Università degli Studi di Salerno
DIPARTIMENTO DI SCIENZE ECONOMICHE E STATISTICHE

Pietro Coretto [1] Giuseppe Storti [2]

MOMENTS BASED INFERENCE IN SMALL SAMPLES[3]

WORKING PAPER 3.159
Marzo 2005

[1]Department of Statistical Sciences — University College London (UCL) — Gower Street, London, WC1E 6BT, UK. Tel. (+44)–(0)20 76.79 18.75 — E-mail: pietro@stats.ucl.ac.uk
[2]Dipartimento di Scienze Economiche e Statistiche — Università di Salerno — Via Ponte Don Melillo; 84084 Fisciano (SA), Italy. Tel: (+39) 089 96.22.12 — E-mail: storti@unisa.it
[3]Giuseppe Storti produced the abstract and the computations in section 8, Pietro Coretto produced all the remaining sections.

Contents

Abstract

In this work we propose a nonparametric estimator for parameters which are embodied in given moment conditions. Here we are interested in small samples problems. We analyze conditions under which it is possible to represent the distribution of data points conditional on their relative frequencies as a multinomial distribution. We derive the posterior distribution for the relative frequencies of the observed data points via Bayes' Law and specifying a Dirichlet objective prior, the latter is obtained matching the prior modes for unknown relative frequencies with the optimal weights computed under a special form of the Empirical Likelihood estimator for the parameters of interest. The prior specification proposed has an interesting interpretation in terms of information-theoretic arguments. The estimators we construct in this paper share the general idea with the Bayesian Bootstrap by Rubin (1981), however its derivation starts from a different point o view, and also the prior specification over the distribution function of the data is objectively derived via Empirical Likelihood methods. We propose a Monte Carlo algorithm to derive a distribution for the parameters of interest based on the derived posterior for the relative frequencies of data points. A simulated toy example shows that in small sample the estimation proposed is more accurate than alternative non-bayesian nonparametric methods. Future works will include an asymptotic analysis of the proposed estimator, as well as more complex simulated validations for small samples.

Keywords

Empirical Likelihood, Generalized Method of Moments, Bayesian nonparametric inference.

1 Introduction

In social sciences empirical analysis we usually need to be very careful about the specification of the model. Specification issues play an important role in economics, where a little is known about the functional relationships among the variables. The practice is to derive an estimating model from the solution of a decision problem under assumptions on the primitives such as preferences, utility, constraints etc. Under conditions that restrict preferences and physical constraints, optimal solutions are usually derived, and these solutions are well known as the decision rules which are usually in the form of a set of moment conditions. The role of the statistician is then to use these moment equations and data to make inference on the parameters of interest. So, from the theory we have a set of equations of the form $\mathsf{E}[g(d,\beta)] = 0$, where $g(d,\beta)$ is a known vector-valued possibly nonlinear function of the vector of observed data d taking value in $\mathbb{R}^q$, and of the parameter of interests $\beta \in \mathbb{R}^p$. There is also need to avoid a complete parametric specification of the data generating process since there is no experimental design such that we can not immagine any probabilistic model for the data generated; thus we aim to reduce the possibly inconsistent specification of distributional assumptions. In the most part of real situations the size of the data vector is limited, i.e. we have few observations. In these critical situations inference is hard and limited in validity, so that possibly sophisticated statistical methods are necessary. For example a common situation is that we observe a finite sequence $\{y_i, x_i\}_{i=1}^n$, we have a solution of a decision problem as an equation of the form $y_i = g(x_i, \gamma) + \varepsilon_i$, $i = 1, 2, \ldots, n$, where γ is a parameter of the model, $g(\cdot)$ is a smooth real

valued function and ε_i is an unobserved random variable. In general no distributional assumption can be made on ε_i and x_i, in most cases ε has the meaning of non systematic error component so that we are willing to assume that $\mathsf{E}[\varepsilon] = 0 = \mathsf{E}[y - \gamma x]$, and $\{\varepsilon\}_{i=1}^{n}$ is an i.i.d. sequence. Under random sampling and some other assumption on the data generating process, from the frequentist point of view, we can use the previous moment condition to give an estimate of γ without any assumption on unknown distributions by using one of the several methods described below.

