Submitted to Bernoulli

Supplement to "Goodness-of-Fit Testing for Copulas: a Distribution-Free Approach"

SAMI UMUT CAN[1,*] JOHN H. J. EINMAHL[2] and ROGER J. A. LAEVEN[1,**]

[1]*Department of Quantitative Economics, University of Amsterdam, P.O. Box 15867, 1001 NJ Amsterdam, The Netherlands.*
E-mail: *s.u.can@uva.nl; **r.j.a.laeven@uva.nl

[2]*Department of Econometrics & OR and CentER, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.*
E-mail: j.h.j.einmahl@tilburguniversity.edu

This text serves as an appendix to the paper "Goodness-of-Fit Testing for Copulas: A Distribution-Free Approach." For context, notation and definitions, see the paper. We provide the proofs of Theorems 2.1 and 3.1.

Proof of Theorem 2.1

By Skorohod's representation theorem (see Billingsley (1999), Theorem 6.7) there is a probability space where all random elements in (5) are defined and the convergence (5) holds in probability. We work on this probability space.

Let us introduce the notation $\widehat{\mathbf{u}} = (\widehat{u}_1, \ldots, \widehat{u}_d)^\mathsf{T}$ for $\mathbf{u} \in (0,1)^d$, with

$$\widehat{u}_j = F_{\boldsymbol{\theta}_j}(Q_{\widehat{\boldsymbol{\theta}}_j}(u_j)), \quad j = 1, \ldots, d.$$

It is then easy to verify that

$$\begin{aligned}\eta_n(\mathbf{u}) &= \sqrt{n}[C_n(\widehat{\mathbf{u}}) - C(\widehat{\mathbf{u}})] + \sqrt{n}[C(\widehat{\mathbf{u}}) - C(\mathbf{u})] + \sqrt{n}[C(\mathbf{u}) - C_{\widehat{\boldsymbol{\lambda}}}(\mathbf{u})] \\ &=: \eta_{1n}(\mathbf{u}) + \eta_{2n}(\mathbf{u}) + \eta_{3n}(\mathbf{u}).\end{aligned} \tag{S.1}$$

We will show that the three components $\eta_{1n}, \eta_{2n}, \eta_{3n}$ produce the three components of the process η in the limit.

Let us first prove the auxiliary result that

$$\sup_{u_j \in [\varepsilon, \tau]} \left| \sqrt{n}(\widehat{u}_j - u_j) - \dot{F}_{\boldsymbol{\theta}_j}(Q_{\boldsymbol{\theta}_j}(u_j))^\mathsf{T} \boldsymbol{\zeta}_j \right| \xrightarrow{P} 0, \quad j = 1, \ldots, d. \tag{S.2}$$

The proof of (S.2) will be identical for each index j, so we will suppress it below for ease of notation. For any $u \in [\varepsilon, \tau]$, we can use the Mean Value Theorem to write

$$\sqrt{n}(\widehat{u} - u) = \sqrt{n}\big[F_{\boldsymbol{\theta}}(Q_{\widehat{\boldsymbol{\theta}}}(u)) - F_{\widehat{\boldsymbol{\theta}}}(Q_{\widehat{\boldsymbol{\theta}}}(u))\big] = \dot{F}_{\breve{\boldsymbol{\theta}}}(Q_{\widehat{\boldsymbol{\theta}}}(u))^\mathsf{T} \sqrt{n}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})$$

for some $\breve{\boldsymbol{\theta}} = \breve{\boldsymbol{\theta}}(u)$ that lies on the line segment connecting $\boldsymbol{\theta}$ to $\widehat{\boldsymbol{\theta}}$ in $\mathbb{R}^m$. It follows that

$$\begin{aligned} &\left|\sqrt{n}(\widehat{u} - u) - \dot{F}_{\boldsymbol{\theta}}(Q_{\boldsymbol{\theta}}(u))^\mathsf{T}\boldsymbol{\zeta}\right| \\ &\quad \le \left|\dot{F}_{\breve{\boldsymbol{\theta}}}(Q_{\widehat{\boldsymbol{\theta}}}(u))^\mathsf{T} - \dot{F}_{\widehat{\boldsymbol{\theta}}}(Q_{\widehat{\boldsymbol{\theta}}}(u))^\mathsf{T}\right| \cdot \left|\sqrt{n}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})\right| \\ &\qquad + \left|\dot{F}_{\widehat{\boldsymbol{\theta}}}(Q_{\widehat{\boldsymbol{\theta}}}(u))^\mathsf{T} - \dot{F}_{\boldsymbol{\theta}}(Q_{\boldsymbol{\theta}}(u))^\mathsf{T}\right| \cdot \left|\sqrt{n}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}})\right| \\ &\qquad + \left|\dot{F}_{\boldsymbol{\theta}}(Q_{\boldsymbol{\theta}}(u))^\mathsf{T}\right| \cdot \left|\sqrt{n}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}) - \boldsymbol{\zeta}\right|, \end{aligned} \tag{S.3}$$

where absolute values of vectors should be interpreted component-wise. Assumptions A1 and A3 ensure that each summand on the right-hand side of (S.3) is $o_P(1)$ uniformly over $u \in [\varepsilon, \tau]$, so the convergence (S.2) is established.

