IWQW

Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

Diskussionspapier
Discussion Papers

No. 10/2015

A multivariate linear rank test of independence based on a multiparametric copula with cubic sections

Benedikt Mangold
University of Erlangen-Nürnberg

ISSN 1867-6707

Friedrich-Alexander-Universität
IWQW
Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

A multivariate linear rank test of independence based on a multiparametric copula with cubic sections

Benedikt Mangold*

University of Erlangen-Nürnberg

September 28, 2015

This paper generalizes the locally optimal linear rank test based on copulæ from Shirahata (1974) resp. Guillén and Isabel (1998) and Genest et al. (2006) to p dimensions and introduces a new χ^2-type test for global independence (Nelsen test). The test is compared to similar nonparametric tests by means of the power under several alternatives and sample sizes. However, the actual strength of the Nelsen test is the fast examination of a test decision due to the closed form expression of the asymptotic distribution of the test statistic which is provided by this paper.

Keywords: Multivariate linear rank test, Copula, Multiparametric copula, Test of independence, Dependogramm, Nonparametric statistics, Dependence

1 Introduction

Testing for independence is an area of relatively huge interest in statistics. A natural case of appliance is testing for independence when there are two samples that each are

*benedikt.mangold@fau.de

independent and identically distributed (i.i.d.). Shirahata (1974) generalizes a nonparametric test of independence based on distribution functions to dimension p and even allows the case of several parameters of dependence. With this procedure, one can derive tests of independence that are locally optimal for a specified alternative hypothesis under certain conditions. However, one disadvantage of the test statistic proposed by Shirahata (1974) in the case of a multiparameteric probability distribution is that components of dependence can cancel each other out across the parameters. The latter can result in wrong test decisions.

This disadvantage and the lag of unique characterization of independence in a multiparametric setting vanishes if only the bivariate case is discussed. In particular, Guillén and Isabel (1998), later Genest and Verret (2005) transferred this test exclusively into the world of bivariate copulæ with one single parameter of dependence. They identify several well known linear rank test statistics to be locally optimal under certain alternatives and compared the power of these tests to other established tests in various scenarios.

This article is concerned with various generalizations. Firstly, a principle of construction for locally optimal rank tests in a multivariate setting with $p \geq 2$ is provided. This principle is applied to a generalization of the copula with cubic sections introduced by Nelsen et al. (1997), p. 83, to p dimensions. Since it possesses $q = 2^p$ dependence parameters, a further generalization is given by the extension of testing for independence in a multiparametric case.

It will be shown that the estimator for the q-dimensional parameter vector is asymptotically multivariate normal distributed. This finding allows to construct a χ^2-type test of independence (Nelsen test). Next, the power of the Nelsen test is compared to similar nonparametric tests under certain alternatives. Those are a multivariate version of Spearman's ρ introduced by Schmid and Schmidt (2007) and a test based on the average squared distance between independence copula and the empirical copula according to Deheuvels (1979) discussed in Genest and Rémillard (2004). Both tests have a distribution of their test statistic that has no closed form expression, which means that the critical values need to be simulated of high computational costs for each sample size n and dimension p.

A key advantage of the Nelsen test is that the asymptotic distribution of the test statistic is available in explicit form and therefore the critical values can be immediately determined. This distinguishes the Nelsen test from other nonparametric tests of independence and can have a crucial impact on real-time systems where calculation time is of importance.

This paper is organized as follows: section 2 initially introduces the bivariate copula as foundation of the Nelsen test and discusses its properties. Further, the copula is generalized to p dimensions and it pointed out that the copula is well-defined along with a reduced special case of it. Section 3 expands the concept of locally optimal rank tests on p dimensions and applies it componentwisely to the reduced Nelsen copula of the previous section. It is shown that the vector consisting of the univariate rank test statistics is multivariate normal distributed and a χ^2-type test of global independence is provided. Section 4 shows one possible application of the Nelsen test and compares its performance to other similar nonparametric rank tests of independence with respect to power and calculation time. This article concludes with a short summary and an outlook in section 5.

2 Copula with cubic sections

Definition 2.1 (Nelsen (2006)). *A copula C with parameter vector $\boldsymbol{\theta} \in \mathbb{R}^p$ is mapping from $[0,1]^p \mapsto [0,1]$ with the following properties:*

- *For all $u_i \in [0,1]$, $i = 1, ..., n$ holds*

$$C_{\boldsymbol{\theta}}(u_1, ..., u_{i-1}, 0, u_{i+1}, ..., u_p) = 0$$
$$C_{\boldsymbol{\theta}}(1, ..., 1, u_i, 1, ..., 1) = u_i.$$

- *$c_{\boldsymbol{\theta}}(\boldsymbol{u}) = \frac{\partial}{\partial \boldsymbol{u}} C_{\boldsymbol{\theta}}(\boldsymbol{u})$ is nonnegative for all $\boldsymbol{u} \in [0,1]^p$.*

If $C_{\boldsymbol{\theta}}$ is a copula, $\boldsymbol{\theta}$ is called admissible.

For the sake of simplicity we will initially introduce the copula that is relevant for this paper for the bivariate case. In Nelsen et al. (1997), p. 83, theorem 3.2.10, the following bivariate copula with cubic sections is introduced:

$$C_{\boldsymbol{\theta}}(u,v) = C_{(A_1,A_2,B_1,B_2)}(u,v) = uv(1 + (1-u)(1-v) \times \\ \times (u(1-v)B_2 + uvB_1 + (1-u)(1-v)A_2 + (1-u)vA_1). \tag{1}$$

A name of this copula has been absent and therefore it shall be named Nelsen copula in the following. The cubic sections imply a simple form of the density function which has quadratic sections and this ensures simple calculations and an intuitive understanding of its properties.

The Nelsen copula includes several copulæ as special case, such as the iterated FGM copula, see Kotz and Johnson (1977), the Lin copula, see Lin (1987), the copula of Kimeldorf and Sampson (1975) and the copula family of Sarmanov, see Sarmanov (1974).

The density function $c_{\boldsymbol{\theta}}(u,v)$ belonging to (1) has the following properties:

$$c_{\boldsymbol{\theta}}(0,0) = 1 + A_2,\ c_{\boldsymbol{\theta}}(0,1) = 1 - A_1,\ c_{\boldsymbol{\theta}}(1,0) = 1 - B_2,\ c_{\boldsymbol{\theta}}(1,1) = 1 + B_1,$$

i. e. the parameters describe the magnitude and the direction of deviations from the independence copula in its $2^p = 4$ vertices, since independence is characterized by

$$C_{\boldsymbol{\theta}}(u,v) = uv \leftrightarrow \boldsymbol{\theta} = (A_1, A_2, B_1, B_2)' = (0,0,0,0)' = \boldsymbol{\theta}_0$$

The Nelsen copula has a strong and weak tail index coefficient of 0 due to its polynomial structure. Kendall's τ takes the value

$$\tau = \frac{A_2\ (B_1 + 25)}{450} + \frac{B_1}{18} - \frac{B_2}{18} - \frac{A_1\ (B_2 + 25)}{450} \in [-0.3, 0.4]$$

which implies that rather weak dependencies can be modeled with the Nelsen copula.

Lemma 2.1. (1) *is admissible, if* (A_2, A_1), (B_1, B_2), (B_1, A_1) *and* (A_2, B_2) *are element of*

$$S := \{[-1,2] \times [-2,1]\} \cap \left\{x, y \in \mathbb{R} | x^2 - xy + y^2 - 3x - 3y \leq 0\right\}$$

Proof. Cf. Nelsen et al. (1997). □

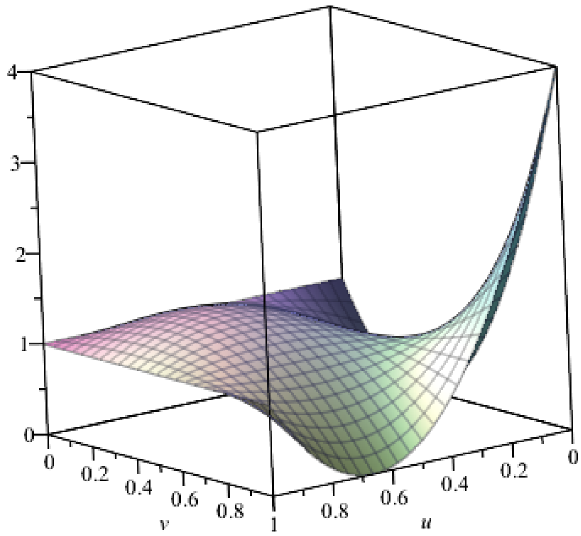

Figure 1: Density function of the Nelsen copula (1) with parameter $\boldsymbol{\theta} = (-3, 0, 0, 0)$

Figure 2 gives a graphical intuition of the set S of lemma 2.1. One has to distinguish between three cases:

- Triangle with vertices $(-1, 1)$, $(1, 1)$, $(-1, -1)$: The minimum of the density function is nonnegative and is attained in a vertex.

