Journal of Statistical Theory and Applications
Vol. **18(2)**, June 2019, *pp.* 113–122
DOI: https://doi.org/10.2991/jsta.d.190514.002; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

Divergence Measures Estimation and Its Asymptotic Normality Theory Using Wavelets Empirical Processes III

Amadou Diadié Bâ[1], Gane Samb Lo[2,*], Diam Bâ[3]

[1] *Unité de Formation et de Recherche des Sciences Appliquées à la Technologie, Laboratoire d'Etudes et de Recherches en Statistiques et Développement, Gaston Berger University, Saint Louis, Sénégal*
[2] *LERSTAD, Gaston Berger University, Saint-Louis, Senegal, Evanston Drive, NW, Calgary, Canada, T3P 0J9, Associate Researcher, LSTA, Pierre et Marie University, Paris, France, Associated Professor, African University of Sciences and Technology, Abuja, Nigeria*
[3] *Unité de Formation et de Recherche des Sciences Appliquées à la Technologie, Laboratoire d'Etudes et de Recherches en Statistiques et Développement, Gaston Berger University, Saint Louis, Sénégal*

ARTICLE INFO

Article History
Received 29 Oct 2018
Accepted 24 Feb 2019

Keywords
Divergence measures estimation

ABSTRACT

In the two previous papers of this series, the main results on the asymptotic behaviors of empirical divergence measures based on wavelets theory have been established and particularized for important families of divergence measures like Rényi and Tsallis families and for the Kullback-Leibler measures. While the proofs of the results in the second paper may be skipped, the proofs of those in paper 1 are to be thoroughly proved since they serve as a foundation to the whole structure of results. We prove them in this last paper of the series. We will also address the applicability of the results to usual distribution functions.

1. INTRODUCTION AND RECALL OF THE RESULTS TO BE PROVED

For a general introduction, we refer the reader to the ten (10) first pages in [1] in which the notation and the assumption are exposed.

Let us recall here the main results we exposed in. The first is related to the empirical process based on wavelets.

Theorem 1.1. *Given the* $(X_n)_{n\geq 1}$, *defined in Condition* (8) *such that* $f \in \mathcal{B}^t_{\infty,\infty}(\mathbb{R})$ *and let* f_n *defined as Formula* (13) *and* $\mathbb{G}^w_{n,X}$ *defined as in Formula* (17). *Then, under* ***Assumptions*** [1–3], *all in* [1] *and for any bounded function h, defined on D, belonging to* $\mathcal{B}^t_{\infty,\infty}(\mathbb{R})$, *we have*

$$\sigma^{-1}_{h,n}\mathbb{G}^w_{n,X}(h) \rightsquigarrow \mathcal{N}(0,1) \quad as\, n \to \infty,$$

where we have

$$\sigma^2_{h,n} = \mathbb{E}_X\left(K_{j_n}(h)(X)\right)^2 - \left(\mathbb{E}_X\left(K_{j_n}(h)(X)\right)^2 \to \mathbb{V}ar(h(X)) \quad as\, n \to \infty.$$

Proof. Suppose that **Assumptions** 1 and 3, in [1], are satisfied and $h \in \mathcal{B}^t_{\infty,\infty}(\mathbb{R})$.

We have

$$\begin{aligned}\int_D \left(f_n(x) - f(x)\right)h(x)dx &= \left(\mathbb{P}_{n,X}\left(K_{j_n}(h)\right) - \mathbb{E}_X(h)\right.\\ &= \left(\mathbb{P}_{n,X} - \mathbb{E}_X\right)\left(K_{j_n}(h)\right) + \mathbb{E}_X\left(\left(K_{j_n}(h)\right)(X) - h(X)\right).\end{aligned}$$

It comes that

$$\mathbb{G}^w_{n,X}(h) = \sqrt{n}\left(\mathbb{P}_{n,X} - \mathbb{E}_X\right)\left(K_{j_n}(h)\right) + \sqrt{n}R_{1,n},$$

where $R_{1,n} = \mathbb{E}_X\left(\left(K_{j_n}(h)\right)(X) - h(X)\right)$.

*Corresponding author. Email: gane-samb.lo@ugb.edu.sn.

To complete the proof, we have to show that (1) $\sqrt{n}\left(\mathbb{P}_{n,X}-\mathbb{E}_X\right)\left(K_{j_n}(h)\right)$ converges in distribution to a centered normal distribution and (2) $\sqrt{n}R_{1,n}$ converges to zero in probability, as $n\rightarrow\infty$. By the way, we will assume that, in the sequel, all the limits as meant as $n\rightarrow\infty$, unless the contrary is specified.

For the first point, we show that

$$\sqrt{n}\left(\mathbb{P}_{n,X}-\mathbb{E}_X\right)\left(K_{j_n}h(X)\right)\rightsquigarrow\mathcal{N}(0,\mathbb{V}ar(h(X)))\text{ as }n\rightarrow\infty,$$

by applying the central theorem for independent randoms variables.

