Generalized Cramér–von Mises goodness-of-fit tests for multivariate distributions

Chiu, Sung Nok; Liu, Kwong Ip

Published in:
Computational Statistics and Data Analysis

DOI:
10.1016/j.csda.2009.04.004

Published: 01/09/2009

Document Version:
Peer reviewed version

Generalized Cramér–von Mises goodness-of-fit tests for multivariate distributions

Sung Nok Chiu*, Kwong Ip Liu

Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong

Abstract

A class of statistics for testing the goodness-of-fit for any multivariate continuous distribution is proposed. These statistics consider not only the goodness-of-fit of the joint distribution but also the goodness-of-fit of all marginal distributions, and can be regarded as generalizations of the multivariate Cramér–von Mises statistic. Simulation shows that these generalizations, using the Monte Carlo test procedure to approximate their finite sample p-values, are more powerful than the multivariate Kolmogorov–Smirnov statistic.

Keywords: goodness-of-fit; multivariate normality; discrepancy; uniformity; Cramér–von Mises statistic; Kolmogorov–Smirnov statistic; Monte Carlo test.

1 Introduction

Given a sample of n independent realizations $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ of an s-dimensional random vector $\mathbf{Y}$ with an unknown distribution function F, it is often desirable to test whether F is equal to a specified distribution F_0, i.e., we would like to test

$$\begin{aligned} \mathrm{H}_0 &: F = F_0, \\ \mathrm{H}_\mathrm{A} &: F \neq F_0. \end{aligned}$$

For discrete distributions, the classical test statistic is the Pearson χ^2-statistic (Greenwood and Nikulin, 1996; Read and Cressie, 1988), which is a measure of the discrepancy between the observed frequencies with the expected frequencies under the null hypothesis.

* Corresponding author.
E-mail address: snchiu@hkbu.edu.hk

Although the Pearson χ^2-statistic can also be applied to continuous distributions by discretization, it is more natural to employ the empirical (cumulative) distribution function for continuous distributions

$$F_n(\mathbf{y}) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(\mathbf{y}_i \leq \mathbf{y}),$$

where $\mathbf{1}(\cdot)$ denotes the indicator function and the partial order $\leq$ in $\mathbb{R}^s$ is defined componentwise. The empirical function F_n is an unbiased and consistent estimator of F. To test the goodness-of-fit for a hypothesized distribution F_0, we can use the discrepancy between the empirical F_n and the hypothesized F_0 as a test statistic.

The aim of this paper is to propose different measures of discrepancy between F_n and F_0 and then compare the powers of them in testing goodness-of-fit for multivariate continuous distributions, including but not restricted to multivariate normal. However, we do not intend to compare the powers of all existing tests for multivariate normality because (i) there are at least 50 testing procedures, including goodness-of-fit type tests as well as tests based on measures of skewness and kurtosis or based on empirical characteristic functions, for multivariate normality (see Mecklin and Mundfrom (2004) for an overview of power studies for testing multivariate normality); (ii) our tests can be used to test for any hypothesized continuous multivariate distributions and so it would not be appropriate to compare ours with tests tailor-made for multivariate normality; (iii) we regard our test statistics as competitors of the multivariate Kolmogorov–Smirnov statistic proposed by Justel et al. (1997).

Some of the discrepancies used in this paper have also been used to construct statistics to test complete spatial randomness hypothesis in spatial point pattern analysis (Ho and Chiu, 2007). These test statistics have been shown to be more powerful in detecting clustering than existing statistics in spatial point process literature.

2 Kolmogorov–Smirnov statistic and Cramér–von Mises statistic

For the univariate case $s = 1$, the two most famous measures of discrepancy between two continuous distributions are the Kolmogorov–Smirnov statistic

$$D^{\mathrm{KS}} = \sup_{y\in\mathbb{R}} |F_n(y) - F_0(y)|,$$

and the Cramér–von Mises statistic

$$D^{\mathrm{CM}} = \int_{y\in\mathbb{R}} [F_n(y) - F_0(y)]^2 \,\mathrm{d}F_0(y).$$

Modifications or generalizations of the Kolmogorov–Smirnov statistic and the Cramér–von Mises statistic include the Watson–Darling statistic (Darling, 1983; Watson, 1976)

$$D^{\mathrm{WD}} = \sup_{y\in\mathbb{R}} \left| F_n(y) - F_0(y) - \int_{z\in\mathbb{R}} [F_n(z) - F_0(z)] \,\mathrm{d}F_0(z) \right|,$$

and the Watson statistic (Watson, 1961)

$$D^{\mathrm{W}} = \int_{y\in\mathbb{R}} \left[F_n(y) - F_0(y) - \int_{z\in\mathbb{R}} \{F_n(z) - F_0(z)\}\, \mathrm{d}F_0(z) \right]^2 \mathrm{d}F_0(y),$$

which are centred versions so that D^{WD} and D^{W} are invariant under translation of the data. Another generalization is obtained by introducing a weight function to the Cramér–von Mises statistic, leading to the famous Anderson–Darling statistic (Anderson and Darling, 1952, 1954)

$$D^{\mathrm{AD}} = \int_{y\in\mathbb{R}} \frac{[F_n(y) - F_0(y)]^2}{F_0(y)[1 - F_0(y)]}\, \mathrm{d}F_0(y),$$

which has been further generalized to phi-divergence test statistics (Jager and Wellner, 2007).

To extend the Kolmogorov–Smirnov statistic and the Cramér–von Mises statistic to multivariate cases $s \geq 2$, Justel et al. (1997) applied a theorem by Rosenblatt (1952), which is given as follows. Consider the $s!$ distinct permutations of $(1, \ldots, s)$. We order these $s!$ permutations in some way and denote them by $(\pi_j(1), \ldots, \pi_j(s))$ for $j = 1, \ldots, s!$. Define the transformation $T_j : \mathbb{R}^s \mapsto \mathbb{R}^s$ by $T_j(\mathbf{y}) = \mathbf{x} = (x_1, \ldots, x_s)$, where

$$\begin{aligned} x_1 &= F_{Y_{\pi_j(1)}}(y_{\pi_j(1)}), \\ x_i &= F_{Y_{\pi_j(i)}|Y_{\pi_j(1)},\ldots,Y_{\pi_j(i-1)}}(y_{\pi_j(i)}|y_{\pi_j(1)}, \ldots, y_{\pi_j(i-1)}), \quad \text{for } i = 2, \ldots, s, \end{aligned}$$

where F_{Y_k} and $F_{Y_k|Y_1,\ldots,Y_{k-1}}$ are, respectively, the marginal distribution and conditional distribution of Y_k, given $Y_1, \ldots, Y_{k-1}$. The Rosenblatt theorem states that for a random vector $\mathbf{y} = (y_{\pi_j(1)}, \ldots, y_{\pi_j(s)})$ with joint density function f, the components x_i's of the random vector $\mathbf{x}$ are independent and identically uniformly distributed in $[0, 1]$. Thus, whenever a distribution has been explicitly given in the null hypothesis, we can always transform the data and use a goodness-of-fit test for uniform distribution.

Transforming the problem of testing multivariate normality into the problem of testing uniformity has already been considered by Hensler et al. (1977), Rincon-Gallardo and Quesenberry (1982) and Rincon-Gallardo et al. (1979). However, the computing requirement of their tests was too high and made their tests unattractive (Bera and John, 1983) when compared with others. Nevertheless, the advancement of computing technology in the last two decades enables us to do extensive computations in (milli)seconds.

For each $j = 1, \ldots, s!$, let

$$U_n^{(j)}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(\mathbf{x}_i \leq \mathbf{x})$$

be the empirical distribution of $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, where $\mathbf{x}_i = T_j(\mathbf{y}_i)$. For $s = 1$, there is only one permutation and so by denoting $U_n^{(1)}(x) =: U_n(x)$, the Kolmogorov–Smirnov statistic and the Cramér–von Mises statistic can be rewritten as

$$\begin{aligned} D^{\mathrm{KS}} &= \sup_{y\in\mathbb{R}} |F_n(y) - F_0(y)| = \sup_{x\in[0,1]} |U_n(x) - x|\,, \\ D^{\mathrm{CM}} &= \int_{y\in\mathbb{R}} [F_n(y) - F_0(y)]^2\, \mathrm{d}F_0(y) = \int_{x\in[0,1]} [U_n(x) - x]^2\, \mathrm{d}x. \end{aligned}$$

Justel et al. (1997) proposed the following multivariate version of the Kolmogorov–Smirnov statistic

$$D_{\max}^{\mathrm{KS}} = \max_{1\le j\le s!} \left\{ \sup_{\mathbf{x}\in[0,1]^s} \left| U_n^{(j)}(\mathbf{x}) - \lambda(\mathbf{x}) \right| \right\},$$

where $\lambda(\mathbf{x}) = x_1 \cdots x_s$ is the volume of the hyper-rectangle $[\mathbf{0}, \mathbf{x}] = [0, x_1] \times \cdots \times [0, x_s]$.

The empirical $U_n^{(j)}(\mathbf{x})$ is piecewise constant and jumps occur only when one of the coordinates of $\mathbf{x}$ is equal to the value of the same coordinate of any one of $\mathbf{x}_i$. Let $\mathbf{x}_i = (x_{i1}, \ldots, x_{is})$. The supremum of the pointwise absolute difference must be the absolute difference at some $(m_1, \ldots, m_s)$ where $m_k = x_{i_k k}$ or $x_{i_k k}^-$ for some $1 \le i_k \le n$, $k = 1, \ldots, s$. Therefore, the computation of the supremum of each permutation requires $O(n^s)$ operations. Justel et al. (1997) showed that for $s = 2$ it is not necessary to evaluate the differences at all of these $2n^2$ combinations but only the differences $U_n^{(j)}(\mathbf{x}) - \lambda(\mathbf{x})$ at $\mathbf{x} \in \{(x_{k1}, x_{i2}) : x_{i1} \le x_{k1}, x_{i2} \ge x_{k2}, i, k = 0, 1, \ldots, n\}$, where $(x_{01}, x_{02}) = (0, 0)$, and the differences $U_n^{(j)}(\mathbf{x}^-) - \lambda(\mathbf{x})$ at $\mathbf{x} \in \{(x_{k1}, x_{i2}) : x_{i1} < x_{k1}, x_{i2} > x_{k2}, i, k = 0, 1, \ldots, n+1\}$, where $(x_{n+1,1}, x_{n+1,2}) = (1, 0)$ and $\mathbf{x}^- = (x_1^-, x_2^-)$. This approach will reduce the number of pointwise differences from $2n^2$ to a number between $3n$ and $3n + \binom{n}{2}$, depending on the sample configuration. Nevertheless, it only works for $s = 2$ and the worst case still requires $O(n^2)$ operations. They also suggested an approximation by considering the differences at $\mathbf{x}_i$ and $\mathbf{x}_i^-$:

$$\tilde{D}_{\max}^{\mathrm{KS}} = \max_{1\le j\le s!} \left[\sup_{1\le i\le n} \left\{ \left| U_n^{(j)}(\mathbf{x}_i) - \lambda(\mathbf{x}_i) \right|, \left| U_n^{(j)}(\mathbf{x}_i^-) - \lambda(\mathbf{x}_i) \right| \right\} \right],$$

which requires only $O(n)$ operations.

Comparing the univariate and the multivariate cases, we can see from this formulation that the univariate Kolmogorov–Smirnov statistic and the Cramér–von Mises statistic, as well as their centred or weighted versions, are distribution-free, but the multivariate Kolmogorov–Smirnov statistic are not, because the covariance structure of F will play a role. In particular, if the components of $\mathbf{y}_i$ are themselves independent, then $U_n^{(j)}$ does not depend on j, but in general if the components are not independent, then different j will give us different empirical $U_n^{(j)}$. Therefore, if we want to test, for example, for multivariate normality $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the distributions of $D_{\max}^{\mathrm{KS}}$ and $\tilde{D}_{\max}^{\mathrm{KS}}$ under the null hypothesis depend on the covariance matrix $\boldsymbol{\Sigma}$.