- Inner area of the ellipse with equation $x^2 - xy + y^2 - 3x - 3y = 0$: The minimum is attained in between two vertices and is nonnegative.

- Black area: The minimum attains a negative value outside the interval $[0, 1]$. However, the density function is nonnegative on $[0, 1]^2$.

The Nelsen copula (1) shall now be extended to p dimensions – at first for a vector of parameters $\boldsymbol{\theta} \in \mathbb{R}^q$; later only a reduced special case $q = 1$ is being discussed.

We use the following notation in order to uniquely identify vertices, even in higher dimensions:

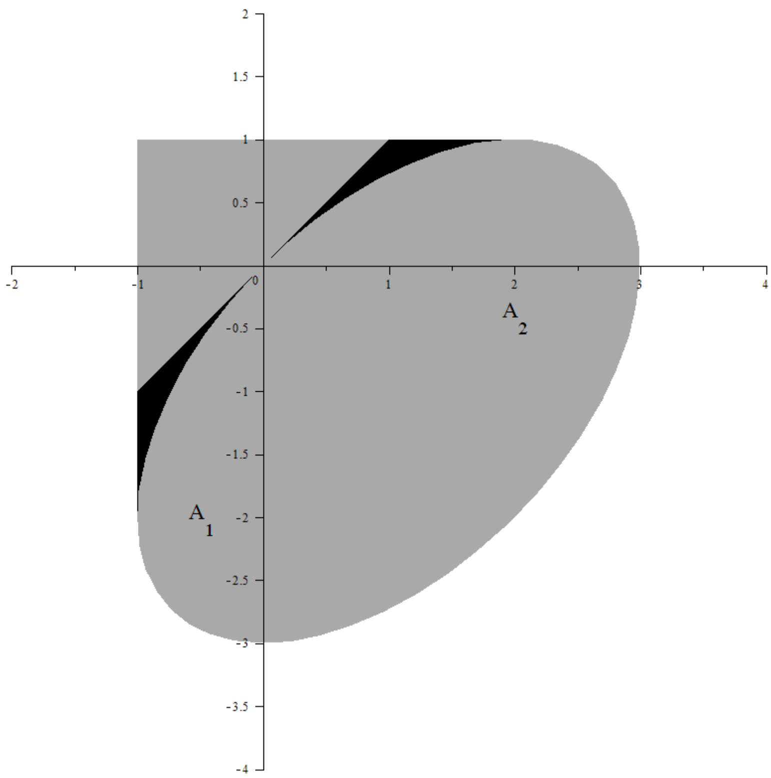

Figure 2: Area S which ensures the bivariate Nelsen copula to be well-defined using the example of (A_2, A_1).

Definition 2.2. *The labeling of a vertex* $\boldsymbol{v}$ *of the cuboid* $[0,1]^p$ *is given by the coordinates as a binary number plus 1. Thereby there are the vertices* $\boldsymbol{v}_i$*,* $i = 1, ..., 2^p$*. A vertex* $\boldsymbol{v} = (u_1, ..., u_p)$*,* $u_i \in \{0,1\}$*,* $i = 1, ..., p$ *is called even, if the amount of* $u_i = 1$ *is even or zero, odd otherwise.*

Example 2.1. *Let* $p = 4$*. The vertex* $(1,0,1,0)'$ *is the eleventh vertex* $\boldsymbol{v}_{10+1=11}$*, since* $1010_2 = 10_{10}$*.* $\boldsymbol{v}_{11}$ *is even, since the value* 1 *appears two times.*

In order to keep the notation simple, we require the following simplifying definition:

Definition 2.3. *Let* $\boldsymbol{u} = (u_1, ..., u_p)'$*. The function*

$$\operatorname{perm}(\boldsymbol{u}) = \left(\begin{Bmatrix} u_1 \\ 1 - u_1 \end{Bmatrix} \cdots \begin{Bmatrix} u_p \\ 1 - u_p \end{Bmatrix} \right)$$

maps to a vector of dimension 2^p, whose components are the products of all possible combinations of the elements in braces.

Example 2.2. *For $\boldsymbol{u} \in [0,1]^2$ we have*

$$\text{perm}(\boldsymbol{u}) = \left(\begin{Bmatrix} u_1 \\ 1-u_1 \end{Bmatrix} \begin{Bmatrix} u_2 \\ 1-u_2 \end{Bmatrix} \right) = \begin{pmatrix} u_1 u_2 \\ u_1(1-u_2) \\ (1-u_1)u_2 \\ (1-u_1)(1-u_2) \end{pmatrix}$$

A natural extension of the concept of the Nelsen copula to p dimensions is given by the following definition:

Definition 2.4 (Extension of Nelsen copula to p dimensions)**.** *For a vector $\boldsymbol{u} = (u_1, ..., u_p)' \in [0,1]^p$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)'$, and $q = 2^p$ we have*

$$C_{\boldsymbol{\theta}}(\boldsymbol{u}) = \prod_{i=1}^{p} u_i \left(1 + \langle \boldsymbol{\theta}, \text{perm}(\boldsymbol{u}) \rangle \prod_{i=1}^{p} (1-u_i) \right) \tag{2}$$

with $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{p} x_i y_i$ for $\boldsymbol{x} = (x_1, ..., x_p)'$ respective $\boldsymbol{y} = (y_1, ..., y_p)'$ as a natural extension of the Nelsen copula to p dimensions.

An important property of the Nelsen copula is given by the following lemma:

Lemma 2.2. *For the density function $c_{\boldsymbol{\theta}}(\boldsymbol{u})$, $\boldsymbol{u} = (u_1, ..., u_p)'$, $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)'$, $q = 2^p$ it holds, that*

$$c_{\boldsymbol{\theta}}(\boldsymbol{v}_i) = 1 + a(\boldsymbol{v}_i)\theta_i, \quad i = 1, ..., 2^p,$$

with

$$a(\boldsymbol{v}_i) = \begin{cases} 1, & \text{if } \boldsymbol{v}_i \text{ is an even vertex} \\ -1, & \text{if } \boldsymbol{v}_i \text{ is an odd vertex} \end{cases}.$$

The parameter θ_i is called associated with the vertex $\boldsymbol{v}_i$.

Even though the Nelsen copula has several dependence parameters, the characterization of independence is unique. The latter is shown by the following proposition:

Proposition 2.1. *The Nelsen copula is equal to the independence copula iff* $\boldsymbol{\theta} = \mathbf{0}$.

Proof. Inserting $\boldsymbol{\theta} = \mathbf{0}$ directly in the definition 2 results in

$$C_{\mathbf{0}}(\boldsymbol{u}) = \prod_{i=1}^{p} u_i = \Pi.$$

Let w.l.o.g. be θ_1 from $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)'$ unequal to 0. Due to lemma 2.2 the value that the density function attains in vertex $\boldsymbol{v}_1$ is unequal to 1. Hence, $C_{\boldsymbol{\theta}}(\boldsymbol{u})$ is not the independence copula. □

Corollary 2.1. *The copula* $C_{\boldsymbol{\theta}}(\boldsymbol{u})$ *is well-defined for* $p \geq 2$*, if the pair of parameters* (θ', θ'') *associated to the vertices of every two dimensional edge of the cuboid* $[0,1]^d$ *lie in the set*

$$S := [-1;2] \times [-2,1] \cup \left\{ \theta''^2 - \theta' \, \theta'' + 3\, \theta'' + \theta'^2 - 3\, \theta' \leq 0 \right\} \tag{3}$$

Proof. $C(u_1, ..., u_{i-1}, 0, u_{i+1}, ..., u_p) = 0$ is true since if any $u_i = 0$ the first product in (2) and hence the whole expression gets 0.
$C(1, ..., 1, u_i, 1, ..., 1) = u_i$ is true since the second product is zero for at least one occurrence of 1 and therefore the whole parenthesis in (2).
Now it is shown that the density function does only attain nonnegative values. The minimum of the density function is attained at a vertex or an edge. If the restriction

$$1 + a(\boldsymbol{v}_i)\theta_i \geq 0. \tag{4}$$

holds for all parameters θ_i, $i = 1, ..., 2^p$, all vertices always have nonnegative function values.