Let us denote $Z_{i,n}=Z_{i,n}^{(h)}=K_{j_n}h(X_i)$, $\mu_{j_n}=\mathbb{E}Z_{i,n}$, and $\sigma_{i,n}^2=\left(\sigma_{i,n}^{(h)}\right)^2=\mathbb{V}ar\left(Z_{i,n}\right)<\infty$. Let

$$T_n=\sum_{i=1}^{n}\left(Z_{i,n}-\mu_{j_n}\right),$$

$$s_n^2=\mathbb{V}ar(T_n)=\sum_{i=1}^{n}\sigma_{i,n}^2.$$

T_n/s_n has mean 0 and variance 1; our goal is to give conditions under which

$$\frac{T_n}{s_n}\rightsquigarrow\mathcal{N}(0,1),\ \text{ as }n\rightarrow+\infty.$$

Such conditions are given in the Lindeberg-Feller-Levy conditions (See [4]), Point B, pp. 292).

We have to check that

$$(L1)\quad s_n^{-1}\max_{1\le i\le n}\left\{\sigma_{i,n}\right\}\rightarrow 0,$$

and for any $\varepsilon>0$,

$$(L2)\quad L(n)=:\frac{1}{s_n^2}\sum_{i=1}^{n}\int_{\left(|Z_{i,n}-\mathbb{E}Z_{i,n}|>\varepsilon s_n\right)}|Z_{i,n}-\mathbb{E}Z_{i,n}|^2d\mathbb{P}\rightarrow 0.$$

To prove this, let us begin to see that $\forall x\in\mathbb{R}$

$$|\left(K_{j_n}h\right)(x)-h(x)|=2^{j_n}\int_D K\left(2^{j_n}x,2^{j_n}t\right)[h(t)-h(x)]\,dt,$$

By Assumption 3 in [1], we have for any $x\in D$,

$$|K_{j_n}h(x)-h(x)|\le\int_D\Phi(u)\left|h\left(x+2^{-j_n}u\right)-h(x)\right|1_{\{(x+2^{-j_n}u)\in D\}}du. \tag{1}$$

Recall that h is uniformly continuous on D and on the compact $\mathcal{K}$ which supports Φ. We have

$$\rho(h,n)=\sup_{(s,t)\in D^2,|t-s|\le C2^{-j_n}}|h(t)-h(s)|\rightarrow 0.$$

For all $p\ge 1$, for all $x\in D$, and for $c=||\Phi||_\infty\lambda(\mathcal{K})$,

$$|K_{j_n}h(x)-h(x)|^p f(x)\,1_D(x)\le c^p\rho(h,n)^p f(x)\,1_D(x) \tag{2}$$

We get that for all $p\ge 1$, for any $n\ge 1$

$$\mathbb{E}|K_{j_n}h(X)-h(X)|^p\le c^p\rho(h,n)^p\rightarrow 0. \tag{3}$$

Then for any $1\le i\le n$,

$$|\mathbb{E}Z_{i,n}-\mathbb{E}h(X)|\le\mathbb{E}|Z_{i,n}-h(X)|\le c\rho(h,n)\rightarrow 0, \tag{4}$$

that is,

$$\max_{1\le i\le n} |\mathbb{E}Z_{i,n} - \mathbb{E}h(X)| \to 0. \tag{5}$$

We have

$$\begin{aligned}|\sigma_{i,n} - \mathbb{V}ar(h(X))^{1/2}| &= |\left[\mathbb{V}ar\left(K_{j_n}h(X_i)\right)\right]^{1/2} - [\mathbb{V}ar(h(X))]^{1/2}| \\ &= |\left[\mathbb{E}\left(Z_{i,n} - \mathbb{E}Z_{i,n}\right)^2\right]^{1/2} - \left[\mathbb{E}(h(X) - \mathbb{E}h(X))^2\right]^{1/2}| \\ &= |\|Z_{i,n} - \mathbb{E}Z_{i,n}\|_2 - \|h(X) - \mathbb{E}h(X)\|_2|.\end{aligned}$$

Hence,

$$|\sigma_{i,n} - \mathbb{V}ar(h(X))^{1/2}| \le |\|\mathbb{E}Z_{i,n} - \mathbb{E}h(X)\|_2 - \|K_{j_n}h(X_i) - h(X)\|_2|,$$

and using (3) and (5), we get that

$$\max_{1\le i\le n} |\sigma_{i,n} - \mathbb{V}ar(h(X))^{1/2}| \to 0. \tag{6}$$

We have from (2)

$$\begin{aligned}\left(Z_{i,n}\right)^2 = \left(\left(K_{j_n}h\right)(X_i)\right)^2 &= \left[h(X) + \left(\left(K_{j_n}h\right)(X_i) - h(X)\right)\right]^2 \\ &\le 2\left[(h(X))^2 + |\left(K_{j_n}h\right)(X_i) - h(X)|^2\right] \\ &\le 2\left[(h(X))^2 + c^2\rho(h,n)^2\right].\end{aligned}$$

