Comportement asymptotique de tests de Sobolev sur la sphère unité.

Eduardo García-Portugués, Davy Paindaveine and Thomas Verdebout

ECARES and Mathematics Department, Université Libre de Bruxelles, Boulevard du Triomphe, CP210, B-1050 Brussels, Belgium, email: tverdebo@ulb.ac.be.

Résumé. L'un des problèmes les plus classiques de la statistique multivariée est le problème de test d'uniformité sur l'hypersphère unité. Plutôt que de se limiter à des tests qui ne peuvent détecter que des types spécifiques d'alternatives, nous considérons la vaste classe des tests de Sobolev. Bien que certains de ces tests soient connus pour être "omnibus", leur comportement asymptotique sous l'alternative ainsi que leurs taux de détection, de manière inattendue, restent largement inexplorés. Pour améliorer cette situation, nous étudions en détail les puissances asymptotiques locales des tests de Sobolev dans le cadre d'alternatives classiques à l'uniformité, à savoir les alternatives à symétrie rotationelle. Nous montrons en particulier que les taux de détection des tests de Sobolev ne dépendent pas seulement des coefficients qui définissent ces tests, mais aussi des dérivées de la fonction angulaire sous-jacente en zéro.

Mots-clés. Données directionnelles, tests d'uniformité, tests de Sobolev.

Abstract. One of the most classical problems in multivariate statistics is considered, namely, the problem of testing isotropy, or equivalently, the problem of testing uniformity on the unit hypersphere. Rather than restricting to tests that can detect specific types of alternatives only, we consider the broad class of Sobolev tests. While these tests are known to allow for omnibus testing of uniformity, their non-null behavior and consistency rates, unexpectedly, remain largely unexplored. To improve on this, we thoroughly study the local asymptotic powers of Sobolev tests under the most classical alternatives to uniformity, namely, under rotationally symmetric alternatives. We show in particular that the consistency rate of Sobolev tests does not only depend on the coefficients defining these tests but also on the derivatives of the underlying angular function at zero.

Keywords. Directional data, uniformity tests, Sobolev tests.

1 Directional data and testing for uniformity

Directional statistics are dealing with observations that belong to the unit hypersphere $\mathbb{S}^{p-1} := \{\mathbf{u} \in \mathbb{R}^p : \|\mathbf{u}\|^2 = \mathbf{u}'\mathbf{u} = 1\}$ of $\mathbb{R}^p$ or more generally on compact Riemannian manifolds. Instances of directional data happen in meteorology (wind directions), astronomy (directions of cosmic rays, positions of stars), paleomagnetism (remanence directions), biology (protein structure, studies of animal navigation), forest sciences (directions of wildfire propagation), medicine (head normal vectors), and text mining (quantitative representation of documents in high-dimensional hyperspheres), to cite but some. Classical monographs on directional statistics are Watson (1983) and Mardia and Jupp (2000); a recent book that overviews the usage of some modern methods in directional statistics is Ley and Verdebout (2017).

When modeling directional data, that is, unit-norm multivariate vectors, a first natural question is to ask whether the directions at hand are uniformly distributed or, on the contrary, whether there exist modes of variation significantly different from uniformity. On the basis of n i.i.d. observations $\mathbf{U}_1, \ldots, \mathbf{U}_n$ with common distribution P on $\mathbb{S}^{p-1}$, the problem we tackle in this work is the problem of testing $\mathcal{H}_0 : \mathrm{P} \equiv \mathrm{Unif}(\mathbb{S}^{p-1})$ against $\mathcal{H}_1 : \mathrm{P} \neq \mathrm{Unif}(\mathbb{S}^{p-1})$, where $\mathrm{Unif}(\mathbb{S}^{p-1})$ stand for the uniform probability measure on $\mathbb{S}^{p-1}$. We study in this work tests belonging to the class of Sobolev tests for this problem. Sobolev tests are introduced in the next Section.

