Available online at www.sciencedirect.com

SCIENCE DIRECT®

C. R. Acad. Sci. Paris, Ser. I 340 (2005) 905–910

COMPTES RENDUS MATHEMATIQUE

http://france.elsevier.com/direct/CRASS1/

Statistics/Probability Theory

Test of homogeneity in semiparametric two-sample density ratio models

Amor Keziou [a], Samuela Leoni-Aubin [b]

[a] *LSTA, université Paris 6, 175, rue du Chevaleret, boîte 158, 75013 Paris, France*
[b] *Dipartimento di Scienze Statistiche, Università di Padova, via 8 Febbraio, 2, 35122 Padova, Italy*

Received 28 February 2005; accepted 18 April 2005

Available online 24 May 2005

Presented by Paul Deheuvels

Abstract

We consider estimation and test problems for some semiparametric two-sample density ratio models. The profile empirical likelihood (EL) poses an irregularity problem under the null hypothesis that the laws of the two samples are equal. We show that a 'dual' form of the profile EL is well defined even under the null hypothesis. A statistical test, based on the dual form of the EL ratio statistic (ELRS), is then proposed. We give an interpretation for the dual form of the ELRS through ϕ-divergences and 'duality' technique. The asymptotic properties of the test statistic are presented both under the null and the alternative hypotheses, and an approximation to the power function is deduced. ***To cite this article: A. Keziou, S. Leoni-Aubin, C. R. Acad. Sci. Paris, Ser. I 340 (2005).***

Résumé

Un test de comparaison de lois pour des modèles à rapport de densités semi-paramétriques. Nous considérons les problèmes d'estimation et de test à deux échantillon dans des modèles à rapport de densités semi-paramétriques. La vraisemblance empirique pose un problème d'irrégularité sous l'hypothèse nulle d'egalité des deux lois. Nous montrons qu'une forme « duale » de la vraisemblance empirique est bien définie. Un test statistique, basé sur la forme duale de la vraisemblance empirique, est ensuite proposé. Les propriétés asymptotiques de la statistique du test sont étudiées sous l'hypothèse nulle et sous l'hypothèse alternative, et une approximation pour la fonction de puissance est déduite. ***Pour citer cet article : A. Keziou, S. Leoni-Aubin, C. R. Acad. Sci. Paris, Ser. I 340 (2005).***

© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

E-mail addresses: keziou@ccr.jussieu.fr (A. Keziou), fragola@stat.unipd.it (S. Leoni-Aubin).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2005.04.034

Version française abrégée

Soient $X_1, \ldots, X_{n_0}$ un échantillon de loi P et $Y_1, \ldots, Y_{n_1}$ un échantillon de loi Q. On considère le modèle à rapport de densités semi-paramétriques suivant $\frac{\mathrm{d}Q}{\mathrm{d}P}(x) := \exp\{\alpha_T + \beta_T^T r(x)\}$ où $\theta_T^T := (\alpha_T, \beta_T^T) \in \Theta$, un ouvert de $\mathbb{R}^{1+d}$. Nous considérons les problèmes d'estimation (du parameter) et de test de l'hypothèse nulle $\mathcal{H}_0 : Q = P$. La vraisemblance empirique pose un problème d'irrégularité sous l'hypothèse nulle $\mathcal{H}_0$. Nous montrons qu'une forme « duale » de la vraisemblance empirique est bien définie, et nous proposons un test statistique basé sur la forme duale de la vraisemblance empirique. Nous obtenons la loi asymptotique de la statistique du test sous l'hypothèse nulle et également sous l'hypothèse alternative. La loi asymptotique sous l'alternative est utilisée pour donner une approximation de la fonction de puissance, ce qui induit une approximation des tailles n_0 et n_1 des échantillons qui garantit une puissance voulue pour une alternative donnée.

