

Article

Testing Group Symmetry of a Multivariate Distribution

Lyudmila Sakhanenko

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824-1027, USA; E-mail: luda@stt.msu.edu, Tel.: +1 517 432-9795, Fax: +1 517 432-1405.

Received: 15 September 2009 / Accepted: 24 November 2009 / Published: 26 November 2009

Abstract: We propose and study a general class of tests for group symmetry of a multivariate distribution, which encompasses different types of symmetry, such as ellipsoidal and permutation symmetries among others. Our approach is based on supremum norms of special empirical processes combined with bootstrap. We show that these tests are consistent against any fixed alternative. This work generalizes the methodology of Koltchinskii and Sakhanenko [7], developed for ellipsoidal symmetry to the case of group symmetry. It also provides a unified approach to testing different types of symmetry of a multivariate distribution.

Keywords: symmetric distributions; Donsker class; linear operator

Classification: MSC 62G10, 60F05, 60F99

1. Introduction

Let \mathcal{S} be a compact group of linear transformations (operators) from \mathbb{R}^d to \mathbb{R}^d . A Borel probability measure P on \mathbb{R}^d is called \mathcal{S} -symmetric if and only if there exist an affine nonsingular transformation A_0 from \mathbb{R}^d onto \mathbb{R}^d and a \mathcal{S} -invariant Borel probability measure Π_0 such that $P = \Pi_0 \circ A_0$. In other words, if X is a random vector with the distribution P , then there exists an affine nonsingular transformation A_0 such that the random vector $Z = A_0X$ is \mathcal{S} -invariant, which in its turn means that $Z \stackrel{d}{=} SZ$ for all transformations S from the group \mathcal{S} . Clearly, an affine transformation A_0 is not unique, as one could take instead SA_0 for an arbitrary transformation S from the group \mathcal{S} . But, on the other hand, one could always fix A_0 by normalizing it in any proper way.

Obviously, we can define $\Pi_0 = \Pi_{0,P}$ as the distribution of the random variable $Z = A_0X$. The couple (A_0, Π_0) will be called parameters (specifiers) of the \mathcal{S} -symmetric probability measure P . We denote by $\mathcal{E}(\mathbb{R}^d)$ the set of all \mathcal{S} -symmetric distributions on \mathbb{R}^d .

We are interested in the following problem. Given an independent identically distributed (i.i.d.) sample (X_1, \dots, X_n) from the distribution P , defined on the probability space $(\Omega, \Sigma, \mathbb{P})$, construct and study tests for \mathcal{S} -symmetry of the distribution P .

Our group approach will unify the theory for different tests for different types of symmetry. Below we give some examples of common in literature types of symmetry \mathcal{S} with corresponding choices of A_0 .

Example 1. Let $\mathcal{S} = \{S : S(x_1, \dots, x_d) = (\pm x_1, \dots, \pm x_d)\}$. Then \mathcal{S} -invariant probability measure Π_0 is a so called sign change symmetrical measure. Probability measure P is \mathcal{S} -symmetric if there exists an affine transformation A_0 in the form $A_0x = x - \theta_0$ with some $\theta_0 \in \mathbb{R}^d$ such that $P = \Pi_0 \circ A_0$. In this case, P is often called diagonally or reflectively symmetrical measure.

Let $|x|$ denote the Euclidean norm of a vector $x \in \mathbb{R}^d$. If $\mathbb{E}|X| < +\infty$, one can define $\theta_0 = \theta_0(P) := \mathbb{E}X = \int_{\mathbb{R}^d} xP(dx)$. Obviously, the parameter θ_0 and the transformation A_0 are uniquely defined for any P such that the corresponding X is integrable (not necessarily \mathcal{S} -symmetric). We can also define in such a generality Π_0 as the distribution of the random variable $Z = A_0X$.

Example 2. Let \mathcal{S} be the group of all orthogonal transformations in \mathbb{R}^d . Then \mathcal{S} -symmetrical probability measure P is often called ellipsoidally symmetric or elliptically symmetric or elliptically contoured measure and Π_0 is spherically symmetric one. In this case, if $\mathbb{E}|X|^2 < +\infty$, one can define $\theta_0 = \theta_0(P)$ as the mean $\mathbb{E}X = \int_{\mathbb{R}^d} xP(dx)$ and $V_0 = V_0(P)$ as the square root of the covariance operator of X , so that $A_0x = A_0(P)x = V_0^{-1}(x - \theta_0)$ for any x from \mathbb{R}^d . Define Π_0 as the distribution of the random variable $Z = A_0X$. As in the previous case, these parameters are defined for any P such that the corresponding X is square-integrable (not necessarily \mathcal{S} -symmetric).

We refer to the papers [1–9] for results on ellipsoidal and spherical symmetry testing.

Example 3. Let $d = 2$ and let \mathcal{S}_k be a group of all transformations translating a regular polygon with k vertices centered at 0 into itself. Clearly, \mathcal{S}_k is a subgroup of the group of all orthogonal transformations. Thus, an affine transformation A_0 can be fixed in the same way as in example 2.

Example 4. Let \mathcal{S} be a group of all reflections about hyperplanes $\{x \in \mathbb{R}^d : x_k = x_l, \} k, l = 1, \dots, d$. Then for each $S \in \mathcal{S}$ there exists a permutation (i_1, \dots, i_d) such that $S(x_1, \dots, x_d) = (x_{i_1}, \dots, x_{i_d})$ for all $x \in \mathbb{R}^d$. In this case \mathcal{S} -invariant probability measure Π_0 is called permutation symmetric measure.

As in examples 2 and 3 we can define $A_0x = A_0(P)x = V_0^{-1}(x - \theta_0)$ for any x from \mathbb{R}^d , where $\theta_0 = \theta_0(P) := \mathbb{E}X = \int_{\mathbb{R}^d} xP(dx)$ and $V_0 = V_0(P)$ is the square root of the covariance operator of X , provided $\mathbb{E}|X|^2 < +\infty$ (also see [10,11]).

Tests for symmetry of a multivariate distribution play an important role in statistics and in various fields of science. To name a few, in finance theory log-returns of assets are assumed to be ellipsoidally symmetric. In genetics it is assumed that gene expression values are diagonally symmetrically distributed. In image analysis components are assumed to be spherically symmetric. In linear programming it is assumed that the distribution of feasible solutions is permutation symmetric. In statistics sliced inverse regression method due to Li, see [12], works for ellipsoidally symmetric distributions. Also since tests for normality are extended to tests for ellipsoidal symmetry, any research field that employs multivariate analysis based on normality assumption can benefit from relaxing this assumption to ellipsoidal symmetry assumption. So clearly symmetry tests are needed in applications. See [13] for a detailed survey on the use of symmetry in various scientific fields.

The rest of the paper is organized in the following way. We give notations and construct test statistics with examples in Section 2. Main results and bootstrapped test statistics are given in Section 3, followed by a detailed example in Section 4. The proofs are in Section 5. The closing remarks are in Section 6, followed by technical details in Appendix.

2. Notations and Preliminaries

Let m denote the uniform distribution (the normalized Haar measure) on the group \mathcal{S} . Given a bounded Borel function f on \mathbb{R}^d , we define

$$m_f(y) := \int_{\mathcal{S}} f(Sy) m(dS), \quad y \in \mathbb{R}^d. \quad (1)$$

It is easy to check that for a \mathcal{S} -symmetric P distribution with specifiers (A_0, Π_0) we have

$$\int_{\mathbb{R}^d} f(A_0x)P(dx) = \int_{\mathbb{R}^d} m_f(y)\Pi_0(dy) \quad (2)$$

for any bounded Borel function f . Indeed, $\int_{\mathbb{R}^d} f(z)\Pi_0(dz) = \int_{\mathbb{R}^d} f(Sy)\Pi_0(dy)$ for any $S \in \mathcal{S}$. Since \mathcal{S} is a compact group, one can integrate the equality over \mathcal{S} with respect to the uniform measure. Thus, $\int_{\mathbb{R}^d} \int_{\mathcal{S}} f(Sy)m(dS)\Pi_0(dy) = \int_{\mathbb{R}^d} m_f(y)\Pi_0(dy)$ implies (2).

As a result, if a class \mathcal{F} characterizes the distribution, i.e.

$$\int_{\mathbb{R}^d} f dQ_1 = \int_{\mathbb{R}^d} f dQ_2 \text{ for all } f \in \mathcal{F}$$

implies that $Q_1 = Q_2$, then P is \mathcal{S} -symmetric if and only if (2) holds for all $f \in \mathcal{F}$. In general, we call P a \mathcal{F} -asymmetric distribution if and only if there exists a function $f \in \mathcal{F}$ such that (2) does not hold. This observation is the key idea behind the tests that we construct and study. Naturally, a class \mathcal{F} should be rich enough and possess good properties for further analysis. Let us describe it. Let \mathcal{F} be a *semialgebraic subgraph* class as introduced in [7]. Basically, for a function from such a class, its subgraph can be constructed from a union of intersections of a finite number of subgraphs of polynomials of a finite degree in \mathbb{R}^d . The same should be true for a product of two functions from such a class. For

instance, one can use polynomials of bounded degree or trigonometric functions of bounded frequency as \mathcal{F} . The precise definition can be found in Appendix. We will provide a few examples later on.