2 Sobolev tests

The class of so-called *Sobolev tests* has been introduced by Beran (1968, 1969) and Gine (1975). Sobolev tests are obtained using the eigenfunctions of the *Laplace–Beltrami operator* (or *Laplacian*) Δ acting on $\mathbb{S}^{p-1}$. Using the n-tuple of observations $\mathbf{U}_1, \ldots, \mathbf{U}_n$, a Sobolev test rejects the null hypothesis of uniformity $\mathcal{H}_0$ for large values of

$$S_n := \frac{1}{n} \sum_{i,j=1}^{n} \sum_{k=1}^{\infty} v_k^2 \langle \mathrm{t}_k(\mathbf{U}_i), \mathrm{t}_k(\mathbf{U}_j) \rangle, \tag{2.1}$$

where $\mathbf{u} \to \mathrm{t}_k(\mathbf{u})$ is a mapping from $\mathbb{S}^{p-1}$ to the space of eigenfunctions associated with the kth non-zero eigenvalue of the Laplacian, the v_k's are weights and $\langle f, g \rangle := \int_{\mathbb{S}^{p-1}} f(\mathbf{u}) g(\mathbf{u}) \, \mathrm{d}\mu(\mathbf{u})$ denotes the inner product on $L^2(\mathbb{S}^{p-1}, \mu)$ (μ is the surface area measure on $\mathcal{S}^{p-1}$). An explicit form for $\langle \mathrm{t}_k(\mathbf{U}_i), \mathrm{t}_k(\mathbf{U}_j) \rangle$ on $\mathbb{S}^{p-1}$ exists. More precisely, given $\mathbf{u}, \mathbf{v} \in \mathbb{S}^{p-1}$,

$$\langle \mathrm{t}_k(\mathbf{u}), \mathrm{t}_k(\mathbf{v}) \rangle = \begin{cases} 2\cos(k\angle(\mathbf{u}, \mathbf{v})), & \text{if } p = 2, \\ \left(1 + \frac{2k}{p-2}\right) C_k^{(p-2)/2}(\mathbf{u}'\mathbf{v}), & \text{if } p > 2, \end{cases} \tag{2.2}$$

where $\cos \angle(\mathbf{u}, \mathbf{v}) = \mathbf{u}'\mathbf{v}$ and C_k^{α} denote the Gegenbauer polynomial of index α and order k. Well-known Sobolev tests are

- the *Rayleigh test*. Taking $v_1 = 1$ and $v_k = 0$ for $k \geq 2$ in (2.1) we obtain the Rayleigh test statistic on $\mathbb{S}^{p-1}$ given by

  $$R_n = \frac{p}{n} \sum_{i,j=1}^{n} \mathbf{U}_i' \mathbf{U}_j. \tag{2.3}$$

  Under $\mathcal{H}_0$, R_n is asymptotically χ_p^2 distributed.

- the *Bingham test*. When $\mathbf{U} \sim \text{Unif}(\mathbb{S}^{p-1})$, then $\mathbb{E}[\mathbf{U}\mathbf{U}'] = \frac{1}{p}\mathbf{I}_p$. The Bingham test evaluates this latter sphericity property of $\mathbf{U}$ by the test statistic
$$B_n := \frac{np(p+2)}{2}\left(\text{tr}(\mathbf{S}^2) - \frac{1}{p}\right),$$
where $\mathbf{S} := \frac{1}{n}\sum_{i=1}^n \mathbf{U}_i\mathbf{U}_i'$ is the empirical covariance matrix of the $\mathbf{U}_i$'s. Under $\mathcal{H}_0$, B_n is asymptotically $\chi^2_{(p-1)(p+2)/2}$ distributed. The statistic B_n is obtained by letting $v_2 = 1$ and $v_k = 0$ for $k \neq 2$ in (2.1).

While much is known about the asymptotic behavior of several Sobolev tests under the null hypothesis of uniformity and when the dimension is fixed, less is known about the asymptotic behaviour of such tests under local alternatives, even under the rotationally symmetric alternatives defined in the next section.