Let $\boldsymbol{e}_i$ be an arbitrary, twodimensional edge of $[0,1]^p$ with the respective vertices $\boldsymbol{v}'$ and $\boldsymbol{v}''$, w.l.o.g. $\boldsymbol{e}_i = (0, ..., 0, u, 1, ..., 1)$ with $\boldsymbol{v}' = (0, ..., 0, 0, 1, ..., 1)$ and $\boldsymbol{v}'' = (0, ..., 0, 1, 1, ..., 1)$ resp. the associated parameters θ' and θ''. Then we have

$$c_{\boldsymbol{\theta}}(\boldsymbol{e}_i) = (3\theta' - 3\theta'')^2 u^2 + (2\theta'' - 4\theta')u + \theta' + 1. \tag{5}$$

For $\theta' > \theta''$ the minimum of (5) is greater than 0 if

$$-\frac{1}{3}\frac{\theta'^2 - \theta'\,\theta'' + 3\,\theta'' + \theta''^2 - 3\,\theta'}{\theta' - \theta''} \geq 0.$$

The latter expression implies, that the pair (θ', θ'') needs to satisfy the ellipsoid inequality

$$\theta'^2 - \theta'\,\theta'' + 3\,\theta'' + \theta''^2 - 3\,\theta' \leq 0.$$

For $0 < \theta' \leq \theta''$ and $\theta' \geq -1$ resp. $\theta'' \leq 1$ the density function $c_{\boldsymbol{\theta}}(\boldsymbol{k_i})$ is concave and the minimum is either attained at $\boldsymbol{v}'$ or at $\boldsymbol{v}''$ and is in either case nonnegative.

In the remaining black area in figure 2 one has $\theta' > \theta''$ and therefore $c_{\boldsymbol{\theta}}(\boldsymbol{k_i})$ is a convex parabola. Although the minimum has in this case a negative value, it is attained outside of the interval $[0, 1]$, and the nonnegativity of (5) at both vertices implies the nonnegativity on $[0, 1]$. □

The linear locally optimal rang tests for copulæ from Guillén and Isabel (1998) and Genest and Verret (2005) have been developed for $\theta \in \mathbb{R}$. Conversely, the Nelsen Copula is equipped with 2^p parameters. In order to transfer the required concept of positive quadrant dependency to dimensions $p > 2$, firstly an auxiliary copula is defined:

Definition 2.5. *Let $C_{\boldsymbol{\theta}_{[k]}}$ be the generalized Nelsen copula of dimension p from definition 2.4 with*

$$\boldsymbol{\theta}_{[k]} = (\underbrace{0, ..., 0}_{k-1}, \theta_k, \underbrace{0, ..., 0}_{q-k}).$$

$C_{\boldsymbol{\theta}_{[k]}}$ has only one remaining parameter and is called k-reduced Nelsen copula of dimension p.

The rules of admissibility for the k-reduced Nelsen copula are even simpler than in the general case:

Lemma 2.3. *For $k = 1, ..., p$, the k-reduced Nelsen copula is well-defined for $\theta_k \in [-1, 3]$, if $\boldsymbol{v}_k$ is an even vertex and $\theta_k \in [-3, 1]$ if $\boldsymbol{v}_k$ is an odd vertex.*

Proof. Follows from corollary 2.1 when one of the parameters is taking the value 0. □

Clearly, the k-reduced Nelsen copula itself will not find application in practical work since it can only consider dependencies in the single vertex k. However, a joint analysis of all vertices can be used to determine a deviation from independence.

Therefore, in the following section the concept of locally optimal rank tests with one parameter of dependence is extended to p dimensions.

3 Multivariate locally optimal linear rank tests

Sklar's theorem (Sklar (1959)) states that a copula describes the functional relation between continuous marginal distributions and the joint distribution of several random variables in a unique manner. For $C(u_1, ..., u_p) = \Pi(u_1, ..., u_p) = \prod_{i=1}^{p} u_i$, the joint distribution is equal to the product of the margin distributions – the respective random variables are stochastically independent. Hence, a simple idea for a test of independence is to examine whether the empirical dependency structure of a sample of size n and dimension p corresponds to an independence copula of dimension p. In this section, the tests based on bivariate copulæ, developed by Guillén and Isabel (1998) and Genest and Verret (2005), are generalized to p dimensions. An important requirement is the generalization of the concept of positive quadrant dependency to higher dimensions:

Definition 3.1 (Positive orthant dependency of Joe (1997))**.** *A distribution function $H(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, ..., x_p)' \in \mathbb{R}^p$ with margins $X_i \sim F_i(x_i)$ is called positive lower orthant dependend (PLOD), if*

$$H(\boldsymbol{x}) - \prod_{i=1}^{p} F_i(x_i) \geq 0 \quad \textit{for all } x_i \in \mathbb{R}, \quad i = 1, ..., p.$$

A copula is called PLOD, if $C \geq \Pi$, where Π denotes the independence copula. If C is parametrized by $\theta \in \mathbb{R}$, $\theta' > \theta \Rightarrow C_{\theta'}(\boldsymbol{u}) \geq C_\theta(\boldsymbol{u})$ for $u \in [0,1]^p$ can be derived by PLOD.

To keep the notation simple, another simplifying definition is required:

Definition 3.2. *Let $\boldsymbol{u} = (u_1, ..., u_p)'$ and $\boldsymbol{c} = (c_1, ..., c_p)'$. Then it is, that*

$$\text{pow}(\boldsymbol{u}, \boldsymbol{c}) = (u_1^{c_1}, ..., u_p^{c_p})' \quad \textit{and} \quad \text{pow}_\pi(\boldsymbol{u}, \boldsymbol{c}) = \prod_{i=1}^{p} u_i^{c_i}.$$

The following lemma ensures the PLOD property for the k-reduced Nelsen copula:

Lemma 3.1. *The k-reduced Nelsen copula from Definition 2.5 is PLOD.*

Proof. Let $\theta'_k > \theta_k$. Let $\boldsymbol{v}_k$ be the vertex associated to θ'_k resp. θ_k. Then we have

$$C_{\boldsymbol{\theta}'_{[k]}}(u_1, ..., u_p) - C_{\boldsymbol{\theta}_{[k]}}(u_1, ..., u_p) =$$

$$\underbrace{(\theta'_k - \theta_k)}_{>0} \times \underbrace{\text{pow}_\pi(\boldsymbol{u}, \boldsymbol{v}_k)}_{\geq 0} \times \underbrace{\text{pow}_\pi(1 - \boldsymbol{u}, 1 - \boldsymbol{v}_k)}_{\geq 0} \times \underbrace{\prod_{i=1}^{p} u_i(1 - u_i)}_{\geq 0} \geq 0, \tag{6}$$

where $1 - \boldsymbol{x} = (1 - x_1, ..., 1 - x_p)$. □

Now the concept of locally optimal rang tests from Guillén and Isabel (1998) and Genest and Verret (2005) shall be generalized to p dimensions. Note, that until know independence is still characterized by one single parameter.

Definition 3.3 (LoR test)**.** *Let $\mathcal{M}_\alpha$ be the set of all rank tests with level α. Let $\boldsymbol{X}_1, ..., \boldsymbol{X}_n$ be i.i.d. continuous random variables of dimension p. A test T_{opt} for the hypotheses*

$$H_0 : \theta = \theta_0 = 0 \textit{ vs. } H_1 : \theta > \theta_0 \tag{7}$$

is called locally optimal rank test (LORT)

$$\forall T \in \mathcal{M}_\alpha \,\exists\, \epsilon > 0 \,\forall\, 0 < \theta < \epsilon : 1 - \beta_{T_{\text{opt}}}(\theta) > 1 - \beta_T(\theta)$$

The requirements for proposition 3.1 are:

- A1 The parameter space Θ is a closed interval and there exists a $\theta_0 \in \Theta$, such that $C_{\theta_0}(\boldsymbol{u}) = \Pi(\boldsymbol{u}) = \prod_{i=1}^{p} u_i$.
- A2 The family C_θ is PLOD.
- A3 For all $\theta \in \Theta$, C_θ and the respective density function $c_\theta(\boldsymbol{u})$ are absolutely continuous in θ for all $\boldsymbol{u} \in (0, 1)^p$.

