On Estimation and Testing by Means of ϕ-disparities Based on m-spacings

Igor Vajda, Edward C. van der Meulen

Abstract: We consider ϕ-divergences and ϕ-disparities $D_\phi(F_0, F)$ of hypothetical and true distributions F_0 and F on the real line. We are interested in estimation of $D_\phi(F_0, F)$ and testing the hypothesis $\mathcal{H}_0 : F = F_0$ on the basis of ϕ-disparity statistics $D_{\phi,n} = D_\phi(\boldsymbol{p}_0, \boldsymbol{p}_n)$ where $\boldsymbol{p}_0$ and $\boldsymbol{p}_n$ are discrete distributions obtained by finite quantizations of F_0 and the empirical distribution F_n corresponding to F-distributed i.i.d. sample $X_1, \ldots, X_n$. The quantization is defined in such a manner that the components p_{nj} of $\boldsymbol{p}_n$ are the m-spacings $X_{n:j+m} - X_{n:j}$. We prove a limit law for the statistics $D_{\phi,n}$.

MSC 2000: 62G05, 62G30
Key words: Estimation, testing, m-spacings

1 Introduction and auxiliary results

In this paper $F(x)$ denotes an absolutely continuous distribution function on $\mathbb{R}$ with a density $f(x)$ a. s. positive on an interval $(a, b) \subseteq \mathbb{R}$ and $X_1, \ldots, X_n$ denote independent observations distributed by $F(x)$. We consider the statistical problems of testing the hypothesis $\mathcal{H}_0 : F = F_0$ for a given absolutely continuous distribution function $F_0(x)$ with a density $f_0(x)$ a. s. positive on (a, b) and estimation of the ϕ-disparities

$$D_\phi(F_0, F) = \int_a^b f(x)\, \phi\left(\frac{f_0(x)}{f(x)}\right) \mathrm{d}x, \quad \phi \in \mathbf{\Phi}. \tag{1.1}$$

Here $\mathbf{\Phi}$ denotes the class of all continuous functions $\phi(t) : (0, \infty) \mapsto \mathbb{R}$ twice differentiable locally around $t = 1$ with $\phi''(1) > 0$, $\phi(1) = 0$ and $\phi(t) - \phi'(1)\,(t - 1)$ monotone on the intervals $(0, 1)$ and $(1, \infty)$. If $\phi : (0, \infty) \mapsto \mathbb{R}$ is convex with $\phi''(1) > 0$ and $\phi(1) = 0$ then it belongs to $\mathbf{\Phi}$ and defines the ϕ-divergence of F_0 and F (cf. Csiszár [1] or Liese and Vajda [5]). Otherwise it measures the divergence in a weaker sense motivated by robustness considerations (cf. Lindsay [6] or Morales et al. [8]). The hypothesis $\mathcal{H}_0$ can be rejected when the estimate of $D_\phi(F_0, F)$ based on the observations $X_1, \ldots, X_n$ exceeds a critical value. Such estimates of $D_\phi(F_0, F)$ can also be used for a minimum disparity selection of F_0 from a given hypothetical class.

It is well known that in both the above considered problems we can assume without loss of generality that the observation space is $(0, 1]$ and $F_0(x) = x$ on

---

Supported by the AS CR grant A 1075403 and the MSMT grant 1 M 0572.

$(0,1]$. We shall do this and therefore (1.1) will be reduced to the form

$$D_\phi(F_0,F) = \int_0^1 f(x)\,\phi\left(\frac{1}{f(x)}\right)\mathrm{d}x, \quad \phi \in \mathbf{\Phi}. \tag{1.2}$$

Since the distribution function

$$F_n(x) = \frac{1}{n}\sum_{n=1}^{n} \boldsymbol{I}(x \geq X_i)$$

is not absolutely continuous, we shall replace $D_\phi(F_0,F_n)$ by the ϕ-disparity of distributions induced by F_0 and F_n on finite partitions $\mathcal{P} = \{A_1,\ldots,A_k\}$ of $(0,1]$.

