Statistical Inference Procedures for Bivariate Archimedean Copulas

CHRISTIAN GENEST and LOUIS-PAUL RIVEST*

A bivariate distribution function $H(x, y)$ with marginals $F(x)$ and $G(y)$ is said to be generated by an Archimedean copula if it can be expressed in the form $H(x, y) = \phi^{-1}[\phi\{F(x)\} + \phi\{G(y)\}]$ for some convex, decreasing function ϕ defined on (0, 1] in such a way that $\phi(1) = 0$. Many well-known systems of bivariate distributions belong to this class, including those of Gumbel, Ali-Mikhail-Haq-Thélot, Clayton, Frank, and Hougaard. Frailty models also fall under that general prescription. This article examines the problem of selecting an Archimedean copula providing a suitable representation of the dependence structure between two variates X and Y in the light of a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$. The key to the estimation procedure is a one-dimensional empirical distribution function that can be constructed whether the uniform representation of X and Y is Archimedean or not, and independently of their marginals. This semiparametric estimator, based on a decomposition of Kendall's tau statistic, is seen to be $\sqrt{n}$-consistent, and an explicit formula for its asymptotic variance is provided. This leads to a strategy for selecting the parametric family of Archimedean copulas that provides the best possible fit to a given set of data. To illustrate these procedures, a uranium exploration data set is reanalyzed. Although the presentation is restricted to problems involving a random sample from a bivariate distribution, extensions to situations involving multivariate or censored data could be envisaged.

KEY WORDS: Asymptotic distribution; Dependence function; Empirical process; Frailty model; Kendall's tau; U statistic.

The need to develop families of bivariate distributions to model nonnormal variations has been felt since the first treatment of the multidimensional normal law by Galton and Dickson in the late 19th century. Early strategies to tackle this problem included attempts to find bivariate generalizations of the differential equations defining Pearson's system of curves and the development of Edgeworth series expansions around the bivariate normal densities. In his classical review of these methods, Pretorius (1930) underscored their statistical limitations. Despite 40 years of additional research, similar views were expressed by Mardia (1970, p. 80), who concluded that "from the point of view of fitting, no system can yet be claimed to be adequate." This article aims to fill some of this gap by providing statistical inference procedures for a wide class of bivariate distributions that are generated by so-called "Archimedean copulas" (Genest and MacKay 1986a,b; Schweizer and Sklar 1983, chap. 5).

The problem of specifying a probability model for independent observations $(X_1, Y_1), \ldots, (X_n, Y_n)$ from a bivariate population with nonnormal distribution function $H(x, y)$ can be simplified by expressing H in terms of its marginals, $F(x)$ and $G(y)$, and its associated *dependence function,* C, implicitly defined through the identity $C\{F(x), G(y)\} = H(x, y)$. As its name suggests, the mapping C, which is uniquely determined on the unit square whenever F and G are continuous, captures the essential features of the dependence between the random variables X and Y (Deheuvels 1978; Kimeldorf and Sampson 1975). A natural way of analyzing bivariate data thus consists of estimating the dependence function and the marginals separately. This two-step approach to stochastic modeling is often convenient, because many tractable models are readily available for the marginal distributions. It is clearly appropriate in situations where the marginals are already known, such as those described by Vitale (1979), but it is also invaluable as a general strategy for data analysis in that it allows one to investigate the dependence structure independently of marginal effects.

One of the earliest families of dependence functions suitable for statistical analysis was proposed by Plackett (1965). Further illustrations of stochastic modeling based on dependence functions have been provided by Clayton (1978), Cook and Johnson (1981, 1986), Oakes (1982, 1989), Tawn (1988), and others. In those works, estimation was restricted to situations where the dependence function was known to belong to a specific class of bivariate distributions indexed by a one- or two-dimensional parameter. What is proposed here is a more general solution to the problem of choosing an appropriate parametric family of dependence functions. This article's basic assumption is that the data can be suitably modeled by an *Archimedean copula,* which implies that on the unit square, the appropriate dependence function is of the form $C_\phi(x, y) = \phi^{-1}\{\phi(x) + \phi(y)\}$ for some convex decreasing function ϕ defined on (0, 1] in such a way that $\phi(1) = 0$. As illustrated by Genest and MacKay (1986a,b), this class of dependence functions is wide and mathematically tractable, and its elements have stochastic properties that make these functions attractive for the statistical treatment of data.

A general estimation procedure for Archimedean copulas should be of interest, because this class of dependence functions encompasses many well-known systems of bivariate distributions, including those of Gumbel (Gumbel 1960), Ali–Mikhail–Haq–Thélot (Ali, Mikhail, and Haq 1978; Thélot 1985), Clayton (Clayton 1978; Cook and Johnson 1981; Oakes 1982), Frank (Frank 1979; Nelsen 1986; Genest 1987), and Hougaard (1984, 1986). Interest in dependence functions of this form was further enhanced by Oakes (1989),

* Christian Genest and Louis-Paul Rivest are Professors, Département de mathématiques et de statistique, Université Laval, Cité universitaire, Québec, Canada G1K 7P4. This work was supported in part by individual research grants from the Natural Sciences and Engineering Research Council of Canada and by a team grant from the Fonds pour la formation de chercheurs et l'aide à la recherche du Gouvernement du Québec. The authors thank Dennis R. Cook and Mark E. Johnson for graciously making their data available to them, as well as Claude Nadeau for his critical comments on a preliminary version of this article and for some programming assistance in the early stages of this research.

who showed that Archimedean copulas arose naturally in the construction of a large class of models for the joint distribution of two survival times, S and T, whose marginal survivor functions might be denoted by $\bar{F}(s) = \Pr(S > s)$ and $\bar{G}(t) = \Pr(T > t)$. These so-called *frailty models* are constructed by assuming the existence of an unobserved, latent variable W such that, conditional on $W = w$, the joint survivor function of S and T is equal to $\{\bar{F}^*(s)\bar{G}^*(t)\}^w$ for some continuous baseline survivor functions $\bar{F}^*(s)$ and $\bar{G}^*(t)$. As it turns out, the unconditional survivor function of the pair (S, T) is then equal to $C[\bar{F}(s), \bar{G}(t)]$ for some Archimedean copula $C(x, y) = \phi^{-1}\{\phi(x) + \phi(y)\}$ whose generator, ϕ, is related to the Laplace transform of W. The relevance of dependence functions of the Archimedean variety in the context of mixture models has also been emphasized by Marshall and Olkin (1988).

To estimate an Archimedean copula, advantage will be made of the fact that $C_\phi(x, y) = \phi^{-1}\{\phi(x) + \phi(y)\}$ is uniquely determined by the function $K(v) = v - \phi(v)/\phi'(v)$ defined on the unit interval. In Section 1 it is shown that because of the nature of ϕ, this mapping is in fact a distribution function on (0, 1). In Section 2 this result is then used to construct a nonparametric estimator K_n of K, based on a decomposition of Kendall's tau statistic. This estimator is seen to be $\sqrt{n}$-consistent, and an explicit formula for its asymptotic variance is obtained. Though $K_n(v)$ could conceivably be used to construct an empirical bivariate Archimedean copula directly, it often proves more convenient in application to use it as a tool for identifying the parametric family of Archimedean copulas that provides the best possible fit to the data. This calls for a method of choosing an appropriate representative from each parametric family of Archimedean copulas under consideration. In Section 3 a method-of-moments estimation technique based on Kendall's tau is proposed. The procedures are illustrated in Section 4, where a uranium exploration data set described by Cook and Johnson (1981, 1986) is reanalyzed. Avenues for further research are briefly outlined in the concluding paragraphs.

1. THEORETICAL BACKGROUND

Bivariate distribution functions with uniform marginals on the unit square play an important role in the study of stochastic dependence. They have been rediscovered many times and used in various contexts under different names, including "copulas" (Schweizer and Sklar 1983, chap. 5), "dependence functions" (Deheuvels 1978), and "uniform representations" (Kimeldorf and Sampson 1975). Here, these expressions will be used interchangeably.