1. Introduction and notation

In this Note, we consider the following problems: two-sample test for comparing two populations and estimation of the parameters for some semiparametric density ratio models. We dispose of two samples: $X_1, \ldots, X_{n_0}$ with distribution P and $Y_1, \ldots, Y_{n_1}$ with distribution Q. We consider the following semiparametric density ratio model

$$\frac{\mathrm{d}Q}{\mathrm{d}P}(x) := \exp\{\alpha_T + \beta_T^T r(x)\}, \tag{1}$$

where $\theta_T^T := (\alpha_T, \beta_T^T)$ is the true unknown value of the parameter which we suppose to belong to some open set $\Theta \subset \mathbb{R}^{1+d}$. For simplicity, we sometimes write $m(\theta, x)$ instead of $\exp\{\alpha + \beta^T r(x)\}$. $r(\cdot)$ is a known function with values in $\mathbb{R}^d$. It often takes the form $r(x) = (x, x^2, \ldots, x^d)^T$, and the model (1) is sometimes called 'log-linear model' in this case. The supports of the two laws Q and P may be known or unknown, discrete or continuous. For statistical examples and motivations of the model (1), see e.g. Qin [15], Kay and Little [8] and Cox and Ferry [4] and the references therein. When the two samples $X_1, \ldots, X_{n_0}$ and $Y_1, \ldots, Y_{n_1}$ are independent, Fokianos et al. [5] present a statistical test, for the null hypothesis $\mathcal{H}_0 : Q = P$ or equivalently $\mathcal{H}_0 : \beta_T = 0$, where the test statistic is based on a 'constrained' empirical likelihood estimate of the parameter β_T (see Qin [15]) and an empirical estimate of the limit variance. In the case when the semiparametric assumption (1) fails, the test commonly used is the nonparametric Wilcoxon rank-sum test (see e.g. Hollander and Wolfe [7]). We expect it not to be powerful, since it does not use the model (1). The empirical likelihood ratio statistic is not well defined under the null hypothesis $\mathcal{H}_0 : Q = P$ (see Section 1.1 below). This problem has been observed also by Zou et al. [17] in the context of a semiparametric mixture models with known weights (see Zou et al. [17] Theorem 1). We propose to use, instead of the empirical likelihood ratio statistic, its 'dual' form (see (8)) (to perform a test of the null hypothesis $\mathcal{H}_0 : Q = P$) which is well defined regardless of the null hypothesis. Simulation results show that the observed level of the test based on the statistic (8) converges (to the nominal level) better than the observed level of the test proposed by Fokianos et al. [5]. Using ϕ-divergences and 'duality' technique, we give an interpretation for the statistic (8), which allows us to give the asymptotic law of the proposed test statistic under the alternative hypothesis. We apply this result to give an approximation to the power function in a similar way to Morales and Pardo [11] who gave some approximations to power functions of ϕ-divergences tests in parametric models. Duality technique has been used by Broniatowski [1] in order to estimate the Kullback–Leibler divergence without making use of any partitioning nor smoothing. It has been used also by Keziou [9] and Broniatowski and Keziou [2] in order to estimate ϕ-divergences between probability measures (without smoothing), and to introduce a new class of estimates and test statistics for discrete or continuous parametric models extending maximum likelihood approach; the use of the duality technique in the context of ϕ-divergences allows us also to study the asymptotic properties of the test statistics (including the likelihood ratio one) both under the null and the alternative hypotheses. Recall that a ϕ-divergence between two probability measures Q and P, when Q is absolutely continuous with respect to P, is defined by $\phi(Q,P) := \int \varphi(\mathrm{d}Q/\mathrm{d}P)\,\mathrm{d}P$ where φ is a real nonnegative convex function satisfying $\varphi(1) = 0$. Note that $\phi(Q,P)$ is nonnegative, $\phi(Q,P) = 0$ when $Q = P$. Further, if φ is strictly convex on a neighborhood of one, then $\phi(Q,P) = 0$ if and only if $Q = P$; we refer to Liese and Vajda [10] for a systematic theory of ϕ-divergences.