Inference in such situations have been done using the generalised method of moments (GMM) by Hansen (1982). Recently the GMM approach has been successfully combined with the empirical likelihood approach (EL) (see Owen, 1988; Qin & Lawless, 1994; Owen, 2001), the exponential tilting (ET) (see Efron, 1981; Imbens, 1997; Kitamura & Stutzer, 1997; Imbens et al., 1998), the generalised method of moments with continuous updating (see Hansen et al., 1996), and generalised empirical likelihood (see Smith, 1997; Newey & Smith, 2004). In GMM the vector of parameters is estimated solving the moment conditions if the model is just identified (i.e. $p = q$), or via minimum distance methods when dealing with over-identifications (i.e. $q > p$); Hansen (1982) showed that for large samples the GMM estimator is consistent and asymptotically normal. On the other hand, EL estimation defines a function which is an analogue to a parametric likelihood function and yet enables inference that does not require distributional assumptions. In large samples EL estimators have been shown to preserve many of the features that maximum likelihood estimators have. Also Kitamura & Stutzer (1997) showed that combining EL with moment conditions provides the most efficient moment condition testing procedure. We should also remember that the existence of GMM estimators, as well as other M-estimators, is based on a set of assumptions which are based on the topological structure of the parameter space, these assumption ore often not easy to verify.

An increasing number of properties that are known to hold for parametric likelihood have been shown to be steel valid for empirical likelihood. It allows for the construction of likelihood ratio tests that are often Bartlett correctable (DiCiccio et al., 1991; Lazar & Mykland, 1999). EL estimators provide frequentist point estimators which after bias-corrections are higher-order efficient (see Newey & Smith, 2004), emulating many of the properties

of the parametric maximum likelihood estimators. The results about EL just described are only valid for large samples. Though in small samples simulations showed that the behaviour (in terms of bias) of the EL procedure is better than the behaviour of the GMM, we still do not have any distributions for the sufficient statistics of the parameters so that reasonable standard classical inference would be impossible.

Bayesian statisticians have proposed some alternative methods which enable us to derive posterior distributions for the parameters even when the sample size is small. On the other hand, bayesian analysis without a specified likelihood function is hard to interpret, and deriving methods of inference based on emprical distributions of data alone is hard. The two main proposals are: The Bayesian Methods of Moment (BMOM) proposed by Zellner (1997) and recently the Bayesian Exponentially Tilted Empirical Likelihood (Schennach, 2005). Despite the name, BMOM is not considered a rigorous bayesian methods, it solves the moment conditions deriving the so called “post-data” moments of the parameters, and than it uses these moments as constraints to derive the posterior distributions for the parameters via maximum entropy. Bayesian Exponentially Tilted Empirical Likelihood (Schennach, 2005) tries to interpret the use of empirical likelihood together with side moment conditions to derive posterior distributions for the parameters of interests. This procedure was first conceptually proposed by Lazar (2004) while in Shennach (2005) a formal interpretation of the empirical likelihood has been given showing interesting properties. Lazar (2004) introduced connections between EL and Bayesian Bootstrap, and he shows that under the coverage probability criterion the inference derived using EL leads to valid bayesian inference. While the BMOM does not require a prior specification for the parameters of interest, the Exponentially Tilted Empirical Likelihood does. The problem is that usually formulating reasonable priors for these parameters is hard, in fact, very often these parameters are reduced form parameters so that they do not have physical or economic content, and specifying a distribution for them results to be questionable.

In this work will propose a non parametric estimator for parameters which are embodied in given moment conditions. We will analyze conditions

under which to represent the distribution of data points conditional on their relative frequencies as a multinomial distribution then we will derive the posterior distribution for the relative frequencies of the observed data points via Bayes' Law specifying a Dirichlet process objective priors obtained matching the prior modes for unknown relative frequencies with the optimal weights computed under a special form of the Empirical Likelihood estimator for the parameters of interest. Trough a Monte Carlo algorithm we derive a distribution for the parameters of interest based on the derived posterior for the relative frequencies of data points. This estimation methods shares some similarities with the bayesian bootstrap, but in Bayesian Bootstrap (see Rubin, 1981) there is no derivation of an objective prior which take into account the given moment condition.

2 The meaning of EL estimator in a nutshell

Given a sample $d_n = \{d_1, d_2, \ldots, d_n\}$ of i.i.d. observations, and a set of specified moment conditions $\mathsf{E}[g(d_i, \beta)] = 0$, where $g(d_i, \beta)$ is a vector-valued smooth and continuous function of the data and the parameter of interest $\beta \in \mathbb{R}^p$, the empirical likelihood estimator of β is defined as solving the following problem:

$$\begin{aligned} \max_{\{w_1, w_2, \ldots, w_n\}} \quad & \sum_{i=1}^{n} \log(w_i / n^{-1}), \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i g(d_i, \beta) = 0, \\ & \sum_{i=1}^{n} w_i = 1, \quad w_i \in (0, 1); \end{aligned} \tag{2.1}$$