Now recall the decomposition (S.1) of η_n. The first component, $\eta_{1n}(\mathbf{u})$, is equal to $\alpha_n(\widehat{\mathbf{u}})$ by definition. By virtue of (S.2), the convergence (5) in probability, and the (uniform) continuity of B_C over $[0,1]^d$, we can now conclude that

$$\sup_{\mathbf{u} \in [\varepsilon,\tau]^d} |\eta_{1n}(\mathbf{u}) - B_C(\mathbf{u})| \xrightarrow{P} 0. \tag{S.4}$$

Moreover, by the Mean Value Theorem, the process η_{2n} can be rewritten for each $\mathbf{u} \in [\varepsilon, \tau]^d$ as

$$\eta_{2n}(\mathbf{u}) = \sum_{j=1}^{d} C^{(j)}(\breve{\mathbf{u}})\sqrt{n}(\widehat{u}_j - u_j),$$

for some $\breve{\mathbf{u}}$ lying on the line segment connecting $\mathbf{u}$ to $\widehat{\mathbf{u}}$. Therefore,

$$\begin{aligned} &\left|\eta_{2n}(\mathbf{u}) - \sum_{j=1}^{d} C^{(j)}(\mathbf{u})\dot{F}_{\boldsymbol{\theta}_j}(Q_{\boldsymbol{\theta}_j}(u_j))^\mathsf{T}\boldsymbol{\zeta}_j\right| \\ &\qquad \le \sum_{j=1}^{d} \left|C^{(j)}(\breve{\mathbf{u}})\sqrt{n}(\widehat{u}_j - u_j) - C^{(j)}(\mathbf{u})\dot{F}_{\boldsymbol{\theta}_j}(Q_{\boldsymbol{\theta}_j}(u_j))^\mathsf{T}\boldsymbol{\zeta}_j\right|. \end{aligned} \tag{S.5}$$

Consider the j^{th} summand on the right-hand side of (S.5). It is bounded from above by

$$\begin{aligned} &\left|C^{(j)}(\breve{\mathbf{u}})\right| \left|\sqrt{n}(\widehat{u}_j - u_j) - \dot{F}_{\boldsymbol{\theta}_j}(Q_{\boldsymbol{\theta}_j}(u_j))^\mathsf{T}\boldsymbol{\zeta}_j\right| \\ &\qquad + \left|\dot{F}_{\boldsymbol{\theta}_j}(Q_{\boldsymbol{\theta}_j}(u_j))^\mathsf{T}\boldsymbol{\zeta}_j\right| \left|C^{(j)}(\breve{\mathbf{u}}) - C^{(j)}(\mathbf{u})\right|. \end{aligned} \tag{S.6}$$

The term in the first line of (S.6) is $o_P(1)$ uniformly over $\mathbf{u} \in [\varepsilon, \tau]^d$, because of (S.2) and the fact that $0 \le C^{(j)} \le 1$ on $[0,1]^d$. The term in the second line of (S.6) is also $o_P(1)$ uniformly over $\mathbf{u} \in [\varepsilon, \tau]^d$, because its first factor is a.s. bounded for $u_j \in [\varepsilon, \tau]$ (Assumption A3), and its second factor is uniformly $o_P(1)$ by virtue of (S.2) and the (uniform) continuity of $C^{(j)}$ over $[\varepsilon/2, (\tau+1)/2]^d$ (Assumption A2). We therefore conclude that

$$\sup_{\mathbf{u} \in [\varepsilon,\tau]^d} \left|\eta_{2n}(\mathbf{u}) - \sum_{j=1}^{d} C^{(j)}(\mathbf{u})\dot{F}_{\boldsymbol{\theta}_j}(Q_{\boldsymbol{\theta}_j}(u_j))^\mathsf{T}\boldsymbol{\zeta}_j\right| \xrightarrow{P} 0. \tag{S.7}$$