A copula is said to be *Archimedean* if it can be expressed in the form

$$C_\phi(x, y) = \phi^{-1}\{\phi(x) + \phi(y)\}, \qquad 0 < x, y < 1 \quad (1)$$

for some convex decreasing function $\phi(v)$ satisfying $\phi(1) = 0$. By convention, $\phi^{-1}(v)$ is taken equal to 0 whenever $v \geq \phi(0)$. Observe that these conditions, which are necessary and sufficient for $C_\phi(x, y)$ to be a distribution function (Schweizer and Sklar 1983, thm. 5.4.8), are equivalent to the requirement that $1 - \phi^{-1}(v)$ be a unimodal distribution function on $[0, \infty)$ with mode at 0.

Motivation, elementary properties, and convergence results concerning sequences of Archimedean copulas can be found in Genest and MacKay (1986a,b). The following proposition, which characterizes the probabilistic structure of distribution functions whose uniform representation is Archimedean, casts new light on some of these authors' results.

Proposition 1.1. Let X and Y be uniform random variables whose dependence function $C(x, y)$ is of the form $\phi^{-1}\{\phi(x) + \phi(y)\}$ for some convex decreasing function ϕ defined on (0, 1] with the property that $\phi(1) = 0$. Set $U = \phi(X)/\{\phi(X) + \phi(Y)\}$, $V = C(X, Y)$, and $\lambda(v) = \phi(v)/\phi'(v)$ for $0 < v \leq 1$. Then, (a) U is uniformly distributed on (0, 1), (b) V is distributed as $K(v) = v - \lambda(v)$ on (0, 1), and (c) U and V are independent random variables. In fact, the existence of a function ϕ for which properties (a), (b), and (c) hold implies that $C(x, y) = \phi^{-1}\{\phi(x) + \phi(y)\}$ on its entire domain. (*Note:* The proof of this proposition, as well as those of all subsequent results reported herein, may be found in the Appendix.)

It will be shown in Section 2 how the distribution $K(v)$ of $V = C(X, Y)$ can be estimated nonparametrically from a random sample of X and Y pairs. In the special case where $C = C_\phi$ is Archimedean, this will provide an indirect means of estimating the bivariate dependence function itself, because it is possible to recover ϕ by solving the differential equation $\phi(v)/\phi'(v) = v - K(v)$. This operation yields

$$\phi(v) = \exp\left\{\int_{v_0}^{v} \frac{1}{\lambda(t)}\, dt\right\}, \quad (2)$$

where $0 < v_0 < 1$ is an arbitrarily chosen constant. Interestingly, though, this function is well defined and generates an Archimedean copula whenever $v - K(v)$ is negative and remains bounded away from 0 on the unit interval. This fact, formally stated as Proposition 1.2, can thus be used to define in some sense the "projection" of (almost) any dependence function C within the class of Archimedean copulas.

Proposition 1.2. Let X and Y be uniform random variables with dependence function $C(x, y)$. For $0 \leq v \leq 1$, let $K(v) = \Pr\{C(X, Y) \leq v\}$ and define $K(v^-) = \lim_{t \uparrow v} K(t)$. The function $\phi(v)$ defined by (2) is convex and decreasing and satisfies $\phi(1) = 0$ if and only $K(v^-) > v$ for all $0 < v < 1$.

In view of this result, it is apparent that among all bivariate dependence functions $C(x, y)$ for which the distribution of $V = C(X, Y)$ satisfies $K(v^-) > v$ on its domain, the Archimedean copula $C_\phi(x, y) = \phi^{-1}\{\phi(x) + \phi(y)\}$ is the only one for which parts (a) and (c) of Proposition 1.1 also hold.

Several one-parameter systems of bivariate distributions with fixed marginals can be seen to have Archimedean copulas as their dependence function (Genest and MacKay 1986a,b; Marshall and Olkin 1988; Oakes 1989). Four key examples of generators are presented in Table 1. Thus if $\phi(v) = (v^{-\alpha} - 1)/\alpha$ for some $\alpha > 0$, then $C_\phi(x, y) = \phi^{-1}\{\phi(x) + \phi(y)\}$ reduces to $(x^{-\alpha} + y^{-\alpha} - 1)^{-1/\alpha}$, a system introduced by Clayton (1978) that was later studied and used by Cook and Johnson (1981) and by Oakes (1982).

Table 1. Examples of Families of Archimedean Copulas

Family	ϕ	Range of α and γ	$-\lambda$	τ
Clayton	$(v^{-\alpha} - 1)/\alpha$	$(0, \infty)$	$v(1 - v^{\alpha})/\alpha$	$\alpha/(\alpha + 2)$
Frank	$\log\left(\frac{1 - \exp(-\alpha)}{1 - \exp(-\alpha v)}\right)$	$(-\infty, \infty)$	$\frac{1 - \exp(-\alpha v)}{\alpha \exp(-\alpha v)} \log\left(\frac{1 - \exp(-\alpha)}{1 - \exp(-\alpha v)}\right)$	$1 + 4\{D_1(\alpha) - 1\}/\alpha$[a]
Gumbel	$\{-\log(v)\}^{\alpha+1}$	$[0, \infty)$	$-v \log(v)/(\alpha + 1)$	$\alpha/(\alpha + 1)$
Log-copula	$\{1 - \log(v)/\alpha\gamma\}^{\alpha+1} - 1$	$(0, \infty)$	$\frac{\alpha\gamma v[\{1 - \log(v)/\alpha\gamma\}^{\alpha+1} - 1]}{(\alpha + 1)\{1 - \log(v)/\alpha\gamma\}^{\alpha}}$	$\{\alpha - 2 + 4C(\alpha)\}/(\alpha + 1)$[b]

NOTE: The first three families are indexed by a single real parameter, α; the last family has two positive parameters, α and γ. All distributions in a family are generated as per Equation (1) by convex, decreasing functions ϕ defined on (0, 1] in such a way that $\phi(1) = 0$. In each case, the appropriate generator is given, together with the quantity $-\lambda = -\phi/\phi'$ appearing in Proposition 1.1 and the population value of Kendall's tau, calculated using identity (3).

[a] D_1 is the Debye function of order 1, $D_1(\alpha) = \int_0^{\alpha} \{t/\alpha(e^t - 1)\}\, dt$

[b] C is the exponential integral $C(\alpha) = \int_0^{\infty} \{e^{-2t}/(1 + t)^{\alpha}\}\, dt$

When $\phi(v) = \log[\{1 - \exp(-\alpha)\}/\{1 - \exp(-\alpha v)\}]$ for α real, Frank's system of bivariate distributions obtains (Frank 1979). Its statistical properties were studied by both Nelsen (1986) and Genest (1987). As for the Gumbel and log-copula families, they were considered by Hougaard (1986) and Hougaard, Harvald and Holm (1992) in the context of frailty models. A particularly useful feature of the log-copula family is that it contains both the Gumbel and the Clayton families as limiting cases. This fact is formally recorded in the following proposition.

Proposition 1.3. Let $C_{\alpha\gamma}$ be the dependence function of the log-copula family defined in Table 1. Then, as α tends to infinity, $C_{\alpha\gamma}$ converges to the Clayton dependence function with parameter $1/\gamma$, and as γ approaches 0, it converges to the Gumbel dependence function with parameter α.

A common feature of the four systems listed in Table 1 is that they each include as a special case the independence distribution, $C_{\phi}(x, y) = xy$, generated by $\phi(v) = \log(1/v)$, as well as the so-called upper Fréchet bound (that is, the copula $C(x, y) = \min(x, y)$ corresponding to a pair (X, Y) of uniforms such that $X = Y$ almost surely). Thus independence obtains in the four models when α equals 0 and also when γ goes to infinity with α fixed in the log-copula family. The upper Fréchet bound corresponds to α infinite in the first three models, as well as in the fourth provided that γ tend to 0 at the same rate (i.e., in such a way that $\beta = \alpha\gamma$ remain fixed). With the later convention, the four families are in fact ordered in α under various notions of bivariate stochastic orderings (Bilodeau 1989; Capéraà and Genest 1990), thereby making it possible to interpret this parameter in terms of association between the variables X and Y. In particular, the population version of Kendall's tau, defined in general as $\tau(X, Y) = 4E\{H(X, Y)\} - 1$, takes on all values between 0 and 1 for the four families of Table 1 as α goes from 0 to infinity (in such a way that $\beta = \alpha\gamma$ remain fixed, in the case of the log-copula). Proposition 1.1 implies that for Archimedean copulas, this quantity can be conveniently computed via the identity

$$\tau = 4E(V) - 1 = 4\int_0^1 \lambda(v)\, dv + 1. \qquad (3)$$

Finally, before turning to the estimation problem, note that copulas can be used to model not only the joint distribution of a pair (X, Y) of random variables, but also that of $(-X, Y)$, $(X, -Y)$, and $(-X, -Y)$. In particular, note that modeling the joint distribution of $-X$ and $-Y$ amounts to setting $\Pr(X > x, Y > y) = C\{\bar{F}(x), \bar{G}(y)\}$. For the copulas of Table 1, therefore, working with the survivor function or with the distribution itself will yield two different models—except for Frank's family, where the two are equivalent because of a peculiar symmetry condition reported by Genest (1987). Modeling of joint distributions and of survivor functions is illustrated in Section 4.