Thus,

$$\begin{aligned}\left(Z_{i,n}\right)^2 + \mathbb{E}\left(Z_{i,n}\right)^2 &\le 2\left[|h(X))|^2 + c^2\rho(h,n)^2\right] + \mathbb{E}\left[2(h(X))^2 + 2c^2\rho(h,n)^2\right] \\ &\le 2\left[|h(X))|^2 + 2c^2\rho(h,n)^2\right] + 2\mathbb{E}(h(X))^2 \\ &\le 2(h(X))^2 + \mathbb{E}(h(X))^2) + 2c^2\rho(h,n)^2 \\ &\le \frac{1}{2}(Z + \delta_n)\end{aligned}$$

where

$$Z = 4\left((h(X))^2 + \mathbb{E}(h(X))^2\right) \text{ and } \delta_n = 8c^2\rho(h,n)^2.$$

Besides, the C_2−inequality gives

$$\begin{aligned}|Z_{i,n} - \mathbb{E}Z_{i,n}|^2 &\le 2\left(\left(Z_{i,n}\right)^2 + |\mathbb{E}Z_{i,n}|^2\right) \\ &\le Z + \delta_n.\end{aligned}$$

By the way, we have also

$$Z + \delta_n \le 8\|h\|_\infty^2 + \delta_n = \Delta_n \to 8\|h\|_\infty^2.$$

To prove $(L1)$, put $\alpha(n) = \max_{1\le i\le n}\left\{|\sigma_{i,n} - (\mathbb{V}ar(h(X)))^{1/2}|\right\}$. Then,

$$\sigma_{i,n}^2 \le \max\left(\left[-\alpha(n) + (\mathbb{V}ar(h(X)))^{1/2}\right]^2, \left[\alpha(n) + (\mathbb{V}ar(h(X)))^{1/2}\right]^2\right), \text{ for any } 1\le i\le n.$$

We get

$$s_n^2 \le n\max\left[(\mathbb{V}ar(h(X)))^{1/2} - \alpha(n)\right]^2, \left[(\mathbb{V}ar(h(X)))^{1/2} + \alpha(n)\right]^2,$$

hence

$$\frac{s_n^2}{n\mathbb{V}ar(h(X))} \le \max\left(\left[1 - \frac{\alpha(n)}{(\mathbb{V}ar(h(X)))^{1/2}}\right]^2, \left[1 + \frac{\alpha(n)}{(\mathbb{V}ar(h(X)))^{1/2}}\right]^2\right).$$

By (6), we have

$$\left|\frac{s_n^2}{n\mathbb{V}ar\,(h\,(X))}-1\right| \leq \max\left(\left|\left[1-\frac{\alpha\,(n)}{(\mathbb{V}ar\,(h\,(X)))^{1/2}}\right]^2-1\right|, \left|\left[1+\frac{\alpha\,(n)}{(\mathbb{V}ar\,(h\,(X)))^{1/2}}\right]^2-1\right|\right) \to 0.$$

And then

$$s_n^2 \sim n\mathbb{V}ar\,(h\,(X))\,. \tag{7}$$

Next

$$s_n^{-1}\max_{1\leq i\leq n}\{\sigma_{i,n}\} \leq \frac{(1+\alpha\,(n))\,(\mathbb{V}ar\,(h\,(X)))^{1/2}}{s_n} \sim \frac{(1+\alpha\,(n))}{\sqrt{n}} \to 0,$$

which proves $(L1)$.

We have

$$\begin{aligned} L\,(n) &\leq \frac{1}{s_n^2}\sum_{i=1}^{n}\int_{(Z+\delta_n>\varepsilon^2 s_n^2)}\Delta_n d\mathbb{P} \\ &= \frac{n}{s_n^2}\left(8\|h\|_\infty^2+\delta_n\right)\int_{(Z+\delta_n>\varepsilon^2 s_n^2)} d\mathbb{P} \\ &= \frac{n}{s_n^2}\left(8\|h\|_\infty^2+\delta_n\right)\mathbb{P}\left(Z+\delta_n>\varepsilon^2 s_n^2\right) \\ &\leq \frac{n}{s_n^4}\left(8\|h\|_\infty^2+\delta_n\right)\frac{\mathbb{E}\,(Z+\delta_n)}{\varepsilon^2} \end{aligned}$$

by Chebyshev's inequality. So

$$L\,(n) \sim \frac{\left(8\|h\|_\infty^2+\delta_n\right)}{n\,(\mathbb{V}ar\,(h\,(X)))^2}\frac{\delta_n+\mathbb{E}\,(Z)}{\varepsilon^2} \to 0,$$

since $s_n^4 \sim n^2\,(\mathbb{V}ar\,(h\,(X)))^2$ as $n \to +\infty$.

Which proves $(L2)$.

