<table>
<thead>
<tr>
<th><strong>Statistica Sinica Preprint No:</strong> SS-2014-0082R3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td><strong>Manuscript ID</strong></td>
</tr>
<tr>
<td><strong>URL</strong></td>
</tr>
<tr>
<td><strong>DOI</strong></td>
</tr>
<tr>
<td><strong>Complete List of Authors</strong></td>
</tr>
<tr>
<td><strong>Corresponding Author</strong></td>
</tr>
<tr>
<td><strong>E-mail</strong></td>
</tr>
</tbody>
</table>

Notice: Accepted version subject to English editing.
Minimum-distance statistics for the selection of an asymmetric copula in Khoudraji’s class of models

Jean-François Quessy & Othmane Kortbi

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

Abstract: The modeling of bivariate dependence is usually accomplished with symmetric copula models. However, many examples on real datasets show that this hypothesis of symmetry may frequently fail to hold, so there is a need for inferential methods using asymmetric dependence structures. In this paper, useful tools for modeling non-exchangeable dependence structures are developed under a broad class of asymmetric copulas introduced by Khoudraji (1995). A special attention is given to the testing of the composite hypothesis that the underlying copula of a population belongs to this general class of models. The problem of selecting a specific Khoudraji-type copula via goodness-of-fit testing is considered as well, hence providing a complete set of tools for inference when facing bivariate data exhibiting an asymmetric dependence structure. Monte Carlo simulations show that the newly introduced methodologies work well in small and moderate sample sizes. Their usefulness for copula modeling is illustrated on real data sets exhibiting patterns of asymmetric dependence.

Keywords and phrases: Empirical copula process, multiplier bootstrap, shape hypothesis.

1. Introduction

Let \((X, Y)\) be a random pair such that the joint distribution function \(F(x, y) = \mathbb{P}(X \leq x, Y \leq y)\) has continuous margins \(F_X(x) = \mathbb{P}(X \leq x)\) and \(F_Y(y) = \mathbb{P}(Y \leq y)\). In that case, it is well known that there exists a unique copula \(C\) :
[0, 1]^2 \rightarrow [0, 1] such that the representation \( F(x, y) = C[F_X(x), F_Y(y)] \) holds for all \((x, y) \in \mathbb{R}^2\). The modeling of dependence using copulas has found many applications in diverse areas including finance, actuarial sciences and hydrology. See Joe (1997), Cherubini et al. (2004) and Nelsen (2006) for details on their theoretical aspects.

Most of the copulas commonly used in practice are symmetric, i.e., they satisfy \( C(u, v) = C(v, u) \) for all \((u, v) \in [0, 1]^2\). This property is shared, e.g., by all models in the Archimedean and meta-elliptical families, making them appropriate only in situations where observed random pairs come from a distribution whose underlying copula is symmetric with respect to the main diagonal. Otherwise, conclusions from such models could be misleading.

In order to illustrate a real life situation of asymmetric dependence, consider 47,388 pairs taken from the Walker Lake data set described by Isaaks & Srivastava (1989). As mentioned by these authors, the meaning of these two variables is not revealed for pedagogical reasons, and may be, for example, “thicknesses of a geologic horizon or the concentration of some pollutant”. From the scatter plot of the pairs of standardized ranks on the upper left panel of Figure 5, one may conclude to a strong asymmetry; trying to fit a symmetric copula to these data would clearly be inappropriate. Hence, for such bivariate data, it is crucial to find a suitable asymmetric copula model.

**FIGURE 5 ABOUT HERE**

Potentially useful models for asymmetric dependence modeling are those studied by Khoudraji (1995). Specifically, based on two copulas \( C_1 \) and \( C_2 \), one can build another copula model via

\[
C_{\delta}(u, v) = C_1(u^{\delta_1}, v^{\delta_2}) C_2(u^{1-\delta_1}, v^{1-\delta_2}),
\]

where \( \delta = (\delta_1, \delta_2) \in [0, 1]^2 \); in general, \( C_\delta \) is asymmetric whenever \( \delta_1 \neq \delta_2 \). See also Liebscher (2008) for related constructions. In this work, particular attention
is given to the special case when $C_1(u, v) = uv$ and $C_2 = D$ is symmetric. This yields a rich family of dependence functions of the form

$$C_{\mathbf{\delta}, D}(u, v) = u^{\delta_1} v^{\delta_2} D(u^{1-\delta_1}, v^{1-\delta_2}),$$

(2)

where $\mathbf{\delta} = (\delta_1, \delta_2) \in [0, 1]^2$. For reasons of uniqueness, the cases $\mathbf{\delta} = (1, \delta)$ and $\mathbf{\delta} = (\delta, 1)$ are excluded since they correspond to the independence copula $C_{\mathbf{\delta}, D}(u, v) = uv$ whatever the value of $\delta \in [0, 1)$. As will be seen, the dependence patterns offered by this class of models is somewhat similar to the kind of asymmetry that one observes for the Walker Lake data. A formal methodology is nevertheless needed to assess the appropriateness of such asymmetric models.

The goal of this paper is to develop statistical tools for the modeling of asymmetric dependence via model (2). Rather than trying to fit a particular parametric copula structure of this form, i.e., fixing $D$ up to a parameter to be estimated, the main focus here is testing the composite hypothesis that the unknown copula of a bivariate population admits this representation. Hence, one first assesses the appropriateness of the general representation (2), and then one seeks for a particular model for the symmetric part $D$. This model selection step is also addressed formally, yielding a complete set of tools for asymmetric dependence modeling.

The paper is organized as follows. In Section 2, a characterization of the null hypothesis is given. In Section 3, an empirical process related to this characterization is defined and test statistics build around it are proposed in the case when the parameter that manages the asymmetry is assumed to be known. The methodology is extended in Section 4 to the more realistic situation where the asymmetry parameter is unknown via with the help of minimum-distance statistics. How to formally choose a specific Khoudraji-type model via goodness-of-fit testing is addressed in Section 5. The details of an investigation of the sample properties of the tests via Monte Carlo simulations are given in Section 6. The article ends with illustrations of the newly introduced methodologies on data
exhibiting patterns of asymmetric dependence; details are given in Section 7.

2. Characterization of the null hypothesis

As already stated in the Introduction, the first goal of this article is to provide a formal way to test that the underlying copula $C$ of a bivariate population admits a representation of the form given in equation (2). To this end, denote by $\mathcal{K}$ the class of copulas of the form $C_{\delta,D}$ described in (2), where $\delta = (\delta_1, \delta_2) \in [0,1)^2$ and $D$ is a copula such that

$$u^\xi D(u^{1-\xi}, v) = v^\xi D(v^{1-\xi}, u) \quad \text{if and only if } \xi = 0.$$  

This assumption excludes, for example, the case when $D = \Pi$, where $\Pi(u,v) = uv$ is the independence copula. More importantly, this assumption ensures that the members of $\mathcal{K}$ can be written in such a way that they have a unique representation; this is formally stated in the following lemma.

**Lemma 1** Any copula $C_{\delta,D} \in \mathcal{K}$ admits the unique representation

$$C_{\delta,D}(u,v) = \begin{cases} u^{\delta_1} D^*(u^{1-\delta_1}, v), & \text{if } \delta_1 > \delta_2; \\ v^{\delta_2} D^*(u, v^{1-\delta_2}), & \text{if } \delta_1 < \delta_2; \\ D^*(u,v), & \text{if } \delta_1 = \delta_2, \end{cases}$$  

where $\beta = \beta(\delta) \in [0,1)$ and $D^*$ is a copula that satisfies (3).

The main null and alternative hypotheses of interest in this paper can be stated as $H_0 : C \in \mathcal{K}$ and $H_1 : C \notin \mathcal{K}$. The latter are of a composite nature because the specific form of the copula under $H_0$ is not specified. They then fall into the category of so-called shape hypotheses. As a consequence of Lemma 1, one can focus on models of the form

$$C_{\beta,D}(u,v) = u^\beta D(u^{1-\beta}, v),$$  

(5)
where $\beta \in [0, 1)$ and $D$ satisfies (3). For, if the copula $C$ associated to a random pair $(X, Y)$ belongs to the class $\mathcal{K}$ of models with $\delta_1 < \delta_2$, then it is the copula of $(Y, X)$ that writes in the form (5), according to Lemma 1. In the sequel, $\mathcal{K}'$ denotes the subset of $\mathcal{K}$ that consists of copulas that admit representation (5).

The null and alternative hypotheses can then be reformulated as

$$\mathcal{H}_0 : C \in \mathcal{K}' \quad \text{and} \quad \mathcal{H}_1 : C \notin \mathcal{K}'.$$

A true null hypothesis means that $C(u, v) = u^{\beta_0} D(u^{1-\beta_0}, v)$ for a unique $\beta_0 \in [0, 1)$ and a unique copula $D$ that satisfies (3). Then, because $D$ is symmetric, one has for all $(u, v) \in [0, 1]^2$ that

$$v^{\beta_0} C(u, v^{1-\beta_0}) = (uv)^{\beta_0} D(u^{1-\beta_0}, v^{1-\beta_0}) = (uv)^{\beta_0} D(v^{1-\beta_0}, u^{1-\beta_0}) = u^{\beta_0} C(v, u^{1-\beta_0}).$$

(6)

The converse is also true, so that equation (6) is a characterization of the class $\mathcal{K}'$ of copulas. This is formally established in the next proposition.

**Proposition 1** Let $C$ be a copula that satisfies equation (6) for a unique $\beta_0 \in [0, 1)$. If in addition the bivariate function defined for $(u, v) \in [0, 1]^2$ by

$$D(u, v) = u^{-\beta_0} C(u^{-\beta_0}, v)$$

is a copula, then $C = C_{\beta_0, D} \in \mathcal{K}'$.