If we interpret the observation space $(0,1]$ as a circle of unit circumference then arbitrary cutpoints

$$0 < a_1 < \cdots < a_k < 1 \tag{1.3}$$

define a partition $\mathcal{P}$ of the circle into k intervals where

$$A_j = (a_j, a_{j+1}] \quad \text{for} \quad 1 \leq j \leq k-1 \tag{1.4}$$

and

$$A_k = (a_k, a_1]\,. \tag{1.5}$$

In the Euclidean ordering on $(0,1]$ the A_j of (1.4) remain to be intervals but the set (1.5) becomes the union of intervals

$$A_k = (a_k, 1] \cup (0, a_1]\,. \tag{1.6}$$

Restrictions of the distributions $F_0(x) = x$ and $F_n(x)$ on $\mathcal{P}$ define discrete hypothetical and empirical distributions

$$\boldsymbol{p}_0 = (p_{0j} : 1 \leq j \leq k) \quad \text{and} \quad \boldsymbol{p}_n = (p_{nj} : 1 \leq j \leq k) \tag{1.7}$$

respectively, where

$$p_{0j} = \begin{cases} F_0(a_{j+1}) - F_0(a_j) = a_{j+1} - a_j & \text{for} \quad 1 \leq j \leq k-1 \\ F_0(1) - F_0(a_k) + F_0(a_1) = 1 - a_k + a_1 & \text{for} \quad j = k \end{cases} \tag{1.8}$$

and

$$p_{nj} = \begin{cases} F_n(a_{j+1}) - F_n(a_j) & \text{for} \quad 1 \leq j \leq k-1 \\ 1 - F_n(a_k) + F_n(a_1) & \text{for} \quad j = k. \end{cases} \tag{1.9}$$

Selecting the cutpoints (1.3) so that all probabilities p_{nj} are a. s. positive we get the ϕ-disparities

$$D_\phi(\boldsymbol{p}_0, \boldsymbol{p}_n) = \sum_{j=1}^{k} p_{nj}\,\phi\left(\frac{p_{0j}}{p_{nj}}\right), \quad \phi \in \mathbf{\Phi} \tag{1.10}$$

as functions of observations $X_1, \dots, X_n$ which may serve as statistics for testing $\mathcal{H}_0 : F = F_0$ as well as for the estimation of the ϕ-disparities $D_\phi(F_0, F)$. The very simple formula

$$D_\phi\left(\boldsymbol{p}_0, \boldsymbol{p}_n\right) = \frac{1}{k}\sum_{j=1}^{k} \phi\left(k\, p_{0j}\right) \tag{1.11}$$

is obtained if the empirical distribution is uniform,

$$\boldsymbol{p}_n = \boldsymbol{u}_k = (1/k, \dots, 1/k)\,. \tag{1.12}$$

In order to obtain the uniform distributions (1.12) we choose an arbitrary but fixed $m \geq 1$ and restrict ourselves to the products $n = n_k = mk$ for $k = 1, 2, \dots$ The convergences and asymptotic formulas will be considered for $k \to \infty$ which implies also $n \to \infty$. Further, we consider the ordered observations

$$Y_1 \equiv X_{n:1} \leq \cdots \leq Y_n \equiv X_{n:n}$$

where the inequalities are a. s. strict, and the empirical quantiles

$$F_n^{-1}(\alpha) = \inf\{x \in (0,1) : F_n(x) \geq \alpha\}$$

of orders $\alpha \in (0,1)$. Finally, we take for a fixed $1 \leq r \leq m$

$$a_j^{(r)} = F_n^{-1}\left(\frac{m(j-1)+r}{n}\right) = Y_{m(j-1)+r}, \quad 1 \leq j \leq k \tag{1.13}$$

as the cutpoints considered in (1.3). Then we obtain from (1.9) the uniform empirical distribution (1.12) for $k = n/m$, and from (1.8) the hypothetical distributions $\boldsymbol{p}_0^{(r)}$ given by the m-spacings

$$p_{0j}^{(r)} = Y_{mj+r} - Y_{m(j-1)+r} \quad \text{for } 1 \leq j \leq k \tag{1.14}$$

where

$$Y_{mk+r} \equiv Y_{n+r} = 1 + Y_r. \tag{1.15}$$

From here and (1.12) we get the ϕ-disparity statistics

$$D_\phi\left(\boldsymbol{p}_0^{(r)}, \boldsymbol{p}_n\right) = \frac{m}{n}\sum_{j=1}^{k} \phi\left(\frac{n}{m}(Y_{mj+r} - Y_{m(j-1)+r})\right). \tag{1.16}$$