The rest of the Note is organized as follows: we end this section recalling the estimation method proposed by Qin [15]. In Section 2, we show that the irregularity problem of the profile empirical likelihood can be adjusted in the context of model (1). We next give a regularized version of the profile empirical likelihood using duality techniques. A statistical test, for the null hypothesis $\mathcal{H}_0 : Q = P$, is then proposed. An other point of view at the test statistic is given using ϕ-divergences and 'duality' technique. In Section 3, we study the asymptotic behavior of the proposed test statistic under the null and the alternative hypotheses with independent samples, and we give an approximation to the power function which leads to an approximation to the sample sizes n_0 and n_1 guaranteeing a desired power for a given alternative. In the sequel, we sometimes write Pf instead of $\int f(x)\,\mathrm{d}P(x)$ for any function f and any measure P.

1.1. The profile empirical likelihood (EL) and its irregularity under the null hypothesis $\mathcal{H}_0 : Q = P$

In the present setting, the estimation method proposed by Qin [15], which is based on the empirical likelihood approach (see Owen [14,13,12]), can be summarized as follows. For any $\theta \in \Theta$, the empirical likelihood of the two samples $X_1,\ldots,X_{n_0}$ and $Y_1,\ldots,Y_{n_1}$, if they are independent, is $L(\theta) := \prod_{i=1}^{n_0} p(X_i) \prod_{j=1}^{n_1} q(Y_j)$. For simplicity, denote $(t_1,\ldots,t_n)$ the combined sample $(X_1,\ldots,X_{n_0},Y_1,\ldots,Y_{n_1})$, where $n := n_0 + n_1$. Since $q(x) = m(\theta,x)p(x)$, then $L(\theta)$ writes $L(\theta) = \prod_{i=1}^{n} p(t_i) \prod_{i=n_0+1}^{n} m(\theta,t_i)$. For convenience we write p_i instead of $p(t_i)$. Hence, the log-likelihood writes $l(\theta,p) := \sum_{i=1}^{n} \log p_i + \sum_{i=n_0+1}^{n} \log[m(\theta,t_i)]$. The profile log-likelihood (in θ) is $l(\theta) := \sup_{p\in\mathcal{C}_\theta} l(\theta,p)$ where p is constrained to the set $\mathcal{C}_\theta := \{p \in \mathbb{R}_+^n$ such that $\sum_{i=1}^{n} p_i = 1$ and $\sum_{i=1}^{n} p_i[m(\theta,t_i) - 1] = 0\}$. The EL estimate of θ_T, proposed by Qin [15], is then $\tilde{\theta} := \arg\sup_{\theta\in\Theta} l(\theta)$. Qin [15] has proved that the estimate $\tilde{\theta}$ is optimal (in the sense of Godambe [6]), in the class of all estimates obtained by unbiased estimating functions (see Qin [15] Theorem 3). For a given $\theta \in \Theta$, the profile log-likelihood $l(\theta)$ is well defined (and finite) if and only if

$$\text{there exists } p \in \mathcal{C}_\theta \text{ such that } l(\theta,p) < \infty. \tag{2}$$

So, when $\beta_T \neq 0$ and if P is not degenerate, using similar arguments to those in Zou et al. [17] Theorem 1, we can show that there exists a neighborhood of θ_T, say $N(\theta_T)$, such that for all $\theta \in N(\theta_T)$, the assumption (2) holds as $n_0 \to \infty$. Hence, $\theta \in N(\theta_T) \mapsto l(\theta)$ is well defined for n_0 sufficiently large. However, when $Q = P$ (i.e., when $\beta_T = 0$), then obviously the set $\mathcal{C}_\theta$ is empty for all $\theta = (\alpha,\beta^T)^T \in \Theta$ with $\alpha \neq 0$ and $\beta = 0$. So, when $Q = P$ (i.e., when $\theta_T = 0$), there exists no neighborhood $N(\theta_T)$ of θ_T such that the profile empirical log-likelihood function $\theta \mapsto l(\theta)$ is well defined on all $N(\theta_T)$. Consequently the estimate $\tilde{\theta}$ is not well defined also in this case. In the following section, we will show, using some arguments of duality theory, that this problem can be adjusted in the context of the model (1).