3. Testing $\mathcal{H}_0$ for a fixed asymmetry parameter $\beta_0$

3.1. **An empirical process for $\mathcal{H}_0$**

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be independent random copies of a pair $(X, Y)$ from some joint distribution function $F$ with continuous margins $F_X$ and $F_Y$. Suppose that the unique copula $C$ of $F$ belongs to the class $\mathcal{K}'$ of copulas, which means...
that \( C = C_{\beta_0, D} \) for some \( \beta_0 \in [0, 1) \) and a copula \( D \) that satisfies equation (3). Defining for each \((u, v) \in [0, 1]^2\) and \( \beta \in [0, 1) \) the function

\[
Q_{\beta, C}(u, v) = v^\beta C(u, v^{1-\beta}) - u^\beta C(v, u^{1-\beta}),
\]

it follows from equation (6) that \( Q_{\beta_0, C}(u, v) = 0 \) for all \((u, v) \in [0, 1]^2\). An empirical version of \( Q_{\beta, C} \) arises while replacing \( C \) by the empirical copula first proposed by Rüschendorf (1976), namely

\[
C_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} I(U_{i,n} \leq u, V_{i,n} \leq v),
\]

where for each \( i \in \{1, \ldots, n\} \), \( U_{i,n} = F_{n,X}(X_i) \), \( V_{i,n} = F_{n,Y}(Y_i) \) and \( F_{n,X}, F_{n,Y} \) are the marginal empirical distribution functions. This suggests the study of the empirical process \( Q_{n, \beta} = \sqrt{n} (Q_{\beta, C_n} - Q_{\beta, C}) \). Under \( \mathbb{H}_0 \), \( Q_{\beta_0, C} \equiv 0 \) and then

\[
Q_{n, \beta_0}(u, v) = \sqrt{n} \left\{ v^{\beta_0} C_n(u, u^{1-\beta_0}) - u^{\beta_0} C_n(v, v^{1-\beta_0}) \right\}.
\]

The asymptotic behavior of \( Q_{n, \beta_0} \) under \( \mathbb{H}_0 \) will inherit from the large-sample properties of \( C_n \), which are now well established. Indeed, the asymptotic behavior of the empirical copula process \( C_n = \sqrt{n}(C_n - C) \) was investigated by Deheuvels (1981) under independence. General weak convergence in the space \( D([0, 1]^2) \) of càdlàg functions equipped with the Skorohod topology was investigated by Gaënßler & Stute (1987); van der Vaart & Wellner (1996) show weak convergence in the space \( \ell^\infty([a, b]^2) \) of bounded functions on \([a, b]^2\) for \(0 < a < b < 1\). The result was extended to the space \( \ell^\infty([0, 1]^2) \) by Fermanian et al. (2004) while assuming the existence and continuity of the partial derivatives \( C_{10}(u, v) = \partial C(u, v)/\partial u \) and \( C_{01}(u, v) = \partial C(u, v)/\partial v \) on \([0, 1]^2\). In that case, \( C_n \) converges weakly with respect to the supremum distance to

\[
C(u, v) = B_C(u, v) - C_{10}(u, v) B_C(u, 1) - C_{01}(u, v) B_C(1, v),
\]

where \( B_C \) is a continuous and centered Gaussian process such that

\[
\mathbb{E} \{ B_C(u, v) B_C(u', v') \} = C \{ \min(u, u'), \min(v, v') \} - C(u, v) C(u', v').
\]
As shown by Segers (2012), these requirements on the partial derivatives of \( C \) are not satisfied for many extensively used copula models. Fortunately, it is shown by this author that the result still holds if \( C_{10} \) and \( C_{01} \) exist and are continuous respectively on the sets \((0, 1) \times [0, 1]\) and \([0, 1] \times (0, 1)\).

The result on the weak behavior of \( Q_{n, \beta_0} \) under \( H_0 \) can now be stated.

**Proposition 2** Let \((X_1, Y_1), \ldots, (X_n, Y_n)\) be i.i.d. from a joint distribution \( F \) with continuous margins \( F_X, F_Y \) and unique copula \( C \). If \( C \in K' \) for some \( \beta_0 \in (0, 1) \) and some copula \( D \) that satisfies (3) and such that \( D_{10}(u, v) = \partial D(u, v)/\partial u \), \( D_{01}(u, v) = \partial D(u, v)/\partial v \) exist and are continuous respectively on the sets \((0, 1) \times [0, 1]\) and \([0, 1] \times (0, 1)\), then the empirical process \( Q_{n, \beta_0} \) converges weakly in the space \( \ell^\infty([0, 1]^2) \) to

\[
Q_{\beta_0}(u, v) = v^{\beta_0} C(u, v^{1-\beta_0}) - u^{\beta_0} C(v, u^{1-\beta_0}),
\]

where \( C \) is the Gaussian weak limit of \( C_n \) described in equation (8).

Proposition 2 is a special case of a more general result about the weak behavior of the empirical process \( Q_{n, \beta} = \sqrt{n}(Q_{\beta, C_n} - Q_{\beta, C}) \). In fact, a straightforward adaptation of the proof of Proposition 2 yields the conclusion that for a fixed \( \beta \in [0, 1) \), the process \( Q_{n, \beta} \) converges weakly in the space \( \ell^\infty([0, 1]^2) \) to

\[
Q_{\beta}(u, v) = v^{\beta} C(u, v^{1-\beta}) - u^{\beta} C(v, u^{1-\beta})
\]

as long as \( C_{10}(u, v) = \partial C(u, v)/\partial u \) and \( C_{01}(u, v) = \partial C(u, v)/\partial v \) exist and are continuous respectively on the sets \((0, 1) \times [0, 1]\) and \([0, 1] \times (0, 1)\). More generally still, the result can be shown to be uniform in \( \beta \in [0, 1) \), i.e.,

\[
\sup_{(u, v, \beta) \in [0, 1]^2 \times [0, 1]} |Q_{n, \beta}(u, v) - Q_{\beta}(u, v)| \xrightarrow{P} 0.
\]

**3.2. Test statistics**

Since \( Q_{\beta_0, C} \) vanishes under the null hypothesis, a test of \( H_0 \) against \( H_1 \) could consider some sort of distance function applied to its empirical version \( Q_{\beta_0, C_n} \).
To this end, let \( \mathcal{M} : \ell^\infty([0,1]^2) \to \mathbb{R}^+ \) be a norm on the space of bounded functions on \([0,1]^2\) and define the test statistic

\[
S_n^\mathcal{M}(\beta_0) = \sqrt{n} \mathcal{M}(Q_{\beta_0}, C_n).
\]

Popular candidates for \( \mathcal{M} \) are the Cramér–von Mises and Kolmogorov–Smirnov functionals, namely

\[
\mathcal{M}^{\text{CvM}}(g) = \left( \int_{[0,1]^2} \{g(u, v)\}^2 \, du \, dv \right)^{1/2},
\]

\[
\mathcal{M}^{\text{KS}}(g) = \sup_{(u,v) \in [0,1]^2} |g(u,v)|,
\]

where \( g \in \ell^\infty([0,1]^2) \). In the sequel, the statistics corresponding to these two functionals are denoted \( S_n^{\text{CvM}}(\beta_0) \) and \( S_n^{\text{KS}}(\beta_0) \), respectively. An interesting feature of \( S_n^{\text{CvM}}(\beta_0) \) is that an explicit formula can be derived. Indeed, letting \( a \vee b = \max(a,b) \), one can show that

\[
n\{S_n^{\text{CvM}}(\beta)\}^2 = \frac{2}{\beta+1} \sum_{i,j=1}^n \left\{ \{1 - (U_{i,n} \vee U_{j,n})\} \left\{1 - (V_{i,n} \vee V_{j,n})^{\frac{2}{\beta+1}}\right\} \right.
\]

\[
- \frac{2}{(\beta + 1)^2} \sum_{i,j=1}^n \left\{ \{1 - (U_{i,n} \vee V_{j,n})^{\frac{1}{\beta+1}}\} \left\{1 - (V_{i,n} \vee U_{j,n})^{\frac{1}{\beta+1}}\right\} \right\}.
\]

Observe that whenever the null hypothesis holds, \( S_n^\mathcal{M}(\beta_0) = \mathcal{M}(Q_{n,\beta_0}) \), where \( Q_{n,\beta_0} \) is defined in equation (7). Hence, under the conditions of Proposition 2, the continuous mapping theorem ensures that \( S_n^\mathcal{M}(\beta_0) \) converges in distribution to a random variable having representation

\[
S_n^\mathcal{M}(\beta_0) = \mathcal{M}(Q_{\beta_0}). \tag{10}
\]

On the other hand, one knows from Proposition 1 that equation (6) holds if and only if \( C \in \mathcal{K}' \). As a consequence, if \( C \notin \mathcal{K}' \), then there is a subset \( \mathcal{A} \) of \([0,1]^2\) such that \( Q_{\beta,C}(u, v) \neq 0 \) for \((u, v) \in \mathcal{A}\). As mentioned in the comment after
Proposition 2, \( Q_{n,\beta} \) converges weakly to a tight Gaussian process in \( \ell^\infty([0, 1]^2) \). As a consequence, it follows from the continuity of \( M \) that
\[
\frac{S_n^M(\beta)}{\sqrt{n}} = M \left( \frac{Q_{n,\beta}}{\sqrt{n}} + Q_{\beta,C} \right) \to M(Q_{\beta,C}) > 0.
\]
This entails that \( S_n^M(\beta) \to \infty \) in probability. The same conclusion holds when \( C \in \mathcal{K}' \) and \( \beta \neq \beta_0 \), since equation (6) holds for a unique \( \beta_0 \in [0, 1) \).

### 3.3. Computation of \( \mathbb{P} \)-values

The multiplier bootstrap was successfully employed in several contexts involving shape hypotheses about copulas. The first contribution is by Scaillet (2005) for testing the hypothesis of positive quadrant dependence. The method has proven useful for testing several other hypotheses including equality of copulas (see Rémillard & Scaillet (2009)), extreme-value dependence (see Kojadinovic & Yan (2010), Quessy (2012)), change-point detection (see Bücher & Ruppert (2013), Quessy et al. (2013)) and symmetry (see Genest et al. (2012)), to name only a few.

In order to describe the method, consider the independent random vectors \( (\xi^{(1)}_1, \ldots, \xi^{(1)}_n), \ldots, (\xi^{(H)}_1, \ldots, \xi^{(H)}_n) \), where for each \( h \in \{1, \ldots, H\} \), the random variables \( \xi^{(h)}_1, \ldots, \xi^{(h)}_n \) are independent, positive, and satisfy
\[
\mathbb{E}(\xi^{(h)}_i) = \text{var}(\xi^{(h)}_i) = 1 \quad \text{and} \quad \int_0^\infty \left\{ \mathbb{P}(\xi^{(h)}_i > x) \right\}^{1/2} dx < \infty.
\]
The value of \( H \) corresponds to the number of replicates and in practice is chosen sufficiently large. A valid choice for the law of \( \xi^{(h)}_i \) is the exponential distribution with mean one. The so-called multiplier versions of the empirical copula process are given for each \( h \in \{1, \ldots, H\} \) by
\[
C_n^{(h)}(u, v) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \gamma_i^{(h)} \left\{ I(U_{i,n} \leq u, V_{i,n} \leq v) - C_{n,10}(u, v) I(U_{i,n} \leq u) - C_{n,01}(u, v) I(V_{i,n} \leq v) \right\},
\]
where \( \gamma_i = \frac{\zeta_i^{(h)}}{\xi_i^{(h)}} - 1 \) and \( C_{n,10} \) and \( C_{n,01} \) are data-based estimators of the partial derivatives of \( C \) such that for any \( \epsilon \in (0, 1/2) \),

\[
\sup_{u \in [\epsilon, 1-\epsilon], \ v \in [0,1]} |C_{n,10}(u, v) - C_{10}(u, v)| \quad \text{and} \quad \sup_{u \in [\epsilon, 1-\epsilon], \ v \in [0,1]} |C_{n,01}(u, v) - C_{01}(u, v)|
\]

converge in probability to zero; such estimators based on finite differences are described in Section 6. Under these conditions, a slight adaptation of a result that one can find in Segers (2012) entails that \( (C_n^{(1)}, C_n^{(2)}, \ldots, C_n^{(H)}) \) converges in probability to \( (C^{(1)}, C^{(2)}, \ldots, C^{(H)}) \), where \( C^{(1)}, \ldots, C^{(H)} \) are independent copies of \( C \). This result is useful since one can replicate the asymptotic distribution of any continuous functional \( L : \ell^\infty([0, 1]^2) \rightarrow \mathbb{R} \) of \( C_n \) with \( L(C_n^{(1)}), \ldots, L(C_n^{(H)}) \), from which \( \mathbb{P} \)-values can be computed.