2. A NONPARAMETRIC ESTIMATION PROCEDURE FOR $H(X, Y)$

Suppose that a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ has been drawn from a bivariate distribution $H(x, y)$ with continuous marginals $F(x)$ and $G(y)$ and dependence function $C(x, y)$. Suppose also that it is desired to estimate C under the assumption that it is Archimedean. Because the statistical procedure to be proposed in this section is independent of the marginals, one can assume without loss of generality that the latter are uniform on the unit interval. Henceforth, H and C will thus be confounded wherever no risk of confusion can arise.

Because, in the light of Proposition 1.1, Archimedean copulas are characterized by the stochastic behavior of the random variate $V = H(X, Y)$, a natural way to proceed is to seek first an estimation of the univariate distribution function $K(v) = \Pr\{H(X, Y) \le v\} = \Pr[C\{F(X), G(Y)\} \le v]$ on the interval (0, 1).

This can be accomplished in two steps, whether or not the uniform representation of H is Archimedean:

1. First, construct the empirical bivariate distribution function, $H_n(x, y)$, associated with H in the standard way.
2. Then, compute $H_n(X_i, Y_i)$ for $i = 1, \ldots, n$ and use those pseudoobservations to construct a one-dimensional empirical distribution function for K.

Fortunately, it is not necessary in practice to construct H_n to build an estimator of K. Rather, one can rely directly on

the fact that $H_n(X_i, Y_i)$ represents the proportion of observations in the sample that are less than or equal to (X_i, Y_i) componentwise. But because $H_n(X_i, Y_i)$ is always larger than $1/n$ and it is desirable to have an estimator taking values on the (0, 1) range, it will prove more convenient in the sequel to use the variables

$$V_i = \#\{(X_j, Y_j): X_j < X_i, Y_j < Y_i\}/(n-1), \quad 1 \le i \le n, \quad (4)$$

where the symbol # stands for the cardinality of a set. If $\delta(t)$ denotes the distribution function of a point mass at the origin, then a nonparametric estimator of $K(v)$ is given by

$$K_n(v) = \sum_{i=1}^{n} \delta(v - V_i)/n. \quad (5)$$

Under the assumption that the dependence function associated with H is Archimedean, a natural estimator of the function $\lambda(v) = \phi(v)/\phi'(v)$ can be derived from K_n through the relation $\lambda_n(v) = v - K_n(v)$, $0 < v < 1$. Provided $K_n(v^-) > v$ on its entire domain, Formula (2) then provides an estimator of C within the class of Archimedean copulas, whether or not the uniform representation of the joint distribution function $H(x, y)$ is Archimedean. An alternative estimation procedure for dependence functions of the Archimedean variety, based on local odds ratios, was proposed by Oakes (1989). For multivariate goodness-of-fit tests based on the univariate distribution of $V = H(X, Y)$, the reader may refer to Saunders and Laud (1980).

Before the sampling properties of the estimator K_n can be effectively examined, it is important to realize that the V_i's of Equation (4) are strongly related to the sample value, τ_n, of Kendall's tau. Indeed, recall that the latter is defined in general as the number of concordant pairs of data points minus the number of discordant pairs, divided by $n(n-1)/2$, the total number of pairs (Cases of equality among the X_i's or among the Y_j's can be safely ignored, as the marginal distributions of the two variates are assumed to be continuous).

To be formal, define I_{ij} for $1 \le i, j \le n$ by

$$I_{ij} = 1 \quad \text{if } X_j < X_i \quad \text{and} \quad Y_j < Y_i$$
$$= 0 \quad \text{otherwise.} \quad (6)$$

With this notation, observations i and j are seen to be concordant if and only if $I_{ij} + I_{ji} = 1$. Because $V_i = \Sigma_j I_{ij}/(n-1)$, the number of concordant pairs of data points is thus equal to $\Sigma_{ij} I_{ij} = (n-1)\Sigma_i V_i$ and hence

$$\tau_n = 4\bar{V} - 1. \quad (7)$$

This equation may be recognized as the sample equivalent of (3). The number $\Sigma_j I_{ji}$ of data points (X_j, Y_j) that are concordant with (X_i, Y_i), and such that (X_j, Y_j) is larger than (X_i, Y_i) componentwise, can be expressed in terms of V_i, RX_i, and RY_i, where RX_i and RY_i are the ranks of X_i and Y_i in the X and the Y samples. Explicitly, one has $\Sigma_j I_{ji} = n - RX_i - RY_i + (n-1)V_i + 1$. Note also that if W_i is defined as $\Sigma_j I_{ji}/(n-1)$, then $\tau_n = 4\bar{W} - 1$ as well. The empirical distribution function of the W_i's would be used to estimate $K(v)$ if one were modeling the bivariate survivor function of (X, Y) using an Archimedean copula.

Because the proposed estimator, K_n, of K can be viewed as a decomposition of Kendall's tau, it seems appropriate to call the functional $\sqrt{n}\{K_n(v) - K(v)\}$ *Kendall's process.* This is an empirical process built from a pseudosample $V_1, \ldots, V_n$ of *dependent* and exchangeable observations. In fact, it can be checked that the correlation between any pair (V_i, V_j) is of the order of $1/n$. A formal investigation of the limiting distribution of Kendall's process is beyond this article's scope. For the current purposes, the following derivation of the large-sample variance of $K_n(v)$ will suffice.

Proposition 2.1. Let $C(x, y)$ be a dependence function that is absolutely continuous with respect to Lebesgue measure, and suppose that $(X_1, Y_1), \ldots, (X_n, Y_n)$ form a random sample from C. Under weak regularity conditions, the distribution of V_i converges to $K(v) = \Pr\{C(X, Y) \le v\}$, and the empirical distribution function $K_n(v)$ of the V_i's defined in (5) is a $\sqrt{n}$-consistent estimator of $K(v)$. An $o(1/n)$ approximation of the asymptotic variance of $K_n(v)$ is given by

$$[K(v)\bar{K}(v) + k(v)\{k(v)R(v) - 2v\bar{K}(v)\}]/n,$$

where $\bar{K}(v) = 1 - K(v)$, $k(v) = K'(v)$ is the assumed density of $V = C(X, Y)$, and

$$R(v) = E[C\{\min(X_1, X_2), \min(Y_1, Y_2)\} - v^2 | C(X_1, Y_1) = C(X_2, Y_2) = v]. \quad (8)$$

For general (i.e., not necessarily Archimedean) copulas, the functions $K(v)$, $k(v)$, and $R(v)$ appearing in the large-sample variance of $K_n(v)$ are involved. Accordingly, there is usually no simple expression for the asymptotic variance of Kendall's process. For Archimedean copulas, however, the calculations are often tractable, as illustrated next.

Proposition 2.2. For an Archimedean copula generated by ϕ, the function $R(v)$ defined by Equation (8) has the form

$$R(v) = 2\int_0^1 (1-t)\phi^{-1}\{(1+t)\phi(v)\}\, dt - v^2.$$

For the independence copula, one has $\phi(v) = \log(1/v)$, and hence $R(v) = 2v\{v - \log(v) - 1\}/\log^2(v) - v^2$. In this case, the $o(1/n)$ approximation to the asymptotic variance of Kendall's process has the remarkably simple form $v\{v - \log(v) - 1\}/n$.