2. Adjustment of the profile empirical likelihood

If the assumption (2) holds, then $l(\theta)$ is finite, and the unique 'optimal solution' (i.e., the value of p which yields the supremum of $l(\theta,p)$), as an explicit expression of $l(\theta)$ can be derived by a Lagrange multiplier argument and the Khun–Tucker Theorem (see e.g. Rockafellar [16] Section 28). Hence, under condition (2), the equality

$$l(\theta) := \sup_{p\in\mathcal{C}_\theta} l(\theta,p) = \inf_{\lambda\in\mathbb{R}} \left\{ -\sum_{i=1}^{n} \overline{\log}\big[n\big(1 + \lambda\big[m(\theta,t_i) - 1\big]\big)\big] + \sum_{i=n_0+1}^{n} \log\big[m(\theta,t_i)\big] \right\} \tag{3}$$

holds with finite values, and the unique optimal solution $(\bar{p}_1, \ldots, \bar{p}_n)$ exists and it is given by $\bar{p}_i = n^{-1}(1 + \bar{\lambda}[m(\theta, t_i) - 1])^{-1}$, for all $i = 1, \ldots, n$, where $\bar{\lambda}$ is the unique 'dual' optimal solution in (3). It is the solution (in λ) of the equation $\sum_{i=1}^{n} \frac{1}{n} \frac{m(\theta,t_i)-1}{1+\lambda[m(\theta,t_i)-1]} = 0$. In (3), $\overline{\log}(\cdot)$ is the function defined on $\mathbb{R}$ by $\overline{\log}(x) = \log x$ if $x > 0$ and $\overline{\log}(x) = -\infty$ elsewhere. The EL estimate $\tilde{\theta}$ of θ_T writes then $\tilde{\theta} := \arg\sup_{\theta\in\Theta} \inf_{\lambda\in\mathbb{R}}\{-\sum_{i=1}^{n} \overline{\log}[n(1 + \lambda[m(\theta, t_i) - 1])] + \sum_{i=n_0+1}^{n} \log[m(\theta, t_i)]\}$. By differentiation with respect to α and λ, we can see by simple calculus that the Lagrange multiplier $\bar{\lambda}$ in (3) has the explicit solution $\bar{\lambda}(\tilde{\theta}) = \frac{n_1}{n}$ which does not depend on the data. Hence, the value of the log-likelihood $l(\cdot)$ in $\tilde{\theta}$ is $l(\tilde{\theta}) = -n \log n - \sum_{i=1}^{n} \log(1 + \frac{n_1}{n}[m(\tilde{\theta}, t_i) - 1]) + \sum_{i=n_0+1}^{n} \log[m(\tilde{\theta}, t_i)]$, and the EL estimate $\tilde{\theta}$ can be written as $\tilde{\theta} = \arg\sup_{\theta\in\Theta}\{-n \log n - \sum_{i=1}^{n} \log(1 + \frac{n_1}{n}[m(\theta, t_i) - 1]) + \sum_{i=n_0+1}^{n} \log[m(\theta, t_i)]\}$. Under the null hypothesis $\mathcal{H}_0 : Q = P$, i.e., when $\beta_T = 0$, the profile log-likelihood $l(\theta)$ is not defined for some θ (see Section 1.1 above). Hence, we propose to consider, instead of $l(\theta)$, the 'dual form': $l_d(\theta) := -n \log n - \sum_{i=1}^{n} \log(1 + \frac{n_1}{n}[m(\theta, t_i) - 1]) + \sum_{i=n_0+1}^{n} \log[m(\theta, t_i)]$ which is well defined for all $\theta \in \Theta$ regardless of the null hypothesis $\mathcal{H}_0 : Q = P$, and to redefine the EL estimate as $\hat{\theta} := \arg\sup_{\theta\in\Theta} l_d(\theta)$. Note that, under condition (2), we have $\hat{\theta} = \tilde{\theta}$ and $l_d(\hat{\theta}) = l(\tilde{\theta})$. Now, we give an interpretation to the 'dual form' $S_n := 2W_d(\hat{\theta}) := 2[\sup_{\theta\in\Theta} l_d(\theta) + n \log n]$ of the empirical likelihood ratio statistic $2W(\tilde{\theta}) := 2[\sup_{\theta\in\Theta} l(\theta) + n \log n]$ (associated to the null hypothesis $\mathcal{H}_0 : Q = P$). First, denote $\rho_n := n_1/n_0$, $a_n := n\rho_n(1 + \rho_n)^{-2}$, and let Q_{n_1} and P_{n_0} to be, respectively, the empirical measures associated to the samples $Y_1, \ldots, Y_{n_1}$ and $X_1, \ldots, X_{n_0}$. By simple calculus, we can show that the statistic S_n writes as follows:

$$S_n = 2a_n \sup_{\theta\in\theta}\left\{\int f_{\rho_n}(\theta, x)\,\mathrm{d}Q_{n_1}(x) - \int g_{\rho_n}(\theta, x)\,\mathrm{d}P_{n_0}(x)\right\}, \tag{4}$$

where $f_{\rho_n}(\theta, x) := (1 + \rho_n)\log[m(\theta, x)] - (1 + \rho_n)\log[1 + \rho_n m(\theta, x)] + (1 + \rho_n)\log(1 + \rho_n)$ and $g_{\rho_n}(\theta, x) := \frac{1+\rho_n}{\rho_n}\log[1 + \rho_n m(\theta, x)] - \frac{1+\rho_n}{\rho_n}\log(1 + \rho_n)$. In (4), the sequence a_n is a normalizing term and the second term can be seen as an empirical estimate of

$$\sup_{\theta\in\Theta}\left\{\int f_\rho(\theta, x)\,\mathrm{d}Q(x) - \int g_\rho(\theta, x)\,\mathrm{d}P(x)\right\}, \tag{5}$$

where $\rho := \lim_{n\to\infty} \rho_n$ (which we suppose to be positive), $f_\rho(\theta, x) := (1 + \rho)\log[m(\theta, x)] - (1 + \rho)\log[1 + \rho m(\theta, x)] + (1 + \rho)\log(1 + \rho)$ and $g_\rho(\theta, x) := \frac{1+\rho}{\rho}\log[1 + \rho m(\theta, x)] - \frac{1+\rho}{\rho}\log(1 + \rho)$.

On the other hand, using the so-called 'dual representation of ϕ-divergences' (see Theorem 2.1 in Keziou [9] and Theorem 4.4 in Broniatowski and Keziou [3]) and choosing the class of functions $\mathcal{F} := \{x \mapsto \varphi_\rho^{\star\prime}(m(\theta, x)); \theta \in \Theta\}$, we can prove the equality

$$\sup_{\theta\in\Theta}\left\{\int f_\rho(\theta, x)\,\mathrm{d}Q(x) - \int g_\rho(\theta, x)\,\mathrm{d}P(x)\right\} = \int \varphi_\rho^\star\left(\frac{\mathrm{d}Q}{\mathrm{d}P}\right)\mathrm{d}P =: \phi^\star(Q, P), \tag{6}$$

where $\varphi_\rho^\star$ is the nonnegative real strictly convex function defined on $\mathbb{R}_+$ by

$$\varphi_\rho^\star(x) := (1 + \rho)\left[x \log x - \frac{1 + \rho x}{\rho}\log(1 + \rho x) + \frac{1}{\rho}\log(1 + \rho) + x\log(1 + \rho)\right], \tag{7}$$

which is a member of the class of ϕ-divergences. In other words, by (4), (5) and (6), $a_n^{-1} W_d(\hat{\theta})$ can be seen as an empirical estimate (which we denote $\hat{\phi}^\star(Q, P)$) of $\phi^\star(Q, P)$, the $\phi^\star$-divergence between Q and P, i.e., $\hat{\phi}^\star(Q, P) := (2a_n)^{-1} S_n$. Since $\phi^\star(Q, P)$ is nonnegative and takes value 0 only when $Q = P$, it is reasonable to perform a test that rejects the null hypothesis $\mathcal{H}_0 : Q = P$ when the statistic

$$S_n = 2a_n\hat{\phi}^\star(Q, P) = 2a_n \sup_{\theta\in\theta}\left\{\int f_{\rho_n}(\theta, x)\,\mathrm{d}Q_{n_1}(x) - \int g_{\rho_n}(\theta, x)\,\mathrm{d}P_{n_0}(x)\right\} \tag{8}$$

takes large values.