Natural multiplier bootstrap versions of \( Q_{\beta_0} \) based on its asymptotic representation given in Proposition 2 are, for each \( h \in \{1, \ldots, H\} \),

\[
Q_{n,\beta_0}^{(h)}(u, v) = v^{\beta_0} C_n^{(h)}(u, v^{1-\beta_0}) - u^{\beta_0} C_n^{(h)}(v, u^{1-\beta_0}).
\] (11)

Their asymptotic validity can be established by a straightforward application of the continuous mapping theorem. Indeed, it suffices to note that the functional \( \Phi_{\beta_0} : \ell^\infty([0, 1]^2) \rightarrow \ell^\infty([0, 1]) \) defined by

\[
\Phi_{\beta_0}(g) = v^{\beta_0} g(u, v^{1-\beta_0}) - u^{\beta_0} g(v, u^{1-\beta_0})
\]

is continuous. Multiplier versions of \( S_n^{M}(\beta_0) \) based on its limit representation in (10) are then

\[
S_n^{M,(h)}(\beta_0) = \mathcal{M}\left(Q_{n,\beta_0}^{(h)}\right).
\]

Another application of the continuous mapping theorem ensures that the latter are asymptotically independent copies of \( S_n^{M}(\beta_0) \). In practice, the computation of \( S_n^{M,(1)}(\beta_0), \ldots, S_n^{M,(H)}(\beta_0) \) is facilitated by choosing \( B \in \mathbb{N} \) sufficiently large and by making use of the approximation

\[
Q_{n,\beta_0}^{(h)}(u, v) \approx Q_{n,\beta_0}^{(h)}(\eta_k, \eta_\ell), \quad \text{when} \ (u, v) \in \left[\frac{k-1}{B}, \frac{k}{B}\right] \times \left[\frac{\ell-1}{B}, \frac{\ell}{B}\right],
\]
where \( \eta_k = (k - .5)/B \). In particular, the multiplier versions of the test statistics \( S_n^{\text{CvM}}(\beta_0) \) and \( S_n^{\text{KS}}(\beta_0) \) are given by

\[
S_n^{\text{CvM},(h)}(\beta_0) \approx \frac{1}{B} \left( \sum_{k,\ell=1}^{B} \left\{ Q_{n,\beta_0}^{(h)}(\eta_k, \eta_\ell) \right\}^2 \right)^{1/2}, \\
S_n^{\text{KS},(h)}(\beta_0) \approx \max_{k,\ell \in \{1,\ldots,B\}} \left| Q_{n,\beta_0}^{(h)}(\eta_k, \eta_\ell) \right|.
\]

4. Testing \( H_0 \) for an unknown asymmetry parameter \( \beta_0 \)

4.1. Minimum-distance statistics

The assertion that the asymmetry parameter \( \beta_0 \) is known, as made in the previous section, is rather unrealistic in practice. The purpose of this section is to extend the methodology of Section 3 in order to take into account the fact that the value of \( \beta \) is unknown in representation \( C_{\beta,D} \). At this stage, it is worth noting that whatever the form of \( C \), the bivariate function \( Q_{\beta,C} \) vanishes whenever \( \beta = 1 \). Thus the criteria based on \( Q_{\beta,C} \) previously used must be adjusted here when \( \beta \in [0,1) \) is unknown. An idea is to work with the modified version

\[
\tilde{Q}_{\beta,C} = \frac{Q_{\beta,C}}{1 - \beta}.
\]

As long as (6) holds, there will be a unique \( \beta_0 \in [0,1] \) such that \( \tilde{Q}_{\beta_0,C} \) vanishes whenever \( C \in \mathcal{K}_r \). Some curves of \( |Q_{\beta,C}| \) and \( |	ilde{Q}_{\beta,C}| \) are given in Figure 1 when \( C = C_{\beta_0,D} \) and \( D \) is Clayton’s copula, i.e., \( D = D_{\theta}^{\text{CL}} \), where

\[
D_{\theta}^{\text{CL}}(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta \geq 0.
\]

FIGURE 1 ABOUT HERE

Based on the empirical version of \( \tilde{Q}_{\beta,C} \) given by \( \tilde{Q}_{\beta,C,n} = Q_{\beta,C,n}/(1 - \beta) \), consider for a given norm \( \mathcal{M} \) the minimum-distance test statistic

\[
T_n^{\mathcal{M}} = \sqrt{n} \inf_{\beta \in (0,1)} \mathcal{M} \left( \tilde{Q}_{\beta,C,n} \right).
\]

Statistica Sinica: Newly accepted Paper
(accepted version subject to English editing)
This statistic is related to $S_n^M(\beta)$ via

$$T_n^M = \inf_{\beta \in (0,1)} \frac{S_n^M(\beta)}{1 - \beta}.$$  

Because the case $\beta_0 = 0$ lies on the boundary of the possible values of $\beta$, it results in theoretical complexities. For that reason, it has been excluded in the above definition. Since this special case corresponds to symmetry, it can easily be tested first. Now the asymptotic behavior of $T_n^M$ under $H_0$ is given; this is the main theoretical result of the paper.

**Theorem 1** If $\tilde{Q}_{\beta,C} = \partial Q_{\beta,C}/\partial \beta$ exists and is not singular at $\beta = \beta_0 \in (0,1)$, i.e., $\tilde{Q}_{\beta_0,C}(u,v) \neq 0$ for all $(u,v) \in [0,1]^2$, then under the conditions of Proposition 2, the statistic $T_n^M$ converges in distribution to

$$T^M = \inf_{t \in \mathbb{R}} M \left( \tilde{Q}_{\beta_0} + t \tilde{Q}'_{\beta_0,C} \right),$$

where $\tilde{Q}_{\beta_0} = Q_{\beta_0}/(1 - \beta_0)$ and $Q_{\beta_0}$ is the limit of $Q_{n,\beta_0}$ given in Proposition 2.

### 4.2. Minimum-distance estimator of the asymmetry parameter

An estimator of $\beta_0$ that is implicit in the definition of $T_n^M$ in (13) is

$$\beta_n^M = \arg \min_{\beta \in (0,1)} M \left( \tilde{Q}_{\beta,C_n} \right).$$  

(14)

There is a close relationship between the large-sample behavior of $\beta_n^M$ and the result on the weak convergence of $T_n^M$ stated in Theorem 1. Indeed, following Pollard (1980), suppose that the mapping

$$t \mapsto M \left( \tilde{Q}_{\beta_0} - t \tilde{Q}'_{\beta_0} \right)$$

attains its minimum at a unique $t \in \mathbb{R}$ for almost all sample paths $\tilde{Q}_{\beta_0}$. Letting $\nu$ be the functional that associates with $g \in \ell^\infty([0,1]^2)$ a value $t \in \mathbb{R}$ that minimizes $M(g - t \tilde{Q}'_{\beta_0})$, one can conclude that

$$\sqrt{n} \left( \beta_n^M - \beta_0 \right) \sim \nu \left( \tilde{Q}_{\beta_0} \right)$$
under the conditions stated in Theorem 1. 

The mean-squared error of $\beta_n$ when $M$ is either the Cramér–von Mises or the Kolmogorov–Smirnov functional has been evaluated with the help of simulations. These estimators are noted $\beta_{nCvM}^n$ and $\beta_{nKS}^n$, respectively; the results are in Table 1. Generally speaking, $\beta_0$ is easier to estimate when it is small. Under most of the scenarios considered, the estimator based on the Kolmogorov–Smirnov distance function is more accurate.

### 4.3. Multiplier bootstrap of the minimum-distance statistics

From the conclusion of Theorem 1, it is natural to define the multiplier bootstrap versions of $T_n^M$ by

$$T_n^{M,(h)} = \inf_{t \in \mathbb{R}} \mathcal{M} \left( \frac{Q_n^{(h)} - t \bar{Q}_{\beta_0,C}'}{Q_n^{(h)} + \ell_n} \right), \quad h \in \{1, \ldots, H\},$$

where $Q_n^{(h)} = Q_n^{(h)}/(1 - \beta)$ for $Q_n^{(h)}$ defined in equation (11), and

$$\bar{Q}_{\beta_0,C}'(u,v) = \frac{1}{2\ell_n} \left\{ Q_{n,\beta_n^M+\ell_n,C_n}(u,v) - Q_{n,\beta_n^M-\ell_n,C_n}(u,v) \right\},$$

where $\ell_n = 1/\sqrt{n}$. Note that a slight modification of the estimator of $Q_{\beta_0,C}'$ is needed when $\beta_n^M \leq \ell_n$ or $\beta_n^M \geq 1 - \ell_n$. This estimator is uniformly consistent for $Q_{\beta_0,C}'$. To see that this is indeed the case, observe that

$$\bar{Q}_{\beta_0,C}' = \left( \frac{Q_{n,\beta_n^M+\ell_n} - Q_{n,\beta_n^M-\ell_n}}{2\ell_n} \right) + \left( \frac{Q_{\beta_n^M+\ell_n,C_n} - Q_{\beta_n^M-\ell_n,C_n}}{2\ell_n} \right).$$

From equation (9), $Q_{n,\beta}$ converges uniformly in $\ell^\infty([0,1]^2 \times [0,1])$ to $Q_{\beta}$; hence the first summand on the right converges uniformly to zero in probability. Applying the mean-value theorem, one concludes that the second summand converges to $Q_{\beta_0,C}'$ as $n \to \infty$ because $\beta_n^M \to \beta_0$ in probability. One can then conclude that under the conditions of Theorem 1, $(T_n^M, T_n^{M,(1)}, \ldots, T_n^{M,(H)})$ converges weakly to $(T^M, T^{(1)}, \ldots, T^{(H)})$, where $T^{(1)}, \ldots, T^{(H)}$ are independent copies of the limit $T^M$ of $T_n^M$. 
4.4. Consistency of the tests

Note that

\[ \frac{T_n^M}{\sqrt{n}} = \inf_{\beta \in (0,1)} \mathcal{M} \left( \frac{\bar{Q}_{n,\beta}}{\sqrt{n}} + \bar{Q}_{\beta,C} \right), \]

where \( \bar{Q}_{n,\beta} = \frac{Q_{n,\beta}}{1 - \beta} \). Since the map \( g \mapsto \inf_{\beta \in (0,1)} \mathcal{M}(g(\cdot,\cdot,\beta)) \), for \( g \in \ell([0,1]^2 \times (0,1)) \), is continuous, and because of equation (9), one has in probability that

\[ \frac{T_n^M}{\sqrt{n}} \to \inf_{\beta \in [0,1)} \mathcal{M} \left( \bar{Q}_{\beta,C} \right). \]

In view of Proposition 1, \( \inf_{\beta \in (0,1)} \mathcal{M}(\bar{Q}_{\beta,C}) = 0 \) if and only if \( C \in \mathcal{K}' \). As a consequence, \( T_n^M \to \infty \) in probability as \( n \to \infty \) for \( C \notin \mathcal{K}' \), since \( \mathcal{M}(\bar{Q}_{\beta,C}) > 0 \) in that case. On the other hand, whether the null hypothesis holds or not, one has that \( (T_n^{M,(1)}, \ldots, T_n^{M,(H)}) \) converges to a vector \( (T^{M,(1)}, \ldots, T^{M,(H)}) \) of tight Gaussian processes. As a consequence,

\[ \bar{P}V^M = \frac{1}{H} \sum_{h=1}^{H} \mathbb{1}(T_n^{M,(h)} > T_n^M) \]

is an asymptotically valid \( P \)-value for the test based on \( T_n^M \). In other words, the test that rejects \( \mathbb{H}_0 \) for large values of \( T_n^M \) is consistent.