Among the systems of Archimedean copulas presented in Table 1, Clayton's family is the only one for which an explicit algebraic expression could be found for $R(v)$. In that special case, one gets

$$R(v) = \frac{2\alpha v}{(1-\alpha)(1-2\alpha)(1-v^\alpha)^2} \times \{\alpha(2 - v^\alpha)^{2-1/\alpha} + (1-v^\alpha)(1-2\alpha) - \alpha\} - v^2.$$

For this family, Table 1 yields $K(v) = v\{1 + (1-v^\alpha)/\alpha\}$ and $k(v)$ equals $(\alpha+1)(1-v^\alpha)/\alpha$. The large-sample variance of Kendall's process thus has an explicit form in this case. It will be referred to as $\sigma_\alpha^2(v)$ in the sequel.

3. METHOD-OF-MOMENTS ESTIMATION BASED ON KENDALL'S TAU

Once an estimate $K_n(v)$ of the distribution of $H(X, Y)$ has been computed from a bivariate random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ through Equation (5), one may be tempted to determine directly from K_n the Archimedean copula $C_{\phi_n}(x, y)$ that is closest to $H(x, y)$, in the sense of Proposition 1.2. Though this would be formally possible whenever $v - K_n(v^-) < 0$ for all $0 < v < 1$, it generally will be theoretically more meaningful—as well as computationally more convenient—to use K_n as a tool to help identify the parametric family of Archimedean copulas that provides the best possible fit to the data. This calls for a method of choosing, from each parametric family of Archimedean copulas under consideration, the most appropriate representative lambda function to compare to $\lambda_n(v) = v - K_n(v)$.

Because λ_n is based on a decomposition of Kendall's tau, a natural way to go about this would be to base the estimation procedure on the observed value of that statistic, τ_n. This is especially attractive in the current context, because the population value of Kendall's tau, $\tau(X, Y) = 4E(V) - 1$, is easily computed through Formula (3) for Archimedean copulas. A rough-and-ready strategy thus might be to estimate the parameter, say α, of an Archimedean family of copulas by that value, $\hat{\alpha}$, for which the theoretical value of Kendall's tau is equal to $\tau_n = 4\bar{V} - 1$. This naive method of estimation, which is in the spirit of Pearson's method of moments, is operational for the first three families included in Table 1. Based on extensive simulations, this method was reported by Genest (1987) to perform reasonably well in the case of Frank's family for samples of size 50 or larger. This method's efficiency was also investigated by Oakes (1986) for Clayton's system of distributions. For systems involving more than one parameter, such as the log-copula family, a natural extension of the procedure would be to identify as many as the first few moments of the pseudosample $V_1, \ldots, V_n$ with the corresponding theoretical expressions to be able to solve algebraically or numerically. These point estimation techniques are illustrated in Section 4.

In the case of one-parameter families of distributions, such as those of Clayton, Frank, or Gumbel, associated confidence intervals can also be obtained using the following large-sample approximation to the distribution of τ_n. An exact expression for the variance of τ_n in Clayton's family has been given by Oakes (1982).

Proposition 3.1. As defined at the beginning of Section 2, let $V_i + W_i$ denote the number of observations that are concordant with (X_i, Y_i), divided by $n - 1$. If $S^2 = \sum_{i=1}^{n} (V_i + W_i - 2\bar{V})^2/(n - 1)$, the large-sample distribution of $\sqrt{n}(\tau_n - \tau)/4S$ is then normal with 0 mean and unit variance.

The derivation of the large sample covariance matrix of the first two moments of the V_i's that would enable one to obtain the large-sample distribution of the parameter estimates for the log-copula family was an open research problem at the time of writing.

4. ILLUSTRATIVE EXAMPLE

This section reports the results of the analysis of a large bivariate data set carried out to illustrate the statistical estimation procedures proposed herein. The data, comprised of log-concentration readings of uranium and cesium for 655 petroleum samples, were part of a multivariate data set originally described by Cook and Johnson (1981). The intended purpose of the exercise was to ascertain whether the one-parameter family of distributions that Cook and Johnson used to model these particular variates truly provided the best fit to the unknown population dependence function from among the four systems of Archimedean copulas listed in Table 1.

Table 2 presents a 7×7 cross-classification of the two variables under study. The cell boundaries for the two variables were taken as the order statistics of rank $[655 * j/7]$, for $j = 1, \ldots, 6$, where $[z]$ denotes the integer part of real z. In principle this choice should have made equal, up to ± 1 unit, the marginal row and column totals of the contingency table. The minor discrepancies observed in the column totals were due to equalities in log-concentration readings for cesium. Table 2 was nevertheless regarded as representative of the dependence function of these two variates.

It was established in Section 1 that an Archimedean copula is characterized by the univariate function $\lambda(v) = \phi(v)/\phi'(v)$. Furthermore, it was seen in Section 2 that a sample estimate of this quantity is given by $\lambda_n(v) = v - K_n(v)$, where $K_n(v)$ is the empirical distribution function of the V_i's defined by Equation (4). Proposition 2.1 gives a formula for the large-sample variance of $\lambda_n(v)$, but because the formula depends on a dependence function that was unknown in the current application, the result could not be used. The closed-form expression $\sigma_\alpha^2(v)$ for the asymptotic variance of $\lambda_n(v)$

Table 2. Cross-Classification of Uranium (X) and Cesium (Y) Log-Concentrations for the Uranium Exploration Data Set

Cell upper boundaries	Cell upper boundaries							
	$Y_{(94)}$	$Y_{(187)}$	$Y_{(281)}$	$Y_{(374)}$	$Y_{(468)}$	$Y_{(561)}$	$Y_{(655)}$	Totals
$X_{(94)}$	48	21	14	10	1	0	0	94
$X_{(187)}$	17	22	19	22	8	5	0	93
$X_{(281)}$	10	23	25	19	11	5	0	93
$X_{(374)}$	6	14	20	15	21	11	7	94
$X_{(468)}$	10	10	4	15	17	21	16	93
$X_{(561)}$	4	4	5	6	17	20	38	94
$X_{(655)}$	2	2	1	7	17	32	33	94
Totals	97	96	88	94	92	94	94	655

NOTE: An x or a y value falls in a given cell if it is less than or equal to the cell upper boundary and strictly larger than the upper boundary of the previous cell.

Table 3. Summary of the Statistics for the Models Fitted to the Uranium Exploration Data Set

Family	$\hat{\alpha}$	$\hat{\gamma}$	χ^2 statistic	df
	Models fitted to the joint distribution			
Clayton	1.714 (.066)	—	82.76	25
Frank	5.078 (.074)	—	44.23	24
Gumbel	.857 (.033)	—	90.36	26
Log-copula	1.17	.100	51.94	23
	Models fitted to the joint survivor function			
Clayton	1.714 (.066)	—	131.39	24
Frank	5.078 (.074)	—	44.23	24
Gumbel	.857 (.033)	—	63.20	27
Log-copula	1.346	.146	53.05	23

NOTE: For each model, parameter estimates are given, together with a chi-squared goodness-of-fit test statistic comparing predicted frequencies of Tables 4 and 5 to the observed counts of Table 2. Quantities in parentheses represent the standard deviations of the estimates, in the one-parameter families.

in the Clayton family was thus relied on as a convenient approximation. Specifically, the large-sample variance of $\lambda_n(v)$ was approximated by $\sigma^2_{\hat{\alpha}}(v)$, where $\hat{\alpha}$ stands for the sample estimate of α for the Clayton family presented in Table 3. Confidence bands for the unknown λ were then constructed using $\lambda_n(v) \pm c\,\sigma_{\hat{\alpha}}(v)$. The choice $c = 1.96$ would have ensured pointwise confidence levels of 95% but was deemed too small, because the focus was on experimentwise confidence levels. The value $c = 4.72$ suggested by Kotel'nikova and Chmaladze (1982) was thus selected instead. As explained by Shorack and Wellner (1986, p. 361), this choice is designed to ensure an experimentwise coverage of 95% for one-sided confidence intervals for the standard empirical process in large samples.