Let P_n be the empirical distribution based on the sample (X_1, \dots, X_n) . Assume in what follows that $A_n = A_n(P)$ is a $n^{1/2}$ -consistent estimator of $A_0 = A_0(P)$. And, furthermore, there exists such a function $\gamma_A : \mathbb{R}^d \rightarrow \mathbb{R}$ that $\mathbb{E}\gamma_A^2(X) < +\infty$ and as $n \rightarrow \infty$

$$n^{1/2}(A_n - A_0) = n^{1/2} \int_{\mathbb{R}^d} \gamma_A(x) d(P_n - P) + o_P(1).$$

For example, assuming $\mathbb{E}|X|^2 < +\infty$ let us define θ_n and V_n as

$$\theta_n := \theta_n(X_1, \dots, X_n) := \bar{X}_n := n^{-1} \sum_{i=1}^n X_i$$

and

$$V_n^2 := V_n^2(X_1, \dots, X_n) := n^{-1} \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^T,$$

where all vectors are columns and superscript T denotes transposition. Then one can define $A_n x = A_n(X_1, \dots, X_n)x = V_n^{-1}(x - \theta_n)$ in examples 2-4 and $A_n x = x - \theta_n$ in example 1 for any $x \in \mathbb{R}^d$. Under the condition $\mathbb{E}|X|^4 < +\infty$, V_n is a $n^{1/2}$ -consistent estimator of V_0 . Weaker moment assumptions on P can be imposed if other statistics are considered for estimation of θ_0, V_0 , such as a sample median and an M -estimator for the covariance matrix. See, e.g., [14].

The scaled residuals of the observations (X_1, \dots, X_n) are defined as

$$Z_j := Z_{j,n} := A_n X_j, \quad j = 1, \dots, n.$$

Let Π_n denote the empirical distribution based on the sample (Z_1, \dots, Z_n) .

Our approach to the problem of testing for \mathcal{S} -symmetry will be to use the sup-norms of the stochastic process

$$\begin{aligned} \xi_n(f) &:= n^{1/2} \left(\int_{\mathbb{R}^d} f(A_n x) P_n(dx) - \int_{\mathbb{R}^d} m_f(y) d\Pi_n(y) \right) \\ &= n^{-1/2} \sum_{j=1}^n [f(Z_j) - m_f(Z_j)], \quad f \in \mathcal{F} \end{aligned}$$

as test statistics

$$T_n(\mathcal{F}) := \sup_{f \in \mathcal{F}} |\xi_n(f)|.$$

Such functionals can be viewed as "measures of asymmetry" of the empirical distribution because of the relationship (2).

Note that a nonsingular affine transformation of the data (X_1, \dots, X_n) results in an orthogonal transformation of the scaled residuals. If a class \mathcal{F} is invariant with respect to all orthogonal transformations (i.e. for all $f \in \mathcal{F}$ and any orthogonal transformation O we have $f \circ O \in \mathcal{F}$), then the test statistic defined as the sup-norm of the process ξ_n is affine invariant. This is the case in the following examples.

Example 1.1. Consider \mathcal{S} from example 1. Let

$$\mathcal{H} := \left\{ \{x \in \mathbb{R}^d : \langle x, u \rangle \leq c\} : u \in S^{d-1}, c \in \mathbb{R} \right\}$$

be the class of all half-spaces in \mathbb{R}^d , where S^{d-1} denotes the unit sphere in \mathbb{R}^d . Consider the class

$$\mathcal{F} := \{I_H(x) : H \in \mathcal{H}\}.$$

For $f(\cdot) = I_H(\cdot)$, we have $m_f(y) = \frac{1}{2^d} \sum_{\varepsilon_i = \pm 1} I_H(\varepsilon_1 y_1, \dots, \varepsilon_d y_d) =: \frac{1}{2^d} \sum_{\varepsilon_i = \pm 1} I_H(\varepsilon y)$, $y \in \mathbb{R}^d$. The process ξ_n becomes

$$\xi_n(H) = \frac{n^{-1/2}}{2^d} \sum_{j=1}^n \sum_{\varepsilon_i = \pm 1} (I_H(Z_j) - I_H(\varepsilon Z_j)), \quad H \in \mathcal{H}.$$

The test statistic is represented as

$$T_n(\mathcal{F}) := \sup_{H \in \mathcal{H}} |\xi_n(H)| = \frac{n^{-1/2}}{2^d} \sup_{H \in \mathcal{H}} \left| \sum_{j=1}^n \sum_{\varepsilon_i = \pm 1} (I_H(Z_j) - I_H(\varepsilon Z_j)) \right|.$$

Example 1.2. Consider \mathcal{S} from example 1. Let \mathcal{H} be the class of all half-spaces in \mathbb{R}^d as in the previous example. Let

$$f(x) = \sum_{\varepsilon_i = \pm 1} (-1)^{|\varepsilon|} I_{\varepsilon H}(x)$$

for $\varepsilon H = \{x : \langle x, \varepsilon u \rangle \leq c\}$ with $|\varepsilon| = \varepsilon_1 + \dots + \varepsilon_d$. Then we have $m_f(y) = 0$. The process ξ_n becomes

$$\xi_n(H) = n^{-1/2} \sum_{j=1}^n \sum_{\varepsilon_i = \pm 1} (-1)^{|\varepsilon|} I_{\varepsilon H}(Z_j), \quad H \in \mathcal{H}$$

and the test statistic looks like

$$T_n(\mathcal{F}) = n^{-1/2} \sup_{H \in \mathcal{H}} \left| \sum_{j=1}^n \sum_{\varepsilon_i = \pm 1} (-1)^{|\varepsilon|} I_{\varepsilon H}(Z_j) \right|.$$

In one-dimensional case $d = 1$ this test statistic becomes

$$T_n := n^{-1/2} \sup_{c \in \mathbb{R}} \left| \sum_{j=1}^n (I_{(Z_j \leq c)} - I_{(-Z_j \leq c)}) \right|.$$

One gets the expression that resembles a well known test for symmetry based on the empirical distribution function P_n ,

$$\tilde{T}_n = n^{-1/2} \sup_{x \leq 0} |n(P_n(x) + P_n(-x) - 1)|.$$

See, for instance the discussion in the paper [15].

Example 2.1. Consider \mathcal{S} from example 2. Let

$$\mathcal{C} := \{\{v \in S^{d-1} : \langle v, l \rangle \geq c\} : l \in S^{d-1}, c \in \mathbb{R}_+\}$$

be the class of "caps" on the unit sphere S^{d-1} . Consider the class

$$\mathcal{F} := \{I_C\left(\frac{x}{|x|}\right) I_{\{0 < |x| \leq t\}} : C \in \mathcal{C}, t > 0\}.$$

For $f(\cdot) = I_C(\frac{\cdot}{|\cdot|})I_{\{0 < |\cdot| \leq t\}}$, we have $m_f(\rho) = m(C)I_{(0,t]}(\rho)$, $\rho > 0$. The process ξ_n becomes now

$$\xi_n(C, t) = n^{-1/2} \sum_1^n (I_C(\frac{Z_j}{|Z_j|}) - m(C))I_{\{0 < |Z_j| \leq t\}}, \quad C \in \mathcal{C}, \quad t > 0.$$

The test statistic $T_n(\mathcal{F}) := \sup_{t>0, C \in \mathcal{C}} |\xi_n(C, t)|$ can also be represented as

$$T_n(\mathcal{F}) = n^{-1/2} \max_{1 \leq j \leq n} \sup_{C \in \mathcal{C}} \left| \sum_{k=1}^j (I_C(\frac{Z_{[k]}}{|Z_{[k]|})} - m(C)) \right|,$$

where $Z_{[j]}$, $j = 1, \dots, n$, is the rearrangement of Z_1, \dots, Z_n , such that $|Z_{[1]}| \leq \dots \leq |Z_{[n]}|$. These tests were studied in [7] and [9].

Example 2.2. Consider \mathcal{S} from example 2. Let \mathcal{G}_l denote the linear space of spherical harmonics of degree less than or equal to l in \mathbb{R}^d , and let \mathcal{B}_l be the unit ball in $\mathcal{G}_l \cap L_2(S^{d-1}, dm)$. Denote

$$\mathcal{F} := \{I_{\{0 < |x| \leq t\}} \psi(\frac{x}{|x|}) : \psi \in \mathcal{B}_l, \quad t > 0\}.$$

Then for $f \in \mathcal{F}$, we have $f(x) = I_{\{0 < |x| \leq t\}} \psi(\frac{x}{|x|})$, $x \in \mathbb{R}^d$, and $m_f(\rho) = m(\psi)I_{(0,t]}(\rho)$, where $m(\psi)$ is the average of ψ on S^{d-1} . In this case, the process ξ_n becomes

$$\xi_n(t, \psi) = n^{-1/2} \sum_1^n (I_{\{0 < |Z_j| \leq t\}} (\psi(\frac{Z_j}{|Z_j|}) - m(\psi))), \quad t > 0, \quad \psi \in \mathcal{B}_l.$$

The statistic $T_n(\mathcal{F}) := \sup_{t>0, \psi \in \mathcal{B}_l} |\xi_n(t; \psi)|$ becomes

$$T_n(\mathcal{F}) = n^{-1/2} \max_{1 \leq j \leq n} \left(\sum_{s=1}^{\dim(\mathcal{G}_l)} \left(\sum_{k=1}^j \psi_s(\frac{Z_{[k]}}{|Z_{[k]|})} - \delta_{s1} \right) \right)^2 \Big)^{1/2},$$

where $\{\psi_s, s = 1, \dots, \dim(\mathcal{G}_l)\}$ denotes an orthonormal basis of the space \mathcal{G}_l , $\psi_1 \equiv 1$, $\delta_{ij} = 1$ for $i = j$ and 0 otherwise, and \dim of a set denotes the number of elements of the set. These tests were studied in [7] and [9], where their superiority in level preservation and power performance over other tests both theoretically and in a simulation study, was shown. A similar approach was used to test for multivariate normality in [16]. The authors of [6] developed a different kind of tests for ellipsoidal symmetry based on spherical harmonics.