5. Selection of the symmetric component \( D \)

Once a test based on \( T_n^M \) concludes that the underlying copula of a population belongs to \( \mathcal{K}' \), there remains the issue of determining a specific, suitable symmetric structure for \( D \). To this end, suppose \( D \) belongs to the parametric family of one-parameter symmetric copula models \( \{D_\theta : \theta \in \Theta \subseteq \mathbb{R}\} \), where \( D_\theta \) satisfies (3) for each \( \theta \in \Theta \). Assuming that \( \beta \) is fixed, define Kendall’s dependence measure associated to \( C = C_{\beta,D_\theta} \) as a function of \( \theta \), i.e., let

\[ \kappa(\theta) = 4 \int_{[0,1]^2} C_{\beta,D_\theta}(u,v) dC_{\beta,D_\theta}(u,v) - 1. \]

In practice, \( \beta \) is replaced by the estimator \( \beta_n^M \) described in equation (14).
If $\tau_n$ is the empirical version of Kendall’s tau based on the original data set $(X_1, Y_1), \ldots, (X_n, Y_n)$, then a moment-based estimator $\theta_n$ of $\theta$ is defined as the solution of $\kappa(\theta_n) = \tau_n$. Since $\kappa$ can generally not be inverted explicitly, one must use a numerical root-finding method in order to obtain a solution. However, since the expression for $\kappa$ involves a double integral whose evaluation requires a numerical approach, a simpler Monte Carlo method can be suggested instead:

(i) Discretize $\Theta$ as $\theta_1, \ldots, \theta_N$;

(ii) For each $j \in \{1, \ldots, N\}$, generate a large sample of pairs from model $C_{\beta, D_n}$ and let $\hat{\kappa}(\theta_j)$ be Kendall’s tau for this sample;

(iii) Finally, define $\hat{\theta}_n = \text{argmin}_{j \in \{1, \ldots, N\}} |\hat{\kappa}(\theta_j) - \tau_n|$.

Another possibility for the estimation of $\theta$ is to use the so-called pseudo maximum likelihood method as described by Genest et al. (1995). The latter would however require using the density of $C_{\beta, D_n}$, which can be cumbersome.

Once $\theta$ is estimated from the above inversion of Kendall’s tau, or by maximum likelihood, consider the goodness-of-fit test statistic proposed by Genest et al. (2009), namely

$$V^\beta_{n, N} = n \int_{[0,1]^2} \{C_n(u, v) - C_N(u, v)\}^2 du dv,$$

where $C_n$ is the empirical copula computed from the original data set $(X_1, Y_1), \ldots, (X_n, Y_n)$ and $C_N$ is the empirical copula of an artificially generated data set $(X^*_1, Y^*_1), \ldots, (X^*_n, Y^*_n)$ of size $N$ from model $C_{\beta, D_n}$. One can show that

$$V^\beta_{n, N} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \{1 - (U_{i,n} \lor U_{j,n})\} \{1 - (V_{i,n} \lor V_{j,n})\} + \frac{n}{N^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \{1 - (U^*_{i,n} \lor U^*_{j,n})\} \{1 - (V^*_{i,n} \lor V^*_{j,n})\} - \frac{2}{N} \sum_{i=1}^{n} \sum_{j=1}^{n} \{1 - (U_{i,n} \lor U^*_{j,n})\} \{1 - (V_{i,n} \lor V^*_{j,n})\},$$
where \((U_{1,n}, V_{1,n}), \ldots, (U_{n,n}, V_{n,n})\) are the pairs of ranks divided by \(N\) deduced from the artificial sample. A Kolmogorov–Smirnov statistic could also be defined, but \(V_{n,N}^\beta\) is chosen here for computational convenience and also because it is generally more powerful in a copula goodness-of-fit context (see Genest et al. (2009) for more details on this aspect).

The \(P\)-value of \(V_{n,N}^\beta\) is obtained from a parametric bootstrap procedure, i.e., for a sufficiently large \(H\), the test statistic \(V_{n,N}^\beta\) is computed repeatedly from \(H\) data sets of size \(n\) simulated from model \(C_{\beta,D_{0,n}}\), yielding \(V_{n,N}^{(1)}, \ldots, V_{n,N}^{(H)}\). As thoroughly obtained by Genest & Rémillard (2008),

\[
\overline{PV} = \frac{1}{H} \sum_{h=1}^{H} I\left(V_{n,N}^{(h)} > V_{n,N}^\beta\right)
\]

is an asymptotically valid \(P\)-value as \(n, N \to \infty\).

6. Investigation of the sample properties of the tests

The aim of this section is to study how well the asymptotic results of the previous sections approximate the sampling distributions of the test statistics in small and moderate sample sizes. In the investigations that follow, the partial derivatives of the copula have been estimated by finite-difference estimators, namely

\[
C_{n,01}(u, v) = \begin{cases} 
\frac{C_n(u, 2\ell_n)}{2\ell_n}, & v \in [0, \ell_n), \\
\frac{C_n(u, v + \ell_n) - C_n(u, v - \ell_n)}{2\ell_n}, & v \in [\ell_n, 1 - \ell_n], \\
\frac{C_n(u, 1) - C_n(u, 1 - 2\ell_n)}{2\ell_n}, & v \in (1 - \ell_n, 1], 
\end{cases}
\]

where \(\ell_n = 1/\sqrt{n}\), and similarly for \(C_{n,10}\). In each of the scenarios considered, the probability of rejection of \(H_0\) has been estimated from 1000 replicates and the type I error was set to \(\alpha = .05\). The number of multiplier bootstrap samples for the computation of \(P\)-values was \(H = 1000\) and the computation of \(T_n^{M,(h)}\)
used an approximation of $\hat{Q}_{\beta_0,n}^{(h)} - t \hat{Q}_{\beta_0,C}'$ on a grid of $[0,1]^2$ of size $5 \times 5$, i.e., with $B = 5$.

6.1. Ability of the test to keep its nominal level

The multiplier method for the computation of $\mathbb{P}$-values is valid asymptotically, but it is important to study its behavior under small and moderate sample sizes. To this end, random samples have been drawn under the null hypothesis, i.e., from copula models of the form $C_{\beta,D}(u,v) = u^{\beta}D(u^{1-\beta}, v)$. This task is easily done upon noting that $C_{\beta,D}$ is the joint distribution of $\max(U_1^{1/\beta}, U_2^{1/(1-\beta)})$ and $V_2$, where $U_1 \sim U(0,1)$ and $(U_2, V_2) \sim D$ are independent.

The chosen models for $D$ are the Clayton copula described in (12), as well as the Gumbel–Hougaard extreme-value copula $D_{GH}(u,v) = \exp\{-\left(|\log u|^{1/1-\theta} + |\log v|^{1/(1-\theta)}\right)^{1-\theta}\}$.

The asymmetric models arising from the construction $C_{\beta,D}$ are referred to the Clayton–Khoudraji and Gumbel–Khoudraji copulas in the sequel. Note that the Gumbel–Khoudraji copula is a special case of the logistic model described by Tawn (1988). See the top and bottom panels of Figure 2 for the scatter plot of random pairs drawn from these asymmetric models.

FIGURE 2 ABOUT HERE

Three values of the asymmetry parameter $\beta_0$ were considered, namely $\beta_0 \in \{.20, .35, .50\}$. The symmetric copula $D$ has been parameterized in terms of the value of Kendall’s tau, i.e., $\tau(D) = 4 \int_{[0,1]^2} D(u,v) dD(u,v) - 1$. The values considered are $\tau(D) \in \{.50, .75\}$. The results can be found in Table 2. Generally, the tests are good at keeping their size, considering the computationally intensive minimum-distance nature of the test statistics and their associated bootstrap versions. However, the tests are too conservative when $\beta_0 = 0.5$. The test based on $T_{KS}^n$ is too liberal when $D$ is the Clayton copula with $\tau(D) = .5$ and $\beta_0 = 0.2$. 
TABLE 2 ABOUT HERE

6.2. Power against asymmetric alternatives

It is important to assess that the methods have a good ability to discard models that are not of a Khoudraji-type. To this end, some families of asymmetric models that are not of the form (2) have been considered. The first class consists of the Liouville copulas proposed and investigated by McNeil & Nešlehová (2010). A Liouville copula arises as the survival copula of a random pair

\[(X, Y) \sim R \left( \frac{G_1}{G_1 + G_2}, \frac{G_2}{G_1 + G_2} \right),\]

where \(G_1, G_2\) are independent Gamma random variables with respective parameters \(\xi_1, \xi_2\) and \(R\) is a random variable whose cdf \(F_R\) satisfies \(F_R(0) = 0\). When \(\xi_1 = \xi_2\), the copula is symmetric and belongs to the Archimedean family. In the simulation study, \(R\) follows a Gamma(5) or an inverse Gamma(5) distribution; the corresponding models are referred to the Gamma–Liouville and Inverse–Gamma–Liouville copulas, respectively. The scatter plots of random pairs from these models are to be found in Figure 3.

FIGURE 3 ABOUT HERE

The results of a power investigation involving these two copulas are to be found in Table 3. The main features are listed next.

(i) The tests are generally good at detecting departures from \(H_0\), except when \((\xi_1, \xi_2) = (1/2, 1)\), which correspond to a model that is hard to distinguish from symmetry;

(ii) The power generally increases as \(n\) increases, as a consequence of the consistency of the tests;

(iii) Generally speaking, \(T_n^{CV}\) is slightly more powerful than \(T_n^{KS}\).

Note that an apparent irregularity occurs when the data come from the Inverse–Gamma–Liouville copula with \((\xi_1, \xi_2) = (1, 1/3)\). Indeed, one can see that the
power is slightly lower when $n = 800$ than when $n = 400$. This behavior can be explained by the discretization used to compute the infimum over $\beta \in (0, 1)$ in the definition of the minimum-distance statistics and the fact that $\beta$ that minimizes $\tilde{Q}_{\beta,C}$ is hard to distinguish in that case.

**TABLE 3 ABOUT HERE**

A second class of alternatives to $H_0$ will be based on a construction of asymmetric copulas. To this end, let $(U, V)$ be a random pair from a symmetric copula $D$ and define, for $\delta_1, \delta_2 \in [0, 1]$, $X = |U - \delta_1|$ and $Y = |V - \delta_2|$. In general, when $\delta_1 \neq \delta_2$, the copula of $(X, Y)$ will be asymmetric. See Figure 3 for the scatter plots of random pairs from this model when $D$ is the Clayton or the Gumbel–Hougaard copula. Note in passing that $\delta_1 = \delta_2 = 1$ yields the (symmetric) survival copula associated to $D$.