Now that Conditions $(L1)$ and $(L2)$ have been proved, we have

$$\frac{T_n}{s_n} \sim \mathcal{N}\,(0,1)\,,\ \text{ as } n \to +\infty. \tag{8}$$

But we have

$$\begin{aligned} \frac{T_n}{s_n} &= \frac{1}{s_n}n\left(\frac{1}{n}\sum_{i=1}^{n}(Z_{i,n}-\mu_{j_n})\right) \\ &= \frac{n}{s_n}\left(\frac{1}{n}\sum_{i=1}^{n}(Z_{i,n}-\mu_{j_n})\right) \\ &= \frac{n}{s_n}\left(\frac{1}{n}\sum_{i=1}^{n}K_{j_n}h(X_i)-\mathbb{E}K_{j_n}h(X)\right) \\ &= \frac{n}{s_n}\left(\mathbb{P}_{n,X}-\mathbb{E}_X\right)\left(K_{j_n}h\,(X)\right). \end{aligned}$$

Using (7), we get

$$\frac{T_n}{s_n} \sim \frac{n\left(\mathbb{P}_{n,X}-\mathbb{E}\right)\left(K_{j_n}h\right)}{\sqrt{n}\,(\mathbb{V}\,(h\,(X)))^{1/2}}.$$

Finally, from (8), we obtain

$$\frac{\sqrt{n}\left(\mathbb{P}_{n,X}-\mathbb{E}\right)\left(K_{j_n}h\right)}{\left(\mathbb{V}\left(h\left(X\right)\right)\right)^{1/2}} \rightsquigarrow \mathcal{N}\left(0,1\right), \text{ as } n\to+\infty,$$

that is,

$$\sqrt{n}\left(\mathbb{P}_{n,X}-\mathbb{E}\right)\left(K_{j_n}h\right) \rightsquigarrow \mathcal{N}\left(0,\mathbb{V}\left(h\left(X\right)\right)\right), \text{ as } n\to+\infty. \tag{9}$$

This ends the first point.

As to the second point, we apply Theorem 9.3 in [3] to have

$$\begin{aligned}\left|\mathbb{E}_X\left(K_{j_n}h-h\right)(X)\right| &\leq \int_D \left|(K_{j_n}h)(x)-h(x)\right|f(x)dx\\ &\leq C_3\left|\left|\left(K_{j_n}h\right)-h\right|\right|_\infty \left|\left|f\right|\right|_\infty\\ &\leq \kappa_2 C_3 2^{-j_n t}.\end{aligned}$$

Therefore, we have

$$\sqrt{n}R_{1,n}\left(h\right) \leq \kappa_2 C_3\sqrt{n}2^{-j_n t} = \kappa_2 C_3 n^{(1-2t)/8} = o_{\mathbb{P}}\left(1\right),$$

for any $1/2 < t < T$.

The two others main results are related to the asymptotics of class of the ϕ-divergence measures. They concern the almost sure efficiency of them.

Theorem 1.2. *Under* ***Assumptions*** [1–3], *C-A, C-h, C1-ϕ, C2-ϕ, and (BD) all in* [1], *we have*

$$\limsup_{n\to+\infty}\frac{\left|J\left(f_n,g\right)-J\left(f,g\right)\right|}{a_n} \leq A_1, a.s \tag{10}$$

$$\limsup_{n\to+\infty}\frac{\left|J\left(f,g_n\right)-J\left(f,g\right)\right|}{b_n} \leq A_2, a.s \tag{11}$$

$$\limsup_{(n,m)\to(+\infty,+\infty)}\left|\frac{J\left(f_n,g_m\right)-J\left(f,g\right)}{c_{n,m}}\right| \leq A_1+A_2 \;\; a.s \tag{12}$$

where a_n, b_n and c_n are as in Formulas (16) in [1].

Proof. In the proofs, we will systematically use the mean values theorem. In the multivariate handling, we prefer to use the Taylor-Lagrange-Cauchy as stated in [5], page 230. The assumptions have already been set up to meet these two rules. To keep the notation simple, we introduce the two following notations:

$$a_n = ||\Delta_n f||_\infty \;\text{ and }\; b_n = ||\Delta_n g||_\infty.$$

Recall that

$$\mathbb{G}^w_{n,X}\left(h\right) = \sqrt{n}\int_D \Delta_n f(x)\,h\left(x\right)dx \;\text{ and }\; \mathbb{G}^w_{n,Y}\left(h\right) = \sqrt{n}\int_D \Delta_n g\left(x\right)h\left(x\right)dx,$$

We start by showing that (10) holds.

We have

$$\phi\left(f_n\left(x\right),g\left(x\right)\right) = \phi\left(f(x)+\Delta_n f(x),g\left(x\right)\right).$$

So by applying the mean value theorem to the function $u_1\left(x\right)\mapsto\phi\left(u_1\left(x\right),g\left(x\right)\right)$, we have

$$\begin{aligned}\phi\left(f_n\left(x\right),g\left(x\right)\right) &= \phi\left(f(x),g\left(x\right)\right)\\ &\quad+\Delta_n f(x)\,\phi_1^{(1)}\left(f(x)+\theta_1\left(x\right)\Delta_n f(x),g\left(x\right)\right)\end{aligned} \tag{13}$$

where $\theta_1(x)$ is some number lying between 0 and 1. In the sequel, any θ_i satisfies $|\theta_i| < 1$. By applying again the mean values theorem to the function $u_2(x) \mapsto \phi_1^{(1)}\left(u_2(x), g(x)\right)$, we have