The estimate $\lambda_n(v)$ for the joint distribution of uranium and cesium log-concentrations is depicted in Figure 1, along with its associated confidence band. Figure 2 presents the

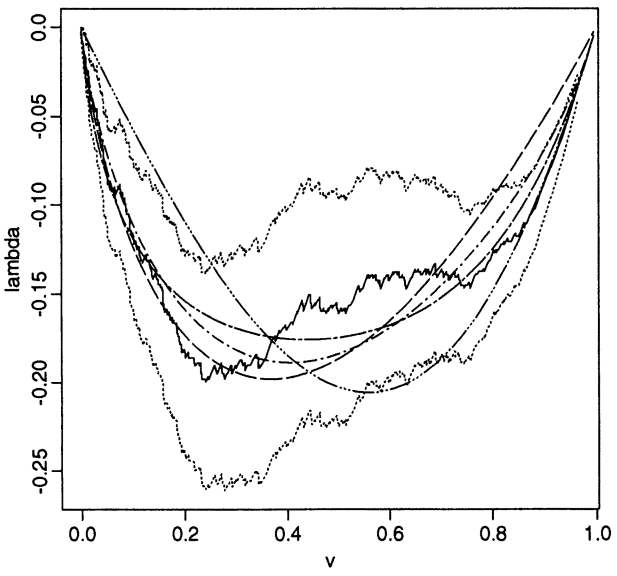

Figure 1. Summary of the Models Fitted to the Joint Distribution of Uranium and Cesium Log-Concentrations. The empirical lambda function and its confidence bands are presented together with the fitted lambda functions for the four families of Table 1. — · —, Log-Copula; — –, Gumbel; — · · ·, Cook and Johnson; ——, Frank.

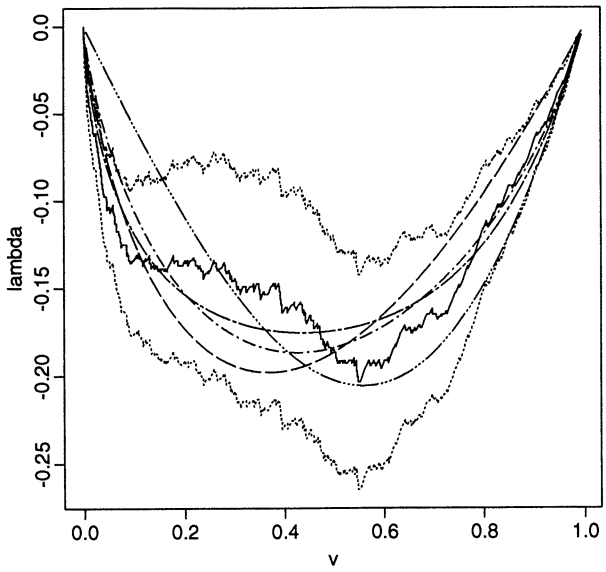

Figure 2. Summary of the Models Fitted to the Joint Survivor Distribution of Uranium and Cesium Log-Concentrations. The empirical lambda function and its confidence bands are presented together with the fitted lambda functions for the four families of Table 1. — · —, Log-Copula; — –, Gumbel; — · · ·, Cook and Johnson; ——, Frank.

$\lambda_n(v)$ obtained when modeling the joint survivor function of the same two variables. Recall that in the latter case, $K_n(v) = v - \lambda_n(v)$ is the empirical distribution function of the W_i's, as defined in Section 2. For illustrative purposes, the four families of Table 1 were fitted to both the joint distribution and the joint survivor function of uranium and cesium log-concentrations, using the method of moments based on Kendall's tau described in Section 3. For each one-parameter family, the value of the parameter α thus was selected so that the theoretical value of Kendall's tau, $\tau = 4E\{C(X, Y)\} - 1$, would match that of the observed value of that statistic, namely $\tau_n = 0.4615$. Naturally, this led to the same model for both the distribution function and the survivor function. In the case of the log-copula, however, different models were obtained. This is because the two parameters were estimated by equating the first two theoretical moments of $C(X, Y)$ to their sample counterparts and, although one had $m_v = m_w = (\tau_n + 1)/4 = .3654$, $s_v^2 = .074$ was, as expected, different than $s_w^2 = .072$.

Table 3 gives parameter estimates for all the fitted models and, for one-parameter families, standard errors in parentheses. These were derived by the delta method from the asymptotic standard deviation of τ_n, which was found to equal .00964 via Proposition 3.1. To assess the fit of the various models, predicted values for the frequencies of all the cells of Table 2 were computed (see Tables 4 and 5) and a Pearson chi-squared goodness-of-fit statistic was then computed for each family. In the calculation of the chi-squared statistic, it was decided to pool together all the cells whose expected frequency was strictly less than 5. The degrees of freedom associated with the goodness-of-fit statistic were thus calculated as $df = 36 - p - (q - 1)$, where p is the number of parameters estimated ($p = 1$ or 2) and q is the number of cells pooled together. The results of this analysis, presented in Table 3, lead to the same conclusion as do Figures 1 and

2: among the classes of Archimedean copulas considered, Frank's family provides the best fit to the contingency table of Figure 2.

That Frank's family provides a better fit than Clayton's system of distributions in this example may come as somewhat of a surprise. For this particular data set was initially introduced by Cook and Johnson (1981) as an instance of an "asymmetric relationship" between subsets of two variables, in the sense that (X, Y) and $(-X, -Y)$ do not have the same distribution; that is, the contingency table $\{h_{ij}\}$ of Table 2 is such that $\{h_{ij}\}$ and $\{h_{(7-j)(7-i)}\}$ are significantly different. A careful look at that table, however, reveals that the data representing the joint distribution of uranium and cesium log-concentrations are more clustered in the left tail than in the right tail. The predicted values associated with Clayton's one-parameter model exhibit the same type of asymmetry, but to a much stronger degree. It thus would appear that the asymmetry of Clayton's original family is too severe for these two variables, as implied by the authors in their 1986 paper, which considered a two-parameter extension of that model.

Interestingly, the asymmetry of the predicted values associated with Gumbel's and Clayton's models appear to be different. With Gumbel's copula there is a clustering effect in the right tail of the joint distribution, but this asymmetry is not as severe as in Clayton's one-parameter system. As a consequence, Gumbel's family could be considered to model the survivor function of the log-concentration of uranium and cesium. This explains why in Table 3, Gumbel's copula yields smaller chi-squared statistics than does Clayton's copula.

Table 4. Predicted Frequencies Under Four Models Fitted to the Joint Distribution of Uranium and Cesium Log-Concentrations

Cell upper boundaries						
$Y_{(94)}$	$Y_{(187)}$	$Y_{(281)}$	$Y_{(374)}$	$Y_{(468)}$	$Y_{(561)}$	$Y_{(655)}$
Clayton predicted frequencies						
64	18	6	3	1	1	1
19	30	18	11	7	5	3
7	20	20	17	13	10	7
3	12	16	18	17	15	12
2	8	12	17	18	19	18
1	5	9	15	18	22	24
1	3	7	12	17	24	29
Gumbel predicted frequencies						
40	23	13	9	5	3	1
23	23	17	14	9	5	2
15	19	18	17	13	8	3
9	15	16	19	17	13	5
6	10	12	17	20	19	9
3	5	8	13	19	27	19
1	2	3	5	8	19	56
Frank predicted frequencies						
41	25	13	8	4	2	1
26	25	18	12	7	4	2
15	20	19	17	11	7	4
8	13	16	20	17	12	8
4	7	11	17	20	19	14
2	4	7	12	19	25	25
1	2	4	8	14	25	40
Log-copula predicted frequencies						
44	23	12	8	4	2	1
23	24	18	13	8	5	2
14	19	18	17	13	8	4
8	14	17	19	17	13	6
4	9	12	17	20	19	12
2	5	8	13	19	25	23
1	2	3	6	11	23	47

Table 5. Predicted Frequencies Under Four Models Fitted to the Joint Survivor Function of Uranium and Cesium Log-Concentrations