For the results presented in Table 4, $D$ is either the Clayton or the Gumbel–Hougaard copula; the values taken by Kendall’s measure of association are $\tau(D) \in \{1/2, 3/4\}$. The asymmetry parameters have been set to $(\delta_1, \delta_2) \in \{(0.4, 0.6), (0.6, 0.4)\}$. Here, both tests are very powerful. However, the Cramér–von Mises statistics is generally more powerful than the Kolmogorov–Smirnov. The observed powers are larger when $D$ is the Clayton copula, a consequence of the fact that the departures from $H_0$ are more pronounced in that case than under the construction using the Gumbel–Hougaard copula.

**TABLE 4 ABOUT HERE**

Another possibility for models under $H_1$ are those of the form given in equation (1) when $C_1$ and $C_2$ are symmetric copulas and $C_1 \neq \Pi$. However, based on several investigations not presented here, the tests hardly detect departure from $H_0$ unless the sample size is very large, say $n = 3000$. This behavior of the test statistics can easily be explained by the fact that models of the form (1) are indeed very close to $C_{\beta,D}$. It somewhat gives an argument in favor of using the simpler models (2) for asymmetric copula modeling.
7. Illustrations

7.1. Walker Lake data

The statistical tools described in the previous sections will now be illustrated on a sub-sample of the Walker Lake data whose scatter plot of the standardized ranks has already been presented on the upper left panel of Figure 5. This data set was described by Isaaks & Srivastava (1989) and is classical in geostatistics. The observations come from the Walker Lake area in Nevada, USA. As mentioned in the Introduction, the meaning of the two variables is not revealed for pedagogical reasons. In the following analysis, a random sub-sample of size \( n = 1150 \) of the complete data set consisting of those 47,388 pairs for which the third variable (an indicator function) equals 1 is considered. The scatter plot of the pairs \((U_{1,n}, V_{1,n}), \ldots, (U_{n,n}, V_{n,n})\) of standardized ranks involved in the computation of \( C_n \) is to be found on the top right panel of Figure 5.

The first step before entering into more sophisticated investigations is to test for the independence between the two random variables. While the departure from independence is quite obvious from the scatter plot, it is also confirmed by a simple test of independence based on the estimation \( \tau_n \) of Kendall’s tau. Here, \( \sqrt{n} \tau_n = \sqrt{1150} \times .5897 = 20.00 \), so that the null hypothesis \( H_0 : \tau = 0 \) is clearly rejected in favor of \( H_1 : \tau > 0 \) based on the well-known result that \( \sqrt{n} \tau_n \sim N(0, 4/9) \) under independence (see Lee, 1990, for instance).

Even if symmetric dependence structures are special cases of the general Khoudraji-type copulas in equation (2), it may be advisable to specifically test for symmetry. Here, the asymmetry is rather clear from the scatter plot of the standardized ranks. This conclusion is confirmed by performing the tests based on \( S_n^{CvM}(\beta_0) \) and \( S_n^{KS}(\beta_0) \) when \( \beta_0 = 0 \). In that case, one has \( S_n^{CvM}(\beta_0) = 0.3536 \) (\( \widehat{PV} < .001 \)) and \( S_n^{KS}(\beta_0) = 1.2680 \) (\( \widehat{PV} < .001 \)), clearly indicating a rejection.
of the hypothesis of a symmetric copula.

Now everything is in place for testing the general Khoudraji-type copula structure of the form \( C_{\beta,D} \equiv C_{\beta} \). To this end, the test based on the minimum-distance statistics have been performed. One computes \( T_n^{\text{CvM}} = .0209 \) (\( P_{\text{V}} = .246 \)) and \( T_n^{\text{KS}} = .4304 \) (\( P_{\text{V}} = .505 \)); the \( P \)-values were estimated from \( H = 1,000 \) multiplier bootstrap samples and a grid of \([0,1]^2\) of size \( 20 \times 20 \), i.e., \( B = 20 \). Thus, one concludes that the underlying copula \( C \) of the population can reasonably be taken as belonging to the family \( \mathcal{K}' \) of asymmetric models. In other words, \( C(u, v) = u^\beta D(u^{1-\beta}, v) \). The curves for \( S_n^{\text{CvM}}(\beta)/(1 - \beta) \) and \( S_n^{\text{KS}}(\beta)/(1 - \beta) \) are to be found in the bottom panels of Figure 5.

In order to determine the form of \( D \) that best fits the data, consider as possible candidates the symmetric one-parameter copula families of Clayton, Frank, Gumbel–Hougaard, Plackett, Normal and Student with \( \nu \in \{3, 5, 7, 9\} \). Details on these models can be found in the monographs by Nelsen (2006) and Salvadori et al. (2007). From the previous analysis, \( \hat{\beta}_n^{\text{CvM}} = .235 \) and \( \hat{\beta}_n^{\text{KS}} = .273 \). In the following analysis, one assumes that the value of the asymmetry parameter is \( \beta \approx .24 \). The results of the parameter estimation and goodness-of-fit testing are to be found in Table 5.

One can see that the model with the highest \( P \)-value is Clayton; other models that were not rejected at the 10% level are the Frank and \( T_3 \) copulas, but only by a small amount. A reasonable model is then \( C_{\beta,D} \) with \( \beta = .24 \) and \( D \) is Clayton’s copula with \( \theta = 7.52 \). For completeness, the goodness-of-fit test based on \( V_n^\beta \) when \( \beta = 0 \), i.e., when one assumes symmetry, was also performed. All models were clearly rejected, showing the inadequacy of trying to fit a symmetric model to these data.

**TABLE 5 ABOUT HERE**
7.2. Nutrient data

Consider the pairs (Ca, Fe) and (Ca, Pr) in the nutrient data set that consists of the daily intake in calcium (Ca), iron (Fe), protein (Pr), vitamin A (vA) and vitamin C (vC) for \( n = 747 \) women; these observations come from a 1985 survey by the United States Department of Agriculture. In their statistical analysis, both Genest et al. (2012) and Quessy & Bahraoui (2013) concluded to a significant copula asymmetry.

A copula structure of the form \( C_{\beta,D} \) seems appropriate for (Ca, Fe) since \( T_{n}^{\text{CvM}} = 0.0418 \) (\( P \bar{V} = 0.280 \)) and \( T_{n}^{\text{KS}} = 0.4436 \) (\( P \bar{V} = 0.435 \)). The estimation of the asymmetry parameter yielded \( \beta_{n}^{\text{CvM}} = 0.397 \) and \( \beta_{n}^{\text{KS}} = 0.402 \). The same models as in Table 5 for the Walker Lake data have been tested in order to find an appropriate model for \( D \). All these models were not rejected at the 10% level.

For the pair (Ca, Pr), the tests clearly reject the null hypothesis since \( T_{n}^{\text{CvM}} = 0.0981 \) (\( P \bar{V} = 0.001 \)) and \( T_{n}^{\text{KS}} = 0.5841 \) (\( P \bar{V} = 0.006 \)). However, it may be for the pair (Pr, Ca) that the copula has a Khoudraji-type structure. It is indeed the case since \( T_{n}^{\text{CvM}} = 0.0142 \) (\( P \bar{V} = 0.828 \)) and \( T_{n}^{\text{KS}} = 0.2023 \) (\( P \bar{V} = 0.942 \)). The asymmetry parameter was estimated by \( \beta_{n}^{\text{CvM}} = \beta_{n}^{\text{KS}} = 0.470 \). Again, several models for \( D \) are acceptable. It just illustrates that for small values of Kendall’s tau, e.g., inferior to 0.5, models of the form \( C_{\beta,D} \) are quite similar even for different symmetric structures for \( D \). From an inferential point-of-view, choosing one of these models would be appropriate in that case.

Acknowledgements

Thanks are due to a referee whose valuable comments led to an improved version of this manuscript. This research was supported in part by an individual grant from the Natural Sciences and Engineering Research Council of Canada, and by a team grant from the Fonds de recherche Nature et technologies.
References


Segers, J. (2012). Weak convergence of empirical copula processes under non-
*Biometrika* 75, 397–415.
With applications to statistics.

**Appendix A: Proofs**

**A.1. Proof of Lemma 1**

Let \( \delta^\wedge = \min(\delta_1, \delta_2) \in [0,1) \) and note that

\[
C_{\delta, D}(u, v) = u^{1-\delta^\wedge} D\left( u^{1-\delta^\wedge}, v^{1-\delta^\wedge} \right) v^{\delta^\wedge} D^*\left( u^{1-\delta^\wedge}, v^{1-\delta^\wedge} \right),
\]

where \( D^*(u, v) = (uv)^{\delta^\wedge} D(u^{1-\delta^\wedge}, v^{1-\delta^\wedge}) \). While \( D^* \) is clearly a symmetric copula, it can also be shown that it satisfies equation (3) as well. Indeed, easy algebra enables to show that the copula \( D^*_\xi(u, v) = u^\xi D^*(u^{1-\xi}, v) \) is symmetric if and only if the equation

\[
u^\xi(1-\delta^\wedge) D\left( u^{(1-\xi)(1-\delta^\wedge)}, v^{1-\delta^\wedge} \right) = \nu^\xi(1-\delta^\wedge) D\left( u^{1-\xi}(1-\delta^\wedge), v^{1-\delta^\wedge} \right)
\]

holds for all \((u, v) \in [0,1]^2\). Letting \( \tilde{u} = u^{1-\delta^\wedge} \) and \( \tilde{v} = v^{1-\delta^\wedge} \), this can equivalently be written \( \tilde{u}^\delta D(\tilde{u}^{1-\delta^\wedge}, \tilde{v}^{1-\delta^\wedge}) = \tilde{v}^\delta D(\tilde{v}^{1-\delta^\wedge}, \tilde{u}^{1-\delta^\wedge}) \), which is true for all \((\tilde{u}, \tilde{v}) \in [0,1]^2\) if and only if \( \xi = 0 \), because \( D \) satisfies (3). Finally note that equation (4) obtains with \( \beta = (\delta^\vee - \delta^\wedge)/(1 - \delta^\wedge) \in [0,1) \), where \( \delta^\vee = \max(\delta_1, \delta_2) \).

In order to show this representation’s uniqueness, suppose there exist \( \tilde{\beta} \in (0, 1] \) and a copula \( \tilde{D}^* \) that satisfies equation (3) such that equation (4) holds. When \( \delta_1 > \delta_2 \), this would imply that

\[
u^\beta D^*(u^{1-\beta}, v) = \nu^{\tilde{\beta}} \tilde{D}^*(u^{1-\tilde{\beta}}, v) \quad \forall (u, v) \in [0,1]^2.
\]
Assuming that $\tilde{\beta} \leq \beta$, one deduces
\[ \tilde{D}^*(u, v) = u^\xi D^*(u^{1-\xi}, v) \quad \forall (u, v) \in [0, 1]^2, \]
where $\xi = (\beta - \tilde{\beta})/(1 - \tilde{\beta})$. Because $\tilde{D}^*$ must be symmetric and since $D^*$ satisfies (3), it will be true if and only if $\xi = 0$, i.e., $\beta = \tilde{\beta}$, which also implies that $D^* = \tilde{D}^*$. The case $\tilde{\beta} \geq \beta$ would give $D^* (u, v) = u^\xi \tilde{D}^*(u^{1-\xi}, v)$, bringing the same conclusion. The proof for $\delta_1 < \delta_2$ is identical.