Example 2.3. Consider \mathcal{S} from example 2. Let \mathcal{H} be the class of all half-spaces in \mathbb{R}^d as in the example 1.1. For $f := I_H$, where $H := \{x \in \mathbb{R}^d : \langle x, u \rangle \leq c\}$, we have $m_f(\rho) := \gamma(\frac{c}{\rho})$, where $\gamma(c) := m\{v : \langle v, u \rangle \leq c\}$. The process ξ_n in this case is

$$\xi_n(u, c) := n^{-1/2} \sum_{j=1}^n (I_{\{\langle Z_j, u \rangle \leq c\}} - \gamma(\frac{c}{|Z_j|})), \quad u \in S^{d-1}, \quad c \in \mathbb{R}$$

and the test statistic can be defined as

$$T_n := \sup_{c \in \mathbb{R}} \sup_{u \in S^{d-1}} |\xi_n(u, c)|.$$

This type of test statistics was systematically studied in papers [7,9,10].

Example 2.4. Consider \mathcal{S} from example 2. Let

$$\mathcal{F} := \left\{ \frac{|\cdot - t| - |\cdot|}{|t|} : t \in \mathbb{R}^d, t \neq 0 \right\}.$$

For $f(\cdot) = \frac{|\cdot - t| - |\cdot|}{|t|}$, we have $m_f(\rho) = \rho^{\frac{\beta(t)-1}{|t|}}$, where $\beta(t) := \int_{S^{d-1}} |v - t| m(dv)$. Thus, the process ξ_n becomes

$$\xi_n(t) := n^{-1/2} |t|^{-1} \sum_{i=1}^n (|Z_i - t| - |Z_i| \beta(t)), \quad t \in \mathbb{R}^d$$

and the test statistic can be chosen as

$$T_n(\mathcal{F}) := \sup_{t \in \mathbb{R}^d, t \neq 0} |\xi_n(t)|.$$

Example 2.5. Consider \mathcal{S} from example 2. Consider the class

$$\mathcal{F} := \{e^{i\langle t, x \rangle}, x \in \mathbb{R}^d : |t| \leq 1\}.$$

For $f(\cdot) = e^{i\langle t, \cdot \rangle}$ we have $m_f(\rho) = c_d \frac{J_{d/2-1}(\rho|t|)}{(\rho|t|)^{d/2-1}}$, $\rho > 0$, where J_l denotes the Bessel function of the l -th order, the constant c_d depends only on d . The process ξ_n becomes

$$\xi_n(t) = n^{-1/2} \sum_{j=1}^n \left(\exp\{i\langle t, Z_j \rangle\} - c_d \frac{J_{d/2-1}(|Z_j||t|)}{(|Z_j||t|)^{d/2-1}} \right), \quad |t| \leq 1$$

and the test statistic can be chosen as

$$T_n := \sup_{|t| \leq 1} |\xi_n(t)|.$$

Consider \mathcal{S} from example 3. Due to similarity between examples 2 and 3, one can choose the same classes of functions for \mathcal{S}_k from example 3. We give just one of the examples as an illustration.

Example 3.1. Consider the class \mathcal{F} from example 2.1. Then for $f(\cdot) = I_C(\frac{\cdot}{|\cdot|}) I_{\{0 < |\cdot| \leq t\}}$, we have $m_f(y) = \frac{1}{k} \sum_{l=0}^{k-1} I_C(O_l y / |y|) I_{(0,t]}(|y|)$ for all $y \in \mathbb{R}^d$, where O_l is the rotation on angle $2\pi l/k$, $l = 0, \dots, k-1$. In this case, the process ξ_n is

$$\xi_n(C, t) = n^{-1/2} \sum_{j=1}^n \left(I_C\left(\frac{Z_j}{|Z_j|}\right) - \frac{1}{k} \sum_{l=0}^{k-1} I_C(O_l Z_j / |Z_j|) \right) I_{\{0 < |Z_j| \leq t\}}, \quad C \in \mathcal{C}, \quad t > 0.$$

The test statistic $T_n := \sup_{t > 0, C \in \mathcal{C}} |\xi_n(C, t)|$ can be also represented as

$$T_n(\mathcal{F}) = n^{-1/2} \max_{1 \leq j \leq n} \sup_{C \in \mathcal{C}} \left| \sum_{m=1}^j \left(I_C\left(\frac{Z_{[m]}}{|Z_{[m]|}}\right) - \frac{1}{k} \sum_{l=0}^{k-1} I_C(O_l Z_{[m]} / |Z_{[m]|}) \right) \right|,$$

where $Z_{[j]}$, $j = 1, \dots, n$ is the rearrangement of Z_1, \dots, Z_n , such that $|Z_{[1]}| \leq \dots \leq |Z_{[n]}|$.

Example 4.1. Consider \mathcal{S} from example 4. Let \mathcal{H} be the class of all half-spaces in \mathbb{R}^d as in the example 1.1. Denote

$$\mathcal{F} := \{I_H(y)I(y_{j_1} \geq \dots \geq y_{j_k}) : H \in \mathcal{H}, (j_1, \dots, j_k) \text{ is any combination from } (1, \dots, d)\}.$$

Then for $f \in \mathcal{F}$ such that $f(y) := I_H(y)I(y_{j_1} \geq \dots \geq y_{j_k})$ for $y \in \mathbb{R}^d$ we have

$$m_f(y) = m_{H, i_1, \dots, i_k}(y) = \frac{1}{d!} \sum_{(i_1, \dots, i_d)} I_H((y_{i_1}, \dots, y_{i_d}))I(y_{i_{j_1}} \geq \dots \geq y_{i_{j_k}}),$$

where the summation is over all permutations (i_1, \dots, i_d) of $(1, \dots, d)$. In this case the process ξ_n is

$$\xi_n(H, i_1, \dots, i_k) := n^{-1/2} \sum_{j=1}^n (I_H(Z_j)I((Z_j)_{j_1} \geq \dots \geq (Z_j)_{j_k}) - m_{H, i_1, \dots, i_k}(Z_j)),$$

$$H \in \mathcal{H}, (j_1, \dots, j_k) \text{ is any combination from } (1, \dots, d)$$

and the test statistic can be defined as

$$T_n := \sup_{H \in \mathcal{H}} \sup_{2 \leq k \leq d} \sup_{(j_1, \dots, j_k)} |\xi_n(H, j_1, \dots, j_k)|,$$

where the last supremum is taken over all combinations (j_1, \dots, j_k) out of $(1, \dots, d)$. Well known and frequently used Friedman's rank tests are based on the similar choice of a class \mathcal{F} . For reference see the papers [17] and [11].

It is not hard to see that the function classes defined in examples 1.1, 1.2, 2.1–2.4, 3.1, 3.2, and 4.1 are semialgebraic subgraph. In addition, classes \mathcal{F} characterize the distribution in the case of examples 1.1, 2.1, 2.3, 2.4, 3.1, 4.1 above.

We say the class of transformations \mathcal{S} preserves the semialgebraic property if for any polynomial p on \mathbb{R}^d of degree less than or equal to r the set $\{(x, t) \in \mathbb{R}^{d+1} : p(Sx, t) \geq 0\}$ belongs to $\mathcal{SA}_{q, d+1, l}$ for some q and l (see Appendix for the definition). Classes \mathcal{S} , defined in examples 1–4, preserve the semialgebraic property.

Let

$$\mathbb{E}(f; A) := \int_{\mathbb{R}^d} [f(Ax) - m_f(Ax)] P(dx).$$

It follows from (2) that, for a \mathcal{S} -symmetric distribution P and for all f

$$\mathbb{E}(f; A_0) = 0.$$

Let W_P° denote the P -Brownian bridge, i.e. a centered Gaussian process indexed by functions in $L_2(\mathbb{R}^d; dP)$ with the covariance

$$\mathbb{E}W_P^\circ(f)W_P^\circ(g) = P(fg) - P(f)P(g).$$

We will frequently use integral notation for $W_P^\circ(f)$:

$$W_P^\circ(f) = \int_{\mathbb{R}^d} f(x)W_P^\circ(dx).$$

As always, $\ell^\infty(\mathcal{F})$ denotes the space of all uniformly bounded functions on \mathcal{F} with the sup-norm $\|Y\|_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |Y(f)|$, $Y \in \ell^\infty(\mathcal{F})$.