$$\begin{aligned}\Delta_n f(x)\, \phi_1^{(1)}\left(f(x) + \theta_1(x)\,\Delta_n f(x), g(x)\right) &= \Delta_n f(x)\, \phi_1^{(1)}\left(f(x), g(x)\right)\\ &\quad + \theta_1(x)\left(\Delta_n f(x)\right)^2 \phi_1^{(2)}\left(f(x) + \theta_2(x)\,\Delta_n f(x), g(x)\right),\end{aligned}$$

where $\theta_2(x)$ is some number lying between 0 and 1. We can write (13) as

$$\begin{aligned}\phi\left(f_n(x), g(x)\right) &= \phi\left(f(x), g(x)\right) + \Delta_n f(x)\, \phi_1^{(1)}\left(f(x), g(x)\right)\\ &\quad + \theta_1(x)\left(\Delta_n f(x)\right)^2 \phi_1^{(2)}\left(f(x)\right.\\ &\quad \left. + \theta_2(x)\,\Delta_n f(x), g(x)\right).\end{aligned}$$

Now we have

$$\begin{aligned}J\left(f_n, g\right) - J\left(f, g\right) &= \int_D \Delta_n f(x)\, \phi_1^{(1)}\left(f(x), g(x)\right) dx\\ &\quad + \int_D \theta_1(x)\left(\Delta_n f(x)\right)^2 \phi_1^{(2)}\left(f(x) + \theta_2(x)\Delta_n f(x), g(x)\right) dx,\end{aligned} \tag{14}$$

and hence,

$$|J\left(f_n, g\right) - J\left(f, g\right)| \leq a_n \int_D |\phi_1^{(1)}\left(f(x), g(x)\right)| dx + a_n^2 \int_D |\phi_1^{(2)}\left(f(x) + \theta_2(x)\,\Delta_n f(x), g(x)\right)| dx.$$

Therefore,

$$\limsup_{n\to\infty} \frac{|J\left(f_n, g\right) - J\left(f, g\right)|}{a_n} \leq A_1 + a_n \int_D \phi_1^{(2)}\left(f(x) + \theta_2(x)\,\Delta_n f(x), g(x)\right) dx.$$

Under the **Boundedness Assumption** (6) in [1], we know that $A_1 < \infty$ and that **condition** (19), [1], is satisfied, that is

$$\int_D \phi_1^{(2)}\left(f(x) + \theta_2(x)\,\Delta_n f(x), g(x)\right) dx \to \int_D \phi_1^{(2)}\left(f(x), g(x)\right) dx < \infty \quad \text{as} \quad n \to \infty.$$

This proves (10).

Formula (11) is obtained in a similar way. We only need to adapt the result concerning the first coordinate to the second.

The proof of (12) comes by splitting $\int_D \left(\phi\left(f_n(x), g_m(x)\right) - \phi\left(f(x), g(x)\right)\right) dx$, into the following two terms:

$$\begin{aligned}\int_D \left(\phi\left(f_n(x), g_m(x)\right) - \phi\left(f(x), g(x)\right)\right) dx &= \int_D \left(\phi\left(f_n(x), g_m(x)\right) - \phi\left(f(x), g_m(x)\right)\right) dx\\ &\quad + \int_D \left(\phi\left(f(x), g_m(x)\right) - \phi\left(f(x), g(x)\right)\right) dx\\ &\equiv I_{n,1} + I_{n,2}\end{aligned}$$

We already know how to handle $I_{n,2}$. As to $I_{n,1}$, we may still use the Taylor-Lagrange-Cauchy formula since we have

$$\| \left(f_n(x), g_m(x)\right) - \left(f(x), g_m(x)\right) \|_\infty = \| \left(f_n(x) - f(x), 0\right) \|_\infty = a_n \to 0.$$

By the Taylor-Lagrange-Cauchy (see [5], page 230), we have

$$\begin{aligned}I_{n,1} &= \int_D \Delta f_n(x) \phi\left(f_n(x) + \theta \Delta_n f(x), g_m(x)\right) dx\\ &\leq a_n \int_D \phi\left(f_n(x) + \theta \Delta f_n(x), g_m(x)\right) dx\\ &= a_n \left(A_2 + o(1)\right).\end{aligned}$$

From there, the combination of these remarks directs to the result.

The second main result concerns the asymptotic normality of the $\phi-$divergence measures.

Theorem 1.3. *Under **Assumptions** [1–3], C-A, C-h, C1-ϕ, C2-ϕ, and (BD) all in [1], we have*

$$\sqrt{n}\left(J\left(f_n, g\right) - J\left(f, g\right)\right) \rightsquigarrow \mathcal{N}\left(0, \mathbb{V}ar\left(h_1\left(X\right)\right)\right), \text{ as } n \to +\infty \tag{15}$$

$$\sqrt{n}\left(J\left(f, g_n\right) - J\left(f, g\right)\right) \rightsquigarrow \mathcal{N}\left(0, \mathbb{V}ar\left(h_2\left(Y\right)\right)\right), \text{ as } n \to +\infty \tag{16}$$

and as $n \to +\infty$ and $m \to +\infty$,

$$\left(\frac{nm}{m\mathbb{V}ar\left(h_1\left(X\right)\right) + n\mathbb{V}ar\left(h_2\left(Y\right)\right)}\right)^{1/2}\left(J\left(f_n, g_m\right) - J\left(f, g\right)\right) \rightsquigarrow \mathcal{N}\left(0, 1\right). \tag{17}$$