Cell upper boundaries						
$Y_{(94)}$	$Y_{(187)}$	$Y_{(281)}$	$Y_{(374)}$	$Y_{(468)}$	$Y_{(561)}$	$Y_{(655)}$
Clayton predicted frequencies						
30	24	17	12	7	3	1
24	22	17	15	9	5	1
18	19	17	17	13	7	2
13	15	16	19	17	12	3
8	10	12	17	20	19	6
3	5	7	12	19	30	18
1	1	2	3	6	18	63
Gumbel predicted frequencies						
57	19	8	5	3	2	1
20	27	18	12	8	5	3
9	20	19	17	13	9	5
5	13	17	19	17	14	9
3	8	13	17	19	19	14
2	5	9	14	18	23	23
1	3	5	9	14	23	39
Frank predicted frequencies						
41	25	13	8	4	2	1
26	25	18	12	7	4	2
15	20	19	17	11	7	4
8	13	16	20	17	12	8
4	7	11	17	20	19	14
2	4	7	12	19	25	25
1	2	4	8	14	25	40
Log-copula predicted frequencies						
49	23	11	6	3	2	1
24	26	18	12	7	4	2
12	20	20	18	12	8	4
6	13	17	20	18	13	6
3	8	12	18	20	19	12
2	4	8	13	19	25	23
1	2	3	7	12	23	46

For Kendall's tau equal to .4615, the log-copula provides a whole class of models. As indicated in Proposition 1.3, the Clayton and Gumbel copulas are both extreme members of this class. This is reflected in Figures 1 and 2, in which the log-copula lambda functions stand between those of the Clayton and Gumbel systems of distributions. In view of the preceding discussion, the log-copula family can thus accommodate clustering in either the right tail or the left tail of a joint distribution. In our current application, however, the predicted values for uranium and cesium log-concentrations associated with the log-copula do not exhibit much asymmetry, as can be judged from Tables 4 and 5. This might be caused, at least in part, by the estimation procedure proposed in Section 3. It would appear that the variances of the V_i's and of the W_i's, namely $s_v^2 = .074$ and $s_w^2 = .072$, do not capture the asymmetry of the joint distribution. In fact, these values may well be compatible with a symmetric distribution

such as Frank's, because their theoretical values both equal .073 when the data are modeled by that family, as in Table 3. This casts doubts on the efficiency of extending the method-of-moments estimation procedure of Section 3 to handle multiparameter families, such as the log-copulas, and suggests that alternative approaches should be investigated.

5. DISCUSSION

The purpose of this article has been to suggest a nonparametric method for estimating the dependence function of a pair of random variables under the assumption that their uniform representation is Archimedean. The proposed estimator, whose definition is based on a simple decomposition of Kendall's tau, can be computed whether or not this assumption holds. This estimation procedure thus can be used to effectively guide the selection of a suitable parametric family of Archimedean copulas and, as illustrated in Section 4, goodness-of-fit techniques can be applied to measure a family's appropriateness for describing the relationship between two sample variates.

Some important issues in the application of Archimedean copulas to sample data need further investigation. In particular, estimating the parameters in parametric families of dependence functions is still an open problem: How efficient are the simple moment-based procedures proposed in Section 3 compared to methods based on bivariate risk sets, such as those described by Oakes (1986)? Also, how important are the gains in precision that can be made by first selecting a model for the marginal distributions and then estimating the association parameter using a full likelihood, as propounded by Cook and Johnson (1981, 1986)? Another avenue worth exploring would be the use of pseudo-observations $U_i = \phi_n(RX_i/n)/\{\phi_n(RX_i/n) + \phi_n(RY_i/n)\}$, $1 \le i \le n$, to check the fit of an Archimedean copula induced by an estimated generator ϕ_n. When the fit is good, Proposition 1.1 implies that these pseudo-observations should be approximately uniformly distributed on the interval (0, 1). Finally, it would also be of interest to investigate modifications of the proposed procedure to handle censoring in the X and the Y variable.

APPENDIX: SKETCH OF PROOFS

Proof of Proposition 1.1. Consider the transformed variables $S = \phi(X)$ and $T = \phi(Y)$, whose joint distribution is $1 - \phi^{-1}(s) - \phi^{-1}(t) + \phi^{-1}(s + t)$ for s and t in the range of ϕ. For all $0 \le u \le 1$ and $v > 0$ in that range, one has

$$\Pr(U \le u, V \le v) = \Pr(S \le uT/(1-u), S \ge \phi(v) - T). \quad \text{(A.1)}$$

Because ϕ is convex, ϕ' exists except possibly at countably many points in the interval (0, 1). Thus T admits a density $-1/\phi'\{\phi^{-1}(t)\}$ on the range of ϕ, and $\Pr(S \le s \mid T = t) = 1 - \phi'\{\phi^{-1}(t)\}/\phi'\{\phi^{-1}(s + t)\}$. Conditioning on the value of T, Equation (A.1) thus can be reexpressed as

$$\int_{(1-u)\phi(v)}^{\phi(v)} -\Pr\{\phi(v) - t \le S \le ut/(1-u) \mid T = t\}/\phi'\{\phi^{-1}(t)\}\, dt + \int_{\phi(v)}^{\phi(0)} -\Pr\{S \le ut/(1-u) \mid T = t\}/\phi'\{\phi^{-1}(t)\}\, dt.$$

On evaluating these integrals, one finds that $\Pr(U \le u, V \le v) = uK(v)$ for all values of $0 \le u \le 1$ and all points v of continuity of $\phi'(v)$ on (0, 1). The characterization obtains, once it is realized that the joint distribution of (X, Y) can be retrieved from that of U and V.

Proof of Proposition 1.2. If the mapping $\phi(v)$ defined by (2) is convex and decreasing and satisfies $\phi(1) = 0$, then it is necessarily nonnegative and continuous. In addition, its derivative $\phi'(v)$ is well defined and equals $\phi(v)/\{v - K(v)\}$, except possibly at countably many points. Because $\phi'(v)$ is increasing and negative wherever it exists, $K(v^-) > v$ clearly must hold true everywhere on (0, 1). To show the converse, observe that the condition on K implies that $1/\lambda(v)$ is integrable on any interval of the form $(\varepsilon, 1 - \varepsilon)$ with $\varepsilon > 0$. As a consequence, ϕ is well defined, positive, and decreasing on its domain. In addition, one must have $\phi(1) = 0$ because

$$\int_{v_0}^{v} \frac{1}{\lambda(t)}\, dt \le -\int_{v_0}^{v} \frac{1}{1-t}\, dt,$$

and the latter integral diverges as v approaches 1. Finally, the convexity of ϕ will be established if one can check that $\phi'(y)/\phi'(x) \le 1$ in all points of definition $x < y$. As the difference of two increasing functions, $\lambda(v) = \phi(v)/\phi'(v)$ is of bounded variation on its domain, which allows one to rewrite this inequality as

$$\phi'(y)/\phi'(x) = \exp\left\{\int_x^y \frac{1}{\lambda(v)}\, dv - \int_x^y \frac{1}{\lambda(v)}\, d\lambda(v)\right\}.$$

Because $d\lambda(v) = dv - dK(v)$, the expression on the right side reduces to $\exp\{\int_x^y 1/\lambda(v)\, dK(v)\}$, which is less than 1 because $\lambda(v)$ is negative, K is increasing, and $x < y$ by hypothesis.

Proof of Proposition 1.3. This proof is a straightforward exercise in calculus based on an application of proposition 4.2 in Genest and MacKay (1986a). The details are left to the reader.

Proof of Proposition 2.1. From the definition of $K_n(v)$, one has $E\{K_n(v)\} = E\{\delta(v - V_1)\} = \Pr(V_1 \le v)$ and

$$\operatorname{var}\{K_n(v)\} = \operatorname{var}\{\delta(v - V_1)\}/n + (n-1)\operatorname{cov}\{\delta(v - V_1), \delta(v - V_2)\}/n.$$

Because $\delta(v - V_1)$ and $\delta(v - V_2)$ are identically distributed, dependent Bernoulli random variables with parameter $p(v) = \Pr(V_1 \le v)$, the variance of $K_n(v)$ can also be written as

$$p(v)\{1 - p(v)\}/n + (n-1)\{\Pr(V_1 \le v, V_2 \le v) - p(v)^2\}/n. \quad \text{(A.2)}$$

To establish the result, one thus must find $o(1/n)$-approximations for both the distribution of V_1 and the joint distribution of V_1 and V_2. This will be done using moment-generating function techniques, taking a conditional approach.