Because of the dependence, general goodness-of-fit tests for multivariate distribution have not been fully explored. To overcome this difficulty, Justel et al. (1997) suggested to compare each observed $\sup_{\mathbf{x}\in[0,1]^s} \left| U_n^{(j)}(\mathbf{x}) - \lambda(\mathbf{x}) \right|$ or $\sup_{1\le i\le n} \left\{ \left| U_n^{(j)}(\mathbf{x}_i) - \lambda(\mathbf{x}_i) \right|, \left| U_n^{(j)}(\mathbf{x}_i^-) - \lambda(\mathbf{x}_i) \right| \right\}$ for $j = 1, \ldots, s!$, to the percentiles of the distribution of $D_{\max}^{\mathrm{KS}}$ or $\tilde{D}_{\max}^{\mathrm{KS}}$ of n independent and uniformly distributed points in $[0, 1]^s$ respectively, and reject the null hypothesis whenever any one of the $s!$ observed suprema exceeds the $100(1 - \alpha/s!)$th percentile, where α is the nominal significance level. We divide α by $s!$ because this testing procedure is equivalent to a multiple test and the Bonferroni adjustment is applied.

3 Discrepancy and Cramér–von Mises Statistics

The Kolmogorov–Smirnov statistic and the Cramér–von Mises statistic for testing uniformity are special cases of the so-called L_p-star discrepancy (Hickernell, 1998a), defined by

$$D_p^* = \left[\int_{[0,1)^s} |U_n(\mathbf{x}) - \lambda(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \right]^{\frac{1}{p}},$$

where for $p \to \infty$, we take the sup-norm

$$D_\infty^* = \sup_{\mathbf{x} \in [0,1)^s} |U_n(\mathbf{x}) - \lambda(\mathbf{x})|,$$

so that when $s = 1$, $D_2^* = \sqrt{D^{\mathrm{CM}}}$ and $D_\infty^* = D^{\mathrm{KS}}$. We change the integration domain from $[0,1]^s$ to $[0,1)^s$ because some of our generalizations below will involve periodic boundary conditions.

The idea of discrepancy has been used extensively in the quasi-Monte Carlo methods in numerical integration (Hua and Wang, 1981; Niederreiter, 1992). Suppose an integral of the function f over the domain $[0,1)^s$ is approximated by the arithmetic mean of the values of f at n distinct locations in $[0,1)^s$. The absolute error of this approximation is bounded above by the product of the variation of the function f and a function, defined according to the definition adopted for the variation of a function, of those n distinct locations. This function is called the discrepancy and any such a discrepancy can be regarded as a goodness-of-fit statistic (Hickernell, 1999a) for testing the uniform distribution, where the n locations are the data $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. In addition to the star discrepancy, a number of other discrepancies were proposed by e.g. Hickernell (1998a,b, 1999a); Niederreiter (1992), and in this paper we investigate six discrepancies that have simple formulae for computation when $p = 2$.

We start with the notations required for the precise definition of the discrepancies. Let $S = \{1, 2, \ldots, s\}$ and for $u \subset S$ and $\mathbf{z} = (z_1, \ldots, z_s) \in \mathbb{R}^s$, let $\mathbf{z}^u$ denote a $|u|$-dimensional vector formed by the components of $\mathbf{z}$ that are indexed by the elements of u, and let $\tilde{\mathbf{z}}^u$ denote the vector $(\tilde{z}_1, \ldots, \tilde{z}_s)$, where $\tilde{z}_k$ is equal to z_k if $k \in u$ and equal to 1 otherwise. Furthermore, let $[0,1)^u$ denote the $|u|$-dimensional unit cube that is the projection of $[0,1)^s$ into the coordinates indexed by the elements of u. Let $\mathbf{x} = (x_1, \ldots, x_s)$ and $\mathbf{x}' = (x_1', \ldots, x_s')$ denote two generic points in the data set $\mathcal{P} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

The original star discrepancy in Warnock (1972) is,

$$\begin{aligned} D_2^* &= \left[\int_{[0,1)^s} |U_n(\mathbf{z}) - \lambda(\mathbf{z})|^2 \, \mathrm{d}\mathbf{z} \right]^{\frac{1}{2}} \\ &= \left[\left(\frac{1}{3}\right)^s - \frac{2}{n} \sum_{\mathbf{x} \in \mathcal{P}} \prod_{k=1}^{s} \left(\frac{1 - x_k^2}{2} \right) + \frac{1}{n^2} \sum_{\mathbf{x}, \mathbf{x}' \in \mathcal{P}} \prod_{k=1}^{s} \{1 - \max(x_k, x_k')\} \right]^{\frac{1}{2}}. \end{aligned}$$

This discrepancy measures the uniformity of the points in the s-dimensional hypercube, i.e., the goodness-of-fit of the joint distribution. In order to be more powerful, we would

like to consider the goodness-of-fit of all marginal distributions. That is, we would like to measure not only the uniformity in the s-dimensional hypercube, but also the uniformity of the projections of the points onto all lower dimensional hypercubes. A straightforward generalization of the L_2-star discrepancy is the modified L_2-star discrepancy (Hickernell, 1998a, 1999a):

$$\begin{aligned} MD_2 &= \left[\sum_{\emptyset \subset u \subseteq S} \int_{[0,1)^u} |U_n(\tilde{\mathbf{z}}^u) - \lambda(\tilde{\mathbf{z}}^u)|^2 \, \mathrm{d}\mathbf{z}^u \right]^{\frac{1}{2}} \\ &= \left[\left(\frac{4}{3}\right)^s - \frac{2}{n} \sum_{\mathbf{x} \in \mathcal{P}} \prod_{k=1}^{s} \left(\frac{3 - x_k^2}{2} \right) + \frac{1}{n^2} \sum_{\mathbf{x}, \mathbf{x}' \in \mathcal{P}} \prod_{k=1}^{s} \{2 - \max(x_k, x_k')\} \right]^{\frac{1}{2}}. \end{aligned}$$

Because U_n and λ are in fact distribution functions, it is very natural to measure from zero $(0, 0, \ldots, 0)$, and we say that the hyper-rectangle $[\mathbf{0}, \mathbf{x}]$ is anchored at zero. From geometrical point of view, however, a measure of the uniformity of a set of points in $[0, 1)^s$ should not depend on such an arbitrarily chosen anchor. One possible modification is to measure not from the zero but, in some sense, from the nearest vertex of the hypercube so that the discrepancy value will be invariant under rotation and reflection. This idea led to the L_2-centred discrepancy (Hickernell, 1998a, 1999a):

$$\begin{aligned} CD_2 &= \left[\sum_{\emptyset \subset u \subseteq S} \int_{[0,1)^u} |U_n^c(\tilde{\mathbf{z}}^u) - \lambda^c(\tilde{\mathbf{z}}^u)|^2 \, \mathrm{d}\mathbf{z}^u \right]^{\frac{1}{2}} \\ &= \left[\left(\frac{13}{12}\right)^s - \frac{2}{n} \sum_{\mathbf{x} \in \mathcal{P}} \prod_{k=1}^{s} \left(1 + \frac{1}{2} \left| x_k - \frac{1}{2} \right| - \frac{1}{2} \left| x_k - \frac{1}{2} \right|^2 \right) \right. \\ &\quad \left. + \frac{1}{n^2} \sum_{\mathbf{x}, \mathbf{x}' \in \mathcal{P}} \prod_{k=1}^{s} \left(1 + \frac{1}{2} \left| x_k - \frac{1}{2} \right| + \frac{1}{2} \left| x_k' - \frac{1}{2} \right| - \frac{1}{2} |x_k - x_k'| \right) \right]^{\frac{1}{2}}, \end{aligned}$$

in which $U_n^c(\mathbf{z})$ is the empirical distribution of the random vector $\mathbf{X}$ taking some value in the hyper-rectangle formed by $\mathbf{z}$ and its nearest vertex in $[0, 1]^s$, where $\mathbf{X}$ follows the distribution from which $\mathbf{x}_i$'s are sampled, and $\lambda^c(\mathbf{z})$ is the volume of this hyper-rectangle.

In the L_2-centred discrepancy, for each fixed $\mathbf{z}$, the empirical distribution $U_n^c(\mathbf{z})$ anchors at only one (the nearest) vertex. A further modification is to anchor at multiple vertices. We divide the 2^s vertices of the hypercube $[0, 1)^s$ into two groups. For each vertex, if the sum of its coordinates is even, then we call it an even vertex; otherwise it is odd. If we anchor the empirical distribution at all even vertices, we have the L_2-symmetric discrepancy

(Hickernell, 1998a, 1999a):

$$\begin{aligned} SD_2 &= \left[\sum_{\emptyset\subset u\subseteq S}\int_{[0,1)^u}\left|U_n^e(\tilde{\mathbf{z}}^u)-\lambda^e(\tilde{\mathbf{z}}^u)\right|^2\,\mathrm{d}\mathbf{z}^u\right]^{\frac{1}{2}} \\ &= \left[\left(\frac{4}{3}\right)^s-\frac{2}{n}\sum_{\mathbf{x}\in\mathcal{P}}\prod_{k=1}^{s}\left(1+2x_k-2x_k^2\right)+\frac{2^s}{n^2}\sum_{\mathbf{x},\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left(1-|x_k-x_k'|\right)\right]^{\frac{1}{2}}, \end{aligned}$$

in which $U_n^e(\mathbf{z})$ is the empirical distribution of the above mentioned random vector $\mathbf{X}$ taking some value in the hyper-rectangles formed by $\mathbf{z}$ and all even vertices, and $\lambda^e(\mathbf{z})$ is the total volume of these hyper-rectangles.

The L_2-centred discrepancy and L_2-symmetric discrepancy still anchor at the vertices of the hypercube $[0,1)^s$. A further modification is to denote $\tilde{\mathbf{z}}^u$ by $\tilde{\mathbf{z}}_1^u$, then replace its nearest vertex or even vertices by another variable $\tilde{\mathbf{z}}_2^u$ and finally integrate over all possible values of $\tilde{\mathbf{z}}_2^u$, which gives us the L_2-unanchored discrepancy (Hickernell, 1998a, 1999a; Niederreiter, 1992)

$$\begin{aligned} UD_2 &= \left[\sum_{\emptyset\subset u\subseteq S}\int\int_{[0,1)^u\times[0,1)^u,\ \mathbf{z}_1^u\le\mathbf{z}_2^u}\left|U_n([\tilde{\mathbf{z}}_{\mathbf{1}}^u,\tilde{\mathbf{z}}_{\mathbf{2}}^u))-\lambda([\tilde{\mathbf{z}}_{\mathbf{1}}^u,\tilde{\mathbf{z}}_{\mathbf{2}}^u))\right|^2\mathrm{d}\mathbf{z}_1^u\mathrm{d}\mathbf{z}_2^u\right]^{\frac{1}{2}} \\ &= \left[\left(\frac{13}{12}\right)^s-\frac{2}{n}\sum_{\mathbf{x}\in\mathcal{P}}\prod_{k=1}^{s}\left\{1+\frac{x_k(1-x_k)}{2}\right\}+\frac{1}{n^2}\sum_{\mathbf{x},\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left\{1+\min(x_k,x_k')-x_kx_k'\right\}\right]^{\frac{1}{2}}, \end{aligned}$$

where $U_n([\mathbf{z_1},\mathbf{z_2}))$ is the empirical distribution of the above mentioned $\mathbf{X}$ taking some value in the hyper-rectangle $[\mathbf{z_1},\mathbf{z_2})$, the volume of which is $\lambda([\mathbf{z_1},\mathbf{z_2}))$.