A sequence of stochastic processes $\zeta_n : \mathcal{F} \mapsto \mathbb{R}$ is said to converge weakly in $\ell^\infty(\mathcal{F})$ (in the sense of Hoffmann-Jørgensen) to the stochastic process $\zeta : \mathcal{F} \mapsto \mathbb{R}$ if and only if there exists a Radon probability measure γ on $\ell^\infty(\mathcal{F})$ such that γ is the distribution of ζ and, for all bounded and $\|\cdot\|_{\mathcal{F}}$ -continuous functionals $\Phi : \ell^\infty(\mathcal{F}) \mapsto \mathbb{R}$, we have $\mathbb{E}^* \Phi(\zeta_n) \rightarrow \int \Phi(x) \gamma(dx)$, where \mathbb{E}^* stands for the outer expectation, which is defined as $\mathbb{E}^* \Psi = \inf\{\mathbb{E}U : U \geq \Psi, \Psi : \Omega \rightarrow [-\infty, \infty]$ measurable and $\mathbb{E}U$ exists} for a $\Psi : \Omega \rightarrow [-\infty, \infty]$. See for instance [18].

We assume in what follows that the class \mathcal{F} satisfies standard measurability assumptions used in the theory of empirical processes (see [19] or [18]). We also need smoothness conditions (S) on P and \mathcal{F} , which are given in Appendix.

3. Main Results

Theorem 1 *Suppose that \mathcal{F} is a semialgebraic subgraph class, the smoothness conditions (S) hold and $\mathbb{E}\gamma_A^2(X) < +\infty$. Define a Gaussian stochastic process*

$$\xi_P(f) := W_P^\circ(f(A_0 \cdot) - m_f(A_0 \cdot)) + \mathbb{E}'_A(f; A_0) \left(\int_{\mathbb{R}^d} \gamma_A(x) W_P^\circ(dx) \right),$$

whose distribution is a Radon measure in $\ell^\infty(\mathcal{F})$. Then the sequence of stochastic processes

$$\{\xi_n(f) - n^{1/2} \mathbb{E}(f; A_0) : f \in \mathcal{F}\}$$

converges weakly in the space $\ell^\infty(\mathcal{F})$ to the process ξ_P . In particular, if P is \mathcal{S} -symmetric with specifiers (A_0, Π_0) , then the sequence ξ_n converges weakly in the space $\ell^\infty(\mathcal{F})$ to the process ξ_P .

Define the test statistics

$$T_n := \|\xi_n\|_{\mathcal{F}}.$$

Given $\alpha > 0$, let

$$t_\alpha := \inf \left\{ t : \mathbb{P}\{\|\xi_P\|_{\mathcal{F}} \geq t\} \leq \alpha \right\}.$$

Let H_0 be the hypothesis that $P \in \mathcal{E}(\mathbb{R}^d)$ and let H_a be the alternative that $P \notin \mathcal{E}(\mathbb{R}^d)$. Also, denote by $H_a(\mathcal{F})$ the alternative that P is \mathcal{F} -asymmetric.

Theorem 1 and the well-known theorem of Cirel'son on continuity of the distribution of the sup-norm of Gaussian processes, see [20], imply the following.

Corollary 1 *Suppose all conditions of Theorem 1 hold. Under the hypothesis H_0*

$$\mathbb{P}\{T_n \geq t_\alpha\} \rightarrow \alpha$$

and under the alternative $H_a(\mathcal{F})$

$$\mathbb{P}\{T_n \geq t_\alpha\} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

In particular, if \mathcal{F} characterizes the distribution, then under the alternative H_a , i.e. for a fixed \mathcal{S} -asymmetric distribution P ,

$$\mathbb{P}\{T_n \geq t_\alpha\} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

In most cases, however, the limit distributions of such statistics as T_n depend on the unknown parameters of the distribution P . Thus, to implement the test one has to evaluate the distribution of the test statistic using, for instance, a bootstrap method. We describe below a version of the conditional bootstrap for \mathcal{S} -symmetry testing. It is a generalization of the bootstrap method proposed in [7].

Given P , let P^s denote the \mathcal{S} -symmetric distribution with specifiers (A_0, Π_0) . It will be called *the \mathcal{S} -symmetrization of P* . Denote by P_n^s the \mathcal{S} -symmetric distribution with specifiers (A_n, Π_n) . Let (X_1^s, \dots, X_n^s) be an i.i.d. sample from the distribution P_n^s , defined on a probability space $(\hat{\Omega}, \hat{\Sigma}, \hat{P})$. One can construct such a sample using the following procedure. Take an i.i.d. sample (Y_1, \dots, Y_n) from Π_n , which is a resampling from (Z_1, \dots, Z_n) . Define

$$X_j^s := A_n^{-1}Y_j, \quad j = 1, \dots, n.$$

Then conditionally on (X_1, \dots, X_n) , (X_1^s, \dots, X_n^s) is an i.i.d. sample from the \mathcal{S} -symmetric distribution P_n^s .

In particular, for \mathcal{S} from example 1

$$X_j^s := \theta_n + \varepsilon_j Y_j, \quad j = 1, \dots, n,$$

where $(\varepsilon_1, \dots, \varepsilon_n)$ is a Rademacher i.i.d. sample, that is $\varepsilon_j = -1$ or 1 with probability $1/2$, $j = 1, \dots, n$, independent of (Y_1, \dots, Y_n) .

For \mathcal{S} from example 2 one can take an i.i.d. sample (U_1, \dots, U_n) uniformly distributed on S^{d-1} and an i.i.d. sample (Y_1, \dots, Y_n) from $\tilde{\Pi}_n$, the empirical distribution based on $(|Z_1|, \dots, |Z_n|)$, independent of (U_1, \dots, U_n) . In other words, (Y_1, \dots, Y_n) is the resampling from the sample $(|Z_1|, \dots, |Z_n|)$. Then

$$X_j^s := \theta_n + V_n U_j Y_j, \quad j = 1, \dots, n.$$

For \mathcal{S} from example 3 let $(\varepsilon_1, \dots, \varepsilon_n)$ be an i.i.d. sample uniformly distributed on $\{0, 1, \dots, k-1\}$ independent of (Y_1, \dots, Y_n) , then

$$X_j^s := \theta_n + V_n O_j Y_j, \quad j = 1, \dots, n,$$

where O_j is a rotation on the angle $2\pi \frac{\varepsilon_j}{k}$ about 0 .

Finally, for \mathcal{S} from example 4 consider n independent permutations $\{i_1^{(j)}, \dots, i_d^{(j)}\}$ of $(1, \dots, d)$, $j = 1, \dots, n$, independent of (Y_1, \dots, Y_n) . Then

$$X_j^s := \theta_n + R_j Y_j, \quad j = 1, \dots, n,$$

where R_j is a reflection transformation such that $R_j(x_1, \dots, x_d) = (x_{i_1^{(j)}}, \dots, x_{i_d^{(j)}})$ for $x \in \mathbb{R}^d$.

Let \hat{P}_n denote the empirical measure based on the sample (X_1^s, \dots, X_n^s) , and let

$$\hat{A}_n := A_n(X_1^s, \dots, X_n^s).$$

Define *the bootstrapped scaled residuals* as

$$\hat{Z}_j := \hat{Z}_{j,n} := \hat{A}_n X_j^s, \quad j = 1, \dots, n.$$

Let $\hat{\Pi}_n$ denote the empirical distribution based on the sample $(\hat{Z}_1, \dots, \hat{Z}_n)$.

The bootstrap version of ξ_n is the process

$$\begin{aligned}\hat{\xi}_n(f) &:= n^{1/2} \left(\int_{\mathbb{R}^d} f(\hat{A}_n x) \hat{P}_n(dx) - \int_{\mathbb{R}^d} m_f(y) d\hat{\Pi}_n(y) \right) \\ &= n^{-1/2} \sum_{j=1}^n [f(\hat{Z}_j) - m_f(\hat{Z}_j)], \quad f \in \mathcal{F}.\end{aligned}$$

Let $BL_1(\ell^\infty(\mathcal{F}))$ denote the set of all functionals $\Phi : \ell^\infty(\mathcal{F}) \mapsto \mathbb{R}$ such that $|\Phi(Y)| \leq 1$ for all $Y \in \ell^\infty(\mathcal{F})$ and $|\Phi(Y_1) - \Phi(Y_2)| \leq \|Y_1 - Y_2\|_{\mathcal{F}}$ for all $Y_1, Y_2 \in \ell^\infty(\mathcal{F})$. Given two stochastic processes $\zeta_1, \zeta_2 : \Omega \times \hat{\Omega} \times \mathcal{F} \mapsto \mathbb{R}$, we define the following bounded Lipschitz distance:

$$d_{BL}(\zeta_1, \zeta_2) := \sup_{\Phi \in BL_1(\ell^\infty(\mathcal{F}))} |\hat{\mathbb{E}}^* \Phi(\zeta_1) - \hat{\mathbb{E}}^* \Phi(\zeta_2)|,$$

where \mathbb{E}^* denotes the outer expectation.

Now we are going to consider a bootstrap version of Theorem 1.

Theorem 2 *Suppose that \mathcal{F} is a semialgebraic subgraph class, the smoothness conditions (S) hold and $\mathbb{E}\gamma_A^2(X) < \infty$. Then the sequence of stochastic processes $\{\hat{\xi}_n\}$ converges weakly in the space $\ell^\infty(\mathcal{F})$ to a version $\hat{\xi}_{P^s}$ of the process ξ_{P^s} (defined on the probability space $(\hat{\Omega}, \hat{\Sigma}, \hat{\mathbb{P}})$) in probability \mathbb{P} . More precisely,*

$$d_{BL}(\hat{\xi}_n; \hat{\xi}_{P^s}) \rightarrow 0 \text{ as } n \rightarrow \infty \text{ in probability } \mathbb{P}.$$

In particular, if P is \mathcal{S} -symmetric, $\hat{\xi}_n$ converges weakly to a version of the process ξ_P .