Instead of the m different statistics (1.16), each of them employing only $\frac{1}{m}$-th of the available observations, it is convenient to use their average

$$D_{\phi,n} = \frac{1}{m}\sum_{r=1}^{m} D_\phi\left(\boldsymbol{p}_0^{(r)}, \boldsymbol{p}_n\right) = \frac{1}{n}\sum_{i=1}^{n} \phi\left(\frac{n}{m}(Y_{i+m} - Y_i)\right). \tag{1.17}$$

Morales et al. [8] studied the ϕ-disparity statistics $D_{\phi,n}$ as alternatives to the m-spacings statistics

$$U_{\phi,n} = \frac{1}{n}\sum_{i=1}^{n} \phi\left(\frac{n+1}{m}(Y_{i+m} - Y_i)\right)$$

introduced by Hall ([3]). Both these papers investigated the asymptotics of $nD_{\phi,n}$ and $nU_{\phi n}$ respectively for $m = m_k$ increasing to ∞ for $k \to \infty$.

In this paper we study the asymptotics of the ϕ-disparity statistics $D_{\phi,n}$ for fixed $m \geq 1$. In fact, we extend to $m > 1$ one of the results proved recently in Vajda and van der Meulen [9] for $m = 1$. Our results are based on the paper of Hall [2] and extend previous results of Khasimov [4], van Es [10], Misra and van der Meulen [7] and some others cited there.

2 General results

We represent the statistics $D_{\phi,n}$ defined by (1.17) for all $\phi \in \mathbf{\Phi}$ as the sum

$$D_{\phi,n} = S_{\phi,n} + T_{\phi,n} \tag{2.1}$$

where

$$S_{\phi,n} = \frac{1}{n}\sum_{i=1}^{n-m} \phi\left(\frac{n}{m}\left(Y_{i+m} - Y_i\right)\right) \tag{2.2}$$

and

$$T_{\phi,n} = \frac{1}{n}\sum_{i=n-m+1}^{n} \phi\left(\frac{n}{m}\left(Y_{i+m} - Y_i\right)\right). \tag{2.3}$$

In the first theorem we show that under mild assumptions about the alternative density $f(x)$, $x \in (0,1)$

$$T_{\phi,n} = o_p(1) \quad \text{for all } \phi \in \mathbf{\Phi}. \tag{2.4}$$

Our second theorem is based on the equality

$$S_{\phi,n} = \tilde{S}_{h_m,n} \quad \text{for} \quad h_m(t) = \phi\left(\frac{t}{m}\right) \tag{2.5}$$

where

$$\tilde{S}_{h,n} = \frac{1}{n}\sum_{i=1}^{n-m} h\left(n\left(Y_{i+m} - Y_i\right)\right) \tag{2.6}$$

is a statistic of the form studied in Theorem 1 of Hall [2] for $h : (0,\infty) \mapsto \mathbb{R}$.

Theorem 2.1. *Let there exist limits*

$$f(0) = \lim_{x \downarrow 0} \frac{F(x)}{x} > 0 \quad \text{and} \quad f(1) = \lim_{x \uparrow 1} \frac{1 - F(x)}{1 - x} > 0. \tag{2.7}$$

Then the statistics $D_{\phi,n}$ of (1.17) *and $S_{\phi,n}$ of* (2.2) *are asymptotically equivalent in the sense that their difference $T_{\phi,n}$ satisfies* (2.4).

Proof. If the index i in the sum (2.3) is of the form $i = n - m + r$ then we get from (1.15) for each $1 \le r \le m$

$$Y_{i+m} - Y_i = Y_r + 1 - Y_{n-m+r} = F^{-1}(W_r) + 1 - F^{-1}(W_{n-m+r})$$

where

$$W_s = \frac{Z_1 + \cdots + Z_s}{Z_1 + \cdots + Z_{n+1}}, \quad 1 \le s \le n$$

and Z_i are independent standard exponential random variables (see, e. g. Hall [3], p. 208). Since for all fixed r and s under consideration

$$W_r = o_p(1), \quad W_{n-s} = 1 + o_p(1) \quad \text{and} \quad nW_s = O_p(1),$$