### A.2. Proof of Proposition 1

As a special case of (4) with $\delta_1 = 0$ and $\delta_2 = \beta_0$, the bivariate function
\[ \tilde{D}(u, v) = v^{\beta_0} C(u, v^{1-\beta_0}) \]
is a copula. Since $C$ satisfies equation (6), one has
\[ \tilde{D}(u, v) = u^{\beta_0} C(v, u^{1-\beta_0}). \]

For an arbitrary $\xi \in [0, 1)$, the equality $u^\xi \tilde{D}(u^{1-\xi}, v) = v^\xi \tilde{D}(v^{1-\xi}, u)$ holds if and only if
\[ u^{\xi + \beta_0(1-\xi)} C(v, u^{(1-\beta_0)(1-\xi)}) = v^{\xi + \beta_0(1-\xi)} C(u, v^{(1-\beta_0)(1-\xi)}). \]

Letting $\alpha = \xi + \beta_0(1 - \xi)$, this can be written
\[ u^{\alpha} C(v, u^{1-\alpha}) = v^{\alpha} C(u, v^{1-\alpha}). \]

Because by assumption, $C$ satisfies equation (6) for a unique $\beta_0 \in [0, 1)$, one must have $\alpha = \beta_0$, which is true if and only if $\xi = 0$. One concludes that $\tilde{D}$ satisfies Assumption (3). Next, note that
\[ \tilde{D}(u, v) = (uv)^{\beta_0} D(u^{1-\beta_0}, v^{1-\beta_0}) \]
and suppose $u^\xi D(u^{1-\xi}, v) = v^\xi D(v, u^{1-\xi})$ for some arbitrary $\xi \in [0, 1)$. Making the change of variable $s = u^{1-\beta_0}$, $t = v^{1-\beta_0}$, this writes
\[ s^{\xi(1-\beta_0)} D \left\{ \left( s^{1-\xi} \right)^{1-\beta_0}, t^{1-\beta_0} \right\} = t^{\xi(1-\beta_0)} D \left\{ \left( t^{1-\xi} \right)^{1-\beta_0}, s^{1-\beta_0} \right\}. \]
which can further be expressed as

\[ s^\xi (s^{1-\xi} t)^{\beta_0} D \left\{ \left(s^{1-\xi}\right)^{1-\beta_0}, t^{1-\beta_0} \right\} = t^\xi (t^{1-\xi} s)^{\beta_0} D \left\{ \left(t^{1-\xi}\right)^{1-\beta_0}, s^{1-\beta_0} \right\} . \]

Equivalently, one has

\[ s^\xi \tilde{D}(s^{1-\xi}, t) = t^\xi \tilde{D}(t^{1-\xi}, s), \]

which holds if and only if \( \xi = 0 \) because \( \tilde{D} \) satisfies Assumption (3). As a consequence, one has the representation \( C = C_{\beta_0, D} \) with \( D \) that satisfies Assumption (3). This completes the proof that \( C \in K' \).

**A.3. Proof of Proposition 2**

Since \( C \in K' \) with \( \beta = \beta_0 \), equation (6) holds and one can write

\[ Q_{n, \beta_0}(u, v) = v^{\beta_0} C_n(u, v^{1-\beta_0}) - u^{\beta_0} C_n(v, u^{1-\beta_0}). \]

Upon noting that

\[ C_{10}(u, v) = \frac{\beta D(u^{1-\beta}, v)}{u^{1-\beta}} + (1 - \beta) D_{10}(u^{1-\beta}, v) \]

and \( C_{01}(u, v) = u^{\beta} D_{01}(u^{1-\beta}, v) \), it is easy to see that the assumption on \( D_{10}, D_{01} \) imply that \( C_{10}, C_{01} \) exist and are continuous respectively on \((0, 1) \times [0, 1]\) and \([0, 1] \times (0, 1)\). Proposition 3.1 in Segers (2012) then entails that

\[ \sup_{(u, v) \in [0, 1]^2} |C_n(u, v) - C(u, v)| \xrightarrow{P} 0. \]

As a consequence,

\[
\begin{align*}
\sup_{(u, v) \in [0, 1]^2} |Q_{n, \beta_0}(u, v) - Q_{\beta_0}(u, v)| & \leq \sup_{(u, v) \in [0, 1]^2} \left| v^{\beta_0} C_n(u, v^{1-\beta_0}) - u^{\beta_0} C(u, v^{1-\beta_0}) \right| \\
& \quad + \sup_{(u, v) \in [0, 1]^2} \left| u^{\beta_0} C_n(v, u^{1-\beta_0}) - u^{\beta_0} C(v, u^{1-\beta_0}) \right| \\
& \leq \sup_{(u, v) \in [0, 1]^2} |C_n(u, v) - C(u, v)| \\
& \quad + \sup_{(u, v) \in [0, 1]^2} |C_n(v, u) - C(v, u)| \\
& = 2 \sup_{(u, v) \in [0, 1]^2} |C_n(u, v) - C(u, v)|.
\end{align*}
\]
Thus
\[ \sup_{(u,v)\in[0,1]^2} |Q_{\beta_0}(u,v) - Q_{\beta_0}(u,v)| \xrightarrow{p} 0, \]
which completes the proof.

A.4. Proof of Theorem 1

The proof will proceed in three steps.

Step I. It will be shown that the minimum in the definition of
\[ T_n = \inf_{\beta \in (0,1)} M(\tilde{Q}_{\beta,C_n}) \]
is necessarily attained, asymptotically, in any arbitrarily small neighborhood of the true value \( \beta_0 \). First note that the triangular inequality entails
\[ M(\tilde{Q}_{\beta,C_n}) \geq M(\tilde{Q}_{\beta,C}) - M(\tilde{Q}_{\beta,C_n} - \tilde{Q}_{\beta,C}) \]
for any \( \beta \in (0,1) \). Hence,
\[ M(\tilde{Q}_{\beta,C_n}) - M(\tilde{Q}_{\beta_0,C_n}) \geq M(\tilde{Q}_{\beta,C}) - M(\tilde{Q}_{\beta_0,C_n}) - M(\tilde{Q}_{\beta,C_n} - \tilde{Q}_{\beta,C}). \]
Consequently, for any neighborhood \( N \) of \( \beta_0 \),
\[ \inf_{\beta \notin N} \left\{ M(\tilde{Q}_{\beta,C_n}) - M(\tilde{Q}_{\beta_0,C_n}) \right\} \geq \inf_{\beta \notin N} \left\{ M(\tilde{Q}_{\beta,C}) - M(\tilde{Q}_{\beta_0,C_n}) - M(\tilde{Q}_{\beta,C_n} - \tilde{Q}_{\beta,C}) \right\}. \]
Since \( C_n \) is uniformly consistent for \( C \), it follows that
\[ \sup_{(u,v)\in[0,1]^2} |\tilde{Q}_{\beta_0,C_n}(u,v) - \tilde{Q}_{\beta_0,C}(u,v)| \xrightarrow{p} 0 \]
and thus in probability, \( M(\tilde{Q}_{\beta_0,C_n}) \to M(\tilde{Q}_{\beta_0,C}) = 0 \). Next, note that
\[ M(\tilde{Q}_{\beta,C_n} - \tilde{Q}_{\beta,C}) = M\left(\tilde{Q}_{\beta_0,C_n}^2 \sqrt{n}\right) \xrightarrow{p} 0, \]
since \( \tilde{Q}_{n,\beta} = \tilde{Q}_{n,\beta'} / (1 - \beta) \) converges to the tight Gaussian process \( \tilde{Q}_\beta \) on \( \ell^\infty([0,1]^2) \). As a consequence,

\[
\inf_{\beta \in \mathbb{N}} \left\{ \mathcal{M} \left( \tilde{Q}_{\beta, C_n} \right) - \mathcal{M} \left( \tilde{Q}_{\beta_0, C_n} \right) \right\} \xrightarrow{\mathbb{P}} \inf_{\beta \in \mathbb{N}} \mathcal{M} \left( \tilde{Q}_{\beta, C} \right) > 0,
\]

where the strict inequality follows from Proposition 1. Hence,

\[
\lim_{n \to \infty} \mathbb{P} \left( \inf_{\beta \in \mathbb{N}} \mathcal{M} \left( \tilde{Q}_{\beta, C_n} \right) > \mathcal{M} \left( \tilde{Q}_{\beta_0, C_n} \right) \right) = \lim_{n \to \infty} \mathbb{P} \left( \inf_{\beta \in (0,1]} \mathcal{M} \left( \tilde{Q}_{\beta, C_n} \right) = \inf_{\beta \in \mathbb{N}} \mathcal{M} \left( \tilde{Q}_{\beta, C_n} \right) \right) = 1.
\]

Because \( \beta_0 \in \mathbb{N} \), this can be equivalently written as

\[
\lim_{n \to \infty} \mathbb{P} \left( \inf_{\beta \in (0,1]} \mathcal{M} \left( \tilde{Q}_{\beta, C_n} \right) = \inf_{\beta \in \mathbb{N}} \mathcal{M} \left( \tilde{Q}_{\beta, C_n} \right) \right) = 1.
\]

Step II. By assumption, \( \tilde{Q}'_{\beta, C} \) exists and is non-singular at \( \beta = \beta_0 \). This entails that for any \( \beta \in [0,1) \),

\[
\tilde{Q}_{\beta, C} = (\beta - \beta_0) \tilde{Q}'_{\beta_0, C} + R(\beta),
\]

where the remainder term is such that

\[
\mathcal{M} \{ R(\beta) \} \leq |\beta - \beta_0| \Delta (|\beta - \beta_0|) \tag{16}
\]

for some increasing function \( \Delta \) that satisfies \( \Delta(\epsilon) = o(1) \) as \( \epsilon \to 0 \). Hence,

\[
\tilde{Q}_{\beta, C_n} = \left( \tilde{Q}_{\beta, C_n} - \tilde{Q}_{\beta, C} \right) + (\beta - \beta_0) \tilde{Q}'_{\beta_0, C} + R(\beta).
\]