Finally, another application of the Mean Value Theorem gives us

$$\eta_{3n}(\mathbf{u}) = \dot{C}_{\check{\boldsymbol{\lambda}}}(\mathbf{u})^{\mathsf{T}}\sqrt{n}(\boldsymbol{\lambda}_0 - \widehat{\boldsymbol{\lambda}}),$$

for some $\check{\boldsymbol{\lambda}} = \check{\boldsymbol{\lambda}}(\mathbf{u})$ between $\boldsymbol{\lambda}_0$ and $\widehat{\boldsymbol{\lambda}}$. But then,

$$\begin{aligned} \left|\eta_{3n}(\mathbf{u}) - \dot{C}(\mathbf{u})^{\mathsf{T}}\boldsymbol{\zeta}_0\right| &\le \left|\dot{C}_{\check{\boldsymbol{\lambda}}}(\mathbf{u})^{\mathsf{T}}\right|\left|\sqrt{n}(\boldsymbol{\lambda}_0 - \widehat{\boldsymbol{\lambda}}) - \boldsymbol{\zeta}_0\right| \\ &\quad + \left|\dot{C}_{\check{\boldsymbol{\lambda}}}(\mathbf{u})^{\mathsf{T}} - \dot{C}(\mathbf{u})^{\mathsf{T}}\right|\left|\boldsymbol{\zeta}_0\right|. \end{aligned} \tag{S.8}$$

Assumptions A1 and A2 ensure that both terms on the right-hand side of (S.8) are $o_P(1)$ uniformly over $\mathbf{u} \in [\varepsilon, \tau]^d$, so we obtain

$$\sup_{\mathbf{u}\in[\varepsilon,\tau]^d} \left|\eta_{3n}(\mathbf{u}) - \dot{C}(\mathbf{u})^{\mathsf{T}}\boldsymbol{\zeta}_0\right| \xrightarrow{P} 0. \tag{S.9}$$

Combining (S.4), (S.7) and (S.9) finishes the proof. □

Proof of Theorem 3.1

Before presenting the proof of Theorem 3.1, we first state a general transformation result, Theorem S.1 below, which is simply a d-variate version of the bivariate Theorem 3.1 in Can et al. (2015). The proof of the bivariate result applies almost verbatim to the d-variate case as well, after obvious notational changes.

Note that the process η in (8) is of the general form

$$\xi(\mathbf{u}) = V_C(\mathbf{u}) + \sum_{j=1}^{\nu} K_j(\mathbf{u})Z_j, \quad \mathbf{u} \in (0,1)^d, \tag{S.10}$$

where V_C is a C-Wiener process, ν is a positive integer, $K_1, \dots, K_\nu$ are deterministic functions mapping $(0,1)^d$ into $\mathbb{R}$ and $Z_1, \dots, Z_\nu$ are random variables.

As in Can et al. (2015), we will consider the set-indexed version of (S.10) restricted to $[\delta, \tau]^d$, for $0 < \delta < \tau < 1$:

$$\xi(B) = V_C(B) + \sum_{j=1}^{\nu} K_j(B)Z_j = V_C(B) + \mathbf{K}^{\mathsf{T}}(B)\mathbf{Z}. \tag{S.11}$$

Here, B is a Borel subset of $[\delta, \tau]^d$, V_C is a set-indexed C-Wiener process, and $K_1, \dots, K_d$ are deterministic signed measures. $\mathbf{K}$ denotes the column vector of $K_1, \dots, K_\nu$ and $\mathbf{Z}$ denotes the column vector of $Z_1, \dots, Z_\nu$.

Let us call a collection of subsets $\{A_s : 0 \le s \le 1\}$ of $[\delta, \tau]^d$ a *scanning family* over $[\delta, \tau]^d$ if the following hold:

(i) $\mathrm{Leb}(A_0) = 0$, $\mathrm{Leb}(A_1) = (\tau - \delta)^d$,
(ii) $A_s \subset A_{s'}$ if $s \le s'$,

(iii) $\mathrm{Leb}(A_{s'} \backslash A_s) \to 0$ if $s' \downarrow s$,

with Leb denoting Lebesgue measure. Note that for any $j \in \{1, \ldots, \nu\}$ and Borel subset B of $[\delta, \tau]^d$, the function $s \mapsto K_j(B \cap A_s)$ generates a signed measure on $[0, 1]$.

***Theorem* S.1.** Let ξ be a set-indexed process of the form (S.11). Suppose there are functions $k_j : [\delta, \tau]^d \to \mathbb{R}$, $1 \le j \le \nu$, that are square-integrable with respect to C and that satisfy

$$K_j(B) = \int_B k_j(\mathbf{u}) \, \mathrm{d}C(\mathbf{u}), \quad 1 \le j \le \nu,$$

for any Borel set $B \subset [\delta, \tau]^d$. Let $\{A_s : 0 \le s \le 1\}$ be a scanning family over $[\delta, \tau]^d$. Then, the process

$$W_C(B) = \xi(B) - \int_0^1 \mathbf{K}^\mathsf{T}(B \cap A_{\mathrm{d}s}) \mathbf{I}^{-1}(A_s^c) \int_{A_s^c} \mathbf{k}(\mathbf{u}) \, \mathrm{d}\xi(\mathbf{u}) \tag{S.12}$$

is a C-Wiener process, where $\mathbf{k}$ denotes the column vector consisting of $k_1, \ldots, k_\nu$, and the matrices $\mathbf{I}(A_s^c)$ are defined by

$$\mathbf{I}(A_s^c) = \int_{A_s^c} \mathbf{k}(\mathbf{u}) \mathbf{k}^\mathsf{T}(\mathbf{u}) \, \mathrm{d}C(\mathbf{u}), \qquad s \in [0, 1),$$

and are assumed to be invertible.