A4 $\dot{c}_\theta(u_1,...,u_p) := \frac{\partial c_\theta(u_1,...,u_p)}{\partial \theta}$ is continuous in an environment around θ_0 with respect to θ and it holds, that

$$\lim_{\theta\to\theta_0} \int_{(0,1)^p} |\dot{c}_\theta(u_1,...,u_p)| \mathrm{d}u_{i_1}\cdots \mathrm{d}u_{i_p} < \infty,$$

for all $\{i_1,...,i_p\} \in \mathcal{S}_p$, where $\mathcal{S}_p$ denotes the set of all permutations of $\{1,...,p\}$.

Proposition 3.1 (Generalization of proposition 1, Genest and Verret (2005))**.** *Let $\boldsymbol{R}_i = (R_{1i},...,R_{pi})$, $i = 1,...,n$ be the ranks associated with a sample $\boldsymbol{U}_i = (U_{1i},...,U_{pi})$, $i = 1,...,n$ where U_{ji} are i.i.d. uniform distributed random variables. Let $\boldsymbol{U}_i$ be from a population that follows a copula from the class C_θ satisfying the requirements A1 to A4. Then the following test statistic T_n^* is LORT of the level of significance α:*

$$T_n^* = \frac{1}{n}\sum_{i=1}^{n} T(R_{1i},...,R_{pi}), \tag{8}$$

with

$$T(r_1,...,r_p) = \mathbb{E}\left[\frac{\partial}{\partial\theta}\log c_\theta(B_{r_1},...,B_{r_p})\bigg|_{\theta=\theta_0}\right],$$

where B_{r_i}, $i = 1,...,p$ are independent random variables with $B_{r_i} \sim \beta(r_i, n-r_i+1)$.

Proof. Straightforward generalization to p dimensions of the proof in Guillén and Isabel (1998), Genest and Verret (2005) in the sense of Shirahata (1974) for $q = 1$. □

In order to show the asymptotic normality of (8), we need the following lemma:

Lemma 3.2. *Let $\varphi(u_1,...,u_p)$ be element of $\mathcal{L}^2([0,1]^p)$ and continuously differentiable. Then*

$$\lim_{n\to\infty} \mathbb{E}\left[\left(a_n^\varphi(R_{n11},...,R_{n1p}) - \varphi(U_{11},...,U_{1p})\right)^2\right] = 0$$

where

$$a_n^\varphi(i_1,...,i_p) = \mathbb{E}\left[\varphi(U_{11},...,U_{1p})|R_{n11} = i_1,...,R_{1p} = i_p\right].$$

Proof. Along (Hájek et al., 1999, p. 189) with the addition, that continuously differentiable functions of mensurable functions are mensurable again. □

Further requirements for proposition 3.2 are:

A5 Let $\dot{c}_{\theta_0}$ be such that for all $\boldsymbol{u} = (u_1, ..., u_p) \in (0,1)^p$ we have:

$$\int_0^1 \dot{c}_{\theta_0}(u_1, ..., u_p)\mathrm{d}u_i = 0,\ i = 1, ..., p \text{ and } \int_{[0,1]^p} \dot{c}_{\theta_0}(\boldsymbol{u})\mathrm{d}\boldsymbol{u} \geq 0.$$

A6 $\dot{c}_{\theta_0}$ can be expressed by a finite sum of squared integrable functions that are monotone in every argument and it holds:

$$\mathbb{E}\left[\dot{c}_{\theta_0}\left(\frac{R_{1i}}{n+1}, ..., \frac{R_{pi}}{n+1}\right)\dot{c}_{\theta_0}\left(\frac{R_{1j}}{n+1}, ..., \frac{R_{pj}}{n+1}\right)\right] = \mathrm{o}\left(\frac{1}{n}\right), \quad i \neq j,$$

Proposition 3.2 (Generalization of proposition 2, Genest and Verret (2005))**.** *If $\dot{c}_{\theta_0}$ satisfies the requirements A5 and A6, $\sqrt{n}T_n^*$ converges to a normal distribution with expectation 0 and variance $\sigma^2(\dot{c}_{\theta_0})$ if the H_0 hypothesis is true, where*

$$\sigma^2(\dot{c}_{\theta_0}) = \int_{(0,1)^p} |\dot{c}_{\theta_0}(\boldsymbol{u})|^2 \,\mathrm{d}\boldsymbol{u}.$$

Further T_n^* and the statistic

$$T_n = \frac{1}{n}\sum_{i=1}^{n} \dot{c}_{\theta_0}\left(\frac{R_{1i}}{n+1}, ..., \frac{R_{pi}}{n+1}\right)$$

are asymptotically equivalent.

Proof. Following (Behnen, 1971, theorem 1) using the lemma 3.2. □

Example 3.1. *We discuss the bivariate k-reduced Nelsen copula from definition 2.5 (see Figure 1) for $k = 1$:*

$$C_{A_1}(u_1, u_2) := u\,v + A_1(1-u_1)^2\, u_1\ (1-u_2)\ {u_2}^2 \tag{9}$$

for $A_1 \in [-3,1]$ and $u, v \in [0,1]$. C_{A_1} satisfies the properties A1 - A6, therefore the test based on the statistic

$$T_n = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{R_{1i}}{n+1} - 1 \right) \left(\frac{3R_{1i}}{n+1} - 1 \right) \left(2 - \frac{3R_{2i}}{n+1} \right) \frac{R_{2i}}{n+1} \tag{10}$$

is LORT for (9). $\sqrt{n}T_n$ is asymptotically normal distributed with expectation 0 and variance $\frac{4}{225}$. H_0 from (7) is rejected at a significance level α if $t_n \geq \lambda_{1-\alpha}$.

The LORT from proposition 3.1 is only locally optimal if the alternative hypothesis is the k-reduced Nelsen copula, meaning that a deviation of θ_k from zero can be detected. This finding implies the detection of a deviation of the empirical copula in the vertex $\boldsymbol{v}_k$ and the independence copula. Hence, in order to obtain a test of global independence that considers deviations in all vertices, one has to analyze all test statistics of all k reduced Nelsen copulæ simultaneously for $k = 1, ..., q$.

Proposition 3.3. *Let $\boldsymbol{B} = (B_1, ..., B_p)$ be a p dimensional random vector with $B_i \sim \beta(u_i, n - u_i + 1)$, $i = 1, ..., p$. A nonparametric, p-dimensional rank test for independence based on the copula from definition 2.5 is given by*

$$T = n\boldsymbol{T}'_{p,n} \Sigma(\dot{c}_{\boldsymbol{\theta}_0})^{-1} \boldsymbol{T}_{p,n} \overset{a}{\sim} \chi^2(q) \tag{11}$$

with

$$\boldsymbol{T}_{p,n} = \mathbb{E}\left[\frac{\partial}{\partial \boldsymbol{\theta}} \log c_{\boldsymbol{\theta}}(\boldsymbol{B}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right]$$

and

$$\Sigma(\dot{c}_{\mathbf{0}})_{i,j} = \int_{[0,1]^p} \left(\frac{\partial c_{\boldsymbol{\theta}}(\boldsymbol{u})}{\partial \theta_i} \Big|_{\boldsymbol{\theta}=\mathbf{0}} \right) \times \left(\frac{\partial c_{\boldsymbol{\theta}}(\boldsymbol{u})}{\partial \theta_j} \Big|_{\boldsymbol{\theta}=\mathbf{0}} \right) \mathrm{d}\boldsymbol{u}$$

for $i, j = 1, ..., q$.