Proof. We start by proving (15). By going back to (14), we have

$$\begin{aligned}\sqrt{n}\left(J\left(f_n, g\right) - J\left(f, g\right)\right) &= \sqrt{n}\int_D \Delta_n f(x)\phi_1^{(1)}\left(f(x), g(x)\right)dx \\ &\quad + \int_D \theta_1(x)\sqrt{n}\left(\Delta_n f(x)\right)^2\phi_1^{(2)}\left(f(x) + \theta_2(x)\Delta_n f(x), g(x)\right)dx. \\ &= \mathbb{G}_{n,X}^{w}\left(h_1\right) + \sqrt{n}R_{2,n}\end{aligned}$$

where $R_{2,n} = \int_D \theta_1\left(x\right)\sqrt{n}\left(\Delta_n f\left(x\right)\right)^2\phi_1^{(2)}\left(f\left(x\right) + \theta_2\left(x\right)\Delta_n f\left(x\right), g\left(x\right)\right)dx.$

Now by theorem 1.1, one knows that $\mathbb{G}_{n,X}^{w}\left(h_1\right) \rightsquigarrow \mathcal{N}\left(0\right), \mathbb{V}ar\left(h_1\left(X\right)\right)$ as $n \to \infty$ provided that $h_1 \in B_{\infty,\infty}^{t}\left(\mathbb{R}\right)$. Thus, (15) will be proved if we show that $\sqrt{n}R_{2,n} = 0_{\mathbb{P}}\left(1\right)$. We have

$$|\sqrt{n}R_{2,n}| \le \sqrt{n}a_n^2\int_D \phi_1^{(2)}\left(f\left(x\right) + \theta_2\left(x\right)\Delta_n f\left(x\right), g\left(x\right)\right)dx. \tag{18}$$

Let show that $\sqrt{na_n^2} = o_{\mathbb{P}}\left(1\right)$. By the Bienaymé-Tchebychev inequality, we have, for any $\epsilon > 0$

$$\mathbb{P}\left(\sqrt{na_n^2} > \epsilon\right) = \mathbb{P}\left(a_n > \frac{\sqrt{\epsilon}}{n^{1/4}}\right) \le \frac{n^{1/4}}{\sqrt{\epsilon}}\mathbb{E}_X\left[a_n^2\right].$$

From Theorem 3 in [2], we have

$$\begin{aligned}\left(\mathbb{E}_X\left[a_n^2\right]\right)^{1/2} &= O\left(\sqrt{\frac{j_n 2^{j_n}}{n}} + 2^{-tj_n}\right) \\ &= O\left(\sqrt{\frac{1}{4\log 2}\frac{\log n}{n^{3/4}}} + n^{-t/4}\right)\end{aligned}$$

where we use the fact that $2^{j_n} \approx n^{1/4}$. Thus,

$$\left(\mathbb{P}\left(\sqrt{na_n^2} > \epsilon\right)\right)^2 = O\left(\sqrt{\frac{1}{4\log 2}\frac{\log n}{n^{1/2}}} + n^{(1-2t)/8}\right)$$

Finally $\sqrt{na_n^2} = o_{\mathbb{P}}\left(1\right)$ since

$$\sqrt{\frac{1}{4\log 2}\frac{\log n}{n^{1/2}}} + n^{(1-2t)/8} \to 0 \text{ as } n \to +\infty$$

for any $t > 1/2$. Finally from (18) and using Condition (19) in [1], we have $\sqrt{n}R_{2,n} \to_{\mathbb{P}} 0$ as $n \to +\infty$.

This ends the proof of (15).

The result (16) is obtained by a symmetry argument by swapping the role of f and g.

Now, it remains to prove Formula (17) of the theorem. Let us use bivariate Taylor-Lagrange-Cauchy formula to get,

$$\begin{aligned}
&J\left(f_n, g_m\right) - J\left(f, g\right) \\
&= \int_D \Delta_n f(x)\, \phi_1^{(1)}\left(f(x), g(x)\right) dx + \int_D \Delta_m g(x)\, \phi_2^{(1)}\left(f(x), g(x)\right) dx \\
&\frac{1}{2}\int_D \left(\Delta_n f(x)^2\, \phi_1^{(2)} + \Delta_n f(x)\, \Delta_n g(x)\, \phi_{1,2}^{(2)} + \Delta_n g(x)^2\, \phi_2^{(2)}\right)\left(u_n(x), v_n(y)\right) dx.
\end{aligned}$$

We have

$$\left(u_n(x), v_n(y)\right) = \left(f(x)\right) + \theta\Delta_n f(x),\ g(x) + \theta\Delta_n g(x).$$

Thus we get

$$J\left(f_n, g_m\right) - J\left(f, g\right) = \frac{1}{\sqrt{n}}\mathbb{G}_{n,X}^{w}\left(h_1\right) + \frac{1}{\sqrt{m}}\mathbb{G}_{m,Y}^{w}\left(h_2\right) + R_{n,m},$$

where $R_{n,m}$ is given by

$$\frac{1}{2}\int_D \left(\Delta_n f(x)^2\, \phi_1^{(2)} + \Delta_n f(x)\, \Delta_m g(x)\, \phi_{1,2}^{(2)} + \Delta_m g(x)^2\, \phi_2^{(2)}\right)\left(u_n(x), v_n(y)\right) dx.$$