The L_2-unanchored discrepancy integrates over $\mathbf{z}_1^u\le\mathbf{z}_2^u$ only because we have to form hyper-rectangle $[\tilde{\mathbf{z}}_{\mathbf{1}}^u,\tilde{\mathbf{z}}_{\mathbf{2}}^u)$. If we use periodic boundary conditions, this restriction can be removed and we have the L_2-wraparound discrepancy (Hickernell, 1998a, 1999a):

$$\begin{aligned} WD_2 &= \left[\sum_{\emptyset\subset u\subseteq S}\int_{[0,1)^u}\int_{[0,1)^u}\left|U_n(J(\tilde{\mathbf{z}}_1^u,\tilde{\mathbf{z}}_2^u))-\lambda(J(\tilde{\mathbf{z}}_1^u,\tilde{\mathbf{z}}_2^u))\right|^2\mathrm{d}\mathbf{z}_1^u\mathrm{d}\mathbf{z}_2^u\right]^{\frac{1}{2}} \\ &= \left[-\left(\frac{4}{3}\right)^s+\frac{1}{n^2}\sum_{\mathbf{x},\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left\{\frac{3}{2}-|x_k-x_k'|(1-|x_k-x_k'|)\right\}\right]^{\frac{1}{2}}, \end{aligned}$$

where $J(\mathbf{z}_1,\mathbf{z}_2)$ is a hyper-rectangle under period boundary conditions:

$$\begin{aligned} J(z_k',z_k) &= \begin{cases}[z_k',z_k), & z_k'\le z_k,\\ [0,z_k)\cup[z_k',1), & z_k<z_k',\end{cases} \\ J(\mathbf{z}_1,\mathbf{z}_2) &= \bigotimes_{k=1}^{s}J(z_{1k},z_{2k}). \end{aligned}$$

The L_2-star discrepancy, D_2^*, is the square root of the Cramér–von Mises goodness-of-fit statistic testing the uniform distribution. The other five discrepancies MD_2, CD_2, SD_2, UD_2 and WD_2, using the L_2-norm, can be considered as generalizations of the Cramér–von Mises statistic; they measure the discrepancy between not only the joint distribution but also all marginal distributions of the empirical distribution and the hypothesized distribution by taking the sums of the discrepancies of all possible projections. It can be believed, and we will confirm by simulation, that these generalizations would lead to more powerful tests. We choose the L_2-norm in these discrepancies because we have simple computational formulae involving $O(n^2 s)$ operations; for the supremum of each permutation in the calculation of the multivariate Kolmogorov–Smirnov statistic, the computation requires $O(n^s)$ operations.

4 Variants of goodness-of-fit statistics

As we mentioned above, such discrepancies have already been suggested (Hickernell, 1999a) as statistics for testing the uniform distribution. However, Liang et al. (2000, Corollary 2.4) showed that the square of any one of the above L_2-type discrepancies, even multiplied by $\sqrt{n}$, converges in probability to zero, as $n \to \infty$, and consequently they suggested that these discrepancies would not be as appealing as the following two variants, denoted by A_n and T_n, where

$$A_n = \sqrt{n}\left\{\frac{(U_1 - M^s) + 2(U_2 - M^s)}{5\sqrt{\zeta_1}}\right\}, \tag{1}$$

$$T_n = n\left(U_1 - M^d, U_2 - M^d\right)\boldsymbol{\Sigma}_n^{-1}\left(U_1 - M^d, U_2 - M^d\right)', \tag{2}$$

in which

$$\boldsymbol{\Sigma}_n = \begin{pmatrix} \zeta_1 & 2\zeta_1 \\ 2\zeta_1 & \frac{4(n-2)}{n-1}\zeta_1 + \frac{2}{n-1}\zeta_2 \end{pmatrix},$$

and the parameters U_1, U_2, M, ζ_1 and ζ_2 depend on the choice of the discrepancy. They argued that these two variants would be more appealing by showing that A_n and T_n converge weakly to the standard normal and the χ^2-distribution with 2 degrees of freedom, respectively.

The parameters in A_n and T_n arisen from MD_2, CD_2 and SD_2 have been given (with two misprints) in Liang et al. (2000, p. 345) and we derived them for those arisen from UD_2 and WD_2. Suppose the points in $\mathcal{P}$ are ordered in some way, the formulae of these parameters are given as follows:

1. the modified star discrepancy MD_2:

$$U_1 = \frac{1}{n}\sum_{\mathbf{x}\in\mathcal{P}}\prod_{k=1}^{s}\left(\frac{3 - x_k^2}{2}\right),$$

$$U_2 = \frac{2}{n(n-1)}\sum_{\mathbf{x}<\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left\{2 - \max(x_k, x_k')\right\},$$

$$M = 4/3, \;\; \zeta_1 = (9/5)^s - (16/9)^s, \;\; \zeta_2 = (11/6)^s - (16/9)^s;$$

2. the centred discrepancy CD_2:

$$U_1 = \frac{1}{n}\sum_{\mathbf{x}\in\mathcal{P}}\prod_{k=1}^{s}\left(1+\frac{1}{2}\left|x_k-\frac{1}{2}\right|-\frac{1}{2}\left|x_k-\frac{1}{2}\right|^2\right),$$
$$U_2 = \frac{2}{n(n-1)}\sum_{\mathbf{x}<\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left(1+\frac{1}{2}\left|x_k-\frac{1}{2}\right|+\frac{1}{2}\left|x'_k-\frac{1}{2}\right|-\frac{1}{2}\left|x_k-x'_k\right|\right),$$
$$M = 13/12,\ \ \zeta_1 = (47/40)^s-(13/12)^{2s},\ \ \zeta_2 = (57/48)^s-(13/12)^{2s};$$

3. the symmetric discrepancy SD_2:

$$U_1 = \frac{1}{n}\sum_{\mathbf{x}\in\mathcal{P}}\prod_{k=1}^{s}\left(1+2x_k-2x_k^2\right),$$
$$U_2 = \frac{2^s}{n^2}\sum_{\mathbf{x}<\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left(1-\left|x_k-x'_k\right|\right),$$
$$M = 4/3,\ \ \zeta_1 = (9/5)^s-(16/9)^s,\ \ \zeta_2 = 2^s-(16/9)^s;$$

4. the unanchored discrepancy UD_2:

$$U_1 = \frac{1}{n}\sum_{\mathbf{x}\in\mathcal{P}}\prod_{k=1}^{s}\left\{1+\frac{x_k(1-x_k)}{2}\right\},$$
$$U_2 = \frac{2}{n(n-1)}\sum_{\mathbf{x}<\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left\{1+\min(x_k,x'_k)-x_kx'_k\right\}$$
$$M = 13/12,\ \ \zeta_1 = (47/40)^s-(169/149)^s,\ \ \zeta_2 = (19/16)^s-(169/144)^s;$$

5. the wraparound discrepancy WD_2:

$$U_1 = (4/3)^s,$$
$$U_2 = \frac{2}{n(n-1)}\sum_{\mathbf{x}<\mathbf{x}'\in\mathcal{P}}\prod_{k=1}^{s}\left\{\frac{3}{2}-\left|x_k-x'_k\right|\left(1-\left|x_k-x'_k\right|\right)\right\}$$
$$M = 4/3,\ \ \zeta_1 = 0,\ \ \zeta_2 = (161/90)^s-(16/9)^s.$$

However, for the wraparound discrepancy UD_2, the matrix $\mathbf{\Sigma}_n$ is singular and $\zeta_1 = 0$, implying that equations (1) and (2) contain undefined terms and are not valid. We follow the generic definitions given in Liang et al. (2000) and obtain the following corresponding formulae for A_n and T_n:

$$\begin{aligned} A_n &= \sqrt{n}\left\{\frac{U_2-M^s}{\sqrt{2\zeta_2/(n-1)}}\right\}, \\ T_n &= A_n^2, \end{aligned}$$

where now the limiting distribution of T_n is χ^2 with only 1 degree of freedom.

5 Test statistics, Multiple tests and Monte Carlo tests

When $s > 1$, the values of the discrepancies of the transformed sample depend on which permutation of $(1, \ldots, s)$ is used in the Rosenblatt transformation, unless the components are independent. The multiple test formulation using Bonferroni adjustment, adopted in Justel et al. (1997), can also be used for all these discrepancies. The advantage of this formulation is that each permutation of a test statistic is compared with the percentile of the distribution, either the asymptotic one or an approximate obtained by Monte Carlo simulation, of that statistic calculated from independent uniformly distributed points, making the test distribution-free. The actual type I error rate, however, may be much lower than the nominal significance level α, leading to conservative tests. The price to pay for conservativeness is a loss in power.

A better way is to use Monte Carlo tests (Davison and Hinkley, 1997, pp. 140–143), which are exact in their own right. We generate R Monte Carlo samples of size n consisting of s-dimensional random vectors simulated according to the hypothesized distribution. Thus, in total we have $R + 1$ samples, namely, the sample of the observed data and the Monte Carlo samples simulated under the null hypothesis. Denote by τ the value of a test statistics calculated from the observed sample and by τ_r^* the value of the same statistic obtained from the rth Monte Carlo sample, $r = 1, \ldots, R$. The sequence $\{\tau, \tau_1^*, \ldots, \tau_R^*\}$ forms a random sample of the distribution of the chosen test statistic under the null hypothesis. The one-sided p-value can be estimated by the sample proportion

$$p_{\text{mc}} = \frac{\#\{\tau_r^* \geq \tau\} + 1}{R + 1}.$$

The null hypothesis will be rejected if p_{mc} is less than or equal to the nominal significance level α.

Hope (1968) showed that the power loss, compared with the corresponding uniformly most powerful test, resulting from using Monte Carlo tests is slight and so R is not necessary to be large. Marriott (1979) suggested that for $\alpha = 0.05$, $R = 99$ is adequate, whilst Davison and Hinkley (1997, p. 156) suggested, for $\alpha \geq 0.05$, that the loss of power with $R = 99$ is not serious and $R = 999$ should generally be safe.

If we adopt the multiple tests with Bonferroni adjustment, a test statistic has to be the maximum over all permutations; if we adopt the Monte Carlo test, we may take either the maximum or the sum. Therefore, for testing multivariate distribution, we have quite a number of possible test statistics, which can be generically denoted by

$$\begin{aligned} D_{\max} &= \max_{1 \leq j \leq s!} D^{(j)}, \\ D_{\text{sum}} &= \sum_{j=1}^{s!} D^{(j)}, \end{aligned}$$

where $D^{(j)}$ is one of the following eight measures of discrepancy of $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} = \{T_j(\mathbf{y}_1),$

$\ldots, T_j(\mathbf{y}_n)\}$: D^*_∞, $\tilde{D}^*_\infty$, D^*_2, MD_2, CD_2, SD_2, UD_2 and WD_2. The discrepancy

$$\tilde{D}^*_\infty = \sup_{1\le i\le n} \left\{ \left| U^{(j)}_n(\mathbf{x}_i) - \lambda(\mathbf{x}_i) \right|, \left| U^{(j)}_n(\mathbf{x}^-_i) - \lambda(\mathbf{x}_i) \right| \right\}$$

is the approximation suggested in Justel et al. (1997) for D^*_∞. In particular, if $D^{(j)}$ is D^*_∞ and $\tilde{D}^*_\infty$, then $D_{\max}$ is equal to $D^{\mathrm{KS}}_{\max}$ and $\tilde{D}^{\mathrm{KS}}_{\max}$, respectively.

Moreover, $D^{(j)}$ can also be the A_n or T_n statistics arisen from the above five discrepancies, which will be denoted below by MA_n and MT_n for MD_2, CA_n and CT_n for CD_2, and similarly for SD_2, UD_2 and WD_2.

However, the goodness-of-fit test based on the statistic A_n is a two-sided test, and the finite-sample distribution of it is not necessarily symmetric. Thus, when we discuss the statistic $D_{\max}$ for A_n below, we in fact, with a slight abuse of notation for ease of presentation, consider not only the maximum but also the minimum over all permutations, and the type I error rate and the power are the sizes of the critical region, which is the union of two intervals:

$$\left\{ \min_{1\le j\le s!} A^{(j)}_n \le c_1 \right\} \bigcup \left\{ \max_{1\le j\le s!} A^{(j)}_n \ge c_2 \right\},$$

under the null and the alternative hypothesis, respectively, where, at the nominal level α, the critical values c_1 and c_2 are the (50α)th and $50(1-\alpha)$th percentiles of the null distributions of $\min_{1\le j\le s!} A^{(j)}_n$ and $\max_{1\le j\le s!} A^{(j)}_n$, respectively. For the Monte Carlo test procedure, these critical values are approximated by R simulated samples.

Moreover, it is not necessarily that the A_n values for the different data sets obtained by different Rosenblatt transformations of the same data $\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ will have the same sign. The two-sided test based on the statistic $\sum_{j=1}^{s!} A^{(j)}_n$ may not be as powerful as the one-sided test based on $\sum_{j=1}^{s!} |A^{(j)}_n|$. By the same spirit, we may also consider the one-sided test based on $\max_{1\le j\le s!} |A^{(j)}_n|$. Thus, we also include $|A_n|$ as one possible form of $D^{(j)}$.