it follows from (2.7) and from the law of large numbers for the standard exponential Z_i

$$nF^{-1}(W_r) = nW_r \frac{F^{-1}(W_r)}{W_r} = O_p(1),$$

and similarly

$$n\left(1 - F^{-1}(W_{n-m+r})\right) = O_p(1).$$

Hence for every i between $n - m$ and n

$$\phi\left(\frac{n}{m}(Y_{i+m} - Y_i)\right) = O_p(1)$$

so that (2.4) follows from the definition of $T_{\phi,n}$ in (2.3). □

In the following theorem we consider the subspace of the functions $\phi \in \mathbf{\Phi}$ satisfying for some ξ, $\eta : (0, \infty) \mapsto \mathbb{R}$ and all s, $t > 0$ the functional equation

$$\phi(st) = \xi(s)\,\phi(t) + \phi(s) + \eta(s)\,(t - 1) \tag{2.8}$$

and the linear space $\boldsymbol{H}_m$ of continuous functions $h : (0, \infty) \mapsto \mathbb{R}$ satisfying for some constants a, $c > 0$ and $0 < b < m$ the condition

$$|h(t)| \le c\left(t^a + t^{-b}\right). \tag{2.9}$$

It is easy to verify (cf. Lemma 3.1 in [9]) that the functions ξ and η satisfying (2.8) are continuous with

$$\xi(1) = 1 \quad \text{and} \quad \eta(1) = 0 \tag{2.10}$$

and

$$h(t) \in \boldsymbol{H}_m \Rightarrow h\left(\frac{t}{m}\right) \in \boldsymbol{H}_m. \tag{2.11}$$

As examples of functions $\phi \in \boldsymbol{\Phi} \cap \boldsymbol{H}_m$ satisfying (2.8) for ξ, $\eta \in \boldsymbol{H}_m$ one can take

$$\phi(t) = \phi_\alpha(t) = \frac{t^\alpha - 1}{\alpha(\alpha-1)} \quad \text{with} \quad \xi(t) = \xi_\alpha(t) = t^\alpha \quad \text{and} \quad \eta(t) = \eta_\alpha(t) = 0$$

for $\alpha > -m$ different from 0 and 1 or $\phi(t) = \eta(t) = t \ln t$ and $\xi(t) = t$. These functions define well known ϕ-divergences by (1.1), (1.2). The corresponding ϕ-divergence statistics $D_{\phi,n}$ are obtained from (1.17).

In the rest of the paper we consider on $(0, \infty)$ the gamma density

$$g_m(t) = \frac{t^{m-1} e^{-t}}{\Gamma(m)} \tag{2.12}$$

defining the linear functional

$$\langle h, m \rangle = \int_0^\infty h\left(\frac{t}{m}\right) g_m(t)\, \mathrm{d}t \tag{2.13}$$

on $\boldsymbol{H}_m$.

Theorem 2.2. *Let the density $f(x)$ of $F(x)$ be piecewise continuous and bounded away from 0 and ∞ on $(0,1)$. Then for all $\phi \in \boldsymbol{\Phi} \cap \boldsymbol{H}_m$ satisfying* (2.8) *for some ξ, $\eta \in \boldsymbol{H}_m$ takes place the stochastic convergence*

$$D_{\phi,n} \xrightarrow{p} \mu_\phi(f) \tag{2.14}$$

to the constant

$$\mu_\phi(f) = \langle \xi, m \rangle\, D_\phi(F_0, F) + \langle \phi, m \rangle \quad (\text{cf. } (2.13)). \tag{2.15}$$

Proof. The distributions under consideration satisfy the assumptions of Theorem 2.1 so that it suffices to prove (2.14) with $D_{\phi,n}$ replaced by $S_{\phi,n}$ of (2.2). Further, by Theorem 1 in Hall [2], these distributions and all $h \in \boldsymbol{H}_m$ satisfy the limit relation

$$\frac{1}{n} \sum_{i=1}^{n-m} h\left(n(Y_{i+m} - Y_i)\right) \xrightarrow{p} \tilde{\mu}_h(f)$$

where

$$\tilde{\mu}_h(t) = \frac{1}{\Gamma(m)} \int_0^1 f(x)^{m+1} \int_0^\infty s^{m-1} h(s)\, e^{-sf(x)}\, \mathrm{d}s \mathrm{d}x.$$