Proof. Proposition 3.2 states that every component of $\boldsymbol{T}_{p,n}$ is asymptotically normal distributed. The general central limit theorem together with lemma 3.2 provides the joint normal distribution of the vector $\boldsymbol{T}_{p,n}$ with expectation $\boldsymbol{\mu} = \mathbf{0}$, covariance matrix $\Sigma(\dot{c}_{\mathbf{0}})$. Hence, it follows that T is χ^2 distributed with q degrees of freedom. □

Example 3.2 (Test of independence for a bivariate sample)**.** *For $p = 2$ we have under the null hypothesis $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 = \mathbf{0}$ that*

$$\boldsymbol{T}_{2,n} = \mathbb{E}\left[\frac{\partial}{\partial \boldsymbol{\theta}} \log c_{\boldsymbol{\theta}}(B_{r_1}, B_{r_2}) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \right]$$
$$= \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} \left(\frac{R_{1i}}{n+1} - 1\right)\left(\frac{3\,R_{1i}}{n+1} - 1\right)\left(2 - \frac{3\,R_{2i}}{n+1}\right)\frac{R_{2i}}{n+1} \\ \frac{1}{n}\sum_{i=1}^{n} \left(\frac{R_{1i}}{n+1} - 1\right)\left(\frac{3\,R_{1i}}{n+1} - 1\right)\left(\frac{R_{2i}}{n+1} - 1\right)\left(\frac{3\,R_{2i}}{n+1} - 1\right) \\ \frac{1}{n}\sum_{i=1}^{n} \left(\frac{3\,R_{1i}}{n+1} - 2\right)\left(\frac{3\,R_{2i}}{n+1} - 2\right)\frac{R_{1i}}{n+1}\frac{R_{2i}}{n+1} \\ \frac{1}{n}\sum_{i=1}^{n} \frac{R_{1i}}{n+1}\left(\frac{3\,R_{1i}}{n+1} - 2\right)\left(1 - \frac{3\,R_{2i}}{n+1}\right)\left(\frac{R_{2i}}{n+1} - 1\right) \end{pmatrix}$$

is normal distributed with expectation $\boldsymbol{\mu}' = (0, 0, 0, 0)'$ and

$$\Sigma(\dot{c}_{\mathbf{0}})_{i=1\ldots4, j=1\ldots4} = \begin{pmatrix} 64 & -16 & -16 & 4 \\ -16 & 64 & 4 & -16 \\ -16 & 4 & 64 & -16 \\ 4 & -16 & -16 & 64 \end{pmatrix}^{-1}.$$

Therefore, we have

$$T = n\boldsymbol{T}_n'\Sigma(\dot{c}_{\boldsymbol{\theta}_0})^{-1}\boldsymbol{T}_n \overset{a}{\sim} \chi^2(4)$$

and H_0 is rejected at a level of significance α, if $t > q_{1-\alpha;\chi^2(4)}$.

A more detailed examination of the test of proposition 3.3 shows that the statistic measures the sum of the squared deviations from zero of θ_i, $i = 1, ..., 2^p$. Lemma 2.2 states that this is exactly the deviation between the empirical dependency structure and the independence copula in the respective associated vertices $\boldsymbol{v}_i$, $i = 1, ..., 2^p$. Therefore, it is actually tested whether the joint occurrence of p combinations of extreme high or low ranks occur more or less often than in the case of independence.

The vector $\boldsymbol{t}_{p,n}$ could find a possible application to financial stock market data, where a certain pattern of joint extreme directions of the returns is desirable i. e. to identify a portfolio that has an opposing movement of its partial returns out of a pool of many assets.

The following section illustrates the theoretical findings of T in a short simulation study and compares its power to other nonparametric tests.

4 Application

In this section an implementation in the language `R` (R Core Team (2015)) is provided and compared by means of power and calculation time to several other nonparametric rank tests of independence. The main focus thereby is the implementation of an dependogram mimicking Genest and Rémillard (2004), which tests simultaneously for every possible partial dependency.

4.1 Dependogram

Definition 4.1 (Dependogram)**.** *The initial point is a sample from a p-variate distribution with w.l.o.g. uniformly distributed margins. The dependogram provides (graphical) information if every partial sample, indexed by*

$$\mathcal{P}_{>1}(\{1,...,p\}) = \{U \subseteq \{1,...,p\} : |U| > 1\}, \tag{12}$$

is stochastically independent by calculating and plotting all test statistics T and critical values of the Nelsen test applied to every partial sample. Thereby, the level of significance α is corrected via Bonferroni correction – all in all one has to perform $2^p - p - 1$ tests.

Initially, the dependogram has been introduced by Genest and Rémillard (2004) and used a test statistic that is based on a Cramer-von Mises distance between the empirical copula by Deheuvels (1979) and the independence copula. An implementation to create a dependogram is available by the command `dependogram` provided by the `R` package `copula` (Hofert et al. (2015), Yan (2007), Kojadinovic and Yan (2010) and Hofert and Mächler (2011)).

Note that the distribution of the used test statistic has no closed form expression under the null hypotheses which implies that the critical values have to be obtained via Monte-Carlo simulation for every sample size n and dimension p. Since the effort in the sense of calculation time is huge, the application of this method can be cumbersome, especially for large sample sizes and/or higher dimensions.

The implementation of the dependogram based on the Nelsen test is now introduced and its usage presented exemplarily by a dependency structure used in Genest and Rémillard (2004). Information about the required calculation time is provided in section 4.3.

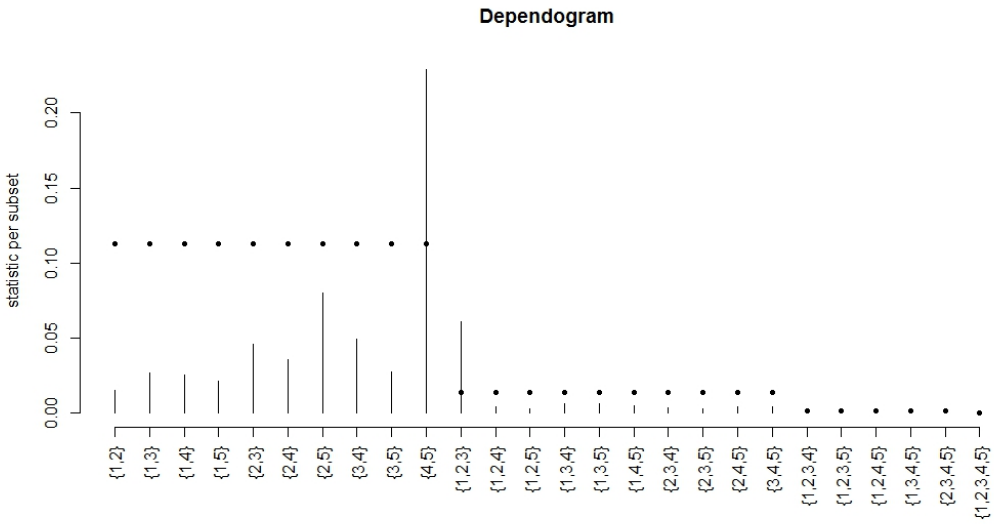

Figure 3: Dependogram based on the test statistic of Genest and Rémillard (2004) from the `R` package `copula`.

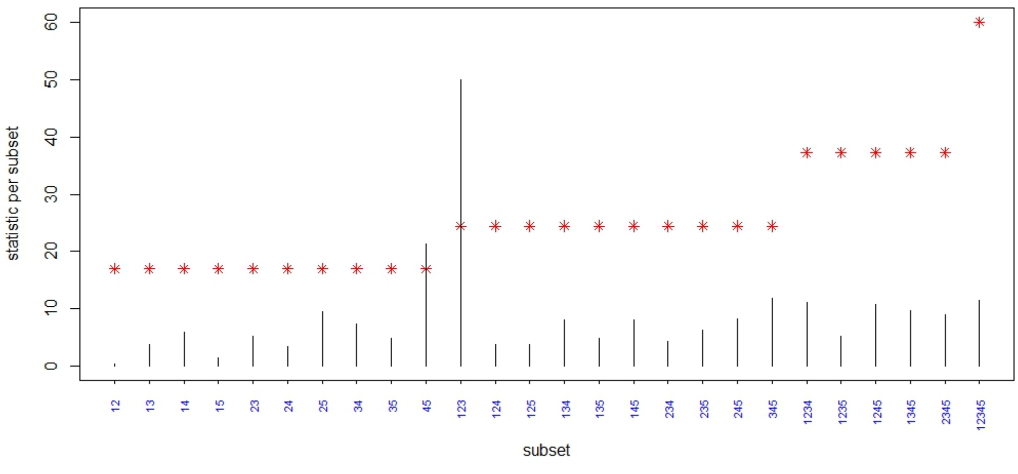

Figure 4: Dependogram based on the Nelsen test statistic T from equation (11).