6 Simulation

The actual type I error rates and the powers of these possible test statistics were compared in the two scenarios considered in Justel et al. (1997). In the first scenario, the null hypothesis was the bivariate normal distribution with mean $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}$, where

$$\begin{aligned} \boldsymbol{\mu}_0 &= (0,0), \\ \boldsymbol{\Sigma} &= \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}. \end{aligned}$$

Let $0 < \varepsilon < 1$. The sample point set was generated according to the alternative model, which is the bivariate normal distribution with the mean $(1-\varepsilon)\boldsymbol{\mu}_0 + \varepsilon\boldsymbol{\mu}_1$ and the same covariance matrix $\boldsymbol{\Sigma}$ as in the null hypothesis. We compared the powers for the following cases: $\boldsymbol{\mu}_1 = (3,3)$ and $(3,-1)$, $\varepsilon = 0.1$, 0.2 and 0.4 and $n = 15$, 25, 50 and 100. In each

Table 1: Estimated power, by 1000 simulations, of the Monte Carlo test with $R = 999$ of the test statistics D_{sum} and $D_{\max}$ using the discrepancy $D^{(j)}$ at the 0.05 significance level for a sample of size n. The null hypothesis is the $\mathrm{N}((0,0), \boldsymbol{\Sigma})$ and the alternative is the mixture $(1-\varepsilon)\mathrm{N}((0,0), \boldsymbol{\Sigma}) + \varepsilon \mathrm{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$. The values in parentheses are the powers of the multiple tests suggested by Justel et al. (1997) with the Bonferroni adjustment.

			$n = 15$		$n = 25$		$n = 50$		$n = 100$	
$\boldsymbol{\mu}_1$	ε	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3,3)$	0.1	MD_2	0.187	0.177	0.318	0.294	0.643	0.628	0.902	0.894
		CD_2	0.181	0.179	0.308	0.292	0.634	0.615	0.905	0.892
		SD_2	0.184	0.185	0.308	0.303	0.621	0.616	0.898	0.893
		UD_2	0.151	0.136	0.214	0.191	0.430	0.369	0.727	0.661
		WD_2	0.150	0.125	0.212	0.183	0.420	0.362	0.734	0.659
		D_2^*	0.149	0.123	0.279	0.238	0.583	0.547	0.861	0.841
		D_∞^*	0.137	0.120	0.224	0.198	0.491	0.448	0.754	0.717
				(0.109)		(0.179)		(0.372)		(0.675)
		$\tilde{D}_\infty^*$	0.118	0.100	0.195	0.175	0.461	0.407	0.743	0.693
				(0.092)		(0.158)		(0.390)		(0.659)
		XD	0.190	0.185	0.309	0.303	0.615	0.616	0.877	0.889
		MA_n	0.190	0.187	0.322	0.300	0.596	0.574	0.847	0.830
		$\lvert MA_n \rvert$	0.158	0.143	0.276	0.249	0.559	0.535	0.831	0.816
		MT_n	0.225	0.213	0.353	0.331	0.671	0.655	0.909	0.899
		CA_n	0.099	0.090	0.090	0.092	0.106	0.095	0.089	0.080
		$\lvert CA_n \rvert$	0.100	0.092	0.096	0.091	0.102	0.103	0.087	0.078
		CT_n	0.195	0.185	0.305	0.283	0.620	0.602	0.891	0.888
		SA_n	0.118	0.121	0.155	0.156	0.286	0.287	0.515	0.512
		$\lvert SA_n \rvert$	0.139	0.137	0.177	0.180	0.310	0.312	0.541	0.537
		ST_n	0.192	0.181	0.297	0.293	0.614	0.603	0.889	0.888
		UA_n	0.079	0.082	0.093	0.098	0.128	0.121	0.232	0.221
		$\lvert UA_n \rvert$	0.095	0.092	0.111	0.107	0.142	0.136	0.251	0.231
		UT_n	0.072	0.075	0.092	0.092	0.153	0.149	0.344	0.334
		WA_n	0.097	0.091	0.146	0.126	0.309	0.250	0.625	0.533
		$\lvert WA_n \rvert$	0.146	0.125	0.202	0.183	0.409	0.362	0.717	0.659
		WTn	0.140	0.125	0.194	0.183	0.388	0.362	0.689	0.659
		XA_n	0.058	0.128	0.064	0.187	0.103	0.461	0.219	0.787
		$\lvert XA_n \rvert$	0.136	0.134	0.190	0.186	0.385	0.461	0.710	0.793
		XT_n	0.191	0.202	0.318	0.308	0.632	0.635	0.907	0.900

Table 2: Same as Table 1 but with different parameter values.

			$n = 15$		$n = 25$		$n = 50$		$n = 100$	
$\boldsymbol{\mu}_1$	ε	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3,3)$	0.2	MD_2	0.725	0.704	0.959	0.945	0.999	0.997	1.000	1.000
		CD_2	0.727	0.705	0.955	0.943	0.999	0.997	1.000	1.000
		SD_2	0.727	0.728	0.951	0.946	0.998	0.998	1.000	1.000
		UD_2	0.489	0.437	0.764	0.688	0.985	0.974	1.000	1.000
		WD_2	0.493	0.435	0.771	0.687	0.986	0.975	1.000	1.000
		D_2^*	0.616	0.564	0.904	0.881	0.996	0.997	1.000	1.000
		D_∞^*	0.500	0.455	0.817	0.756	0.992	0.988	1.000	1.000
				(0.430)		(0.731)		(0.983)		(1.000)
		$\tilde{D}_\infty^*$	0.476	0.407	0.775	0.722	0.987	0.981	1.000	1.000
				(0.396)		(0.701)		(0.978)		(1.000)
		XD	0.714	0.728	0.939	0.946	0.999	0.998	1.000	1.000
		MA_n	0.686	0.669	0.920	0.908	0.995	0.995	1.000	1.000
		$\lvert MA_n\rvert$	0.641	0.584	0.901	0.867	0.994	0.992	1.000	1.000
		MT_n	0.765	0.750	0.969	0.955	0.999	0.998	1.000	1.000
		CA_n	0.320	0.297	0.457	0.432	0.682	0.663	0.874	0.852
		$\lvert CA_n\rvert$	0.345	0.342	0.471	0.463	0.691	0.679	0.879	0.867
		CT_n	0.701	0.680	0.937	0.931	0.999	0.999	1.000	1.000
		SA_n	0.358	0.359	0.577	0.578	0.895	0.896	0.993	0.991
		$\lvert SA_n\rvert$	0.408	0.401	0.618	0.617	0.912	0.910	0.995	0.994
		ST_n	0.684	0.680	0.930	0.927	0.999	0.999	1.000	1.000
		UA_n	0.070	0.073	0.084	0.081	0.127	0.124	0.241	0.229
		$\lvert UA_n\rvert$	0.089	0.082	0.095	0.091	0.137	0.129	0.252	0.236
		UT_n	0.117	0.115	0.228	0.232	0.700	0.681	0.991	0.988
		WA_n	0.383	0.332	0.666	0.584	0.970	0.950	1.000	1.000
		$\lvert WA_n\rvert$	0.486	0.435	0.759	0.687	0.982	0.975	1.000	1.000
		WT_n	0.470	0.435	0.732	0.687	0.977	0.975	1.000	1.000
		XA_n	0.206	0.501	0.388	0.844	0.787	0.995	0.988	1.000
		$\lvert XA_n\rvert$	0.484	0.474	0.778	0.816	0.986	0.996	1.000	1.000
		XT_n	0.720	0.727	0.947	0.953	0.999	0.998	1.000	1.000

Table 3: Same as Table 1 but with different parameter values.

			$n = 15$		$n = 25$		$n = 50$		$n = 100$	
$\boldsymbol{\mu}_1$	ε	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3,3)$	0.4	MD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		CD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		SD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		UD_2	0.994	0.992	1.000	1.000	1.000	1.000	1.000	1.000
		WD_2	0.995	0.992	1.000	1.000	1.000	1.000	1.000	1.000
		D_2^*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		D_∞^*	0.999	0.996	1.000	1.000	1.000	1.000	1.000	1.000
				(0.996)		(1.000)		(1.000)		(1.000)
		$\tilde{D}_\infty^*$	0.997	0.993	1.000	1.000	1.000	1.000	1.000	1.000
				(0.993)		(1.000)		(1.000)		(1.000)
		XD	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		MA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert MA_n\rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		MT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		CA_n	0.996	0.997	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert CA_n\rvert$	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
		CT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		SA_n	0.946	0.949	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert SA_n\rvert$	0.959	0.959	1.000	1.000	1.000	1.000	1.000	1.000
		ST_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		UA_n	0.180	0.179	0.262	0.280	0.537	0.535	0.818	0.811
		$\lvert UA_n\rvert$	0.143	0.158	0.230	0.243	0.507	0.515	0.808	0.804
		UT_n	0.905	0.895	1.000	1.000	1.000	1.000	1.000	1.000
		WA_n	0.990	0.986	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert WA_n\rvert$	0.995	0.992	1.000	1.000	1.000	1.000	1.000	1.000
		WT_n	0.993	0.992	1.000	1.000	1.000	1.000	1.000	1.000
		XA_n	0.945	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert XA_n\rvert$	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
		XT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: Same as Table 1 but with different parameter values.

			$n = 15$		$n = 25$		$n = 50$		$n = 100$	
$\boldsymbol{\mu}_1$	ε	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3,-1)$	0.1	MD_2	0.217	0.208	0.371	0.353	0.724	0.703	0.970	0.963
		CD_2	0.265	0.253	0.443	0.430	0.777	0.763	0.983	0.979
		SD_2	0.262	0.260	0.433	0.424	0.769	0.766	0.983	0.982
		UD_2	0.185	0.165	0.262	0.229	0.517	0.467	0.854	0.798
		WD_2	0.181	0.156	0.263	0.224	0.517	0.469	0.858	0.795
		D_2^*	0.096	0.098	0.145	0.142	0.374	0.366	0.712	0.715
		D_∞^*	0.112	0.122	0.193	0.203	0.466	0.465	0.857	0.858
				(0.106)		(0.185)		(0.423)		(0.822)
		$\tilde{D}_\infty^*$	0.090	0.102	0.139	0.166	0.375	0.385	0.777	0.783
				(0.093)		(0.156)		(0.366)		(0.752)
		XD	0.243	0.260	0.411	0.424	0.746	0.766	0.961	0.959
		MA_n	0.047	0.083	0.045	0.087	0.050	0.186	0.063	0.314
		$\lvert MA_n\rvert$	0.039	0.063	0.042	0.072	0.048	0.167	0.061	0.295
		MT_n	0.193	0.187	0.319	0.317	0.658	0.653	0.947	0.948
		CA_n	0.111	0.100	0.121	0.115	0.140	0.128	0.198	0.188
		$\lvert CA_n\rvert$	0.115	0.122	0.123	0.123	0.144	0.137	0.204	0.199
		CT_n	0.242	0.235	0.423	0.403	0.767	0.741	0.982	0.978
		SA_n	0.125	0.127	0.167	0.169	0.307	0.315	0.559	0.565
		$\lvert SA_n\rvert$	0.149	0.146	0.203	0.200	0.334	0.329	0.584	0.587
		ST_n	0.237	0.230	0.406	0.403	0.750	0.740	0.982	0.981
		UA_n	0.077	0.076	0.076	0.071	0.118	0.106	0.174	0.165
		$\lvert UA_n\rvert$	0.087	0.090	0.088	0.080	0.123	0.119	0.187	0.179
		UT_n	0.078	0.075	0.089	0.088	0.162	0.156	0.411	0.408
		WA_n	0.116	0.097	0.183	0.150	0.408	0.334	0.766	0.698
		$\lvert WA_n\rvert$	0.172	0.156	0.252	0.224	0.504	0.469	0.846	0.795
		WT_n	0.167	0.156	0.240	0.224	0.493	0.469	0.834	0.795
		XA_n	0.103	0.106	0.155	0.154	0.334	0.290	0.666	0.650
		$\lvert XA_n\rvert$	0.122	0.144	0.153	0.191	0.301	0.352	0.622	0.691
		XT_n	0.215	0.204	0.378	0.356	0.729	0.699	0.972	0.967

Table 5: Same as Table 1 but with different parameter values.