Define test statistics

$$\hat{T}_n := \|\hat{\xi}_n\|_{\mathcal{F}}.$$

Given $\alpha > 0$, let

$$\hat{t}_{n,\alpha} := \inf \left\{ t : \hat{\mathbb{P}}\{\hat{T}_n \geq t\} \leq \alpha \right\}.$$

In other words, $\hat{t}_{n,\alpha}$ is a $(1 - \alpha)$ -quantile of the distribution of \hat{T}_n conditional on the sample (X_1, \dots, X_n) .

Then Theorems 1 and 2 imply the following.

Corollary 2 *Suppose all the conditions of Theorems 1 and 2 hold. Under the hypothesis H_0*

$$\mathbb{P}\{T_n \geq \hat{t}_{n,\alpha}\} \rightarrow \alpha$$

and under the alternative $H_a(\mathcal{F})$

$$\mathbb{P}\{T_n \geq \hat{t}_{n,\alpha}\} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

In particular, if \mathcal{F} characterizes the distribution, the bootstrap test is consistent against any asymmetric alternative (subject to the smoothness conditions (S)): under the alternative H_a

$$\mathbb{P}\{T_n \geq \hat{t}_{n,\alpha}\} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

Thus, our method provides tests that are consistent against any \mathcal{S} -asymmetric alternative.

4. Detailed example

In this section we provide an example for which we verify all the assumptions and supply a step-by-step computational algorithm. Let $d = 2$ and consider the problem of testing whether P is elliptically contoured measure (Example 2). For a vector $x \in \mathbb{R}^2$ let $(r, \varphi) \in [0, \infty) \times [0, 2\pi)$ be its polar coordinates. For a fixed integer l let

$$\mathcal{F} := \text{lin}(\{I_{\{0 < r \leq t\}} \cos(k\varphi), I_{\{0 < |x| \leq t\}} \sin(k\varphi) : 1 \leq k \leq l, t > 0\}),$$

where $\text{lin}(\mathcal{G})$ denotes the linear span of \mathcal{G} with all the functions bounded by 1, see Example 2.2. This class \mathcal{F} satisfies the following assumptions.

1. It characterizes distribution only for $l = \infty$. For a finite l it does not characterize the distribution, since one might find two different distributions $Q_{1,2}$ such that

$$\int_0^t \int_0^{2\pi} \cos(k\varphi) dQ_1(r \cos \varphi, r \sin \varphi) = \int_0^t \int_0^{2\pi} \cos(k\varphi) dQ_2(r \cos \varphi, r \sin \varphi)$$

and

$$\int_0^t \int_0^{2\pi} \sin(k\varphi) dQ_1(r \cos \varphi, r \sin \varphi) = \int_0^t \int_0^{2\pi} \sin(k\varphi) dQ_2(r \cos \varphi, r \sin \varphi)$$

for all $t > 0, 1 \leq k \leq l$.

2. \mathcal{F} is a semialgebraic subgraph class. Indeed, for (x_1, x_2) the sets

$$\begin{aligned} & \{(x_1, x_2, s) : I(0 < x_1^2 + x_2^2 \leq t^2) \cos(k \arctan(x_2/x_1)) \geq s \geq 0 \\ & \text{or } I(0 < x_1^2 + x_2^2 \leq t^2) \cos(k \arctan(x_2/x_1)) \leq s \leq 0\} \end{aligned}$$

can be represented as unions of finite number of intersections of polynomial sets of finite degree. For instance, for $k = 1$ we have have the following representation

$$\begin{aligned} & \{1 - s^2 \geq 0, t^2 - x_1^2 - x_2^2 \geq 0, s \geq 0, x_1^2 - s^2(x_1^2 + x_2^2) \geq 0\} \\ & \cup \{1 - s^2 \geq 0, t^2 - x_1^2 - x_2^2 \geq 0, -s \geq 0, s^2(x_1^2 + x_2^2) - x_1^2 \geq 0\}. \end{aligned}$$

The representations for any $1 \leq k \leq l$ can be obtained similarly using trigonometric identities. Obviously, similar arguments work for sines, linear combinations of sines and cosines and products of any two functions from \mathcal{F} .

3. \mathcal{F} is invariant with respect to all orthogonal transformations, which are rotations on the unit circle. Indeed, for any rotation on an angle τ a vector (r, φ) is transformed into the vector $(r, \varphi + \tau)$. So for any $t > 0, 1 \leq k \leq l$ we have $f(r, \varphi + \tau) = I_{\{0 < r \leq t\}} \cos(k(\varphi + \tau))$ or $f(r, \varphi + \tau) = I_{\{0 < r \leq t\}} \sin(k(\varphi + \tau))$, where both functions belong to the linear span of \mathcal{F} and are bounded by 1. So any linear combination of such functions, which is bounded by 1, would also lie in \mathcal{F} .

4. Condition (S2) holds. Indeed, \mathcal{F} is uniformly bounded by 1. For any $\delta > 0$ and any $f \in \mathcal{F}$ we have $\omega_f(\cdot; \delta) = C\delta$ for some constant $C > 0$, so that the measure of the set defined in (S2) is zero for $\delta < \varepsilon$.

Also note that the group \mathcal{S} of all orthonormal transformations of \mathbb{R}^2 preserves semialgebraic property. Indeed, for any rotation on an angle τ the sets

$$\begin{aligned} & \{(x_1, x_2, s) : I(0 < x_1^2 + x_2^2 \leq t^2) \cos(k \arctan(x_2/x_1) + \tau) \geq s \geq 0 \\ & \text{or } I(0 < x_1^2 + x_2^2 \leq t^2) \cos(k \arctan(x_2/x_1) + \tau) \leq s \leq 0\} \end{aligned}$$

are semialgebraic. The same holds for sines and linear combinations of sines and cosines.

We also require the following two conditions on P : (S1) holds and $\mathbb{E}\gamma_A^2(X) < \infty$. The first condition is satisfied for absolutely continuous P with the uniformly bounded and continuously differentiable Lebesgue density with the corresponding derivative approaching zero at infinity faster than $|x|^{-3}$, $x \in \mathbb{R}^2$. For example, distributions with densities on a finite support and normal distributions satisfy (S1). The last condition is satisfied if $\mathbb{E}|X|^4 < \infty$.

Given a random sample (X_1, \dots, X_n) from a distribution P let us describe a step-by-step testing algorithm.

1. Obtain θ_n, V_n . In our example θ_n is the sample mean and V_n is the square root of the sample covariance for the sample (X_1, \dots, X_n) .
2. Calculate residuals $Z_j = V_n^{-1}(X_j - \theta_n)$, $j = 1, \dots, n$.
3. Find the test statistics, which can be simplified as follows

$$T_n(\mathcal{F}) = n^{-1/2} \max_{1 \leq j \leq n} \left(\sum_{k=1}^l \left[\left(\sum_{i=1}^j \cos(k\varphi_i) \right)^2 + \left(\sum_{i=1}^j \sin(k\varphi_i) \right)^2 \right] \right)^{1/2},$$

where φ_i is the polar coordinate of Z_i .

4. Choose a number of bootstrap repetitions, say M . On practice we often take a large number, for instance $M = 10000$. Then the next four steps are repeated M times.

4.1. Generate a sample (X_1^s, \dots, X_n^s) . In this example, first, generate a sample (U_1, \dots, U_n) from a uniform distribution on the unit circle, independent of $(|Z_1|, \dots, |Z_n|)$. Secondly, resample with replacement from $(|Z_1|, \dots, |Z_n|)$ to obtain (R_1, \dots, R_n) . Thirdly, $X_j^s = \theta_n + V_n R_j U_j$, $j = 1, \dots, n$.

4.2. Obtain $\hat{\theta}_n, \hat{V}_n$. In our example $\hat{\theta}_n$ is the sample mean and \hat{V}_n is the square root of the sample covariance for the sample (X_1^s, \dots, X_n^s) .

4.3. Calculate residuals $\hat{Z}_j = \hat{V}_n^{-1}(X_j^s - \hat{\theta}_n)$, $j = 1, \dots, n$.

- 4.4. Find the bootstrapped test statistics, which can be simplified as follows

$$\hat{T}_n(\mathcal{F}) = n^{-1/2} \max_{1 \leq j \leq n} \left(\sum_{k=1}^l \left[\left(\sum_{i=1}^j \cos(k\hat{\varphi}_i) \right)^2 + \left(\sum_{i=1}^j \sin(k\hat{\varphi}_i) \right)^2 \right] \right)^{1/2},$$

where $\hat{\varphi}_i$ is the polar coordinate of \hat{Z}_i .

5. Based on $(\hat{T}_n^{(1)}, \dots, \hat{T}_n^{(M)})$ find the empirical $(1 - \alpha)$ -quantile of the distribution of $\hat{T}_n(\mathcal{F})$, conditional on (X_1, \dots, X_n) . Let us denote it as $\hat{t}_{n,\alpha,M}$.

6. If $T_n(\mathcal{F}) \geq \hat{t}_{n,\alpha,M}$ then reject $H_0 : P$ is elliptically contoured distribution, at the significance level α .

