Journal of Statistical Planning and Inference 142 (2012) 2554–2573

Contents lists available at SciVerse ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Divergences and duality for estimation and test under moment condition models

Michel Broniatowski [a], Amor Keziou [a,b,*]

[a] LSTA, Université Pierre et Marie Curie, Paris 6, France
[b] Laboratoire de Mathématiques de Reims, EA 4535, Université de Reims Champagne-Ardenne, France

ARTICLE INFO

Article history:
Received 8 April 2011
Received in revised form
29 February 2012
Accepted 9 March 2012
Available online 26 March 2012

Keywords:
Empirical likelihood
Generalized empirical likelihood
Minimum divergence
Efficiency
Power function
Duality
Divergence projection

ABSTRACT

We introduce estimation and test procedures through divergence minimization for models satisfying linear constraints with unknown parameter. These procedures extend the empirical likelihood (EL) method and share common features with generalized empirical likelihood approach. We treat the problems of existence and characterization of the divergence projections of probability distributions on sets of signed finite measures. We give a precise characterization of duality, for the proposed class of estimates and test statistics, which is used to derive their limiting distributions (including the EL estimate and the EL ratio statistic) both under the null hypotheses and under alternatives or misspecification. An approximation to the power function is deduced as well as the sample size which ensures a desired power for a given alternative.

1. Introduction and notation

Statistical models are often defined through estimating equations

$$\mathbb{E}[g(X,\theta)] = 0,$$

where $\mathbb{E}[\cdot]$ denotes the mathematical expectation, $g := (g_1,\ldots,g_l)^\top \in \mathbb{R}^l$ is some specified vector valued function of a random vector $X \in \mathbb{R}^m$ and a parameter vector $\theta \in \Theta \subset \mathbb{R}^d$. Examples of such models are numerous, see e.g. Qin and Lawless (1994), Haberman (1984), Sheehy (1987), McCullagh and Nelder (1983), Owen (2001) and the references therein. Denote P_0 the probability distribution of the random vector X. Then the above estimating equations can be written as

$$\int_{\mathbb{R}^m} g(x,\theta)\, dP_0(x) = 0.$$

Denoting M^1 the collection of all probability measures (p.m.) on the measurable space $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, the submodel $\mathcal{M}_\theta^1$, associated to a given value θ of the parameter, consists of all distributions Q satisfying l linear constraints induced by the vector valued function $g(.,\theta)$, namely

$$\mathcal{M}_\theta^1 := \left\{ Q \in M^1 \text{ such that } \int g(x,\theta)\, dQ(x) = 0 \right\},$$

* Corresponding author at: Laboratoire de Mathématiques de Reims, EA 4535, Université de Reims Champagne-Ardenne, France.
E-mail addresses: michel.broniatowski@upmc.fr (M. Broniatowski), amor.keziou@univ-reims.fr (A. Keziou).

0378-3758/$ - see front matter
http://dx.doi.org/10.1016/j.jspi.2012.03.013

with $l \geq d$. The statistical model which we consider can be written as

$$\mathcal{M}^1 := \bigcup_{\theta \in \Theta} \mathcal{M}^1_\theta. \tag{1.1}$$

Let $X_1, \ldots, X_n$ denote an i.i.d. sample of X with unknown distribution P_0. We denote θ_0, if it exists, the value of the parameter such that P_0 belongs to $\mathcal{M}^1_{\theta_0}$, namely the value satisfying

$$\mathbb{E}[g(X,\theta_0)] = 0,$$

and we assume obviously that θ_0 is unique. This paper addresses the two following natural questions:

Problem 1. Does P_0 belong to the model $\mathcal{M}^1$?

Problem 2. When P_0 is in the model, which is the value θ_0 of the parameter for which $\mathbb{E}[g(X,\theta_0)] = 0$? Also can we perform tests about θ_0? Can we construct confidence areas for θ_0?

When $m = d = l$, and $g(x,\theta) = x - \theta$, then the model is the same as those of Owen (1988, 1990), and in this case our interest concerns interval estimation or confidence areas construction for the parameter θ. The main interest, however, is the case where $l > d$. We introduce some examples for illustration; see Qin and Lawless (1994), Guggenberger and Smith (2005) and Owen (2001).