3. Asymptotic behavior of the test statistic and power approximation

In this section, for independent samples, we give the asymptotic properties of the estimate $\hat{\theta}$ (of the parameter θ_T) and the test statistic (8) both under the null and the alternative hypotheses. As an application, we obtain an approximation to the power function for a given alternative. In all the sequel, $f'(\theta, x)$ and $f''(\theta, x)$ denote respectively the gradient and the Hessian of f at the point θ, for all x and any function f. $|\cdot|$ denotes the Euclidean norm. Let $\rho_{n_1} := n_1/n$ and $\rho_{n_0} := n_0/n$, and assume that $\rho_{n_1} \to \rho_1 > 0$ and $\rho_{n_0} \to \rho_0 > 0$ when $n = n_0 + n_1 \to \infty$. Denote also $l_{\phi^\star}(\theta) := a_n[Q_{n_1} f_{\rho_n}(\theta) - P_{n_0} g_{\rho_n}(\theta)]$. In all the sequel, for simplicity, we write f and g instead of f_ρ and g_ρ defined in (5). We give our results under the following assumptions: (A.1) There exists a neighborhood $N(\theta_T)$ of θ_T such that the third order partial derivative functions $\{x \mapsto (\partial^3/\partial\theta_i\partial\theta_j\partial\theta_k) f(\theta, x);\ \theta \in N(\theta_T)\}$ (resp. $\{x \mapsto (\partial^3/\partial\theta_i\partial\theta_j\partial\theta_k) g(\theta, x);\ \theta \in N(\theta_T)\}$) are dominated by some function Q-integrable (resp. some function P-integrable); (A.2) The integrals $Q|f'(\theta_T)|^2$, $P|g'(\theta_T)|^2$, $Q|f''(\theta_T)|$ and $P|g''(\theta_T)|$ are finite, and the matrix $[Qf''(\theta_T) - Pg''(\theta_T)]$ is nonsingular.

Theorem 3.1. *Assume that assumptions* (A.1)–(A.2) *hold.*

(a) *Let* $B(\theta_T, n^{-1/3}) := \{\theta \in \Theta;\ |\theta - \theta_T| \leqslant n^{-1/3}\}$. *Then as* $n \to \infty$, *with probability one,* $l_{\phi^\star}(\theta)$ *attains its maximum value at some point* $\hat{\theta}$ *in the interior of the ball* $B(\theta_T, n^{-1/3})$, *and the estimate* $\hat{\theta}$ *satisfies* $l'_{\phi^\star}(\hat{\theta}) = 0$.

(b) $\sqrt{n}(\hat{\theta} - \theta_T)$ *converges in distribution to a centered multivariate normal random variable with covariance matrix* $LCM = S^{-1}VS^{-1}$ *where* $S = -Qf''(\theta_T) + Pg''(\theta_T)$ *and* $V = \rho_1^{-1}(Qf'(\theta_T)f'(\theta_T)^T - Qf'(\theta_T)Qf'(\theta_T)^T) + \rho_0^{-1}(Pg'(\theta_T)g'(\theta_T)^T - Pg'(\theta_T)Pg'(\theta_T)^T)$.

(c) *Under the null hypothesis* $\mathcal{H}_0 : Q = P$, *the statistic* S_n *converges in distribution to a* χ^2 *random variable with* d *degrees of freedom.*

In order to give the asymptotic properties of the test statistic S_n under the alternative hypothesis $\mathcal{H}_1 : Q \neq P$, we need the following additional assumption pertaining to the function f and g defined in (5)

(A.3) The integrals $Q(f(\theta_T)^2)$ and $P(g(\theta_T)^2)$ are finite.

Theorem 3.2. *Assume that assumptions* (A.1)–(A.3) *hold. Then, under the alternative hypothesis* $\mathcal{H}_1 : Q \neq P$, *we have* $\sqrt{a_n}\,[(2a_n)^{-1}S_n - \phi^\star(Q, P)]$ *converges in distribution to a centered normal random variable with variance* $\sigma^2(\theta_T) = \rho_0[Q(f^2) - (Qf)^2] + \rho_1[P(g^2) - (Pg)^2]$.