Proof of Theorem 3.1. We apply Theorem S.1 to the process η in (8). We let $\tau = 1 - \delta/2$, and instead of arbitrary Borel sets B, we consider hyperrectangles $[\boldsymbol{\delta}, \mathbf{u}] \subset [\delta, \tau]^d$. We also introduce the scanning family $A_s = [\delta, \tau]^{d-1} \times [\delta, (1-s)\delta + s\tau]$ for $0 \le s \le 1$. It is now a direct consequence of Theorem S.1 that

$$\widetilde{W}_C(\mathbf{u}) = \int_{[\boldsymbol{\delta}, \mathbf{u}]} \mathrm{d}\eta(\mathbf{s}) - \int_{[\boldsymbol{\delta}, \mathbf{u}]} \mathbf{k}(\mathbf{s})^\mathsf{T} \left(\mathbf{I}_\delta^{-1}(s_d) \int_{S_\delta(s_d)} \mathbf{k}(\mathbf{s}') \, \mathrm{d}\eta(\mathbf{s}') \right) \mathrm{d}C(\mathbf{s})$$

is a C-Wiener process on $[\delta, 1-\delta]^d$.

From standard theory of multivariate Gaussian processes (see, e.g., the Lemma preceding Theorem 3 in Khmaladze (1988)), we know that the normalized process

$$\begin{aligned} \widetilde{W}(\mathbf{u}) &= \int_{[\boldsymbol{\delta}, \mathbf{u}]} \frac{1}{\sqrt{c(\mathbf{s})}} \, \mathrm{d}\widetilde{W}_C(\mathbf{s}) \\ &= \int_{[\boldsymbol{\delta}, \mathbf{u}]} \frac{1}{\sqrt{c(\mathbf{s})}} \, \mathrm{d}\eta(\mathbf{s}) \\ &\qquad - \int_{[\boldsymbol{\delta}, \mathbf{u}]} \mathbf{k}(\mathbf{s})^\mathsf{T} \left(\mathbf{I}_\delta^{-1}(s_d) \int_{S_\delta(s_d)} \mathbf{k}(\mathbf{s}') \, \mathrm{d}\eta(\mathbf{s}') \right) \sqrt{c(\mathbf{s})} \, \mathrm{d}\mathbf{s} \end{aligned}$$

is a standard Wiener process on $[\delta, 1-\delta]^d$. Using the scale invariance property of the standard Wiener process, we conclude that the re-scaled process

$$\begin{aligned} W(\mathbf{u}) &= \frac{1}{(1-2\delta)^{d/2}} \widetilde{W}(\boldsymbol{\delta} + (1-2\delta)\mathbf{u}) \\ &= \frac{1}{(1-2\delta)^{d/2}} \Bigg[\int_{[\boldsymbol{\delta}, \boldsymbol{\delta}+(1-2\delta)\mathbf{u}]} \frac{1}{\sqrt{c(\mathbf{s})}} \,\mathrm{d}\eta(\mathbf{s}) \\ &\qquad - \int_{[\boldsymbol{\delta}, \boldsymbol{\delta}+(1-2\delta)\mathbf{u}]} \mathbf{k}(\mathbf{s})^\mathsf{T} \left(\mathbf{I}_\delta^{-1}(s_d) \int_{S_\delta(s_d)} \mathbf{k}(\mathbf{s}') \,\mathrm{d}\eta(\mathbf{s}') \right) \sqrt{c(\mathbf{s})} \,\mathrm{d}\mathbf{s} \Bigg] \end{aligned}$$

is indeed a standard Wiener process on $[0, 1]^d$. □

References

Billingsley, P. (1999). *Convergence of Probability Measures*, second ed. Wiley.

Can, S. U., Einmahl, J. H. J., Khmaladze, E. V. and Laeven, R. J. A. (2015). Asymptotically distribution-free goodness-of-fit testing for tail copulas. *Ann. Statist.* **43** 878–902.

Khmaladze, E. V. (1988). An innovation approach to goodness-of-fit tests in R^m. *Ann. Statist.* **16** 1503–1516.