5. Proofs

We use ideas and methods of the work [7]. Their technique was developed for ellipsoidal symmetry and is needed to be adjusted for group symmetry. Basically one should change $V^{-1}(x - \theta)$ to Ax throughout the proofs. However, there are technical difficulties associated with using transformations A instead of (θ, V) , they are hidden in the proofs of lemmas. We give a few details for completeness.

Let \mathcal{SP} denote a subset of all nonsingular linear transformations in \mathbb{R}^d . Given a transformation $A \in \mathcal{SP}$ denote $\tau_A f(\cdot) := f(A\cdot)$. For a function f on \mathbb{R}^d , let

$$\tilde{f}(x) := f(x) - m_f(x), \quad x \in \mathbb{R}^d.$$

Given a class \mathcal{G} of functions on \mathbb{R}^d , define $\tilde{\mathcal{G}} := \{\tilde{g} : g \in \mathcal{G}\}$,

$$\mathcal{G}_{\text{aff}} := \{\tau_A g : A \in \mathcal{SP}\}.$$

Now the process ξ_n is represented as

$$\xi_n(f) = n^{1/2} \int_{\mathbb{R}^d} \tau_{A_n} \tilde{f} dP_n, \quad f \in \mathcal{F}$$

and the process $\hat{\xi}_n$ as

$$\hat{\xi}_n(f) = n^{1/2} \int_{\mathbb{R}^d} \tau_{\hat{A}_n} \tilde{f} d\hat{P}_n, \quad f \in \mathcal{F}.$$

Clearly,

$$\mathbb{E}(f; A) = \int_{\mathbb{R}^d} \tau_A \tilde{f} dP.$$

Define

$$\mathbb{E}^s(f; A) = \int_{\mathbb{R}^d} \tau_A \tilde{f} dP^s.$$

Given a function g on \mathbb{R}^d , we can write

$$\begin{aligned} \int_{\mathbb{R}^d} g(x) P_n^s(dx) &= \int_{\mathbb{R}^d} \int_{\mathcal{S}} g(A_n^{-1} S y) m(dS) \Pi_n(dy) \\ &= \int_{\mathbb{R}^d} M_g(A_n; x) P_n(dx), \end{aligned} \quad (3)$$

where

$$M_g(A; x) := \int_{\mathcal{S}} g(A^{-1} S A x) m(dS).$$

A similar computation shows that

$$\int_{\mathbb{R}^d} g(x) P^s(dx) = \int_{\mathbb{R}^d} M_g(A_0; x) P(dx) = \int_{\mathbb{R}^d} M_g(A_0; x) P^s(dx). \quad (4)$$

Let

$$\Gamma(g; A) := \int_{\mathbb{R}^d} M_g(A; x) P(dx).$$

Given a class \mathcal{G} of functions, define

$$\mathcal{G}^2 := \{gh : g \in \mathcal{G}, h \in \mathcal{G}\},$$

$$\mathcal{M}(\mathcal{G}) := \{M_g(A, \cdot) : g \in \mathcal{G}, A \in \mathcal{SP}\}.$$

We reformulate the following versions of lemmas from [7] that describe smoothness properties of the functions introduced above and Donsker properties of the classes of functions given above. The convergence of transformations is with respect to the operator norm on the set of all linear transformations. Smoothness conditions are used in the proof of Lemma 1. Properties of Vapnik-Chervonenkis subgraph classes are used in the proof of Lemma 2. See [18] for details on Vapnik-Chervonenkis, Glivenko-Cantelli, and Donsker classes of functions.

Lemma 1 Suppose that P and \mathcal{F} satisfy the smoothness conditions (S). Then the following statements hold:

(C1) If $A \rightarrow A_0$ then

$$\sup_{f \in \mathcal{F}} \int_{\mathbb{R}^d} |\tilde{f}_A - \tilde{f}_{A_0}|^2 P(dx) \rightarrow 0$$

and

$$\sup_{f \in \mathcal{F}} \int_{\mathbb{R}^d} |\tilde{f}_A - \tilde{f}_{A_0}|^2 P^s(dx) \rightarrow 0.$$

(C2) The function $\mathbb{E}(f; A)$ is differentiable at the point A_0 for any $f \in \mathcal{F}$, and the Taylor expansion of the first order

$$\mathbb{E}(f; A) = \mathbb{E}(f; A_0) + \mathbb{E}'_A(f; A_0)(A - A_0) + o(|A - A_0|)$$

holds uniformly in $f \in \mathcal{F}$.

(C3) Similarly, the function $\mathbb{E}^s(f; A)$ is differentiable at the point A_0 for any $f \in \mathcal{F}$, and the Taylor expansion of the first order

$$\mathbb{E}^s(f; A) = \mathbb{E}^s(f; A_0) + (\mathbb{E}^s)'_A(f; A_0)(A - A_0) + o(|A - A_0|)$$

holds uniformly in $f \in \mathcal{F}$.

(C4) The function $\Gamma(g; A)$ is continuous with respect to A at A_0 uniformly in $g \in ((\tilde{\mathcal{F}})_{\text{aff}})^2$.

(C5) The function $\Gamma(g; A)$ is differentiable at the point A_0 for any $g \in (\tilde{\mathcal{F}})_{\text{aff}}$, and, moreover, the Taylor expansion of the first order

$$\Gamma(g; A) = \Gamma(g; A_0) + \Gamma'_A(g; A_0)(A - A_0) + o(|A - A_0|)$$

holds uniformly in $g \in (\tilde{\mathcal{F}})_{\text{aff}}$. Moreover, the matrix-valued function $A \mapsto \Gamma'_A(\tau_A \tilde{f}; A_0)$ is continuous at A_0 uniformly in $f \in \mathcal{F}$.

(C6) if $A \rightarrow A_0$, then for all $\delta > 0$

$$\sup \left\{ \int_{\mathbb{R}^d} |M_{\tau_B \tilde{f}}(A; x) - M_{\tau_B \tilde{f}}(A_0; x)|^2 P(dx) : B \in \mathcal{SP}, \quad |B - A_0| \leq \delta \right\} \rightarrow 0$$

and

$$\sup_{f \in \mathcal{F}} \int_{\mathbb{R}^d} |M_{\tau_A \tilde{f}}(A_0; x) - M_{\tau_{A_0} \tilde{f}}(A_0; x)|^2 P(dx) \rightarrow 0.$$

Lemma 2 For a uniformly bounded semialgebraic subgraph class \mathcal{F} , the classes $(\tilde{\mathcal{F}})_{\text{aff}}$, $\mathcal{M}((\tilde{\mathcal{F}})_{\text{aff}})$ and $\mathcal{M}((\tilde{\mathcal{F}})_{\text{aff}})^2$ are uniformly Donsker and uniformly Glivenko–Cantelli.

Proof of Theorem 1. Define a process

$$\eta_n(f; A) := n^{1/2}(P_n - P)(\tau_A \tilde{f}), \quad f \in \mathcal{F}, A \in \mathcal{SP}.$$

(C1) and $(\tilde{\mathcal{F}})_{\text{aff}}$ being a P -Donsker class by Lemma 2 imply that we can use asymptotic equicontinuity to obtain

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}^* \left(\left\{ \sup_{|A - A_0| \leq \delta} \sup_{f \in \mathcal{F}} |\eta_n(f; A) - \eta_n(f; A_0)| \geq \varepsilon \right\} \right) = 0$$

for all $\varepsilon > 0$. Clearly,

$$\begin{aligned}\xi_n(f) - n^{1/2}\mathbb{E}(f; A_0) &= \eta_n(f; A_n) + n^{1/2}(\mathbb{E}(f; A_n) - \mathbb{E}(f; A_0)) \\ &= \eta_n(f; A_0) + n^{1/2}(\mathbb{E}(f; A_n) - \mathbb{E}(f; A_0)) + (\eta_n(f; A_n) - \eta_n(f; A_0)).\end{aligned}$$

If $|A_n - A_0| \leq \delta$, we have

$$\begin{aligned}\sup_{f \in \mathcal{F}} |\xi_n(f) - n^{1/2}\mathbb{E}(f; A_0) - \eta_n(f; A_0) - n^{1/2}(\mathbb{E}(f; A_n) - \mathbb{E}(f; A_0))| \\ \leq \sup_{|A-A_0| \leq \delta} \sup_{f \in \mathcal{F}} |\eta_n(f; A) - \eta_n(f; A_0)|.\end{aligned}\quad (5)$$

Note that

$$\eta_n(f; A_0) = n^{1/2} \int_{\mathbb{R}^d} \tau_{A_0} \tilde{f} d(P_n - P).$$

Using (5) and $n^{1/2}$ -consistency of A_n we obtain

$$\xi_n(f) - n^{1/2}\mathbb{E}(f; A_0) = n^{1/2} \int_{\mathbb{R}^d} \tau_{A_0} \tilde{f} d(P_n - P) + n^{1/2}(\mathbb{E}(f; A_n) - \mathbb{E}(f; A_0)) + o_p(1) \quad (6)$$

as $n \rightarrow \infty$ uniformly in $f \in \mathcal{F}$. It follows from (C1) and $n^{1/2}$ -consistency of A_n that

$$n^{1/2}(\mathbb{E}(f; A_n) - \mathbb{E}(f; A_0)) = n^{1/2}\mathbb{E}'_A(f, A_0)(A_n - A_0) + o_p(1), \quad n \rightarrow \infty \quad (7)$$

uniformly in $f \in \mathcal{F}$. Representations (6) and (7), the fact that $(\tilde{\mathcal{F}})_{\text{aff}}$ is a uniformly Donsker class from Lemma 2 and (C1) imply that the sequence $\xi_n(f) - n^{1/2}\mathbb{E}(f; A_0)$ converges weakly in the space $\ell^\infty(\mathcal{F})$ to the Gaussian stochastic process ξ_P . This implies the first statement of the theorem. If P is ellipsoidally symmetric then $\mathbb{E}(f; A_0) = 0$, which concludes the proof of Theorem 1.