Example 4.1 (Application on the example of Genest and Rémillard (2004)). *Let* $\boldsymbol{x}$ *be a* 5*-variate random sample of a multivariate random variable with an identity matrix as covariance matrix. The dependency structure is obtained by:*

```
x <- matrix(rnorm(500),100,5)
x[,1] <- abs(x[,1]) * sign(x[,2] * x[,3])
x[,5] <- x[,4]/2 + sqrt(3) * x[,5]/2
```

The figures 3 and 4 show the dependogram of the two R *functions applied to an example: Both reject the null hypothesis for the pairwise test on* $\{4,5\}$ *and the triple test of* $\{1,2,3\}$ *whereas the global hypotheses of all 5 samples would not have been rejected. Even if the global hypothesis would have been rejected, the finding of the partial dependencies allows a deeper insight into the relationship between the variables.*

4.2 Comparison of power

In a simulation study it shall be investigated how the Nelsen test performs in comparison to the test of Genest and Rémillard (2004) (ff. Genest test) and a third test, based on a multivariate version of Spearman's ρ by Schmid and Schmidt (2007) (ff. Schmid test).

Definition 4.2 (Multivariate version of Spearman's ρ by Schmid and Schmidt (2007)). *Let* $X_{i1},...,X_{in}$ *be an i.i.d. sample of a* p *variate population,* $i = 1,...,p$*. A version of the generalization of the correlation coefficient by Spearman is given by*

$$\rho_{1,p} = h(p)\left(2^p \int_{[0,1]^p} \hat{C}_n(\boldsymbol{u})\mathrm{d}\boldsymbol{u} - 1\right) =$$

$$= h(p)\left(\frac{2^p}{n}\left(\sum_{i=1}^{n}\prod_{j=1}^{p}\left(1-\frac{R_{ij}}{n}\right)\right) - 1\right)$$

whereas R_{ij} *denotes the rank of* X_{ij} *in* $X_{i1},...,X_{in}$*. The distribution of* $\rho_{1,p}$ *itself is dependent on the dimension* p *and the sample size* n *under the null hypothesis and needs to be determined via simulation.*

The power is compared in $N = 1,000$ iterations for the case $p = 5$ for samples from a Gumbel and a Joe copula for various sample sizes n. To determine the critical value a

Monte-Carlo simulation with $N = 1,000$ iterations has been carried out under the null hypothesis. This is the default in the package `copula`.

The magnitude of the dependency is held constant for both alternatives by choosing the parameter such that it results in $\tau = 0.07$. Table 1 gives an overview on the simulated values: For every partial sample $\mathcal{P}_2, \mathcal{P}_3$ and $\mathcal{P}_4$ only one representative subset-index is stated. Please note that a greater value of τ results in a higher power for all of the three tests and converges very quickly to 1. Therefore only $\tau = 0.07$ is presented in table 1.

Table 1: Power of the Nelsen, Genest and Schmid test under H_1: Gumbel/Joe copula with $\tau = 0.07$, $N = 1,000$ repetitions

	Gumbel				**Joe**			
n\subset	12	123	1234	12345	12	123	1234	12345
Genest								
50	0.01	0.01	0.01	0.05	0.31	0.13	0.62	0.85
100	0.02	0.01	0.03	0.10	0.74	0.49	0.95	0.99
200	0.05	0.02	0.06	0.21	0.98	0.95	1.00	1.00
500	0.18	0.07	0.22	0.52	1.00	1.00	1.00	1.00
Nelsen								
50	0.00	0.01	0.11	0.23	0.00	0.01	0.15	0.31
100	0.01	0.03	0.19	0.38	0.01	0.08	0.32	0.57
200	0.04	0.09	0.34	0.61	0.04	0.20	0.58	0.82
500	0.15	0.29	0.71	0.91	0.23	0.62	0.93	0.99
Schmid								
50	0.12	0.15	0.18	0.21	0.10	0.14	0.13	0.12
100	0.14	0.26	0.38	0.41	0.14	0.22	0.29	0.28
200	0.28	0.49	0.59	0.62	0.28	0.42	0.47	0.45
500	0.59	0.86	0.95	0.96	0.56	0.78	0.87	0.88

If the true distribution of the sample is the Gumbel copula, the Nelsen test has a greater power than the test of Genest for elements of $\mathcal{P}_{>2}$ with minor differences for $\mathcal{P}_2$. However, up to one exception both, the Nelsen and the Genest test, are dominated by the Schmid test for $\mathcal{P}_{\leq 4}$. Merely for $\mathcal{P}_5$ the values of the power of the Nelsen test and the Schmid test nearly coincide.

Further, if the true distribution is the Joe copula, the results are diametrically different. The Schmid test dominates the Nelsen test for elements of $\mathcal{P}_2$ and $\mathcal{P}_3$ as against the power is greater for the Nelsen test for elements of $\mathcal{P}_{>3}$. However, both tests are dominated by the Genest test that has the highest power for all subsets and sample sizes. Note, that all examined tests could hold the level of significance $\alpha = 0.05$ quite well.

Since even this simple situation cannot determine an uniquely best test, it shall be concluded that there are situations in which the Nelsen test outperforms the Schmid and the Genest test in the sense of greatest power. A broader class of alternatives can identify further alternatives under which each of the introduced tests is favorable.

Concluding, it shall be mentioned that the described effects are nearly the same for all sample sizes. A large loss in power due to the use of an asymptotic distribution instead of the unknown finite sample distribution has not been detected.

4.3 Calculation time

This section discusses the required calculation time with respect to dimension p and sample size n. Since the variances of the distributions of the test statistics of the Genest and the Schmid test contain a Brownian bridge, their quantiles cannot be calculated directly – they have to be determined via a preceded simulation for every constellation of p and n. Both, quality of the derived critical values and the required calculation time are augmenting with increasing amount of iterations (the default setting in the package `copula` is $N = 1,000$). In addition to that, it is obvious that the required calculation time increases for larger sample sizes and higher dimensions.

Although only the asymptotic distribution of the Nelsen test statistic T is known, it is provided in a closed form expression. For small sample sizes the error of approximation of the critical value can be large. However, with increasing sample size this error diminishes as table 1 suggests. Since the dimension p only has an influence on the degrees of freedom needed to determine the critical value of an univariate distribution, the dimension has virtually no influence on the required calculation time for the critical value compared to the other nonprametric tests. Nevertheless, p influences the time that is needed to make a test decision because one needs to generate a function that evaluates the vector of test statistics of length 2^p.

Table 2 gives an overview on the calculation time that is required in order to perform a test. Obviously, the time that is needed to simulate a critical value is increasing exponentially as the sample sizes or the dimension rises. Especially for large sample

Table 2: Calculation time in seconds (except as noted otherwise) that is needed to make a test decision. Genest: Time to simulate a critical value. Nelsen: Time to generate the calculation routine.

Calculation time	**Genest**					**Nelsen**
n	25	50	100	200	500	∞
p						
2	0.06	0.07	0.15	0.58	3.80	0.00
3	0.04	0.11	0.33	1.29	17.74	0.00
4	0.09	0.20	0.67	2.78	48.48	0.00
5	0.17	0.45	1.62	6.53	132.56	0.00
6	0.33	1.00	3.82	14.83	321.15	0.04
7	0.70	2.23	8.91	34.31	767.48	0.14
...	...	...	...	...	...	...
13	15 [1]	2 m[1]	20 m[1]	1 h[1]	5 d[1]	23 m

[1] Extrapolated

sizes it gets clear that the time needed to simulate the critical values for the Genest test plays an important role for actual applications. Likewise, the performance time augments for the Nelsen test with higher dimension p – hereby the bottleneck is to generate a function that evaluates the 2^p dimensional test statistic and the required covariance matrix. But once this is done, the actual calculation time is barely influenced by the sample size n. The Nelsen test rather profits of increasing sample sizes since the error of approximation is shrinking. The time provided in table 2 is the run-time needed in `R` in order to generate the routine that can be used to calculate the vector of test statistics $\boldsymbol{t}_{p,n}$. Once the function is generated the actual calculation time is negligible.