where $w_1, w_2, \ldots, w_n$ are the mass probabilities placed on the observations in d_n. As we pointed out before this estimator has many desirable asymptotic properties that make easy to derive a chi-square asymptotic likelihood ratio test for testing hypothesis, as well as asymptotic confidence intervals. However these results are only valid in large samples analysis we

want to stress that these results are only valid for infinite samples. We are not interested in review the statistical properties of this estimator since our problem is small sample inference, instead for now, we are interested in understanding the geometrical properties of the EL estimator and why it improves GMM estimator. As we see from (2.1), EL estimator of β is the value of β which solves moment equations under probabilities for data given by $w_1^*, w_2^*, \ldots, w_n^*$, which are obtained maximizing the empirical distribution of the data under the moment constraints. In the GMM setup, the moment equation would be $n^{-1}\sum_{i=1}^{n} g(d_i, \beta) = 0$, that is the expectation of $g(d_i, \beta)$ is taken placing uniform probability masses over all data points. Obviously this is not reasonable in small samples, in fact, while in an infinite number of repetitions we expect that all the support of d_n will be drawn so that putting an equal probability mass over all the point observations will not affect the approximation of the true distribution of the data, in a small sample we cannot expect so. The improvement produced by the EL estimator consists in the fact that it reweights data according with the actual observed proportions and the given moment conditions.

Many particular versions of the empirical likelihood have been proposed. In order to make easier our interpretation of the EL procedure, we prefer to look at a particular version of (2.1). Changing the previous objective function we have the Entropy Empirical Likelihood Estimator (EELE) which is the value of β that solves

$$\begin{aligned} \max_{\{w_1, w_2, \ldots, w_n\}} \quad & \sum_{i=1}^{n} -w_i \log(w_i), \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i g(d_i, \beta) = 0, \\ & \sum_{i=1}^{n} w_i = 1, \quad w_i \in (0, 1); \end{aligned} \tag{2.2}$$

This particular form of likelihood (see Efron, 1981; Lee & Young, 1999; Corcoran, 1998; Jing & Wood, 1996) has a nice interpretation in terms of information theoretic quantities. In fact the objective function in (2.2) is the

entropy associated with probability mass function which assigns weights $\{w_1, w_2, \ldots, w_n\}$. The EELE estimator will be the value of β which solves the moment equations where moments are taken with the respect to the distribution of the data which maximizes the entropy. In the discrete case we can definitively take the entropy as a measure of uncertainty, so that with the EELE we are computing the value of β which is derived from the moment conditions assuming maximum uncertainty in the distribution of the data. Maximizing entropy, i.e. uncertainty, means also minimizing information, so that the EELE computes the value of β assuming that the random sample contents minimum information. This particular form of EL is also proposed by Schennach (2005) to propose an interpretation of empirical likelihood in bayesian methods, she also makes an interesting link between EL methods ad Maximum Entropy methods in bayesian analysis.

These interpretation makes the use of the empirical likelihood machinery compelling, and we want to use these interpretations to derive statistical methods which we can not classify as standard bayesian methods, but they will still share important properties with the bayesian approach; that is our approach is to make probabilistic judgement conditional on observables even if we will not use the Bayes' law as it is usual practice in bayesian statistics.

3 A "quasi-bayeasian" inference

Our problem is to make inference on a finite sample. The object of our analysis is the vector of parameters β for which we would like to derive a distribution in order to make probability statements in the form of hypothesis testing and confidence intervals. The bayesian way seems to be the most reasonable approach to do this. As we pointed out before, we also need a way to skip a full prior specification for the parameters as this specification would be not straightforward. By "quasi-bayesian" we mean that the inference we will propose here shares some features with standard bayesian methods, however it does not follows the usual structure of the bayesian analysis.