Given $(X_1, Y_1) = (x_1, y_1)$, the quantity $(n-1)V_1$ is distributed as a binomial random variable with parameters $(n-1)$ and $C(x_1, y_1)$, the probability that $(X, Y) \le (x_1, y_1)$ componentwise. The conditional moment-generating function of V_1 is thus given by $\{1 - C(x_1, y_1) + C(x_1, y_1)e^{t/(n-1)}\}^{n-1}$, whence

$$E(e^{tV_1}) = E[\{1 - C(X_1, Y_1) + C(X_1, Y_1)e^{t/(n-1)}\}^{n-1}].$$

By Lebesgue's dominated convergence theorem, it follows that $E(e^{tV_1})$ converges to $E\{e^{tC(X_1,Y_1)}\}$ as n tends to infinity. Asymptotically, therefore, V_1 is seen to have the same distribution as $C(X_1, Y_1)$; that is,

$$\lim_{n\to\infty} \Pr(V_1 \le v) = K(v). \quad \text{(A.3)}$$

Next, it will be shown that

$$\Pr(V_1 \le v, V_2 \le v) - \Pr(V_1 \le v)^2 = k(v)\{k(v)R(v) - 2v\bar{K}(v)\}/n + o(1/n). \quad \text{(A.4)}$$

Substituting (A.3) and (A.4) in (A.2) will then yield the stated conclusion. The idea is to compute the term of order $1/n$ in the bivariate Laplace–Stieltjes transform (Abramowitz and Stegun 1972, chap. 29) of

$$\Pr(V_1 \le v_1, V_2 \le v_2) - \Pr(V_1 \le v_1)\Pr(V_2 \le v_2). \quad \text{(A.5)}$$

Inverting this term of order $1/n$ will then yield the desired approximation.

Because Laplace–Stieltjes transforms are just moment-generating functions, up to a sign change, the Laplace–Stieltjes transform of (A.5) can simply be expressed as $M_{12}(t_1, t_2) - M(t_1)M(t_2)$, where $M_{12}(-t_1, -t_2)$ is the moment-generating function of (V_1, V_2) and $M(-t_1)$ is the moment-generating function of V_1. The argument proceeds in two steps.

Step 1. Approximating $M_{12}(t_1, t_2) - M(t_1)M(t_2)$. To compute the bivariate Laplace–Stieltjes transform $M_{12}(t_1, t_2)$, use the indicator random variables I_{im} defined by (2.3). The argument proceeds by conditioning on the values (x_1, y_1) and (x_2, y_2) of the pairs (X_1, Y_1) and (X_2, Y_2). Clearly, the $n - 2$ remaining sample points can then be allocated to the different cells of the following 2×2 table according to a multinomial distribution, namely

	$I_{2m} = 1$	$I_{2m} = 0$
$I_{1m} = 1$	N_{11}	N_{12}
$I_{1m} = 0$	N_{21}	N_{22}

Here N_{ij} denotes the observed frequency associated with cell (i, j), $i, j = 1, 2$. If p_{ij} denotes the conditional probability associated with that cell, one then has $p_{11} = C\{\min(x_1, x_2), \min(y_1, y_2)\}$, $p_{1+} = p_{11} + p_{12} = C(x_1, y_1)$, and $p_{+1} = p_{11} + p_{21} = C(x_2, y_2)$. Furthermore, note that from the definition of V_1 and V_2 given in Equation (4), $(n-1)V_1 = I_{12} + \sum_{m=3}^{n} I_{1m} = I_{12} + N_{11} + N_{12}$ and $(n-1)V_2 = I_{21} + \sum_{m=3}^{n} I_{2m} = I_{21} + N_{11} + N_{21}$.

Because the joint conditional distribution of the N_{ij}'s is multinomial with parameters $n - 2$ and p_{ij}, $i, j = 1, 2$, the conditional Laplace–Stieltjes transform of (V_1, V_2) is given by

$$\exp\{(-t_1 I_{12} - t_2 I_{21})/(n-1)\} \times [p_{11}\exp\{(-t_1 - t_2)/(n-1)\} + p_{12}\exp\{-t_1/(n-1)\} + p_{21}\exp\{-t_2/(n-1)\} + p_{22}]^{n-2}.$$

Taking the expectation with respect to (X_1, Y_1) and (X_2, Y_2) yields the unconditional Laplace–Stieltjes transform $M_{12}(t_1, t_2)$. Straightforward but tedious algebra leads to the following approximation:

$$\begin{aligned} M_{12}(t_1, t_2) = E[&\exp(-t_1 p_{1+} - t_2 p_{+1}) \\ &\times \{1 - t_1(1 + I_{12} - 2p_{1+})/n \\ &- t_2(1 + I_{21} - 2p_{+1})/n + t_1^2 p_{1+}(1 - p_{1+})/2n \\ &+ t_2^2 p_{+1}(1 - p_{+1})/2n \\ &+ t_1 t_2(p_{11} - p_{1+}p_{+1})/n\}] + o(1/n). \end{aligned} \quad \text{(A.6)}$$

Setting $t_2 = 0$ and evaluating the expectation with respect to (X_2, Y_2) yields the following $o(1/n)$ approximation for the Laplace–Stieltjes transform of any $\Pr(V_i \le v_i)$, $1 \le i \le n$:

$$M(t) = E[e^{-tp_{1+}}\{1 - t(1 - p_{1+})/n + t^2 p_{1+}(1 - p_{1+})/2n\}] + o(1/n).$$

Because p_{1+} and p_{+1} are independent random variables with the same distribution, one can write the product $M(t_1)M(t_2)$ as an expectation similar to (A.6). Subtracting $M(t_1)M(t_2)$ from $M_{12}(t_1, t_2)$ yields the term of order $1/n$ of the Laplace–Stieltjes transform of (A.5), namely

$$E[\exp(-t_1 p_{1+} - t_2 p_{+1})\{-t_1(I_{12} - p_{1+}) - t_2(I_{21} - p_{+1}) + t_1 t_2(p_{11} - p_{1+}p_{+1})\}]/n. \quad \text{(A.7)}$$

Step 2. Deriving the function whose Laplace–Stieltjes transform is given by (A.7). To get an $o(1/n)$ approximation for (A.5), it now remains to invert (A.7). To do this, one can adapt classical inversion rules for Laplace–Stieltjes transforms to the current context (Abramowitz and Stegun 1972, chap. 29). Bearing in mind that p_{1+} and p_{+1} are independent random variables with distribution function $K(v)$ and density $k(v)$, which is assumed to exist, the following inversion rules for Laplace–Stieltjes transforms can be used with suitable choices of the arbitrary function $r(v_1, v_2)$:

Rule A. If $r(0, t) = r(1, t) = 0$ for t in $[0, 1]$, then the inverse Laplace–Stieltjes transform of $E\{t_1 r(p_{1+}, p_{+1})e^{-t_1 p_{1+} - t_2 p_{+1}}\}$ is $k(v_1)\int_0^{v_2} k(v)r(v_1, v)\,dv$.

Rule B. If $r(0, t) = r(1, t) = r(t, 0) = r(t, 1) = 0$, for t in $[0, 1]$, then the inverse Laplace–Stieltjes transform of $t_1 t_2 r(p_{1+}, p_{+1})e^{-t_1 p_{1+} - t_2 p_{+1}}$ is $k(v_1)k(v_2)r(v_1, v_2)$.

Define $R(v_1, v_2) = E(p_{11} - v_1 v_2 | p_{1+} = v_1, p_{+1} = v_2)$, $S_1(v_1, v_2) = E(I_{12} - v_1 | p_{1+} = v_1, p_{+1} = v_2)$, and $S_2(v_1, v_2) = E(I_{21} - v_2 | p_{1+} = v_1, p_{+1} = v_2)$. Because S_1 satisfies the assumption of Rule A, the inverse Laplace–Stieltjes transform of the linear terms in t_1 appearing in (A.7) is $k(v_1)\int_0^{v_2} k(v)S_1(v_1, v)\,dv/n$. A similar reasoning with the term in t_2 yields $k(v_2)\int_0^{v_1} k(v)S_2(v, v_2)\,dv/n$ as its inverse Laplace–Stieltjes transform.