Thus, the Nelsen test can be a valid alternative in situations where the time it lasts to perform a test of independence is relevant. This finding is especially important if the sample size is varying, which would require several cumbersome simulations of the critical values of other nonparametric tests, or if the sample size is very large.

5 Summary and Outlook

In this article, a nonparametric multivariate rank test for independence has been developed. Basis for the test has been an extension of bivariate locally optimal rank tests to p dimensions. In a next step, the bivariate Nelsen copula with 4 parameters has been extended to dimension p with 2^p parameters that characterize independence in an unique manner.

Based on this generalization, a simple test of independence has been developed. Its key feature is an asymptotic distribution that has a closed form expression. Especially when testing $p > 3$ samples for independence, the introduced test appears to have an equivalent or better performance in the sense of a higher power than other nonparametric rank tests in the present setting. A broader simulation study considering further alternative hypotheses could illuminate the differences of the examined tests in a more detailed manner. If one had knowledge of the type of dependence, this would give a guideline on how to choose the nonparametric test with the greatest power.

One key advantage of the Nelsen test is the performance for large sample sizes n, since the used distribution is asymptotic and the the critical value is therefore not affected by n. The introduced alternatives need a cumbersome simulation of the critical values preceding the test. This time-consuming procedure can be omitted using the Nelsen test which makes it attractive in areas where the computation time is crucial. The loss in power which arises by the use of an asymptotic instead of a finite sample distribution has been minor in the simulation.

The decoding of the violation of independence into the single vertices (e.g. the more frequent joint occurrence of high ranks as in the case of independence) gives insight into the nature of dependence. This finding could be used e.g. in the area of financial market data – if one wants to build a portfolio of p assets from a pool of N assets, one could identify the one out of $\binom{N}{p}$ combinations that provides favorable joint movement of the returns. In this way, one could search for a combination of assets, whose returns move in the most diametrical of parallel manner. The former is of interest for diversification, the latter for an investment strategy. Both shall be the subject of future research.

References

Behnen, K. (1971). Asymptotic optimality and are of certain rank-order tests under contiguity. *The Annals of Mathematical Statistics*, 42(1):325–329.

Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. un test non paramétrique d'indépendance. *Acad. Roy. Belg. Bull. Cl. Sci.(5)*, 65(6):274–292.

Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Local efficiency of a cramér–von mises test of independence. *Journal of Multivariate Analysis*, 97(1):274–294.

Genest, C. and Rémillard, B. (2004). Test of independence and randomness based on the empirical copula process. *Test*, 13(2):335–369.

Genest, C. and Verret, F. (2005). Locally most powerful rank tests of independence for copula models. *Journal of Nonparametric Statistics*, 17(5):521–539.

Guillén, G. and Isabel, A. (1998). *Dependencia y tests de rangos para leyes bidimensionales*. Universidad de Granada, Granada.

Hájek, J., Šidák, Z., and Sen, P. K. (1999). *Theory of rank tests*. Probability and mathematical statistics. Academic Press, San Diego, Calif., 2nd ed. edition.

Hofert, M., Kojadinovic, I., Maechler, M., and Yan, J. (2015). copula: Multivariate dependence with copulas.

Hofert, M. and Mächler, M. (2011). Nested archimedean copulas meet r: The nacopula package. *Journal of Statistical Software*, 39(9):1–20.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.

Kimeldorf, G. and Sampson, A. (1975). Uniform representations of bivariate distributions. *Communications in Statistics–Theory and Methods*, 4(7):617–627.

Kojadinovic, I. and Yan, J. (2010). Modeling multivariate distributions with continuous margins using the copula r package. *Journal of Statistical Software*, 34(9):1–20.

Kotz, S. and Johnson, N. L. (1977). Propriétés de dépendance des distributions itérées, généralisées à deux variables farlie-gumbel-morgenstern. *CR Acad Sci Paris Sér I Math*, 285:277–280.

Lin, G. D. (1987). Relationships between two extensions of farlie-gumbel-morgenstern distribution. *Annals of the Institute of Statistical Mathematics*, 39(1):129–140.

Nelsen, R. B. (2006). *An introduction to copulas*. Springer series in statistics. Springer, New York, NY, 2. ed edition.

Nelsen, R. B., Quesada-Molina, J. J., and Rodriguez-Lallena, J. A. (1997). Bivariate copulas with cubic sections. *Journal of Nonparametric Statistics*, 7(3):205–220.

R Core Team (2015). R: A language and environment for statistical computing.

Sarmanov, I. O. (1974). New forms of correlation relationships between positive quantities applied in hydrology. *Mathematical Models in Hydrology*, 1.

Schmid, F. and Schmidt, R. (2007). Multivariate extensions of spearman's rho and related statistics. *Statistics & Probability Letters*, 77(4):407–416.

Shirahata, S. (1974). Locally most powerful rank tests for independence. *Bulletin of Mathematical Statistics*, 16(1):11–21.

Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8.

Yan, J. (2007). Enjoy the joy of copulas: With a package copula. *Journal of Statistical Software*, 21(4):1–21.

Diskussionspapiere 2015
Discussion Papers 2015

01/2015 **Seebauer, Michael**: Does Direct Democracy Foster Efficient Policies? An Experimental Investigation of Costly Initiatives

02/2015 **Bünnings, Christian, Schmitz, Hendrik, Tauchmann, Harald and Ziebarth, Nicolas R.**: How Health Plan Enrollees Value Prices Relative to Supplemental Benefits and Service Quality

03/2015 **Schnabel, Claus:** United, yet apart? A note on persistent labour market differences between western and eastern Germany

04/2015 **Dieckmann, Anja, Fischbacher, Urs, Grimm, Veronika, Unfried, Matthias, Utikal, Verena and Valmasoni, Lorenzo:** Trust and Beliefs among Europeans: Cross-Country Evidence on Perceptions and Behavior

05/2015 **Grimm, Veronika, Utikal, Verena and Valmasoni, Lorenzo**: In-group favoritism and discrimination among multiple out-groups

06/2015 **Reichert, Arndt R., Tauchmann, Harald and Wübker, Ansgar**: Weight Loss and Sexual Activity in Adult Obese Individuals: Establishing a Causal Link

07/2015 **Klein, Ingo, Mangold, Benedikt:** Cumulative Paired ϕ-Entropy

08/2015 **Erbe, Katharina:** Tax Planning of Married Couples in East and West Germany

09/2015 **Krauss, Christopher:** Statistical Arbitrage Pairs Trading Strategies: Review and Outlook

Diskussionspapiere 2014
Discussion Papers 2014

01/2014 **Rybizki, Lydia**: Learning cost sensitive binary classification rules accounting for uncertain and unequal misclassification costs

02/2014 **Abbiati, Lorenzo, Antinyan, Armenak and Corazzini, Lucca:** Are Taxes Beautiful? A Survey Experiment on Information, Tax Choice and Perceived Adequacy of the Tax Burden

03/2014 **Feicht, Robert, Grimm, Veronika and Seebauer, Michael**: An Experimental Study of Corporate Social Responsibility through Charitable Giving in Bertrand Markets

04/2014 **Grimm, Veronika, Martin, Alexander, Weibelzahl, Martin and Zoettl, Gregor**: Transmission and Generation Investment in Electricity Markets: The Effects of Market Splitting and Network Fee Regimes

05/2014 **Cygan-Rehm, Kamila and Riphahn, Regina**: Teenage Pregnancies and Births in Germany: Patterns and Developments

06/2014 **Martin, Alexander and Weibelzahl, Martin:** Where and when to Pray? - Optimal Mass Planning and Efficient Resource Allocation in the Church

07/2014 **Abraham, Martin, Lorek, Kerstin, Richter, Friedemann and Wrede, Matthias:** Strictness of Tax Compliance Norms: A Factorial Survey on the Acceptance of Inheritance Tax Evasion in Germany

08/2014 **Hirsch, Boris, Oberfichtner, Michael and Schnabel Claus**: The levelling effect of product market competition on gender wage discrimination

09/1014 **Mangold, Benedikt:** Plausible Prior Estimation

10/2014 **Gehrke, Britta:** Fiscal Rules and Unemployment

11/2014 **Gehrke, Britta** and **Yao, Fang**: Phillips Curve Shocks and Real Exchange Rate Fluctuations: SVAR Evidence

12/2014 **Mäder, Miriam, Müller, Steffen, Riphahn, Regina and Schwientek, Caroline**: Intergenerational transmission of unemployment - evidence for German sons

13/2014 **Casal, Sandro, Ploner, Matteo and Sproten, Alec N.**: Fostering the Best Execution Regime. An Experiment about Pecuniary Sanctions and Accountability in Fiduciary Money Management

14/2014 **Lochner, Benjamin**: Employment protection in dual labor markets - Any amplification of macroeconomic shocks?