Example 1.1. Sometimes we have information relating the first and second moments of a random variable X (see e.g. Godambe and Thompson, 1989; McCullagh and Nelder, 1983). Let $X_1, \ldots, X_n$ be an i.i.d. sample of a random variable X with mean $\mathbb{E}(X) = \theta$, and assume that $\mathbb{E}(X^2) = h(\theta)$, where $h(\cdot)$ is a known function. Our aim is to estimate θ. The information about the distribution P_0 of X can be expressed in the form of (1.1) by taking $g(x,\theta) := (x - \theta, x^2 - h(\theta))^\top$.

Example 1.2. Let $(X_{1,1}, X_{2,1}), \ldots, (X_{1,n}, X_{2,n})$ be an i.i.d. sample of a bivariate random vector $X := (X_1, X_2)^\top$ with $\mathbb{E}(X_1) = \mathbb{E}(X_2) = \theta$. In this case, we can take $g(x,\theta) = (x_1 - \theta, x_2 - \theta)^\top$. A somewhat similar problem is when $\mathbb{E}(X_1) = c$ is known and $\mathbb{E}(X_2) = \theta$ is to be estimated, by taking $g(x,\theta) = (x_1 - c, x_2 - \theta)^\top$. Such problems are common in survey sampling (see e.g. Kuk and Mak, 1989; Chen and Qin, 1993).

Example 1.3. Let F_0 be a continuous distribution function of a random variable X with probability distribution P_0 that is symmetric about zero, namely $F_0(t) = 1 - F_0(-t)$ for all $t \in \mathbb{R}$. Consider estimation of the parameter $\theta = F_0(t)$, for a given $t \in \mathbb{R}$, from an i.i.d. sample $X_1, \ldots, X_n$ of X. This problem can be handled in the context of model (1.1) by taking $g(x,\theta) = (\mathbb{1}_{]-\infty,t]}(x) - \theta, \mathbb{1}_{]-t,+\infty[}(x) - \theta)^\top$.

We note that the Problems 1 and 2 have been investigated by many authors. Hansen (1982) considered generalized method of moments (GMM). Hansen et al. (1996) introduced the continuous updating (CU) estimate. The empirical likelihood (EL) approach, developed by Owen (1988, 1990), has been investigated in the context of model (1.1) by Qin and Lawless (1994) and Imbens (1997) introducing the EL estimate. The recent literature in econometrics focusses on such models; Smith (1997) and Newey and Smith (2004) provided a class of estimates called generalized empirical likelihood (GEL) estimates which contains the EL and the CU ones. Schennach (2007) discussed the asymptotic properties of the empirical likelihood estimate under misspecification; the author showed the important fact that the EL estimate may cease to be root n consistent when the functions g_j defining the moments conditions and the support of P_0 are unbounded. Among other results pertaining to EL, Newey and Smith (2004) stated that EL estimate enjoys optimality properties in terms of efficiency when bias corrected among all GEL estimates including the GMM one. Moreover, Corcoran (1998) and Baggerly (1998) proved that in a class of minimum discrepancy statistics (called power divergence statistics), EL ratio is the only one that is Bartlett correctable. Confidence areas for the parameter θ_0 have been considered in the seminal paper by Owen (1990). Problems 1 and 2 have been handled via EL and GEL approaches in Qin and Lawless (1994), Smith (1997) and Newey and Smith (2004) under the null hypothesis $\mathcal{H}_0 : P_0 \in \mathcal{M}^1$; the limiting distributions of the GEL estimates and the GEL test statistics have been obtained only under the model and under the null hypotheses; Imbens (1997) discusses the asymptotic properties of the EL and exponential tilting estimates under misspecification and give the formula of the asymptotic variance, using dual characterizations, without presenting the hypotheses under which their results hold. Chen et al. (2007) give the limiting distribution of the EL estimate under misspecification as well as the EL ratio statistic between a parametric model and a moment condition model. The paper by Kitamura (2007) gives a discussion of duality for GEL estimates under moment condition models. Bertail (2006) uses duality to study, under the model, the asymptotic properties of the EL ratio statistic and its Bartlett correctability; the author extends his results to semiparametric problems with infinite-dimensional parameters.

The main contribution of the present paper is the precise characterization of duality for a large class of estimates and test statistics (including GEL and EL ones) and its use in deriving the limiting properties of both the estimates and the test statistics under misspecification and under alternative hypotheses. Moreover,

(1) The approach which we develop is based on *minimum discrepancy estimates*, which extends the EL method and has common features with minimum distance and GEL techniques, using merely divergences. We propose a wide class of estimates, test statistics and confidence regions for the parameter θ_0 as well as various test statistics for *Problems* 1 and 2, all depending on the choice of the divergence.