Hence, by (2.2), (2.5) and (2.11), $S_{\phi,n} \xrightarrow{p} \mu_\phi(f)$ for

$$\begin{aligned}\mu_\phi(f) &= \frac{1}{\Gamma(m)}\int_0^1 f(x)^{m+1}\int_0^\infty s^{m-1}\phi\left(\frac{s}{m}\right)e^{-sf(x)}\,\mathrm{d}s\mathrm{d}x \\ &= \frac{1}{\Gamma(m)}\int_0^1 f(x)\int_0^\infty t^{m-1}\phi\left(\frac{t}{mf(x)}\right)e^{-t}\,\mathrm{d}t\mathrm{d}x.\end{aligned}$$

By (2.8),

$$\phi\left(\frac{t}{mf(x)}\right) = \xi\left(\frac{t}{m}\right)\phi\left(\frac{1}{f(x)}\right) + \phi\left(\frac{t}{m}\right) + \eta\left(\frac{t}{m}\right)\left(\frac{1}{f(x)} - 1\right)$$

so that

$$\begin{aligned}\mu_\phi(f) &= \langle\xi, m\rangle\int_0^1 f(x)\,\phi\left(\frac{1}{f(x)}\right)\mathrm{d}x + \langle\phi, m\rangle + \langle\eta, m\rangle\int_0^1 (1 - f(x))\,\mathrm{d}x \\ &= \langle\xi, m\rangle\, D_\phi(F_0, F) + \langle\phi, m\rangle \quad \text{(cf. (1.2))}.\end{aligned}$$

□

The following Corollary extends the results formerly established by Khasimov [4], van Es [10], Misra and van der Meulen [7] and other cited there concerning estimation of functionals of densities $f(x)$ by means of statistics based on m-spacings.

Corollary 2.3. *If f and ϕ satisfy the assumptions of Theorem 2.2 and $\langle\xi, m\rangle \neq 0$ then the statistic $(D_{\phi,n} - \langle\phi, m\rangle)/\langle\xi, m\rangle$ consistently estimates the ϕ-disparity $D_\phi(F_0, F)$.*

References

[1] I. Csiszár. Informationstheoretische Ungleichung und ihrer Anwendung auf den Beweis der Ergodisität von Markoffschen Ketten. *Publ. Math. Inst. Hungarian Acad. Sci.*, Ser. A, 8:85–108, 1963.

[2] P. Hall. Limit theorems for sums of general functions of m-spacings. *Math. Proc. Cambridge Philos. Soc.*, 96:517–532, 1984.

[3] P. Hall. On powerful distributional tests based on sample spacings. *J. Multivar. Analysis*, 19:201–224, 1986.

[4] Sh. A. Khasimov. Asymptotic properties of functions of spacings. *Theory Probab. Appl.*, 34:298–307, 1989.

[5] F. Liese and I. Vajda. *Convex Statistical Distances.* Teubner, Leipzig, 1987.

[6] G. G. Lindsay. Efficiency versus robustness. The case of minimum Hellinger distance and other methods. *Ann. Statist.*, 22:1081–1114, 1994.

[7] N. Misra and E. C. van der Meulen. A new test of uniformity based on overlapping sample spacings. *Commun. Statist. – Theory Meth.*, 30:1435–1470, 2001.

[8] D. Morales, L. Pardo, M. C. Pardo, and I. Vajda. Limit laws for disparities of spacings. *Nonparametric Statistics*, 15: 325–342, 2003.

[9] I. Vajda, E. C. van der Meulen. *Goodness-of-fit Tests Based on Observations Quantized by Hypothetical and Empirical Quantiles.* Res. Rep. 2160, Inst. of Inform Theory, Prague, 2006.

[10] B. van Es. Estimating functionals related to density by a class of statistics based on spacings. *Scand. J. Statist.*, 19:61–72, 1992.

Igor Vajda: Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, Prague, 18208, Czech Republic, vajda@utia.cas.cz

Edward C. van der Meulen: Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, Leuven, 3001, Belgium, edward.vandermeulen@wis.kuleuven.ac.be