15/2014 **Herrmann, Klaus, Teis, Stefan and Yu, Weijun**: Components of Intraday Volatility and Their Prediction at Different Sampling Frequencies with Application to DAX and BUND Futures

Diskussionspapiere 2013
Discussion Papers 2013

01/2013 **Wrede, Matthias**: Rational choice of itemized deductions

02/2013 **Wrede, Matthias**: Fair Inheritance Taxation in the Presence of Tax Planning

03/2013 **Tinkl, Fabian**: Quasi-maximum likelihood estimation in generalized polynomial autoregressive conditional heteroscedasticity models

04/2013 **Cygan-Rehm, Kamila**: Do Immigrants Follow Their Home Country's Fertility Norms?

05/2013 **Ardelean, Vlad and Pleier, Thomas**: Outliers & Predicting Time Series: A comparative study

06/2013 **Fackler, Daniel and Schnabel, Claus**: Survival of spinoffs and other startups: First evidence for the private sector in Germany, 1976-2008

07/2013 **Schild, Christopher-Johannes:** Do Female Mayors Make a Difference? Evidence from Bavaria

08/2013 **Brenzel, Hanna, Gartner, Hermann and Schnabel Claus**: Wage posting or wage bargaining? Evidence from the employers' side

09/2013 **Lechmann, Daniel S. and Schnabel Claus**: Absence from work of the self-employed: A comparison with paid employees

10/2013 **Bünnings, Ch. and Tauchmann, H.**: Who Opts Out of the Statutory Health Insurance? A Discrete Time Hazard Model for Germany

Diskussionspapiere 2012
Discussion Papers 2012

01/2012 **Wrede, Matthias**: Wages, Rents, Unemployment, and the Quality of Life

02/2012 **Schild, Christopher-Johannes:** Trust and Innovation Activity in European Regions - A Geographic Instrumental Variables Approach

03/2012 **Fischer, Matthias:** A skew and leptokurtic distribution with polynomial tails and characterizing functions in closed form

04/2012 **Wrede, Matthias**: Heterogeneous Skills and Homogeneous Land: Segmentation and Agglomeration

05/2012 **Ardelean, Vlad**: Detecting Outliers in Time Series

Diskussionspapiere 2011
Discussion Papers 2011

01/2011 **Klein, Ingo, Fischer, Matthias and Pleier, Thomas:** Weighted Power Mean Copulas: Theory and Application

02/2011 **Kiss, David:** The Impact of Peer Ability and Heterogeneity on Student Achievement: Evidence from a Natural Experiment

03/2011 **Zibrowius, Michael:** Convergence or divergence? Immigrant wage assimilation patterns in Germany

04/2011 **Klein, Ingo and Christa, Florian:** Families of Copulas closed under the Construction of Generalized Linear Means

05/2011 **Schnitzlein, Daniel**: How important is the family? Evidence from sibling correlations in permanent earnings in the US, Germany and Denmark

06/2011 **Schnitzlein, Daniel**: How important is cultural background for the level of intergenerational mobility?

07/2011 **Steffen Mueller**: Teacher Experience and the Class Size Effect - Experimental Evidence

08/2011 **Klein, Ingo:** Van Zwet Ordering for Fechner Asymmetry

09/2011 **Tinkl, Fabian and Reichert Katja:** Dynamic copula-based Markov chains at work: Theory, testing and performance in modeling daily stock returns

10/2011 **Hirsch, Boris and Schnabel, Claus:** Let's Take Bargaining Models Seriously: The Decline in Union Power in Germany, 1992 – 2009

11/2011 **Lechmann, Daniel S.J. and Schnabel**, **Claus** : Are the self-employed really jacks-of-all-trades? Testing the assumptions and implications of Lazear's theory of entrepreneurship with German data

12/2011 **Wrede, Matthias**: Unemployment, Commuting, and Search Intensity

13/2011 **Klein, Ingo**: Van Zwet Ordering and the Ferreira-Steel Family of Skewed Distributions

Diskussionspapiere 2010
Discussion Papers 2010

01/2010 **Mosthaf, Alexander, Schnabel, Claus and Stephani, Jens:** Low-wage careers: Are there dead-end firms and dead-end jobs?

02/2010 **Schlüter, Stephan and Matt Davison:** Pricing an European Gas Storage Facility using a Continuous-Time Spot Price Model with GARCH Diffusion

03/2010 **Fischer, Matthias, Gao, Yang and Herrmann, Klaus:** Volatility Models with Innovations from New Maximum Entropy Densities at Work

04/2010 **Schlüter, Stephan and Deuschle, Carola:** Using Wavelets for Time Series Forecasting – Does it Pay Off?

05/2010 **Feicht, Robert and Stummer, Wolfgang:** Complete closed-form solution to a stochastic growth model and corresponding speed of economic recovery.

06/2010 **Hirsch, Boris and Schnabel, Claus:** Women Move Differently: Job Separations and Gender.

07/2010 **Gartner, Hermann, Schank, Thorsten and Schnabel, Claus:** Wage cyclicality under different regimes of industrial relations.

08/2010 **Tinkl, Fabian**: A note on Hadamard differentiability and differentiability in quadratic mean.

Diskussionspapiere 2009
Discussion Papers 2009

01/2009 **Addison, John T. and Claus Schnabel:** Worker Directors: A German Product that Didn't Export?

02/2009 **Uhde, André and Ulrich Heimeshoff:** Consolidation in banking and financial stability in Europe: Empirical evidence

03/2009 **Gu, Yiquan and Tobias Wenzel:** Product Variety, Price Elasticity of Demand and Fixed Cost in Spatial Models

04/2009 **Schlüter, Stephan:** A Two-Factor Model for Electricity Prices with Dynamic Volatility

05/2009 **Schlüter, Stephan and Fischer, Matthias**: A Tail Quantile Approximation Formula for the Student t and the Symmetric Generalized Hyperbolic Distribution

06/2009 **Ardelean, Vlad:** The impacts of outliers on different estimators for GARCH processes: an empirical study

07/2009 **Herrmann, Klaus:** Non-Extensitivity versus Informative Moments for Financial Models - A Unifying Framework and Empirical Results

08/2009 **Herr, Annika:** Product differentiation and welfare in a mixed duopoly with regulated prices: The case of a public and a private hospital

09/2009 **Dewenter, Ralf, Haucap, Justus and Wenzel, Tobias**: Indirect Network Effects with Two Salop Circles: The Example of the Music Industry

10/2009 **Stuehmeier, Torben and Wenzel, Tobias**: Getting Beer During Commercials: Adverse Effects of Ad-Avoidance

11/2009 **Klein, Ingo, Köck, Christian and Tinkl, Fabian**: Spatial-serial dependency in multivariate GARCH models and dynamic copulas: A simulation study

12/2009 **Schlüter, Stephan**: Constructing a Quasilinear Moving Average Using the Scaling Function

13/2009 **Blien, Uwe, Dauth, Wolfgang, Schank, Thorsten and Schnabel, Claus:** The institutional context of an "empirical law": The wage curve under different regimes of collective bargaining

14/2009 **Mosthaf, Alexander, Schank, Thorsten and Schnabel, Claus**: Low-wage employment versus unemployment: Which one provides better prospects for women?

Diskussionspapiere 2008
Discussion Papers 2008

01/2008 **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008 **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008 **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008 **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008 **Wenzel, Tobias**: On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008 **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility

07/2008 **Klein, Ingo and Michael Grottke:** On J.M. Keynes' "The Principal Averages and the Laws of Error which Lead to Them" - Refinement and Generalisation