We can summarize our problem as follows: a researcher observes a se-

quence of n data points, with n finite; this data points are only $k \leq n$ distinct values, in practice the researcher observes a set set $d_n = \{d_1, d_2, \ldots, d_k\}$ of ordered distinct values; also he or she knows that there is an unknown vector of parameters β such that it must obey to a given set of moment conditions $\mathsf{E}_F[g(d, \beta)] = 0$, where the expectation is taken with respect to the unknown probability distribution of the data F. A classical bayesian procedure would be to specify a parametric likelihood $P(d_n|\beta)$, a prior $\pi(\beta)$ for β, and than deriving the posterior probability $P(\beta|d_n) \propto P(d_n|\beta)\pi(\beta)$ via Bayes' law. However we can not think to any reasonable statistical probability model for the data conditional on the the unknown parameter, so that the specification of $P(d_n|\beta)$ is not feasible. Also this methods would not take into account the moment condition $\mathsf{E}[g(d, \beta)] = 0$ which represents the most relevant bit of prior information we dispose. Notice that since we want to treat the unknowns as random variables, in our approach we look at β as an having a probability distribution. Resuming, we have two sources of uncertainty: we do not know the probability distribution for the data, and we do not know the probability distribution for the parameters of interest; on the other hand we have two bits of information: we observed k distinct points data drawn from an unknown F, and we know that the unknown parameter should obey to the moment equation $\mathsf{E}_F[g(d, \beta)] = 0$. Also the true distribution of the data is unknown. But as new data come, we revise our belief about the true distribution of the data, and then looking at the moment conditions $\mathsf{E}[g(d, \beta)] = 0$ we revise our beliefs about β. Our idea is to derive a reasonable posterior distribution for the distribution of data d_n which fully takes into account the main bit of information we have in term of moment condition, and than to use this distribution to draw samples from it, at each drawn we solve the moment condition under the derived posterior distribution for the data; this solution will give us a set of β which solve the moment conditions at each drawn, and then we use Monte Carlo integration to compute summary statistics and quantiles for β. Where EL is? We will use EL to derive prior distribution for the data. This procedure will be understood following next section.

4 Representing the marginal data density

First of all, based on the data collected we want to derive a posterior for the distribution of the data which takes into account the given moment condition. We need some hypothesis to represent the marginal distribution of the data. One way to ho is simply to assume that the observed data points fall into k distinct categories, so that we can write the distribution of these data points as multinomial distribution conditional on the multinomial probabilities, this a common approach in nonparametric bayesian analysis. Another way to go is to discuss more general conditions under which this representation holds possibly for more general experimental design. A general concept useful to generalize the idea of multinomial representation is the the idea of exchangeability.

Let us consider a $k-$dimensional random vector x_i whose jth component $x_{i,j}$ takes value 1 if it belongs to the jth category of the $k+1$ possible categories. At most one of the component of x_i can take value 1, and if all take value 0 this means that the $k+1$th takes value 1. So x_i can be regarded as a 0-1 random vector. If we assume that the sequence $x_1, x_2, \ldots$ is an infinite exchangeable sequence with a probability measure P, than the extension of the De Finetti's representation theorem (see De Finetti, 1930a; 1930b;) states that there exists a distribution function Q such that the joint mass function $p(x_1, x_2, \ldots, x_n)$ has the form

$$p(x_1, x_2, \ldots, x_n) = \int_W \prod_{i=1}^{n} w_1^{x_{i,1}} \cdot w_2^{x_{i,2}} \cdots w_k^{x_{i,k}} \left(1 - \sum_{j=1}^{k} w_j\right)^{1-\sum x_{i,j}} dQ(w), \tag{4.1}$$

where,

$$W := \{w = (w_1, w_2, \ldots w_k)\,; \quad w_i \in [0,1]\}\,,$$
$$Q(w) = \lim_{n\to\infty} P[(\bar{x}_{1,n} \leqslant w_1) \cup \ldots \cup (\bar{x}_{k,n} \leqslant w_k)],$$

and,

$$\bar{x}_{i,n} = \frac{1}{n}\sum_{i=1}^{n} x_{i,j}, \quad w_i = \lim_{n\to\infty} \bar{x}_{i,n}.$$

We are often more interested in the representation of the distribution of the k-dimensional random variable $d_n = \sum_{i=1}^{n} x_i$ whose jth component is total number of occurrences of category j in n observations, from the previous result it follows that $p(d_{n,1}, d_{n,2,} \ldots, d_{n,k})$ can be written as

$$\int_W \left(\begin{array}{c} n \\ d_{n,1} \cdots d_{n,k} \end{array} \right) w_1^{d_{n,1}} \cdot w_2^{d_{n,2}} \cdots w_k^{d_{n,k}} \left(1 - \sum_{j=1}^{k} w_j \right)^{n - \sum d_{n,k}} dQ(w), \tag{4.2}$$

where $Q(w)$ is defined as above. This is the same as saying that the joint sampling distribution for a random samples $\{x_1, x_2, \ldots, x_n\}$ can be represented as has a multinomial distribution with probability mass function Multinomial$_k(x_i|w, 1)$ together with a prior distribution Q over the multinomial parameter vector w, where $Q(w)$ is the long run proportion of occurrences of each category, i.e w_j is the long run limiting relative frequency of the membership to the the jth category. On the other hand this leads to represent observed point data $\{d_1, d_2, \ldots, d_k\}$ as arising from a multinomial distribution with probability mass function given by a Multinomial$_k(d_n|w, n)$, with w distributed according to $Q(w)$ as long run distribution of relative frequencies.