But we have

$$\begin{aligned}
\mathbb{G}_{n,X}^{w}\left(h_1\right) &= N_n(1) + o_{\mathbb{P}}(1) \\
\mathbb{G}_{m,Y}^{w}\left(h_2\right) &= N_n(2) + o_{\mathbb{P}}(1),
\end{aligned}$$

where $N_n(i) \sim \mathcal{N}\left(0, \mathbb{V}ar\left(h_i(X)\right)\right)$, $i \in \{1, 2\}$ and $N_n(1)$ and $N_n(2)$ are independent.

Using this independence, we have

$$f\frac{1}{\sqrt{n}}\mathbb{G}_{n,X}^{w}\left(h_1\right) + \frac{1}{\sqrt{m}}\mathbb{G}_{m,Y}^{w}\left(h_2\right) = N\left(0, \frac{\mathbb{V}\left(h_1(X)\right)}{n} + \frac{\mathbb{V}\left(h_2(Y)\right)}{m}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{m}}\right).$$

Therefore, we have

$$J\left(f_n, g_m\right) - J\left(f, g\right) = \mathcal{N}\left(0, \frac{\mathbb{V}\left(h_1(X)\right)}{n} + \frac{\mathbb{V}\left(h_2(Y)\right)}{m}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{m}}\right) + R_{n,m}.$$

Hence,

$$\begin{aligned}
\frac{1}{\sqrt{\frac{\mathbb{V}(h_1(X))}{n} + \frac{\mathbb{V}(h_2(Y))}{m}}}\left(J\left(f_n, g_m\right) - J\left(f, g\right)\right) &= N(0,1) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\frac{1}{\sqrt{\frac{\mathbb{V}(h_1(X))}{n} + \frac{\mathbb{V}(h_2(Y))}{m}}}\right) \\
&+ o_{\mathbb{P}}\left(\frac{1}{\sqrt{m}}\frac{1}{\sqrt{\frac{\mathbb{V}(h_1(X))}{n} + \frac{\mathbb{V}(h_2(Y))}{m}}}\right) \\
&+ \frac{1}{\sqrt{\frac{\mathbb{V}(h_1(X))}{n} + \frac{\mathbb{V}(h_2(Y))}{m}}}R_{n,m}.
\end{aligned}$$

That leads to

$$\begin{aligned}
\sqrt{\frac{nm}{m\mathbb{V}\left(h_1(X)\right) + n\mathbb{V}\left(h_2(Y)\right)}}\left(J\left(f_n, g_m\right) - J\left(f, g\right)\right) &= N(0,1) + o_{\mathbb{P}}(1) \\
&+ \sqrt{\frac{nm}{m\mathbb{V}\left(h_1(X)\right) + n\mathbb{V}\left(h_2(Y)\right)}}R_{n,m},
\end{aligned}$$

since $m/(m\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y)))$ and $m/(n\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y)))$ are bounded, and then

$$o_{\mathbb{P}}\left(\frac{1}{\sqrt{n}}\frac{1}{\sqrt{\frac{\mathbb{V}(h_1(X))}{n}+\frac{\mathbb{V}(h_2(Y))}{m}}}\right) = o_{\mathbb{P}}\left(\sqrt{\frac{m}{m\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y))}}\right) = o_{\mathbb{P}}(1)$$

and

$$o_{\mathbb{P}}\left(\frac{1}{\sqrt{m}}\frac{1}{\sqrt{\frac{\mathbb{V}(h_1(X))}{n}+\frac{\mathbb{V}(h_2(Y))}{m}}}\right) = o_{\mathbb{P}}\left(\sqrt{\frac{n}{m\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y))}}\right) = o_{\mathbb{P}}(1).$$

It remains to prove that $\left|\sqrt{\frac{nm}{m\mathbb{V}(h_1(X))+n\mathbb{V}(h_2(Y))}}R_{n,m}\right| = o_{\mathbb{P}}(1)$. But we have by the continuity assumptions on ϕ and on its partial derivatives and by the uniform convergence of $\Delta_n f(x)$ and $\Delta_n g(x)$ to zero, that

$$\begin{aligned}\left|\sqrt{\frac{nm}{m\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y))}}R_{n,m}\right| \leq\ & \frac{1}{2}\left(\sqrt{n}a_n^2\left(\int_D \phi_1^{(2)}\left(f(x),g(x)\right)dx + o(1)\right)\right)\left(\sqrt{\frac{m}{m\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y))}}\right)\\ &+\frac{1}{2}\left(\sqrt{m}b_m^2\left(\int_D \phi_2^{(2)}\left(f(x),g(x)\right)dx + o(1)\right)\right)\left(\sqrt{\frac{n}{m\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y))}}\right)\\ &+\frac{1}{2}\left(\sqrt{n}a_m b_m\left(\int_D \phi_2^{(2)}\left(f(x),g(x)\right)dx + o(1)\right)\right)\left(\sqrt{\frac{n}{m\mathbb{V}(h_1(X)) + n\mathbb{V}(h_2(Y))}}\right)\end{aligned}$$

As previously, we have $\sqrt{n}a_n^2 = o_{\mathbb{P}}(1)$, $\sqrt{m}b_m^2 = o_{\mathbb{P}}(1)$ and $\sqrt{n}a_m b_m = o_{\mathbb{P}}(1)$.