			$n=15$		$n=25$		$n=50$		$n=100$	
$\boldsymbol{\mu}_1$	ε	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3,-1)$	0.2	MD_2	0.824	0.809	0.990	0.982	1.000	1.000	1.000	1.000
		CD_2	0.871	0.852	0.995	0.995	1.000	1.000	1.000	1.000
		SD_2	0.860	0.858	0.994	0.994	1.000	1.000	1.000	1.000
		UD_2	0.633	0.574	0.872	0.813	0.997	0.995	1.000	1.000
		WD_2	0.646	0.575	0.875	0.821	0.998	0.996	1.000	1.000
		D_2^*	0.459	0.472	0.774	0.792	0.992	0.993	1.000	1.000
		D_∞^*	0.487	0.527	0.813	0.820	0.997	0.996	1.000	1.000
				(0.478)		(0.805)		(0.995)		(1.000)
		$\tilde{D}_\infty^*$	0.351	0.386	0.679	0.693	0.983	0.982	1.000	1.000
				(0.371)		(0.669)		(0.981)		(1.000)
		XD	0.848	0.858	0.991	0.994	1.000	1.000	1.000	1.000
		MA_n	0.049	0.165	0.049	0.287	0.107	0.641	0.194	0.932
		$\lvert MA_n\rvert$	0.026	0.125	0.044	0.239	0.083	0.601	0.544	0.936
		MT_n	0.774	0.770	0.973	0.966	1.000	1.000	1.000	1.000
		CA_n	0.535	0.509	0.722	0.696	0.929	0.913	0.998	0.997
		$\lvert CA_n\rvert$	0.562	0.559	0.739	0.728	0.934	0.922	0.998	0.997
		CT_n	0.842	0.834	0.991	0.987	1.000	1.000	1.000	1.000
		SA_n	0.428	0.433	0.660	0.664	0.941	0.940	0.998	0.998
		$\lvert SA_n\rvert$	0.475	0.478	0.696	0.695	0.951	0.947	0.998	0.998
		ST_n	0.837	0.828	0.987	0.986	1.000	1.000	1.000	1.000
		UA_n	0.052	0.054	0.038	0.029	0.045	0.052	0.061	0.062
		$\lvert UA_n\rvert$	0.056	0.058	0.039	0.037	0.050	0.056	0.062	0.060
		UT_n	0.165	0.166	0.352	0.350	0.879	0.874	1.000	1.000
		WA_n	0.518	0.458	0.803	0.736	0.995	0.991	1.000	1.000
		$\lvert WA_n\rvert$	0.629	0.575	0.867	0.821	0.998	0.996	1.000	1.000
		WT_n	0.601	0.575	0.846	0.821	0.997	0.996	1.000	1.000
		XA_n	0.477	0.551	0.759	0.764	0.989	0.988	1.000	1.000
		$\lvert XA_n\rvert$	0.519	0.638	0.765	0.824	0.988	0.992	1.000	1.000
		XT_n	0.816	0.798	0.983	0.979	1.000	1.000	1.000	1.000

Table 6: Same as Table 1 but with different parameter values.

			$n = 15$		$n = 25$		$n = 50$		$n = 100$	
$\boldsymbol{\mu}_1$	ε	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3, -1)$	0.4	MD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		CD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		SD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		UD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		WD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		D_2^*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		D_∞^*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
				(0.999)		(1.000)		(1.000)		(1.000)
		$\tilde{D}_\infty^*$	0.966	0.963	1.000	0.999	1.000	1.000	1.000	1.000
				(0.957)		(0.999)		(1.000)		(1.000)
		XD	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		MA_n	0.038	0.444	0.066	0.818	0.245	0.998	0.636	1.000
		$\lvert MA_n\rvert$	0.023	0.335	0.047	0.759	0.890	0.998	1.000	1.000
		MT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		CA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert CA_n\rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		CT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		SA_n	0.978	0.982	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert SA_n\rvert$	0.990	0.989	1.000	1.000	1.000	1.000	1.000	1.000
		ST_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		UA_n	0.533	0.632	0.779	0.862	0.971	0.988	1.000	1.000
		$\lvert UA_n\rvert$	0.478	0.600	0.758	0.838	0.969	0.986	1.000	1.000
		UT_n	0.997	0.997	1.000	1.000	1.000	1.000	1.000	1.000
		WA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert WA_n\rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		WT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		XA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		$\lvert XA_n\rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		XT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

case, the empirical powers were estimated by 1000 simulations with $R = 999$. The results are shown in Tables 1–6.

In the second scenario, the null model was the Morgenstern distribution (Justel et al., 1997), the joint probability density function of which is

$$f(x_1, x_2) = 1 + a(2x_1 - 1)(2x_2 - 1), \qquad 0 \le x_1, x_2 \le 1, \;\; -1 \le a \le 1,$$

where $a = 0.5$ was used in the simulations. The distribution functions are

$$\begin{aligned} F_1(x_1) &= x_1, \\ F_2(x_2|x_1) &= [1 - a(2x_1 - 1)]x_2 + a(2x_1 - 1)x_2^2. \end{aligned}$$

The alternative model was the product of two independent Beta(α, β) distributions. Five cases were considered: $(\alpha, \beta) = (10, 10)$, $(3, 3)$, $(3, 2)$, $(0.5, 1)$ and $(0.5, 0.5)$. In each case, the empirical powers were estimated by 1000 simulations with $R = 999$. The results are shown in Tables 7–11.

In the tables, we have new symbols XD, XA_n and XT_n, which will be explained later in this section. We also reported the powers of the Kolmogorov–Smirnov statistics $D_{\max}^{\mathrm{KS}}$ and its approximation $\tilde{D}_{\max}^{\mathrm{KS}}$ by multiple tests with the Bonferroni adjustment suggested by Justel et al. (1997), in which the percentiles of the distributions of $D_{\max}^{\mathrm{KS}}$ and $\tilde{D}_{\max}^{\mathrm{KS}}$ for n independent and uniformly distributed points in $[0, 1)^2$ can be found. We, however, used our own simulated percentiles, which are nevertheless very close to theirs; our estimates of the powers under the multiple testing formulation, however, are not always close to theirs. Unfortunately, some implementation details, such as the number of simulation, used by them more than ten years ago are no longer available (Justel, 2007) and so, other than random error, we are not able to identify any causes of the differences.

We can see from these tables that for $D_{\max}^{\mathrm{KS}}$ and $\tilde{D}_{\max}^{\mathrm{KS}}$ the powers of the multiple tests are always not higher than those of the corresponding Monte Carlo tests. The reason is clear. The loss in power is the price for the conservativeness caused by the Bonferroni adjustment.

Tables 12 and 13 show the estimated actual type I error rates of the Monte Carlo tests at $\alpha = 0.05$, with $R = 999$, from 10,000 simulations, as well as the estimated actual type I error rates of the multiple tests with the Bonferroni adjustment. The results confirm that the Bonferroni adjustment will lead to conservative and hence less powerful tests. We also obtained but chose not to bother the reader with the powers of other $D_{\max}$ statistics using the Bonferroni adjustment, because, as expected, the powers are all lower than the corresponding Monte Carlo tests. However, the variants A_n and T_n have the merit of the known usual limiting distributions and hence can be conveniently applied, with the Bonferroni adjustment, in practice. Thus, it may be of interest to see the size distortion if the asymptotic distributions are used to approximate the p-values for finite samples. Figure 1 shows the p-value plot (Davidson and MacKinnon, 1998) in which the p-values of the A_n variant of the MD_2 discrepancy for 10,000 independent samples of size $n = 100$ from the null model $\mathrm{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma})$ are approximated by the asymptotic standard normal distribution. It is evident that the Bonferroni adjustment will cause serious negative size distortion, leading to overly

Table 7: Estimated power, by 1000 simulations, of the Monte Carlo test with $R = 999$ of the test statistics D_{sum} and $D_{\max}$ using the discrepancy $D^{(j)}$ at the 0.05 significance level for a sample of size n. The null hypothesis is the Morgenstern distribution and the alternative is the product of two independent Beta(α, β). The values in parentheses are the powers of the multiple tests suggested by Justel et al. (1997) with the Bonferroni adjustment.

		$n = 10$		$n = 20$		$n = 50$		$n = 100$	
(α, β)	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(10, 10)$	MD_2	0.727	0.700	1.000	1.000	1.000	1.000	1.000	1.000
	CD_2	0.531	0.466	1.000	1.000	1.000	1.000	1.000	1.000
	SD_2	0.923	0.915	1.000	1.000	1.000	1.000	1.000	1.000
	UD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	WD_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	D_2^*	0.693	0.636	1.000	1.000	1.000	1.000	1.000	1.000
	D_∞^*	0.867	0.829	1.000	1.000	1.000	1.000	1.000	1.000
			(0.746)		(1.000)		(1.000)		(1.000)
	$\tilde{D}_\infty^*$	0.309	0.363	0.734	0.749	1.000	0.999	1.000	1.000
			(0.322)		(0.735)		(0.998)		(1.000)
	XD	0.988	0.915	1.000	1.000	1.000	1.000	1.000	1.000
	MA_n	0.057	0.043	0.540	0.498	0.999	0.999	1.000	1.000
	$\lvert MA_n \rvert$	0.125	0.107	0.668	0.642	1.000	0.999	1.000	1.000
	MT_n	0.825	0.796	1.000	1.000	1.000	1.000	1.000	1.000
	CA_n	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
	$\lvert CA_n \rvert$	0.997	0.993	1.000	1.000	1.000	1.000	1.000	1.000
	CT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	SA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$\lvert SA_n \rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	ST_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	UA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$\lvert UA_n \rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	UT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	WA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$\lvert WA_n \rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	WT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	XA_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$\lvert XA_n \rvert$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	XT_n	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 8: Same as Table 7 but with different parameter values.

		$n = 10$		$n = 20$		$n = 50$		$n = 100$	
(α, β)	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3,3)$	MD_2	0.047	0.040	0.306	0.293	0.987	0.985	1.000	1.000
	CD_2	0.023	0.022	0.227	0.212	0.976	0.973	1.000	1.000
	SD_2	0.103	0.105	0.489	0.486	0.999	0.999	1.000	1.000
	UD_2	0.705	0.688	0.980	0.975	1.000	1.000	1.000	1.000
	WD_2	0.655	0.622	0.962	0.955	1.000	1.000	1.000	1.000
	D_2^*	0.046	0.040	0.278	0.248	0.985	0.977	1.000	1.000
	D_∞^*	0.150	0.148	0.449	0.439	0.968	0.961	1.000	1.000
			(0.086)		(0.348)		(0.929)		(1.000)
	$\tilde{D}_\infty^*$	0.103	0.124	0.248	0.294	0.824	0.814	1.000	1.000
			(0.106)		(0.257)		(0.789)		(0.999)
	XD	0.224	0.105	0.729	0.486	1.000	0.999	1.000	1.000
	MA_n	0.024	0.024	0.087	0.082	0.359	0.352	0.812	0.798
	$\lvert MA_n\rvert$	0.044	0.046	0.117	0.115	0.408	0.404	0.842	0.835
	MT_n	0.064	0.059	0.343	0.322	0.986	0.986	1.000	1.000
	CA_n	0.641	0.630	0.944	0.941	1.000	1.000	1.000	1.000
	$\lvert CA_n\rvert$	0.532	0.482	0.915	0.905	1.000	0.999	1.000	1.000
	CT_n	0.643	0.639	0.942	0.934	1.000	1.000	1.000	1.000
	SA_n	0.863	0.860	1.000	1.000	1.000	1.000	1.000	1.000
	$\lvert SA_n\rvert$	0.892	0.890	1.000	1.000	1.000	1.000	1.000	1.000
	ST_n	0.783	0.772	0.997	0.997	1.000	1.000	1.000	1.000
	UA_n	0.908	0.907	1.000	1.000	1.000	1.000	1.000	1.000
	$\lvert UA_n\rvert$	0.929	0.926	1.000	1.000	1.000	1.000	1.000	1.000
	UT_n	0.912	0.908	1.000	1.000	1.000	1.000	1.000	1.000
	WA_n	0.528	0.505	0.924	0.912	1.000	1.000	1.000	1.000
	$\lvert WA_n\rvert$	0.655	0.622	0.962	0.955	1.000	1.000	1.000	1.000
	WT_n	0.648	0.622	0.960	0.955	1.000	1.000	1.000	1.000
	XA_n	0.650	0.730	0.985	0.995	1.000	1.000	1.000	1.000
	$\lvert XA_n\rvert$	0.835	0.814	0.999	1.000	1.000	1.000	1.000	1.000
	XT_n	0.702	0.585	0.985	0.971	1.000	1.000	1.000	1.000

Table 9: Same as Table 7 but with different parameter values.