From the triangle inequality, it follows that

\[
\mathcal{M} \left( \tilde{Q}_{\beta, C_n} \right) \geq |\beta - \beta_0| \mathcal{M} \left( \tilde{Q}'_{\beta_0, C} \right) - \mathcal{M} \{ R(\beta) \} - \mathcal{M} \left( \tilde{Q}_{\beta, C_n} - \tilde{Q}_{\beta, C} \right)
\]

\[
= |\beta - \beta_0| \mathcal{M} \left( \tilde{Q}'_{\beta_0, C} \right) - \mathcal{M} \{ R(\beta) \} - \mathcal{M} \left( \frac{\tilde{Q}_{\beta, \beta_n}}{\sqrt{n}} \right).
\]
The non-singularity of \( \tilde{Q}_{\beta_0,C} \) entails \( M(\tilde{Q}_{\beta_0,C}) > W \) for some positive constant \( W \). If one lets \( N_1 \) be the neighborhood of \( \beta_0 \) consisting of those values of \( \beta \) such that \( \Delta(|\beta - \beta_0|) \leq W/2 \), then for all \( \beta \in N_1 \),

\[
M(\tilde{Q}_{\beta,C_n}) \geq \frac{W}{2} |\beta - \beta_0| - M(\tilde{Q}_{\beta_0,C_n}).
\]

One then has

\[
M(\tilde{Q}_{\beta,C_n}) - M(\tilde{Q}_{\beta_0,C_n}) \geq \frac{W}{2} |\beta - \beta_0| - M(\tilde{Q}_{\beta_0,C_n}) - M(\tilde{Q}_{\beta_0,C_n}) = \frac{W}{2} |\beta - \beta_0| - \frac{1}{\sqrt{n}} \{ M(\tilde{Q}_{\beta_0}) + M(\tilde{Q}_{\beta_0}) \}.
\]

Defining

\[
\Lambda_n(\beta) = \frac{2}{W} \sqrt{n} \left\{ M(\tilde{Q}_{\beta_0}) + M(\tilde{Q}_{\beta_0}) \right\},
\]

one can write

\[
M(\tilde{Q}_{\beta,C_n}) - M(\tilde{Q}_{\beta_0,C_n}) \geq \frac{W}{2} \left\{ |\beta - \beta_0| - \Lambda_n(\beta) \right\}.
\]

Note that under \( \mathbb{H}_0 \), the random variable

\[
\Lambda_n = \sqrt{n} \inf_{\beta \in (0,1)} \Lambda_n(\beta) = \frac{2}{W} \sqrt{n} \left\{ \inf_{\beta \in (0,1)} M(\tilde{Q}_{\beta_0}) + M(\tilde{Q}_{\beta_0}) \right\}
\]

converges in distribution to

\[
\Lambda = \frac{2}{W} \left\{ \inf_{\beta \in (0,1)} M(\tilde{Q}_{\beta_0}) + M(\tilde{Q}_{\beta_0}) \right\}.
\]

Then the infimum of \( M(\tilde{Q}_{\beta,C_n}) \) over \( N_1 \) is the same as its infimum on \( \tilde{N}_1 = N_1 \cap \{ \beta : |\beta - \beta_0| \leq \Lambda_n/\sqrt{n} \} \), so from the conclusion of Step I,

\[
\lim_{n \to \infty} P \left( \inf_{\beta \in (0,1)} M(\tilde{Q}_{\beta,C_n}) = \inf_{|\beta - \beta_0| \leq \frac{\Lambda_n}{\sqrt{n}}} M(\tilde{Q}_{\beta,C_n}) \right) = 1.
\]

Step III, Make the change of variable \( t = \sqrt{n}(\beta - \beta_0) \) and define a random neighborhood of \( \beta_0 \) as

\[
J_n = \left\{ t : |t| \leq \Lambda_n \text{ and } \beta_0 + \frac{t}{\sqrt{n}} \in (0,1) \right\}.
\]
Because $\tilde{Q}_{\beta_0,C}$ is non-singular, one can write for $\beta = \beta_0 + t/\sqrt{n}$ that

\[
\tilde{Q}_{\beta,C_n} = \tilde{Q}_{\beta_0,C_n} + \left(\tilde{Q}_{\beta,C} - \tilde{Q}_{\beta_0,C}\right) + \left(\tilde{Q}_{\beta,C_n} - \tilde{Q}_{\beta,C}\right) - \left(\tilde{Q}_{\beta_0,C_n} - \tilde{Q}_{\beta_0,C}\right)
\]

\[
= \frac{\tilde{Q}_{\beta_0,C_n}}{\sqrt{n}} + \frac{t\tilde{Q}_{\beta_0,C}}{\sqrt{n}} + R(\beta) + \left(\frac{\tilde{Q}_{n,\beta} - \tilde{Q}_{n,\beta_0}}{\sqrt{n}}\right).
\]

Then,

\[
\sqrt{n} \tilde{Q}_{\beta,C_n} = \tilde{Q}_{n,\beta_0} + t \tilde{Q}_{\beta_0,C} + \sqrt{n} R(\beta) + \left(\tilde{Q}_{n,\beta} - \tilde{Q}_{n,\beta_0}\right),
\]

so from the triangular inequality again,

\[
\mathcal{M}\left(\sqrt{n} \tilde{Q}_{\beta,C_n} - \tilde{Q}_{n,\beta_0} - t \tilde{Q}_{\beta_0,C}\right) \leq \sqrt{n} \mathcal{M}\{R(\beta)\} + \mathcal{M}\left(\tilde{Q}_{n,\beta} - \tilde{Q}_{n,\beta_0}\right).
\]

In view of the assumption on $R$ stated in equation (16) and of the convergence in distribution of $\Lambda_n$, one has for $t \in J_n$ that

\[
\sqrt{n} \mathcal{M}\{R(\beta)\} = \sqrt{n} \mathcal{M}\left\{R\left(\beta_0 + \frac{t}{\sqrt{n}}\right)\right\}
\]

\[
\leq |t| \Lambda \Delta \left(\frac{t}{\sqrt{n}}\right)\]

\[
\leq \Lambda_n \Delta \left(\frac{\Lambda_n}{\sqrt{n}}\right)
\]

\[
= o_P(1).
\]

Upon noting that the process $\tilde{Q}_{n,\beta}$ is continuous as a function of $\beta \in (0, 1)$, one concludes that $\tilde{Q}_{n,\beta} - \tilde{Q}_{n,\beta_0} = \tilde{Q}_{n,\beta_0 + t/\sqrt{n}} - \tilde{Q}_{n,\beta_0}$ converges uniformly to zero in probability; as a consequence, $\mathcal{M}(\tilde{Q}_{n,\beta} - \tilde{Q}_{n,\beta_0}) \to 0$ in probability. Thus,

\[
\sup_{t \in J_n} \left\{\sqrt{n} \mathcal{M}\left(\tilde{Q}_{\beta,C_n}\right) - \mathcal{M}\left(\tilde{Q}_{n,\beta_0} + t \tilde{Q}_{\beta_0,C}\right)\right\} = o_P(1).
\]

It remains to show that the minimum in the definition of $T_n^{M}$ is indeed achieved inside $J_n$. To this end, first note that from the triangle inequality,

\[
\mathcal{M}\left(\tilde{Q}_{n,\beta_0} + t \tilde{Q}_{\beta_0,C}\right) \geq |t| \mathcal{M}\left(\tilde{Q}_{\beta_0,C}\right) - \mathcal{M}\left(\tilde{Q}_{n,\beta_0}\right).
\]

Then for $t \notin J_n$, i.e., $|t| > \Lambda_n$, one has from the fact that there exists $W > 0$
such that $\mathcal{M}(\tilde{Q}'_{\beta_0,C}) > W$ and from the definition of $\Lambda_n$ in (17) that

$$
\mathcal{M}\left(\tilde{Q}_{n,\beta_0} + t \tilde{Q}'_{\beta_0,C}\right) > W \Lambda_n - \mathcal{M}\left(\tilde{Q}_{n,\beta_0}\right) \\
= 2 \inf_{\beta \in (0,1)} \mathcal{M}\left(\tilde{Q}_{n,\beta}\right) + \mathcal{M}\left(\tilde{Q}_{n,\beta_0}\right) \\
\geq \mathcal{M}\left(\tilde{Q}_{n,\beta_0}\right) \\
= \mathcal{M}\left(\tilde{Q}_{n,\beta_0} + 0 \tilde{Q}'_{\beta_0,C}\right).
$$

Thus, the infimum of $\mathcal{M}(\tilde{Q}_{n,\beta_0} + t \tilde{Q}'_{\beta_0,C})$ cannot be reached outside $J_n$. Hence,

$$
T^\mathcal{M}_n = \sqrt{n} \inf_{|\beta - \beta_0| \leq \frac{1}{\sqrt{n}}} \mathcal{M}\left(\tilde{Q}_{\beta,C_n}\right) = \inf_{t \in \mathbb{R}} \mathcal{M}\left(\tilde{Q}_{n,\beta_0} + t \tilde{Q}'_{\beta_0,C}\right) + o_P(1).
$$

For $g_1, g_2 \in \ell^\infty([0, 1]^2)$,

$$
\left|\inf_{t \in \mathbb{R}} \mathcal{M}\left(g_1 + t \tilde{Q}'_{\beta_0,C}\right) - \inf_{t \in \mathbb{R}} \mathcal{M}\left(g_2 + t \tilde{Q}'_{\beta_0,C}\right)\right| \leq \mathcal{M}(g_1 - g_2),
$$

so that the functional $\inf_{t \in \mathbb{R}} \mathcal{M}$ is continuous. An application of the continuous mapping theorem combined with Slutsky’s lemma then yield

$$
T^\mathcal{M}_n \Rightarrow T^\mathcal{M} = \inf_{t \in \mathbb{R}} \mathcal{M}\left(\tilde{Q}_{\beta_0} + t \tilde{Q}'_{\beta_0,C}\right),
$$

which completes the proof.
**Table 1**

Estimation based on 1,000 replicates of the mean-squared errors ($10^3$) of $\beta_{CvM}^n$ and $\beta_{KS}^n$ for the estimation of the asymmetry parameter $\beta$ under Khoudraji-type dependence structures $C_D$ when $D$ is either the Clayton or Gumbel-Hougaard copula.