From there, the conclusion is immediate.

We finish the series by this section on the applicability of our results for usual *pdf*'s.

2. APPLICABILITY OF THE RESULTS FOR USUAL PROBABILITY LAWS

Here, we address the applicability of our results on usual distribution functions. We have seen that we need to avoid infinite and null values. For example, integrals in the Rényi's and the Tsallis family, we may encounter such problems as signaled in the first pages of paper [1]. To avoid them, we already suggested to used a modification of the considered divergence measure in the following way:

First of all, it does not make sense to compare two distributions of different supports. Comparing a *pdf* with support $\mathbb{R}$, like the Gaussian one, with another with support $[0, 1]$, like the standard uniform one, is meaningless. So, we suppose that the *pdf*'s we are comparing have the same support D.

Next, for each $\varepsilon > 0$, we find a domain D_ε included in the common support D of f and g such that

$$\int_{D_\varepsilon} f(x)\,dx \leq 1-\varepsilon \text{ and } \int_{D_\varepsilon} g(x)\,dx \leq 1-\varepsilon. \tag{19}$$

And there exist two finite numbers $\kappa_1 > 0$ and $\kappa_2 > 0$, such that we have

$$\kappa_1 \leq f1_{D_\varepsilon}, g1_{D_\varepsilon} \leq \kappa_2. \tag{20}$$

Besides, we choose the D_ϵ's increasing to D as ϵ decreases to zero. We define the modified divergence measure

$$\mathcal{D}^{(\varepsilon)}\left(f,g\right) = \mathcal{D}\left(f_\varepsilon, g_\varepsilon\right), \tag{21}$$

where

$$f_\varepsilon = D_1^{-1} f 1_{D_\varepsilon},\quad g_\varepsilon = D_2^{-1} g 1_{D_\varepsilon},$$

$$\text{with } D_1 = \int_D f(x)\,dx \text{ and } D_2 = \int_D g(x)\,dx.$$

Based on the remarks that the D_ϵ's increases to D as ϵ decreases to zero and that the equality between f and g implies that of f_ε and g_ε, we recommend to replace the exact test of $f = g$ by the approximated test $f_\varepsilon = g_\varepsilon$, for ε as small as possible.

So each application should begin by a quick look at the domain D of the two *pdfs* and the founding of the appropriate sub-domain D_ε on which are applied the tests.

Assumption (20) also ensures that the *pdf*'s f_ε and g_ε lie in $\mathcal{B}^t_{\infty,\infty}$ for almost all the usual laws. Actually, according to [3], page 104, we have that $f \in lB^t_{\infty,\infty}$, for some $t > 0$, if and only if

$$\sup_{x\in\mathbb{R}}|f(x)| + \sup_{x\in\mathbb{R}}\sup_{h\neq 0}\frac{|f^{[t]}(x+h) - 2f^{[t]}(x) + f^{[t]}(x-h)|}{|h|^{t-[t]}} < \infty,$$

where $[t]$ stands for the integer part of the real number t, that is the greatest integer less or equal to t and $f^{[p]}$ denotes the $[p]$-th derivative function of f.

Whenever the functions f_ε and g_ε have $([t]+1)$-th derivatives bounded and not vanishing on D_ε, they will belong to $\mathcal{B}^t_{\infty,\infty}$. Assumption (20) has been set on purpose for this. Once this is obtained, all the functions that are required to lie on $\mathcal{B}^t_{\infty,\infty}$ for the validity of the results, effectively are in that space. All examples we will use in this sections satisfy these conditions, including the following random variables to cite a few: Gaussian, Gamma, Hyperbolic, and so on.

3. CONCLUSION

In this last paper of this series, the main results have been proved. Wavelet theory has proved to be a good framework for processing estimates of divergence measures. We believe that having exactly the values of the scaling function will give better results in our work.

ACKNOWLEDGMENTS

The three (1 &2 &3) authors acknowledges support from the World Bank Excellence Center (CEA-MITIC) that is continuously funding his research activities from starting 2014.

REFERENCES

1. A.D. Bâ, G.S. Lô, D. Bâ, J. Stat. Theory Appl. 17 (2018), 158–171.
2. E. Giné, R. Nickl, Ann. Probab. 37 (2009), 1605–1646.
3. W. Hardle, G. Kerkyacharian, D. Picard, A. Tsybakov, Wavelets, Approximation, and Statistical Applications. Lecture Notes in Statistics, Springer, New York, 1998.
4. M. Loève, Probability Theory I, fourth ed., Springer, New York, 1972.
5. G. Valiron, Théorie des fonctions, Masson, Paris, Milan, Melbourne, 1966.