		$n = 10$		$n = 20$		$n = 50$		$n = 100$	
(α, β)	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(3, 2)$	MD_2	0.192	0.179	0.656	0.641	0.998	0.998	1.000	1.000
	CD_2	0.182	0.179	0.618	0.604	0.998	0.998	1.000	1.000
	SD_2	0.299	0.292	0.753	0.751	0.998	0.998	1.000	1.000
	UD_2	0.572	0.553	0.913	0.904	1.000	1.000	1.000	1.000
	WD_2	0.524	0.499	0.902	0.887	1.000	1.000	1.000	1.000
	D_2^*	0.092	0.082	0.525	0.470	0.994	0.991	1.000	1.000
	D_∞^*	0.058	0.053	0.222	0.205	0.904	0.879	1.000	1.000
			(0.034)		(0.145)		(0.807)		(0.999)
	$\tilde{D}_\infty^*$	0.065	0.055	0.242	0.213	0.912	0.876	1.000	1.000
			(0.043)		(0.185)		(0.837)		(1.000)
	XD	0.368	0.292	0.829	0.751	0.998	0.998	1.000	1.000
	MA_n	0.035	0.033	0.061	0.055	0.263	0.252	0.609	0.608
	$\lvert MA_n\rvert$	0.016	0.012	0.036	0.029	0.215	0.203	0.581	0.564
	MT_n	0.226	0.209	0.643	0.630	0.997	0.997	1.000	1.000
	CA_n	0.205	0.197	0.349	0.346	0.626	0.613	0.879	0.874
	$\lvert CA_n\rvert$	0.141	0.128	0.303	0.286	0.586	0.581	0.875	0.860
	CT_n	0.487	0.482	0.894	0.887	1.000	1.000	1.000	1.000
	SA_n	0.700	0.699	0.985	0.985	1.000	1.000	1.000	1.000
	$\lvert SA_n\rvert$	0.753	0.749	0.989	0.989	1.000	1.000	1.000	1.000
	ST_n	0.608	0.603	0.970	0.970	1.000	1.000	1.000	1.000
	UA_n	0.638	0.638	0.938	0.940	1.000	1.000	1.000	1.000
	$\lvert UA_n\rvert$	0.681	0.678	0.951	0.951	1.000	1.000	1.000	1.000
	UT_n	0.608	0.599	0.927	0.926	1.000	1.000	1.000	1.000
	WA_n	0.411	0.384	0.828	0.800	0.999	0.999	1.000	1.000
	$\lvert WA_n\rvert$	0.524	0.499	0.902	0.887	1.000	1.000	1.000	1.000
	WT_n	0.519	0.499	0.900	0.887	1.000	1.000	1.000	1.000
	XA_n	0.388	0.512	0.837	0.922	1.000	1.000	1.000	1.000
	$\lvert XA_n\rvert$	0.612	0.629	0.951	0.958	1.000	1.000	1.000	1.000
	XT_n	0.545	0.452	0.939	0.906	1.000	1.000	1.000	1.000

Table 10: Same as Table 7 but with different parameter values.

		$n = 10$		$n = 20$		$n = 50$		$n = 100$	
(α, β)	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(0.5, 1)$	MD_2	0.629	0.632	0.893	0.893	1.000	0.999	1.000	1.000
	CD_2	0.632	0.622	0.902	0.898	1.000	1.000	1.000	1.000
	SD_2	0.593	0.595	0.888	0.890	1.000	1.000	1.000	1.000
	UD_2	0.327	0.313	0.630	0.615	0.965	0.963	1.000	1.000
	WD_2	0.339	0.320	0.631	0.615	0.965	0.963	1.000	1.000
	D_2^*	0.613	0.619	0.842	0.843	0.996	0.997	1.000	1.000
	D_∞^*	0.481	0.482	0.746	0.743	0.984	0.984	1.000	1.000
			(0.411)		(0.667)		(0.975)		(1.000)
	$\tilde{D}_\infty^*$	0.449	0.422	0.727	0.709	0.983	0.978	1.000	1.000
			(0.382)		(0.674)		(0.973)		(1.000)
	XD	0.600	0.595	0.882	0.890	0.999	1.000	1.000	1.000
	MA_n	0.534	0.536	0.799	0.800	0.990	0.991	1.000	1.000
	$\lvert MA_n \rvert$	0.585	0.594	0.840	0.846	0.993	0.992	1.000	1.000
	MT_n	0.569	0.572	0.863	0.863	0.998	0.998	1.000	1.000
	CA_n	0.539	0.546	0.826	0.821	0.991	0.991	1.000	1.000
	$\lvert CA_n \rvert$	0.599	0.600	0.852	0.854	0.991	0.995	1.000	1.000
	CT_n	0.571	0.573	0.880	0.880	1.000	1.000	1.000	1.000
	SA_n	0.123	0.117	0.152	0.156	0.279	0.284	0.441	0.451
	$\lvert SA_n \rvert$	0.097	0.094	0.136	0.128	0.269	0.272	0.436	0.441
	ST_n	0.638	0.636	0.911	0.913	1.000	1.000	1.000	1.000
	UA_n	0.479	0.482	0.740	0.733	0.971	0.972	1.000	1.000
	$\lvert UA_n \rvert$	0.437	0.441	0.720	0.718	0.968	0.967	1.000	1.000
	UT_n	0.515	0.523	0.801	0.801	0.983	0.983	1.000	1.000
	WA_n	0.229	0.222	0.520	0.503	0.944	0.936	1.000	1.000
	$\lvert WA_n \rvert$	0.338	0.320	0.631	0.615	0.963	0.963	1.000	1.000
	WT_n	0.334	0.320	0.626	0.615	0.963	0.963	1.000	1.000
	XA_n	0.267	0.573	0.494	0.852	0.922	0.999	1.000	1.000
	$\lvert XA_n \rvert$	0.578	0.595	0.858	0.879	1.000	0.999	1.000	1.000
	XT_n	0.627	0.628	0.904	0.892	1.000	0.999	1.000	1.000

Table 11: Same as Table 7 but with different parameter values.

		$n=10$		$n=20$		$n=50$		$n=100$	
(α,β)	$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
$(0.5,0.5)$	MD_2	0.208	0.210	0.319	0.319	0.762	0.753	0.993	0.992
	CD_2	0.236	0.237	0.371	0.373	0.795	0.795	0.995	0.993
	SD_2	0.224	0.229	0.355	0.355	0.815	0.819	0.996	0.995
	UD_2	0.355	0.349	0.684	0.669	0.985	0.985	1.000	1.000
	WD_2	0.366	0.351	0.683	0.672	0.984	0.983	1.000	1.000
	D_2^*	0.155	0.157	0.240	0.236	0.593	0.573	0.952	0.945
	D_∞^*	0.243	0.239	0.372	0.364	0.767	0.760	0.986	0.984
			(0.192)		(0.308)		(0.694)		(0.973)
	$\tilde{D}_\infty^*$	0.179	0.179	0.281	0.263	0.693	0.688	0.984	0.977
			(0.153)		(0.240)		(0.654)		(0.956)
	XD	0.268	0.229	0.492	0.355	0.912	0.819	1.000	1.000
	MA_n	0.169	0.166	0.194	0.196	0.327	0.334	0.510	0.518
	$\lvert MA_n\rvert$	0.139	0.139	0.170	0.168	0.309	0.313	0.505	0.496
	MT_n	0.160	0.160	0.229	0.225	0.641	0.633	0.979	0.977
	CA_n	0.331	0.333	0.656	0.655	0.971	0.963	1.000	1.000
	$\lvert CA_n\rvert$	0.411	0.408	0.691	0.690	0.971	0.974	1.000	1.000
	CT_n	0.318	0.330	0.572	0.574	0.934	0.934	0.999	1.000
	SA_n	0.537	0.540	0.863	0.861	1.000	1.000	1.000	1.000
	$\lvert SA_n\rvert$	0.479	0.483	0.840	0.844	1.000	1.000	1.000	1.000
	ST_n	0.556	0.558	0.829	0.834	0.994	0.993	1.000	1.000
	UA_n	0.681	0.681	0.913	0.912	1.000	1.000	1.000	1.000
	$\lvert UA_n\rvert$	0.639	0.644	0.903	0.903	1.000	1.000	1.000	1.000
	UT_n	0.694	0.700	0.911	0.915	1.000	0.999	1.000	1.000
	WA_n	0.281	0.270	0.596	0.578	0.959	0.957	0.999	0.999
	$\lvert WA_n\rvert$	0.366	0.351	0.682	0.672	0.984	0.983	1.000	1.000
	WT_n	0.366	0.351	0.679	0.672	0.984	0.983	1.000	1.000
	XA_n	0.188	0.595	0.242	0.856	0.192	0.996	0.232	1.000
	$\lvert XA_n\rvert$	0.551	0.508	0.837	0.820	0.997	0.996	1.000	1.000
	XT_n	0.484	0.455	0.775	0.748	0.990	0.988	1.000	1.000

Table 12: Estimated actual type I error rate of the Monte Carlo tests with $R = 999$ of the test statistics D_{sum} and $D_{\max}$ using the discrepancy $D^{(j)}$ at the 0.05 nominal significance level for a sample of size n by 10000 simulations. The null hypothesis is the $\text{N}((0,0), \mathbf{\Sigma})$.

	$n = 15$		$n = 25$		$n = 50$		$n = 100$	
$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
MD_2	0.055	0.058	0.046	0.040	0.044	0.042	0.052	0.053
CD_2	0.044	0.049	0.049	0.045	0.041	0.041	0.054	0.055
SD_2	0.047	0.045	0.046	0.045	0.042	0.042	0.052	0.050
UD_2	0.047	0.051	0.043	0.042	0.040	0.045	0.041	0.046
WD_2	0.049	0.043	0.044	0.042	0.040	0.046	0.042	0.047
D_2^*	0.056	0.057	0.041	0.041	0.040	0.040	0.048	0.051
D_∞^*	0.057	0.054	0.048	0.048	0.043	0.045	0.051	0.049
		(0.041)		(0.038)		(0.039)		(0.045)
$\tilde{D}_\infty^*$	0.050	0.051	0.050	0.048	0.044	0.042	0.050	0.051
		(0.043)		(0.045)		(0.037)		(0.048)
XD	0.049	0.045	0.046	0.045	0.045	0.042	0.054	0.050
MA_n	0.056	0.056	0.044	0.047	0.046	0.044	0.050	0.049
$\lvert MA_n\rvert$	0.054	0.053	0.041	0.045	0.046	0.047	0.049	0.050
MT_n	0.056	0.061	0.044	0.046	0.040	0.044	0.051	0.046
CA_n	0.054	0.047	0.051	0.041	0.035	0.043	0.041	0.042
$\lvert CA_n\rvert$	0.045	0.044	0.050	0.045	0.042	0.044	0.037	0.040
CT_n	0.052	0.052	0.044	0.043	0.042	0.040	0.050	0.050
SA_n	0.059	0.056	0.060	0.060	0.049	0.046	0.042	0.041
$\lvert SA_n\rvert$	0.060	0.058	0.058	0.055	0.050	0.051	0.043	0.045
ST_n	0.053	0.052	0.050	0.048	0.043	0.044	0.046	0.045
UA_n	0.049	0.048	0.056	0.057	0.055	0.050	0.041	0.036
$\lvert UA_n\rvert$	0.052	0.054	0.055	0.052	0.052	0.041	0.037	0.039
UT_n	0.046	0.049	0.055	0.051	0.050	0.046	0.037	0.034
WA_n	0.046	0.053	0.039	0.039	0.037	0.043	0.051	0.052
$\lvert WA_n\rvert$	0.050	0.051	0.045	0.048	0.040	0.043	0.043	0.045
WT_n	0.051	0.043	0.044	0.042	0.040	0.046	0.045	0.047
XA_n	0.061	0.042	0.052	0.042	0.056	0.037	0.047	0.034
$\lvert XA_n\rvert$	0.060	0.056	0.052	0.053	0.041	0.046	0.044	0.048
XT_n	0.050	0.058	0.049	0.042	0.041	0.041	0.049	0.052

Table 13: Estimated actual type I error rate of the Monte Carlo tests with $R = 999$ of the test statistics D_{sum} and $D_{\max}$ using the discrepancy $D^{(j)}$ at the 0.05 nominal significance level for a sample of size n by 10000 simulations The null hypothesis is the Morgenstern with parameter $a = 0.5$ and uniform marginals.