Proof of Theorem 2. Define a process

$$\hat{\eta}_n(f; A) := n^{1/2}(\hat{P}_n - P_n^s)(\tau_A \tilde{f}), \quad f \in \mathcal{F}, \quad A \in \mathcal{SP}.$$

By Lemma 2, the class $\mathcal{M}(((\tilde{\mathcal{F}})_{\text{aff}})^2)$ is uniformly Glivenko–Cantelli. This together with (C4) and representations (3), (4) implies that

$$\int_{\mathbb{R}^d} gh dP_n^s \rightarrow \int_{\mathbb{R}^d} gh dP^s \text{ as } n \rightarrow \infty \text{ a.s.}$$

uniformly in $g, h \in (\tilde{\mathcal{F}})_{\text{aff}}$. Similarly, since the class $\mathcal{M}((\tilde{\mathcal{F}})_{\text{aff}})$ is uniformly Glivenko–Cantelli, by (C5) and representations (3), (4), we obtain

$$\int_{\mathbb{R}^d} g dP_n^s \rightarrow \int_{\mathbb{R}^d} g dP^s \text{ as } n \rightarrow \infty \text{ a.s.}$$

uniformly in $g \in (\tilde{\mathcal{F}})_{\text{aff}}$.

Since $(\tilde{\mathcal{F}})_{\text{aff}}$ is a uniformly Donsker class, we can use Corollary 2.7 in [21] to prove that a.s. $n^{1/2}(\hat{P}_n - P_n^s)$ converges weakly in the space $\ell^\infty((\tilde{\mathcal{F}})_{\text{aff}})$ to the same limit as $n^{1/2}(\tilde{P}_n - P^s)$, where \tilde{P}_n is the empirical measure based on a sample from P^s , i.e. to the P^s -Brownian bridge $W_{P^s}^\circ$. Asymptotic equicontinuity and (C1) yield that for all $\varepsilon > 0$ \mathbb{P} a.s.

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \hat{\mathbb{P}}^*(\{\sup_{|A-A_n| \leq \delta} \sup_{f \in \mathcal{F}} |\hat{\eta}_n(f; A) - \hat{\eta}_n(f; A_n)| \geq \varepsilon\}) = 0. \quad (8)$$

Define

$$\hat{\mathbb{E}}_n(f; A) := \int_{\mathbb{R}^d} \tau_A \tilde{f} dP_n^s.$$

Since $\hat{\mathbb{E}}_n(f; A_n) = 0$, we can write

$$\hat{\xi}_n(f) = \hat{\eta}_n(f; \hat{A}_n) + n^{1/2} \hat{\mathbb{E}}_n(f; \hat{A}_n) = \hat{\eta}_n(f; A_n) + n^{1/2} (\hat{\mathbb{E}}_n(f; \hat{A}_n) - \hat{\mathbb{E}}_n(f; A_n)) + (\hat{\eta}_n(f; \hat{A}_n) - \hat{\eta}_n(f; A_n)).$$

If $|\hat{A}_n - A_n| \leq \delta$, we have

$$\begin{aligned} & \sup_{f \in \mathcal{F}} |\hat{\xi}_n(f) - \hat{\eta}_n(f; A_n) - n^{1/2} (\hat{\mathbb{E}}_n(f; \hat{A}_n) - \hat{\mathbb{E}}_n(f; A_n))| \\ & \leq \sup_{|A - A_n| \leq \delta} \sup_{f \in \mathcal{F}} |\hat{\eta}_n(f; A) - \hat{\eta}_n(f; A_n)|. \end{aligned} \quad (9)$$

Note that

$$\hat{\eta}_n(f; A_n) = n^{1/2} \int_{\mathbb{R}^d} \tau_{A_n} \tilde{f} d(\hat{P}_n - P_n^s).$$

Using (8), (9) and standard asymptotic properties of the estimators A_n we obtain

$$\hat{\xi}_n(f) = n^{1/2} \int_{\mathbb{R}^d} \tau_{A_n} \tilde{f} d(\hat{P}_n - P_n^s) + n^{1/2} (\hat{\mathbb{E}}_n(f; \hat{A}_n) - \hat{\mathbb{E}}_n(f; A_n)) + o_p(1) \quad (10)$$

as $n \rightarrow \infty$ uniformly in $f \in \mathcal{F}$. Here and in what follows the remainder term $o_p(1)$ converges to 0 as $n \rightarrow \infty$ uniformly in $f \in \mathcal{F}$ in probability $\mathbb{P} \times \hat{\mathbb{P}}$.

Applying the asymptotic equicontinuity condition to the process $n^{1/2}(\hat{P}_n - P_n^s)$ and using (C1), we obtain

$$n^{1/2} \int_{\mathbb{R}^d} \tau_{A_n} \tilde{f} d(\hat{P}_n - P_n^s) = n^{1/2} \int_{\mathbb{R}^d} \tau_{A_0} \tilde{f} d(\hat{P}_n - P_n^s) + o_p(1) \quad (11)$$

as $n \rightarrow \infty$. Now we can write

$$\begin{aligned} & n^{1/2} (\hat{\mathbb{E}}_n(f; \hat{A}_n) - \hat{\mathbb{E}}_n(f; A_n)) = n^{1/2} (\hat{\mathbb{E}}_n(f; \hat{A}_n) - \mathbb{E}^s(f; \hat{A}_n)) \\ & - n^{1/2} (\hat{\mathbb{E}}_n(f; A_n) - \mathbb{E}^s(f; A_n)) + n^{1/2} (\mathbb{E}^s(f; \hat{A}_n) - \mathbb{E}^s(f; A_n)). \end{aligned} \quad (12)$$

Note that by (3) and (4)

$$\begin{aligned} & n^{1/2} (\hat{\mathbb{E}}_n(f; A) - \mathbb{E}^s(f; A)) = n^{1/2} (P_n - P)(M_{\tau_A \tilde{f}}(A_n; \cdot)) + n^{1/2} (\Gamma(\tau_A \tilde{f}; A_n) - \Gamma(\tau_A \tilde{f}; A_0)) \\ & = n^{1/2} (P_n - P)(M_{\tau_A \tilde{f}}(A_0; \cdot)) + [n^{1/2} (P_n - P)(M_{\tau_A \tilde{f}}(A_n; \cdot)) - n^{1/2} (P_n - P)(M_{\tau_A \tilde{f}}(A_0; \cdot))] \\ & + n^{1/2} (\Gamma(\tau_A \tilde{f}; A_n) - \Gamma(\tau_A \tilde{f}; A_0)). \end{aligned}$$

Since, by Lemma 2, $\mathcal{M}((\tilde{\mathcal{F}})_{\text{aff}})$ is uniformly Donsker class and since (C5) and (C6) hold, it is easy to prove the weak convergence of the processes

$$\{n^{1/2} (\hat{\mathbb{E}}_n(f; A) - \mathbb{E}^s(f; A)), f \in \mathcal{F}, A \in B(A_0)\}$$

in the space $\ell^\infty(\mathcal{F} \times B(A_0))$, where $B(A_0)$ is a ball in \mathcal{SP} with the center A_0 . Using the asymptotic equicontinuity and (C6), we obtain

$$n^{1/2} (\hat{\mathbb{E}}_n(f; \hat{A}_n) - \mathbb{E}^s(f; \hat{A}_n)) - n^{1/2} (\hat{\mathbb{E}}_n(f; A_n) - \mathbb{E}^s(f; A_n)) = o_p(1) \text{ as } n \rightarrow \infty. \quad (13)$$

It follows from (C3) and standard asymptotic properties of the estimators A_n, \hat{A}_n that

$$\begin{aligned} n^{1/2}(\mathbb{E}^s(f; \hat{A}_n) - \mathbb{E}^s(f; A_n)) &= n^{1/2}(\mathbb{E}^s)'_A(f, A_0)(\hat{A}_n - A_n) + o_p(1) \\ &= (\mathbb{E}^s)'_A(f, A_0)n^{1/2} \int_{\mathbb{R}^d} \gamma_A(x) d(\hat{P}_n - P_n^s) + o_p(1), \quad n \rightarrow \infty, \end{aligned} \quad (14)$$

uniformly in $f \in \mathcal{F}$. Relationships (10)-(14) along with, again, Corollary 2.7 in [21], imply the statement of the theorem.

6. Conclusion

We propose and study a general class of tests for group symmetry, which encompasses different types of symmetry, such as ellipsoidal and permutation symmetries. Our approach is based on supremum norms of special empirical processes combined with bootstrap.

There are several advantages to our methodology. First, the test statistics are indexed by classes of functions that are rich enough and still relatively simple to use. This provides some flexibility in choosing a suitable class of functions, thereby giving an appropriate test. Secondly, these tests are *consistent against all* possible asymmetric alternatives. Thirdly, they enjoy the property of affine invariance. Fourth, these are bootstrap tests, which could be considered as a drawback but it is a way to deal with complex nature of asymptotic null distribution of a non-bootstrap semiparametric test, and these tests have good theoretical properties. Fifth, this approach gathers separate ideas and methods developed for various types of symmetry under one umbrella. It provides a unified theory for studying statistical properties of seemingly different tests for different types of symmetry.

