisfies the regularity assumption R3. If the copula family is indexed by Kendall's tau, then assumption R3 is satisfied. In the latter case, classical non-parametric dependence measures such as Spearman's rho or van der Waerden's coefficient also satisfy R3.