By Rule B, one finds $k(v_1)k(v_2)R(v_1, v_2)/n$ for the inverse Laplace–Stieltjes transform for the remaining term in (A.7). Bringing these various terms together, the following $o(1/n)$ approximation to (A.5) obtains

$$\left\{k(v_1)k(v_2)R(v_1, v_2) - k(v_1)\int_0^{v_2} k(v)S_1(v_1, v)\,dv - k(v_2)\times\int_0^{v_1} k(v)S_2(v, v_2)\,dv\right\}/n. \quad \text{(A.8)}$$

To show that this leads to (A.4) when $v_1 = v_2 = v$, note that

$$\int_0^{v} k(u)S_1(v, u)\,du = v\bar{K}(v). \quad \text{(A.9)}$$

To verify this assertion, simply write the expression on the left side of (A.9) as

$$\Pr\{X_2 \le X_1, Y_2 \le Y_1, C(X_2, Y_2) \le v | C(X_1, Y_1) = v\} - vK(v). \quad \text{(A.10)}$$

Because $\{X_2 \le X_1, Y_2 \le Y_1\}$ implies $\{C(X_2, Y_2) \le v\}$, the first term in (A.10) is equal to v and (A.9) is true. Setting $R(v) = R(v, v)$, the quantity defined in Equation (8), (A.8) thus reduces to $k(v)\{k(v)R(v) - 2v\bar{K}(v)\}/n$.

Proof of Proposition 2.2. For Archimedean copulas, the identity $C(X_1, Y_1) = v$ can be reexpressed as $\phi(X_1) + \phi(Y_1) = \phi(v)$. To evaluate $R(v)$, note that $X_1 = \min(X_1, X_2)$ implies $Y_2 = \min(Y_1, Y_2)$, when $C(X_1, Y_1) = C(X_2, Y_2) = v$. By symmetry,

$$\begin{aligned} E[C\{\min(X_1, X_2), \min(Y_1, Y_2)\} | C(X_1, Y_1) = C(X_2, Y_2) = v] \\ = 2E[C(X_1, Y_2)\delta(X_2 - X_1) | C(X_1, Y_1) = C(X_2, Y_2) = v], \end{aligned} \quad \text{(A.11)}$$

where δ denotes the distribution function of a point mass at the origin. By Proposition 1.1, the conditional distributions of $U_1 = \phi(X_1)/\phi(v)$ and $U_2 = \phi(X_2)/\phi(v)$ are uniform on the unit interval, whence the right side of (A.11) reduces to

$$2\int_0^1 \int_0^{u_1} \phi^{-1}\{u_1\phi(v) + \phi(v) - u_2\phi(v)\}\,du_2\,du_1.$$

Setting $t = u_1 - u_2$ in the second integral and changing the order of integration completes the argument.

Proof of Proposition 3.1. Because τ_n is a U statistic, its asymptotic distribution is normal, according to Lee (1990, pp. 75–76) (Example 4 on page 14 of that reference supplies an exact expression for $\text{var}(\tau_n)$.) Neglecting terms of the order of $1/n^2$ yields the expression $\text{var}(\tau_n) = 16\ \text{var}\{1 - X - Y + 2C(X, Y)\}/n$ for the asymptotic variance of τ_n. Proceeding as in the proof of Proposition 2.1, one can show that in large samples, the distribution of $V_i + W_i$ is approximately equal to that of $1 - X - Y + 2C(X, Y)$. Therefore, the sample variance of the $V_i + W_i$'s is a consistent estimator of $\text{var}\{1 - X - Y + 2C(X, Y)\}$.

[*Received February 1992. Revised September 1992.*]

REFERENCES

Abramowitz, M., and Stegun, I. E. (1972), *Handbook of Mathematical Functions,* New York: Dover.

Ali, M. M., Mikhail, N. N., and Haq, M. S. (1978), "A Class of Bivariate Distributions Including the Bivariate Logistic," *Journal of Multivariate Analysis,* 8, 405–412.

Bilodeau, M. (1989), "On the Monotone Regression Dependence for Archimedean Bivariate Uniform," *Communications in Statistics,* Part A: Theory and Methods, 18, 981–988.

Capéraà, P., and Genest, C. (1990), "Concepts de dépendance et ordres stochastiques pour des lois bidimensionnelles," *The Canadian Journal of Statistics,* 18, 315–326.

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., and Johnson, M. E. (1981), "A Family of Distributions for Modeling Nonelliptically Symmetric Multivariate Data," *Journal of the Royal Statistical Society,* Ser. B, 43, 210–218.

——— (1986), "Generalized Burr–Pareto–logistic Distributions With Applications to a Uranium Exploration Data Set," *Technometrics,* 28, 123–131.

Deheuvels, P. (1978), "Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes," *Publications de l'Institut de statistique de l'Université de Paris,* 23, 1–36.

Frank, M. J. (1979), "On the Simultaneous Associativity of $F(x, y)$ and $x + y - F(x, y)$," *Æquationes Mathematicæ,* 19, 194–226.

Genest, C. (1987), "Frank's Family of Bivariate Distributions," *Biometrika,* 74, 549–555.

Genest, C., and MacKay, R. J. (1986a), "Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données," *The Canadian Journal of Statistics,* 14, 145–159.

——— (1986b), "The Joy of Copulas: Bivariate Distributions With Uniform Marginals," *The American Statistician,* 40, 280–283.

Genest, C., and Rivest, L.-P. (1989), "A Characterization of Gumbel's Family of Extreme Value Distributions," *Statistics and Probability Letters,* 8, 207–211.

Gumbel, E. J. (1960), "Distribution des valeurs extrêmes en plusieurs dimensions," *Publications de l'Institut de statistique de l'Université de Paris,* 9, 171–173.

Hougaard, P. (1984), "Life Table Methods for Heterogeneous Populations: Distributions Describing the Heterogeneity," *Biometrika,* 71, 75–83.

——— (1986), "Survival Models for Heterogeneous Populations Derived from Stable Distributions," *Biometrika,* 73, 387–396.

Hougaard, P., Harvald, B., and Holm, N. V. (1992), Measuring the similarities between the lifetimes of adult Danish twins born between 1881–1930. *Journal of the American Statistical Association,* 87, 17–24.

Kimeldorf, G., and Sampson, A. R. (1975), "Uniform Representations of Bivariate Distributions," *Communications in Statistics,* Part A: Theory and Methods, 4, 617–627.

Kotel'nikova, V. F., and Chmaladze, E. V. (1982), "On Computing the Probability of an Empirical Process not Crossing a Curvilinear Boundary," *Theory of Probability and its Applications,* 27, 640–648.

Lee, A. J. (1990), *U-Statistics: Theory and Practice,* New York: Marcel Dekker.

Mardia, K. V. (1970), *Families of Bivariate Distributions,* London: Griffin.

Marshall, A. W., and Olkin, I. (1988), "Families of Multivariate Distributions," *Journal of the American Statistical Association,* 83, 834–841.

Nelsen, R. B. (1986), "Properties of a One-Parameter Family of Bivariate Distributions With Specified Marginals," *Communications in Statistics,* Part A: Theory and Methods, 15, 3277–3285.

Oakes, D. (1982), "A Model for Association in Bivariate Survival Data," *Journal of the Royal Statistical Society,* Ser. B, 44, 414–422.

——— (1986), "Semi-Parametric Inference in a Model for Association in Bivariate Survival Data," *Biometrika,* 73, 353–361.

——— (1989), "Bivariate Survival Models Induced by Frailties," *Journal of the American Statistical Association,* 84, 487–493.

Plackett, R. L. (1965), "A Class of Bivariate Distributions," *Journal of the American Statistical Association,* 60, 516–522.

Pretorius, S. J. (1930), "Skew Bivariate Frequency Surfaces Examined in the Light of Numerical Illustrations," *Biometrika,* 22, 109–223.

Saunders, R., and Laud, P. (1980), "The Multidimensional Kolmogorov Goodness-of-Fit Test," *Biometrika,* 67, 237.

Schweizer, B., and Sklar, A. (1983), *Probabilistic Metric Spaces,* Amsterdam: North-Holland.

Shorack, G. R., and Wellner, J. A. (1986), *Empirical Processes with Applications to Statistics,* New York: John Wiley.

Tawn, J. A. (1988), "Bivariate Extreme Value Theory: Models and Estimation," *Biometrika,* 75, 397–415.

Thélot, C. (1985), "Lois logistiques à deux dimensions," *Annales de l'INSEE,* 58, 123–149.

Vitale, R. A. (1979), "Regression With Given Marginals," *The Annals of Statistics,* 7, 653–658.