<table>
<thead>
<tr>
<th>True copula</th>
<th>$\tau(D)$</th>
<th>$\beta_0$</th>
<th>$n = 200$</th>
<th>$\beta_{CvM}^n$</th>
<th>$\beta_{KS}^n$</th>
<th>$n = 400$</th>
<th>$\beta_{CvM}^n$</th>
<th>$\beta_{KS}^n$</th>
<th>$n = 800$</th>
<th>$\beta_{CvM}^n$</th>
<th>$\beta_{KS}^n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton</td>
<td>.50</td>
<td>.20</td>
<td>1.226</td>
<td>1.189</td>
<td>1.209</td>
<td>.35</td>
<td>1.258</td>
<td>1.237</td>
<td>.300</td>
<td>1.185</td>
<td>.300</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>1.270</td>
<td>1.167</td>
<td>1.209</td>
<td></td>
<td>1.258</td>
<td>1.237</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>1.280</td>
<td>1.184</td>
<td>1.299</td>
<td></td>
<td>1.323</td>
<td>1.193</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.20</td>
<td>1.311</td>
<td>1.123</td>
<td>1.075</td>
<td>.35</td>
<td>1.228</td>
<td>1.108</td>
<td>.987</td>
<td>1.043</td>
<td>.770</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>1.367</td>
<td>1.206</td>
<td>1.082</td>
<td></td>
<td>1.228</td>
<td>1.108</td>
<td>.987</td>
<td>1.043</td>
<td>.770</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>1.326</td>
<td>1.174</td>
<td>1.190</td>
<td></td>
<td>1.268</td>
<td>1.190</td>
<td>.1082</td>
<td>1.056</td>
<td>.770</td>
</tr>
<tr>
<td>Gumbel–Hougaard</td>
<td>.50</td>
<td>.20</td>
<td>1.268</td>
<td>1.165</td>
<td>1.380</td>
<td>.35</td>
<td>1.382</td>
<td>1.243</td>
<td>1.364</td>
<td>1.220</td>
<td>.769</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>1.358</td>
<td>1.243</td>
<td>1.364</td>
<td></td>
<td>1.382</td>
<td>1.243</td>
<td>1.364</td>
<td>1.220</td>
<td>.769</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>1.305</td>
<td>1.171</td>
<td>1.281</td>
<td></td>
<td>1.438</td>
<td>1.273</td>
<td>1.312</td>
<td>1.194</td>
<td>.769</td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.20</td>
<td>1.286</td>
<td>1.156</td>
<td>1.105</td>
<td>.35</td>
<td>1.321</td>
<td>1.221</td>
<td>1.043</td>
<td>0.913</td>
<td>.769</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>1.467</td>
<td>1.245</td>
<td>1.093</td>
<td></td>
<td>1.321</td>
<td>1.221</td>
<td>1.043</td>
<td>0.913</td>
<td>.769</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>1.416</td>
<td>1.238</td>
<td>1.966</td>
<td></td>
<td>1.328</td>
<td>1.234</td>
<td>0.966</td>
<td>0.931</td>
<td>.769</td>
</tr>
</tbody>
</table>

**Table 2**

Percentages of rejection of the null hypothesis $H_0: C \in K'$ for the test statistics $T_{CvM}^n$ and $T_{KS}^n$ as estimated from 1,000 replicates from various Khoudraji-type dependence structures $C_D$ when $D$ is either the Clayton or Gumbel-Hougaard copula.

<table>
<thead>
<tr>
<th>True copula</th>
<th>$\tau(D)$</th>
<th>$\beta_0$</th>
<th>$n = 200$</th>
<th>$T_{CvM}^n$</th>
<th>$T_{KS}^n$</th>
<th>$n = 400$</th>
<th>$T_{CvM}^n$</th>
<th>$T_{KS}^n$</th>
<th>$n = 800$</th>
<th>$T_{CvM}^n$</th>
<th>$T_{KS}^n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton</td>
<td>.50</td>
<td>.20</td>
<td>4.6</td>
<td>8.3</td>
<td>6.5</td>
<td>9.6</td>
<td>6.6</td>
<td>9.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>3.9</td>
<td>4.8</td>
<td>4.3</td>
<td>6.7</td>
<td>6.7</td>
<td>7.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>1.4</td>
<td>1.9</td>
<td>1.2</td>
<td>1.5</td>
<td>0.7</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.20</td>
<td>5.1</td>
<td>5.8</td>
<td>5.9</td>
<td>8.9</td>
<td>4.1</td>
<td>6.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>3.9</td>
<td>5.3</td>
<td>5.5</td>
<td>6.2</td>
<td>4.9</td>
<td>4.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>1.9</td>
<td>2.2</td>
<td>1.0</td>
<td>0.7</td>
<td>1.2</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gumbel–Hougaard</td>
<td>.50</td>
<td>.20</td>
<td>2.8</td>
<td>5.9</td>
<td>3.8</td>
<td>6.3</td>
<td>4.3</td>
<td>6.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>3.1</td>
<td>4.4</td>
<td>3.4</td>
<td>5.0</td>
<td>3.6</td>
<td>4.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>0.6</td>
<td>0.5</td>
<td>1.1</td>
<td>1.2</td>
<td>1.1</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.75</td>
<td>.20</td>
<td>5.6</td>
<td>5.2</td>
<td>6.0</td>
<td>5.8</td>
<td>5.1</td>
<td>5.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>.35</td>
<td>5.9</td>
<td>6.0</td>
<td>8.8</td>
<td>6.6</td>
<td>5.4</td>
<td>4.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td></td>
<td>4.2</td>
<td>2.9</td>
<td>3.8</td>
<td>2.1</td>
<td>1.7</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3
Percentages of rejection of the null hypothesis $H_0: C \in K'$ for the test statistics $T_{n}^{CvM}$ and $T_{n}^{KS}$ as estimated from 1000 replicates from Gamma–Liouville (Ga) and Inverse–Gamma–Liouville (IGa) copulas

<table>
<thead>
<tr>
<th>True copula</th>
<th>$n = 200$</th>
<th>$n = 400$</th>
<th>$n = 800$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model $(\xi_1, \xi_2)$</td>
<td>$T_{n}^{CvM}$</td>
<td>$T_{n}^{KS}$</td>
<td>$T_{n}^{CvM}$</td>
</tr>
<tr>
<td>Ga</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(1, 1/2)$</td>
<td>26.1</td>
<td>16.0</td>
<td>51.5</td>
</tr>
<tr>
<td></td>
<td>37.9</td>
<td>25.1</td>
<td>47.1</td>
</tr>
<tr>
<td>$(1, 1/3)$</td>
<td>6.8</td>
<td>4.9</td>
<td>16.2</td>
</tr>
<tr>
<td>$(1, 3/1)$</td>
<td>30.0</td>
<td>23.9</td>
<td>47.7</td>
</tr>
<tr>
<td>IGa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(1, 1/2)$</td>
<td>50.6</td>
<td>36.6</td>
<td>84.1</td>
</tr>
<tr>
<td>$(1, 1/3)$</td>
<td>52.5</td>
<td>42.2</td>
<td>65.1</td>
</tr>
<tr>
<td>$(1, 2/1)$</td>
<td>3.9</td>
<td>3.2</td>
<td>5.9</td>
</tr>
<tr>
<td>$(1, 3/1)$</td>
<td>26.1</td>
<td>27.8</td>
<td>28.4</td>
</tr>
</tbody>
</table>

Table 4
Percentages of rejection of the null hypothesis $H_0: C \in K'$ for the test statistics $T_{n}^{CvM}$ and $T_{n}^{KS}$ as estimated from 1000 replicates from $([U - \delta_1], [V - \delta_2])$, where $(U, V) \sim C$

<table>
<thead>
<tr>
<th>True copula</th>
<th>$\delta_1, \delta_2$</th>
<th>$n = 200$</th>
<th>$n = 400$</th>
<th>$n = 800$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C \tau(D)$</td>
<td>$(\xi_1, \xi_2)$</td>
<td>$T_{n}^{CvM}$</td>
<td>$T_{n}^{KS}$</td>
<td>$T_{n}^{CvM}$</td>
</tr>
<tr>
<td>CL</td>
<td>$(1/2, 3/4, 5/6)$</td>
<td>76.2</td>
<td>71.2</td>
<td>97.7</td>
</tr>
<tr>
<td></td>
<td>$(0.4, 0.6)$</td>
<td>64.7</td>
<td>57.4</td>
<td>95.1</td>
</tr>
<tr>
<td>GH</td>
<td>$(1/2, 3/4)$</td>
<td>17.0</td>
<td>15.7</td>
<td>35.1</td>
</tr>
<tr>
<td></td>
<td>$(0.4, 0.6)$</td>
<td>24.9</td>
<td>22.9</td>
<td>46.8</td>
</tr>
</tbody>
</table>

Table 5
Results of the parameter estimation and goodness-of-fit testing based on $V_{0,N}^{\beta}$, $\beta = 24$, and on $V_{n,N}^{\beta}$ for the 1150 pairs in the sub-sample of the Walker Lake data set

<table>
<thead>
<tr>
<th>$D_{\beta}$</th>
<th>$\hat{\theta}_{n}$</th>
<th>$\tau(D_{\beta})$</th>
<th>$V_{0,N}^{\beta}$</th>
<th>$P$-value</th>
<th>$V_{n,N}^{\beta}$</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clayton</td>
<td>7.524</td>
<td>.79</td>
<td>.0247</td>
<td>.572</td>
<td>.1017</td>
<td>.800</td>
</tr>
<tr>
<td>Frank</td>
<td>13.517</td>
<td>.74</td>
<td>.0468</td>
<td>.132</td>
<td>.0875</td>
<td>.000</td>
</tr>
<tr>
<td>Gumbel–Hougaard</td>
<td>0.740</td>
<td>.74</td>
<td>.0862</td>
<td>.036</td>
<td>.1599</td>
<td>.000</td>
</tr>
<tr>
<td>Plackett</td>
<td>78.170</td>
<td>.95</td>
<td>.0730</td>
<td>.026</td>
<td>.4834</td>
<td>.012</td>
</tr>
<tr>
<td>Normal</td>
<td>0.918</td>
<td>.74</td>
<td>.0561</td>
<td>.056</td>
<td>.1154</td>
<td>.000</td>
</tr>
<tr>
<td>Student $\nu = 3$</td>
<td>0.941</td>
<td>.78</td>
<td>.0512</td>
<td>.180</td>
<td>.1001</td>
<td>.004</td>
</tr>
<tr>
<td>Student $\nu = 5$</td>
<td>0.930</td>
<td>.76</td>
<td>.0783</td>
<td>.036</td>
<td>.1718</td>
<td>.000</td>
</tr>
<tr>
<td>Student $\nu = 9$</td>
<td>0.905</td>
<td>.72</td>
<td>.0660</td>
<td>.072</td>
<td>.1235</td>
<td>.000</td>
</tr>
</tbody>
</table>
Fig 1. Functions \(Q_{\beta}(0.25, 0.75)\) (solid line) and \(\tilde{Q}_{\beta}(0.25, 0.75)\) (dashed line) for the Khoudraji–Clayton copula when \(\tau = 0.5, \beta_0 = 0.2\) (top left), \(\tau = 0.75, \beta_0 = 0.2\) (top right), \(\tau = 0.5, \beta_0 = 0.5\) (bottom left) and \(\tau = 0.75, \beta_0 = 0.5\) (bottom right)
Fig 2. 5,000 realizations from Khoudraji-type copulas $C_{5,D}$: top panels: $D$ is the Clayton copula with $\tau = .5$ (left) and $\tau = .9$ (right); bottom panels: $D$ is the Gumbel–Hougaard copula with $\tau = .5$ (left) and $\tau = .9$ (right).

Fig 3. 5,000 realizations from Gamma–Liouville (left) and Inverse–Gamma–Liouville (right) copulas when $(\xi_1, \xi_2) = (1/3, 1)$ and $\theta = 5.$
Section of an asymmetric copula

Fig 4. 5000 realizations from the copula of \( (|U - 4|, |V - 6|) \) when \( (U, V) \) follows the Clayton (left) or the Gumbel-Hougaard (right) copula with a Kendall’s tau of .75.

Fig 5. Top panels: scatter plot of the standardized ranks of the full Walker Lake data set (left) and of the sub-sample of 1150 pairs (right); bottom panels: curves \( S_{n}^{CvM}(\beta) \) (left) and \( S_{n}^{KS}(\beta) \) (right) for \( \beta \in (0, .5) \).