	$n = 10$		$n = 20$		$n = 50$		$n = 100$	
$D^{(j)}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$	D_{sum}	$D_{\max}$
MD_2	0.070	0.069	0.052	0.049	0.046	0.050	0.041	0.039
CD_2	0.063	0.063	0.049	0.051	0.052	0.054	0.046	0.043
SD_2	0.058	0.060	0.050	0.050	0.054	0.053	0.042	0.043
UD_2	0.058	0.057	0.044	0.046	0.064	0.068	0.052	0.051
WD_2	0.058	0.050	0.045	0.047	0.063	0.066	0.051	0.052
D_2^*	0.063	0.064	0.059	0.059	0.048	0.045	0.048	0.046
D_∞^*	0.065	0.060	0.046	0.043	0.051	0.049	0.046	0.049
		(0.041)		(0.029)		(0.037)		(0.039)
$\tilde{D}_\infty^*$	0.066	0.069	0.053	0.052	0.047	0.051	0.050	0.048
		(0.055)		(0.044)		(0.043)		(0.041)
XD	0.057	0.060	0.046	0.050	0.055	0.053	0.048	0.043
MA_n	0.059	0.060	0.050	0.052	0.056	0.051	0.052	0.054
$\lvert MA_n\rvert$	0.059	0.055	0.054	0.058	0.056	0.059	0.050	0.046
MT_n	0.066	0.064	0.050	0.053	0.048	0.049	0.042	0.041
CA_n	0.065	0.067	0.060	0.055	0.056	0.060	0.053	0.051
$\lvert CA_n\rvert$	0.063	0.058	0.058	0.055	0.054	0.053	0.051	0.048
CT_n	0.064	0.062	0.048	0.050	0.057	0.057	0.045	0.044
SA_n	0.053	0.055	0.047	0.048	0.061	0.056	0.044	0.045
$\lvert SA_n\rvert$	0.057	0.059	0.047	0.048	0.061	0.052	0.045	0.048
ST_n	0.063	0.061	0.049	0.052	0.062	0.062	0.042	0.042
UA_n	0.066	0.069	0.046	0.053	0.054	0.057	0.043	0.043
$\lvert UA_n\rvert$	0.066	0.062	0.051	0.052	0.057	0.059	0.044	0.047
UT_n	0.060	0.060	0.050	0.049	0.056	0.058	0.043	0.044
WA_n	0.051	0.053	0.044	0.046	0.061	0.058	0.049	0.045
$\lvert WA_n\rvert$	0.058	0.057	0.045	0.050	0.063	0.061	0.051	0.048
WT_n	0.057	0.050	0.044	0.047	0.062	0.066	0.052	0.052
XA_n	0.054	0.057	0.047	0.038	0.055	0.056	0.049	0.046
$\lvert XA_n\rvert$	0.062	0.060	0.050	0.049	0.058	0.054	0.049	0.048
XT_n	0.067	0.066	0.051	0.046	0.053	0.056	0.042	0.044

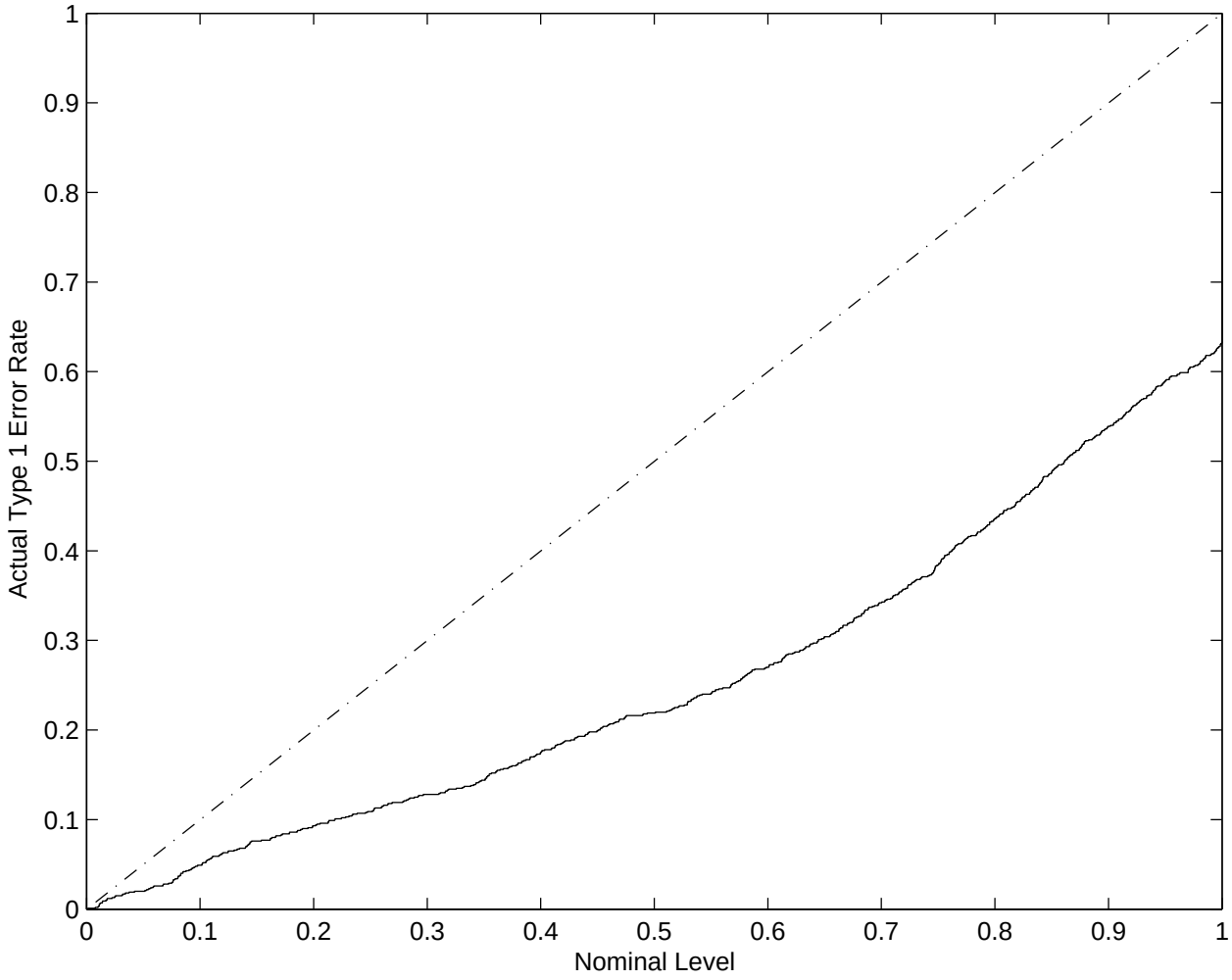

Figure 1: The actual type I error rate, estimated by 10,000 simulation, vs the nominal level of the multiple test using MA_n. The p-values are calculated by the asymptotic standard normal distribution. Each sample contains $n = 100$ data drawn from $\mathrm{N}((0,0), \mathbf{\Sigma})$. The dash-dotted line is a reference line of slope 1.

conservative and hence less powerful tests. The same is true for other discrepancies. Thus, we do not recommend the multiple test formulation.

The approximation $\tilde{D}^{\mathrm{KS}}_{\max}$ of the Kolmogorov–Smirnov statistic, as expected, is less powerful than the Kolmogorov–Smirnov statistic $D^{\mathrm{KS}}_{\max}$ itself, which is neither superior nor inferior to (the square root of) the original Cramér–von Mises statistic. Also, neither the sum D_{sum} nor the maximum $D_{\max}$ over all $s!$ permutation shows a definite advantage over the other, although the sum D_{sum} often gives a slightly higher power.

What is more interesting should be the performance of the generalized Cramér–von Mises statistics by using discrepancies and their variants A_n and T_n, for MD_2, CD_2, SD_2, UD_2 and WD_2, to measure the goodness-of-fit. Let us first compare between different discrepancies and later we will discuss whether we should use the original discrepancies or their variants as the test statistics.

It is found that the powers of at least one of these five discrepancies are better than the Cramér–von Mises statistic and the Kolmogorov–Smirnov statistic in each case. More precisely, the powers of the statistics using MD_2, CD_2 and SD_2 are usually the highest in detecting the difference in means for the bivariate normal distribution (see Tables 1–6). The powers of the Kolmogorov–Smirnov and the original Cramér–von Mises statistics sometimes, especially when $n = 15$, can be much smaller.

The superiority of MD_2, CD_2 and SD_2, however, is not universal; the powers of the statistics using MD_2, CD_2 and SD_2 are much smaller than those using WD_2 and UD_2, when we test the null hypothesis of the Morgenstern distribution against the product of two independent Beta(α, β), where $(\alpha, \beta) = (10, 10)$, $(3, 3)$ and $(3, 2)$ (see Tables 7–9). Nevertheless, For $(\alpha, \beta) = (0.5, 0.5)$ (see Table 11) the powers of the statistics using MD_2, CD_2 and SD_2 are slightly smaller than WD_2 and UD_2, and for $(\alpha, \beta) = (0.5, 1)$ (see Table 10), MD_2, CD_2 and SD_2 outperform WD_2 and UD_2.

The different performance of these discrepancies is caused by the different features of the alternative distributions. We divide the unit square into two regions, namely, the inner square $[1/2 - \sqrt{2}/4, 1/2 + \sqrt{2}/4]^2$ of area 0.5 and its complement. If the points are independent and uniformly distributed, then approximately 50% of the points would be in the inner square. Tables 14 and 15 show the average proportion of points in all transformed data $\{T_j(\mathbf{y}_1), \ldots, T_j(\mathbf{y}_n)\}$ for $j = 1, \ldots, s!$ lying in the inner square. It is found that the test statistics using WD_2 and UD_2 are much more powerful when the transformed points are concentrated on the centre of the unit square $[0, 1)^2$. Thus, if the scatter plot of the transformed data shows a high point density in the centre, we should use WD_2 or UD_2 to measure the discrepancy and construct the test statistic; otherwise we should use MD_2, CD_2 or SD_2.

We empirically put MD_2, CD_2 and SD_2 into one group, and WD_2 and UD_2 into another group. This grouping, interestingly, agrees with the grouping in the context of numerical integration. The discrepancies MD_2, CD_2 and SD_2 belong to the group of discrepancies used to describe the error bounds of the quasi-Monte Carlo integration for non-periodic integrands, whilst WD_2 and UD_2 belong to the one for periodic integrands (Hickernell, 1999b). Let us simply call these two groups "non-periodic" and "periodic". The relationship

Table 14: Average proportion, in 1000 simulations, of n transformed points lying in the inner square $[1/2 - \sqrt{2}/4, 1/2 + \sqrt{2}/4]^2$ of area 0.5, where the n points were generated from the mixture $(1 - \varepsilon)\mathrm{N}((0,0), \boldsymbol{\Sigma}) + \varepsilon\mathrm{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$, and were transformed according to the null distribution $\mathrm{N}((0,0), \boldsymbol{\Sigma})$.

$\boldsymbol{\mu}_1$	ε	$n = 15$	$n = 25$	$n = 50$	$n = 100$
	0	0.50007	0.50003	0.49999	0.50001
(3,3)	0.1	0.54800	0.54450	0.54391	0.54356
	0.2	0.52013	0.51902	0.51828	0.51997
	0.4	0.28387	0.28280	0.28091	0.28704
(3,-1)	0.1	0.53933	0.53796	0.53477	0.53357
	0.2	0.47793	0.48070	0.47687	0.47685
	0.4	0.17823	0.17838	0.17989	0.18022

Table 15: Average proportion, in 1000 simulations, of n transformed points lying in the inner square $[1/2 - \sqrt{2}/4, 1/2 + \sqrt{2}/4]^2$ of area 0.5, where the n points were generated from the distribution function F_A and were transformed according to the Morgenstern distribution with parameter $a = 0.5$ and uniform marginals.

F_A	$n = 10$	$n = 20$	$n = 50$	$n = 100$
Morgenstern	0.50004	0.49990	0.50002	0.50003
$\mathrm{Beta}(10,10) \times \mathrm{Beta}(10,10)$	0.99940	0.99948	0.99942	0.99939
$\mathrm{Beta}(3,3) \times \mathrm{Beta}(3,3)$	0.89165	0.89773	0.89870	0.89808
$\mathrm{Beta}(3,2) \times \mathrm{Beta}(3,2)$	0.77880	0.78350	0.78355	0.78199
$\mathrm{Beta}(0.5,1) \times \mathrm{Beta}(0.5,1)$	0.29470	0.29095	0.29345	0.29286
$\mathrm{Beta}(0.5,0.5) \times \mathrm{Beta}(0.5,0.5)$	0.24830	0.25015	0.24746	0.24636

between the distribution of the locations of the points and the periodicity of the integrand, however, is not clear and will be left for future endeavour.