7. Appendix

Definition of a semialgebraic set. For any polynomial p on \mathbb{R}^m of degree less than or equal to r we will call $\{x \in \mathbb{R}^m : p(x) \geq 0\}$ a *polynomial set of degree less than or equal to r* in \mathbb{R}^m . Let $\mathcal{P}_{r,m}$ denote the class of all polynomial sets in \mathbb{R}^m of degree less than or equal to r . Then any set from the union $\bigcup\{\mathcal{A}(B_1, \dots, B_l) : B_1, \dots, B_l \in \mathcal{P}_{r,m}\}$ is called a *semialgebraic set of degree less than or equal to r and order less than or equal to l* , $\mathcal{A}(B_1, \dots, B_l)$ being the minimal set algebra generated by B_1, \dots, B_l . Let $\mathcal{SA}_{r,m,l}$ denote the class of all semialgebraic sets of degree less than or equal to r and order less than or equal to l in \mathbb{R}^m .

A class \mathcal{G} of functions on \mathbb{R}^d is a *semialgebraic subgraph class* if and only if for some r, l for all functions g from \mathcal{G} the set $\{(x, t) : g(x) \geq t \geq 0 \text{ or } g(x) \leq t \leq 0\}$ belongs to $\mathcal{SA}_{r,d+1,l}$ and for all functions g_1, g_2 from \mathcal{G} the set $\{(x, y, t) : g_1(x)g_2(y) \geq t \geq 0 \text{ or } g_1(x)g_2(y) \leq t \leq 0\}$ belongs to $\mathcal{SA}_{r,2d+1,l}$.

Conditions on P and \mathcal{F} . We also introduce the following smoothness conditions on the distribution P and the class \mathcal{F} :

(S1) P is absolutely continuous with a uniformly bounded and continuously differentiable density p such that for some $C_A > d + 1$

$$\sup_{x \in \mathbb{R}^d} (1 + |x|)^{C_A} |p'(x)| < +\infty,$$

where p' denotes the derivative of the density p .

(S2) The class \mathcal{F} is uniformly bounded and for all $\varepsilon > 0$ and $R > 0$

$$\sup_{f \in \mathcal{F}} \text{mes} \left\{ x \in \mathbb{R}^d : |x| \leq R \text{ and } \omega_f(x; \delta) \geq \varepsilon \right\} \rightarrow 0 \text{ as } \delta \rightarrow 0.$$

Here mes denotes Lebesgue measure in \mathbb{R}^d and

$$\omega_f(x; \delta) := \sup \{ |f(x_1) - f(x_2)| : |x_1 - x| \leq \delta, |x_2 - x| \leq \delta \}.$$

The classes \mathcal{F} characterize the distribution. The classes \mathcal{F} characterize the distribution in the case of the examples 1.1, 2.1, 2.3, 2.4, 3.1, 4.1 above. Indeed, this is a well-known property of the classes used in examples 1.1, 2.3 and 4.1. As to the Example 2.4, we refer, e.g., to the paper [22] for similar statements. To prove that this is the case in Example 2.1 (and in Example 3.1, similarly), consider the map $\mathbb{R}^d \setminus \{0\} \ni x \mapsto (|x|, \frac{x}{|x|}) \in \mathbb{R}_+ \times S^{d-1}$. Since this map is a Borel isomorphism (even a homeomorphism), it suffices to show that for any two finite measures P, Q in $\mathbb{R}_+ \times S^{d-1}$ the condition

$$P((0, t] \times C) = Q((0, t] \times C) \text{ for all } t > 0, C \in \mathcal{C}$$

implies $P = Q$. We will prove that, in fact, for any two finite measures P, Q on $\mathbb{R}_+ \times \mathbb{R}^d$ the condition

$$P((0, t] \times H) = Q((0, t] \times H) \text{ for all } t > 0, H \in \mathcal{H},$$

where \mathcal{H} is the class of all half-spaces in \mathbb{R}^d , implies that $P = Q$ (the previous statement then follows, since one can consider two measures in $\mathbb{R}_+ \times \mathbb{R}^d$ both supported in $\mathbb{R}_+ \times S^{d-1}$). The condition

$$P((0, t] \times H) = Q((0, t] \times H) \text{ for all } t > 0, H \in \mathcal{H}$$

is equivalent to the following one

$$\begin{aligned} & \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} I_{(0,t]}(u) I_{(-\infty, c]}(\langle l, x \rangle) P(du, dx) \\ &= \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} I_{(0,t]}(u) I_{(-\infty, c]}(\langle l, x \rangle) Q(du, dx) \end{aligned}$$

for all $l \in S^{d-1}, t > 0, c \in \mathbb{R}$. Using a standard approximation of Borel functions by simple functions, we extend this to the equality

$$\int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \varphi(u) \psi(\langle l, x \rangle) P(du, dx) = \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \varphi(u) \psi(\langle l, x \rangle) Q(du, dx)$$

that holds for all bounded Borel functions φ, ψ . If we set $\varphi(u) := e^{isu}$ and $\psi(u) := e^{iu}$, we obtain that the characteristic functions of P and Q are equal, which implies that $P = Q$.

Acknowledgments

The author would like to acknowledge the support by NSF grant DMS-0806176 and to thank anonymous referees for a number of comments and suggestions that improved this paper.

References

1. Beran, R. Testing for ellipsoidal symmetry of a multivariate density. *Ann. Statist.*, **1979**, *7*, 150-162.
2. Romano, J. Bootstrap and randomization tests of some nonparametric hypothesis. *Ann. Statist.* **1989**, *17*, 141-159.
3. Baringhaus, L. Testing for spherical symmetry of a multivariate distribution. *Ann. Statist.* **1991**, *19*, 899-917.
4. Heathcote, C.R.; Rachev, S.T.; Cheng, B. Testing Multivariate Symmetry. *J. Multivariate Analysis* **1995**, *54*, 91-112.
5. Koltchinskii, V.; Li, L. Testing for spherical symmetry of a multivariate distribution. *J. Multivariate Analysis* **1998**, *65*, 228-244.
6. Manzotti, A.; Pérez, F.J.; Quiroz, A.J. A procedure for testing the null hypothesis of elliptical symmetry. *J. Multivariate Analysis* **2002**, *81*, 2, 274-285.
7. Koltchinskii, V.; Sakhanenko, L. Testing for ellipsoidal symmetry of a multivariate distribution. In *High Dimensional Probability II.*; Giné, E., Mason, D., Wellner, J., Eds.; Progress in probability, Birkhäuser: Boston, MA, USA, 2000; pp. 493-510.
8. Huffer, F.W.; Park, C. A test for elliptical symmetry. *J. Multivariate Analysis* **2007**, *98*, 256-281.
9. Sakhanenko, L. Testing for Ellipsoidal Symmetry: A comparison study. *Comp. Stat. & Data Analysis* **2008**, *53*, 565-581.
10. Beran, R.; Millar, P. Multivariate symmetry models. In *Festschrift for Lucien le Cam. Research papers in Probability and Statistics*; Pollard, D., Torgersen, E., Yang, G., Eds.; Springer: New York, NY, USA, 1997; pp. 13-42.
11. Wormleighton, R. Some tests of permutation symmetry. *Ann. Math. Statist.* **1959**, *30*, 1005-1017.
12. Li, K.-C. Sliced inverse regression for dimensional reduction (with discussion). *J. Am. Statist. Assoc.* **1991**, *86*, 316-342.
13. Petitjean, M. Chirality and Symmetry Measures: a Transdisciplinary Review. *Entropy* **2003**, *5*, 271-312.
14. Maronna, R. Robust M -estimators of multivariate location and scatter. *Ann. Statist.* **1976**, *4*, 51-67.
15. Hill, D.; Rao, P. Tests of symmetry based on Cramer-von Mises statistics. *Biometrika* **1977**, *64*, 489-494.
16. Quiroz, A.J.; Dudley, R. Some New Tests for Multivariate Normality. *Probab. Theory and Related Fields* **1991**, *87*, 521-546.
17. Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Statist. Assoc.* **1937**, *32*, 675-701.
18. van der Vaart, A.W.; Wellner, J. *Weak Convergence and Empirical Processes with Applications to Statistics*; Springer-Verlag: New York, NY, USA, 1996.
19. Dudley, R.M. *Uniform Central Limit Theorem*; Cambridge University Press: Cambridge, UK, 1999.
20. Cirel'son, B. The density of the distribution of the maximum of a Gaussian process. *Theory of Probability and Appl.* **1975**, *20*, 847-855.
21. Giné, E.; Zinn, J. Gaussian characterization of uniform Donsker classes. *Ann. Probab.* **1991**, *19*, 758-782.

22. Koldobskii, A.L. Inverse problem for potentials of measures in Banach spaces. In *Probab. Theory and Math. Stat.*; Grigelionis, B., Prohorov, Y., Sazonov, V., Eds.; VSP-Mokslas: Vilnius, Lithuania, 1990; Vol. 1, pp. 627-637.

© 2009 by the author; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license <http://creativecommons.org/licenses/by/3.0/>.