Remark 1. Using Theorem 3.1 part (c), we propose to reject the null hypothesis $\mathcal{H}_0 : Q = P$ if $S_n > \chi^2_\epsilon(d)$, where $\chi^2_\epsilon(d)$ is the $(1-\epsilon)$-quantile of the χ^2 distribution with d degrees of freedom. This leads to a test asymptotically of level ϵ. The asymptotic result in Theorem 3.2 allows us to give an approximation to the power function for a given alternative: for a given $\beta_T \neq 0$, we obtain for the power function $\beta(\theta_T) := P_{\theta_T}\{S_n > \chi^2_\epsilon(d)\}$ the following approximation $\beta(\theta_T) \approx 1 - F_{\mathcal{N}}(\frac{\sqrt{a_n}}{\hat{\sigma}(\theta_T)}[(2a_n)^{-1}\chi^2_\epsilon(d) - H_n(\theta_T)])$, where $F_{\mathcal{N}}(\cdot)$ is the cumulative distribution function of a normal random variable with mean zero and variance one, $\hat{\sigma}(\theta_T)^2 := \rho_{n_0}[Q_{n_1}(f(\theta_T)^2) - (Q_{n_1} f(\theta_T))^2] + \rho_{n_1}[P_{n_0}(g(\theta_T)^2) - (P_{n_0} g(\theta_T))^2]$ and $H_n(\theta_T) := Q_{n_1} f(\theta_T) - P_{n_0} g(\theta_T)$. Note also that the power $\beta(\theta_T)$, by the asymptotic result in Theorem 3.2, tends to one, as $n \to \infty$, under the alternative hypothesis $\mathcal{H}_1 : Q \neq P$.

Acknowledgements

The authors thank Professor Michel Broniatowski for his helpful discussions and suggestions leading to improvement of this paper.

References

[1] M. Broniatowski, Estimation of the Kullback–Leibler divergence, Math. Methods Statist. 12 (4) (2004) 391–409.
[2] M. Broniatowski, A. Keziou, Parametric estimation and testing through divergences, Preprint 2004-1, L.S.T. A – Université Paris 6, 2003.
[3] M. Broniatowski, A. Keziou, On the minimization of ϕ-divergences on sets of signed measures, Studia Sci. Math. Hungar. (2005), in press.
[4] T.F. Cox, G. Ferry, Robust logistic discrimination, Biometrika 78 (4) (1991) 841–849.
[5] K. Fokianos, B. Kedem, J. Qin, D.A. Short, A semiparametric approach to the one-way layout, Technometrics 43 (1) (2001) 56–65.
[6] V.P. Godambe, An optimum property of regular maximum likelihood estimation, Ann. Math. Statist. 31 (1960) 1208–1211.
[7] M. Hollander, D.A. Wolfe, Nonparametric Statistical Methods, second ed., Wiley, New York, 1999.
[8] R. Kay, S. Little, Transformations of the explanatory variables in the logistic regression model for binary data, Biometrika 74 (3) (1987) 495–501.
[9] A. Keziou, Dual representation of ϕ-divergences and applications, C. R. Acad. Sci. Paris, Ser. I 336 (10) (2003) 857–862.
[10] F. Liese, I. Vajda, Convex Statistical Distances, Teubner-Texte Math., vol. 95, Teubner, Leipzig, 1987.
[11] D. Morales, L. Pardo, Some approximations to power functions of ϕ-divergences tests in parametric models, Test 10 (2) (2001) 249–269.
[12] A.B. Owen, Empirical Likelihood, Chapman and Hall, New York, 2001.
[13] A.B. Owen, Empirical likelihood ratio confidence regions, Ann. Statist. 18 (1) (1990) 90–120.
[14] A.B. Owen, Empirical likelihood ratio confidence intervals for a single functional, Biometrika 75 (2) (1988) 237–249.
[15] J. Qin, Inferences for case-control and semiparametric two-sample density ratio models, Biometrika 85 (3) (1998) 619–630.
[16] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.
[17] F. Zou, J.P. Fine, B.S. Yandell, On empirical likelihood for a semiparametric mixture model, Biometrika 89 (1) (2002) 61–75.