Because sometimes WD_2 and UD_2 are better and sometimes MD_2, CD_2 and SD_2 are better, it makes sense to combine the five discrepancies by using

$$XD = \begin{cases} XD_{\max} &= \max\{MD_2, CD_2, SD_2, UD_2, WD_2\}, \\ XD_{\mathrm{sum}} &= MD_2 + CD_2 + SD_2 + UD_2 + WD_2, \end{cases}$$

and consider $D^{(j)} = XD = XD_{\max}$ for $D_{\max}$ and $D^{(j)} = XD = XD_{\mathrm{sum}}$ for D_{sum}. The powers of the statistics using XD were also tabulated in Tables 1–11. As we can see from these tables, the powers of the statistics using XD_{sum} are often slightly higher than those using $XD_{\max}$, also often higher than the Kolmogorov–Smirnov statistic and the original Cramér–von Mises statistic, but usually less than the most powerful one among the statistics using

MD_2, CD_2, SD_2, UD_2 and WD_2, especially when MD_2, CD_2 and SD_2 have low powers.

A referee suggested that because $MD_2 \geq CD_2$ with very high probability and $WD_2 \geq UD_2$, we may exclude CD_2 and UD_2 from $XD_{\max}$. Moreover, because MD_2 and CD_2, as well as WD_2 and UD_2, have similar powers, we may also exclude CD_2 and UD_2 from XD_{sum}. Our simulation results (not reported here) reveal that he is quite right with $XD_{\max}$ such that its powers with and without CD_2 and UD_2 are almost identical in the examples considered. For XD_{sum}, if we exclude CD_2 and UD_2, we may occasionally loss a little bit power.

Given a sample of size $n = 100$, we compared the computation times for carrying out one Monte Carlo test with $R = 999$, implemented in MATLAB on a Pentium IV 2.66GHz desktop computer. For any one of MD_2, CD_2, SD_2, UD_2, WD_2 and D_2^*, it took about 30 seconds; for $\tilde{D}_\infty^*$, it took about 50 seconds; for D_∞^*, it took about 25 minutes. Nevertheless, the computation times can be substantially reduced by implementing in, instead of the scripting language MATLAB, a compiled language such as C.

After comparing the discrepancies, we turn to the comparison between the discrepancies and their variants A_n and T_n. From the tables, for each discrepancy, none of the discrepancy itself or its variants A_n, $|A_n|$ and T_n uniformly outperform the others. For example, in Tables 1–6, CD_2 is much more powerful than its variants CA_n and $|CA_n|$, and is usually slightly better than CT_n, but in Table 8, for small n, CD_2 is much worse than its three variants. Similar observations can be made for each discrepancy. However, we can see that our suggested data-driven approach for choosing a more powerful discrepancy is still sound, and if the "non-periodic" group is chosen by the data-driven approach, then it is sensible to use the discrepancies themselves, because even though they are not always the most powerful, their variants, especially MA_n and $|MA_n|$ (see Tables 4–6), can be much worse. If the "periodic" group is chosen, then from the cases we considered, UT_n is usually more powerful (but note that for the cases where the "non-periodic" group should be chosen, UD_2, though should be not used, is often much more powerful than UT_n). These observations also suggest that for the discrepancies considered in this paper, the variants A_n and $|A_n|$ should not be recommended as a goodness-of-fit test statistic.

For the combined statistic XD, it seems that XD itself and its variant XT_n are similarly good in Tables 1–7, in which XD often has a bit higher powers, but XT_n performs better for small sample sizes in Tables 8–11.

7 Conclusion

The Cramér–von Mises statistics and its generalizations can be calculated by simple formulae, requiring $O(n^2 s)$ operations to perform the goodness-of-fit test for the null hypothesis $F = F_0$. The generalizations enhance the measure of uniformity of the distribution of the data points after the Rosenblatt transformation by considering not only the uniformity in s-dimensional hypercube, but also the uniformity in all projections onto lower dimensional hypercubes. In the other words, the generalizations measure the discrepancy between the empirical distribution F_n and the hypothesized distribution F_0 not only in their joint distri-

bution but in all marginal distributions. Because of the dependence between the components y_k of $\mathbf{Y}$, there are $s!$ possible ways to do the Rosenblatt transformation such that, for a given sample, the numerical value of a discrepancy of one transformed point set is not independent of the values of the same discrepancy of other transformed point sets, unless the components y_k are independent. We showed that the multiple test formulation suggested by Justel et al. (1997) is conservative and proposed to use the Monte Carlo tests. Simulation results showed that the generalized Cramér–von Mises statistics by using MD_2, CD_2, SD_2, UD_2 and WD_2, as well as their variants A_n and T_n, are more powerful than the multivariate Kolmogorov–Smirnov and the original Cramér–von Mises statistics. However, the variant A_n and $|A_n|$ are discouraged because they sometimes can perform very poorly. To choose which discrepancy and which statistic (the discrepancy itself or its variant T_n) to use, we recommend the interested reader to use either one of the following two approaches. The first one is to investigate the scatter plot of a transformed point set. If the transformed points are denser at the middle, we should use WD_2 or UD_2 to measure the discrepancy and the preferred test statistic is the T_n came from UD_2. Otherwise, we should use MD_2, CD_2 or SD_2 and the test statistic should be the discrepancy itself, rather than its T_n. The other approach is to use the statistic D_{sum} in which the variant XT_n of the combined discrepancy XD is employed. Either approach improves the power of the multivariate goodness-of-fit test, compared with the Kolmogorov–Smirnov statistic or the Cramér–von Mises statistics.

It is also possible to generalize the discrepancy statistics further by introducing weight functions, just as the Anderson–Darling statistic, but this will be left for future endeavour.

A final remark is about the Lilliefors-type modification. If the parameters $\boldsymbol{\theta}$ of F_0 are not specified, we still can use the Monte Carlo tests in which the nuisance parameters $\boldsymbol{\theta}$ are estimated by, for example, the maximum likelihood estimation. In such a situation the Monte Carlo tests are known as the parametric bootstrap tests (Davison and Hinkley, 1997).

Acknowledgements

We thank the referees for their helpful suggestions and drawing our attention to the work by Liang et al. (2000). This research was partially supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Numbers HKBU200605 and HKBU200807) and an FRG grant of the Hong Kong Baptist University.

References

Anderson, T. W., Darling, D. A., 1952. Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. The Annals of Mathematical Statistics 23, 193–212.

Anderson, T. W., Darling, D. A., 1954. A test of goodness-of-fit. Journal of The American Statistical Association 49, 765–769.

Bera, A., John, S., 1983. Tests for multivariate normality with Pearson alternatives. Communications in Statistics — Theory and Methods 12, 103–117.

Darling, D., 1983. On the asymptotic distribution of Watson's statistic. The Annals of Statistics 11, 1263–1266.

Davidson, R., MacKinnon, J. G., 1998. Graphical methods for investigating the size and power of hypothesis tests. The Manchester School 66, 1–26.

Davison, A. C., Hinkley, D. V., 1997. Bootstrap Methods and their Application. Cambridge University Press, New York.

Greenwood, P. E., Nikulin, M. S., 1996. A Guide to Chi-Squared Testing. John Wiley & Sons, New York.

Hensler, J. L., Mehrorta, K. G., Michalek, J. E., 1977. A goodness-of-fit test for multivariate normality. Communications in Statistics — Theory and Methods 19, 1403–1481.

Hickernell, F. J., 1998a. A generalized discrepancy and quadrature error bound. Mathematics of Computation 67, 299–322.

Hickernell, F. J., 1998b. Lattice rules: How well do they measure up? In: Hellekalek, P., Larcher, G. (Eds.), Random and Quasi-Random Point Sets. Vol. 138 of Lecture Notes in Statistics. Springer-Verlag, New York, pp. 109–166.

Hickernell, F. J., 1999a. Goodness of fit statistics, discrepancy and robust designs. Statistics & Probability Letters 44, 73–78.

Hickernell, F. J., 1999b. What affects the accuracy of quasi-monte carlo quadrature? In: Niederreiter, H., Spanier, J. (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 1998. Springer-Verlag, Berlin.

Ho, L. P., Chiu, S. N., 2007. Testing uniformity of a spatial point pattern. Journal of Computational and Graphical Statistics 16, 378–398.

Hope, A. C. A., 1968. A simplified Monte Carlo significance test procedure. Journal of the Royal Statistical Society Series B (Methodological) 30, 582–598.

Hua, L., Wang, Y., 1981. Applications of Number Theory to Numerical Analysis. Springer-Verlag, Berlin.

Jager, L., Wellner, J. A., 2007. Goodness-of-fit tests via phi-divergences. The Annals of Statistics 35, 2018–2053.

Justel, A., 2007. Personal communication.

Justel, A., Peña, D., Zamar, R., 1997. A multivariate Kolmogorov–Smirnov test of goodness of fit. Statistics & Probability Letters 35, 251–259.

Liang, J. J., Fang, K. T., Hickernell, F. J., Li, R., 2000. Testing multivariate uniformity and its applications. Mathematics of Computation 70, 337–355.

Marriott, F. H. C., 1979. Barnard's Monte Carlo tests: how many simulations? Applied Statistics 28 (1), 75–77.

Mecklin, C. J., Mundfrom, D. J., 2004. An appraisal and bibliography of tests for multivariate normality. International Statistical Review 72, 123–138.

Niederreiter, H., 1992. Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia.

Read, T. R. C., Cressie, N. A. C., 1988. Goodness-of-fit Statistics for Discrete Multivariate Data. Springer-Verlag, New York.

Rincon-Gallardo, S., Quesenberry, C. P., 1982. Testing multivariate normality using several samples: applications and techniques. Communications in Statistics — Theory and Methods 11, 343–358.

Rincon-Gallardo, S., Quesenberry, C. P., O'Reilly, F. J., 1979. Conditional probability integral transformations and goodness-of-fit tests for multivariate normal distributions. The Annals of Statistics 7, 1052–1057.

Rosenblatt, M., 1952. Remarks on a multivariate transformation. The Annals of Mathematical Statistics 23, 470–472.

Warnock, T. T., 1972. Computational investigations of low discrepancy point sets. In: Zaremba, S. K. (Ed.), Applications of Number Theory to Numerical Analysis. Academic Press, pp. 319–343.

Watson, G. S., 1961. Goodness-of-fit tests on a circle, part I. Biometrika 48, 109–114.

Watson, G. S., 1976. Optimal invariant tests for uniformity. In: William, E. J. (Ed.), Studies in Probability and Statistics, Papers in honour of E. J. G. Pitman. North-Holland, Amsterdam, pp. 121–127.

Corrigendum to "Generalized Cramér–von Mises goodness-of-fit tests for multivariate distributions" [Comput. Statist. Data Anal. 53 (2009) 3817-3834]

Sung Nok Chiu*, Kwong Ip Liu

Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong

On page 3824, the second to last line, the factor $2^s/n^2$ on the right-hand side should be corrected to $2^{s+1}/\{n(n-1)\}$. After this correction, the expression of U_2 for the symmetric discrepancy SD_2 should read

$$U_2 = \frac{2^{s+1}}{n(n-1)} \sum_{\mathbf{x}<\mathbf{x}'\in\mathcal{P}} \prod_{k=1}^{s} (1 - |x_k - x'_k|)\,.$$

However, we did use the correct formula in our simulations. Thus, this typo does not affect the results and conclusions presented in the paper.

Finally, the two misprints in Liang et al. (2000, p. 345) we mentioned are: (i) x_k^2 in the expression of U_1 for the modified star discrepancy MD_2, (i.e. the z_{kj} in Liang et al. (2000, p. 345, line 17) should be corrected to z_{kj}^2) and (ii) $(16/9)^s$ in the expression of ζ_1 for the symmetric discrepancy SD_2 (i.e. the $(6/9)^d$ in Liang et al. (2000, p. 345, line 11) should be corrected to $(16/9)^d$).

* Corresponding author.
E-mail address: snchiu@hkbu.edu.hk