3 Asymptotic powers of Sobolev tests

We consider in this work specific alternatives to the null of uniformity over the p-dimensional unit sphere $\mathcal{S}^{p-1}$, namely rotationally symmetric alternatives. A p-dimensional unit vector $\mathbf{U}$ is called *rotationally symmetric about* $\boldsymbol{\theta}(\in \mathcal{S}^{p-1})$ if and only if $\mathbf{O}\mathbf{U}$ is equal in distribution to $\mathbf{U}$ for any orthogonal $p \times p$ matrix $\mathbf{O}$ satisfying $\mathbf{O}\boldsymbol{\theta} = \boldsymbol{\theta}$. We actually restrict to rotationally symmetric densities of the form
$$\mathbf{u} \mapsto c_{p,\kappa,f} f(\kappa\, \mathbf{u}'\boldsymbol{\theta}), \qquad \mathbf{x} \in \mathcal{S}^{p-1}, \tag{3.4}$$
where $\boldsymbol{\theta}(\in \mathcal{S}^{p-1})$ is a location parameter, $\kappa(> 0)$ is a concentration parameter, and the function $f : \mathbb{R} \to \mathbb{R}^+$ is monotone strictly increasing, twice differentiable at 0, and satisfies $f(0) = f'(0) = 1$. Consider triangular arrays of observations $\mathbf{U}_{ni}$, $i = 1, \ldots, n$, $n = 1, 2, \ldots$ where the random vectors $\mathbf{U}_{ni}$, $i = 1, \ldots, n$ take values in $\mathcal{S}^{p_n-1}$. More specifically, for any $\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}$, $\kappa_n > 0$ and f as above, we will denote as $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$ the hypothesis under which $\mathbf{U}_{ni}$, $i = 1, \ldots, n$ are mutually independent and share the common density $\mathbf{u} \mapsto c_{p_n,\kappa_n,f} f(\kappa_n\, \mathbf{u}'\boldsymbol{\theta}_n)$; $\mathrm{P}^{(n)}_0$ will denote triangular arrays of uniformly distributed observations. We have the following result (note that in this result p_n may diverges to ∞).

Proposition 3.1 *Let (p_n) be a sequence in $\{2, 3, \ldots\}$. Let $(\boldsymbol{\theta}_n)$ be a sequence such that $\boldsymbol{\theta}_n \in \mathcal{S}^{p_n-1}$ for all n, (κ_n) be a positive sequence such that $\kappa_n = O(\sqrt{\frac{p_n}{n}})$. Then, the sequence of alternative hypotheses $\mathrm{P}^{(n)}_{\boldsymbol{\theta}_n,\kappa_n,f}$ and the null sequence $\mathrm{P}^{(n)}_0$ are mutually contiguous.*

Proposition 3.1 provides the contiguous alternatives to the uniform. Some Sobolev tests described in the previous are known to allow for omnibus testing of uniformity. Their non-null behavior and consistency rates, unexpectedly, remain largely unexplored. In particular, nothing is know about their potential rate-consistency. A natural question is then "do Sobolev tests detect the contiguous alternatives of Proposition 3.1 above?" To tackle this question, we thoroughly study the local asymptotic powers of Sobolev tests under rotationally symmetric alternatives. We show in particular that the consistency rate of a Sobolev test does not only depend on the coefficients defining the test but also on the derivatives of the underlying angular function at zero.

Bibliographie

Beran, R. J. (1968). Testing for uniformity on a compact homogeneous space. *Journal of Applied Probability*, 5, pp. 177-195.

Beran, R. J. (1969). Asymptotic theory of a class of tests for uniformity of a circular distribution. *Annals of Mathematical Statistics*, 40, pp. 1196-1206.

Cutting, C., Paindaveine, D., and Verdebout, T. (2017). Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. *Annals of Statistics*, 45(3), pp. 1024-1058.

Ley, C. and Verdebout, T. (2017). Modern Directional Statistics. Chapman and Hall/CRC.

Mardia, K. V. and Jupp, P. E. (2000). *Directional Statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester.

Watson, G. S. (1983). *Statistics on Spheres*, volume 6 of University of Arkansas Lecture Notes in the Mathematical Sciences. John Wiley & Sons, New York.