In non parametric bayesian statistics such a representation is often used for non parametric density estimation. We can think the set of the n observed data as belonging to k distinct categories identified as the k possible values $\{d_1, d_2, \ldots, d_k\}$, and we assume that the sequence observed is the result of an exchangeable sequence. The problem is that we have finite sample so in principle we are not allow to assume infinite exchangeability so that the representation above is impossible since in this case we could not approximate $Q(w)$ with the long run proportions. However, in general if we have a finite sequence $(y_1, y_2, \ldots, y_n) \in S \subseteq Y^n$, where this sequence can be seen as a subpart of o finite but larger exchangeable sequence of observables it is possible to show (see Diaconis & Freedman 1980; Jaynes 1986) that it is still possible to represent the probability $p(y_1, y_2, \ldots, y_n)$ as

$$P_Q(y_1, y_2, \ldots, y_n) = \int F^n(E) dQ(F), \tag{4.3}$$

where F is the long run limiting distribution of the empirical distribution function F^n, notice that this notation do not imply multinomial distribution for F^n, here we are referring to the generalized representation theorem (see Chow and Teicher; 1978) of which the multinomial representation above is a special case. Under what assumption can we have the representation above? If $P(S_N)$ would be the corresponding probability assigned to an extended sequence of N observations, with $N > n$,

$$\sup_{S_N} \|P(S_N) - P_Q(y_1, y_2, \ldots, y_n)\|,$$

would be the deviation we face by assuming infinite exchangeability when we only observe a finite sequence, comparing the finite sequence with the larger one; for larger N this measure is more appropriate. Here the supremum is taken with the respect of all possibile events of in the σ-field on the power set Y^n. The representation (4.3) holds when there exists Q and N such that

$$\sup_{S_N} \|P(S_N) - P_Q(y_1, y_2, \ldots, y_n)\| \leq \frac{nt(n)}{N},$$

where N is the length of the extended sequences and $t(n) = \#\{Y^n\}$ if Y^n is finite and $t(n) = (n-1)$ otherwise.

Let us now come back to the possibility of assuming the multinomial representation as in (4.2), since our sample is finite we can only assume finite exchangeability, and to get the multinomial representation as in (4.2), we need to assume that $\{x_1, x_2, \ldots, x_n\}$ is a finite sequence of exchangeable 0-1 random vectors, also for $\{d_1, d_2, \ldots, d_k\} \subseteq \Delta^N \subseteq d_n^n$, where $N > n$, d_n^n is the set of all possible samples with values in d_n, and $\Delta^N \subseteq d_n^n$ is an extended sample with size larger than n,

$$\sup_{\Delta_N} \|P(\Delta_N) - P_Q(d_1, d_2, \ldots, d_k)\| \leq \frac{nt(n)}{N}. \tag{4.4}$$

Even if the mathematical machinery of such assumption seems to bee too much to be assumed, coming back to the example at the beginning of this writing everything will be more clear. In that example we observed data (y_i, x_i), $i = 1, 2, \ldots, n$. Let suppose that in our finite sample we observe k

distinct pairs we label $\{d_1, d_2, \ldots, d_k\}$, the model implied $y_i = g(x_i, \gamma) + \varepsilon_i$ with $\mathsf{E}[\epsilon_i] = 0 = \mathsf{E}[y_i - \gamma x_i]$ for each i. Now suppose we observe an infinite sequence of data points, the realization of this sequence is determined by two unknown components: γ and the i.i.d sequence $\{\varepsilon\}_{i=1}^{n}$ which we do not observe, in this situation exchangeability about the observables is not too much to assume as it means complete ignorance about the data generating process which we do not know. The problem is still that the representation as in (4.2) holds if the condition (4.4) holds. Roughly speaking, from the practical point of view condition (4.4) means that given our finite sample, and given our prior $Q(w)$ for the small sample, for any larger sample which takes values into d_n^n we can always bound the distance between the probability of the sample derived under the prior $Q(w)$ and the probability of the larger sample. That is the same as saying that we choose a prior such that the derived probability of the sample computed on small number of observations is such that its distance from the probability of the sample computed on larger sample is bounded, and this bound depend on the size of the sample and the size of the extended sample, for some given size n, this bound become tighter as the size of the extended sample increases.

We believe that such assumptions are not too much demanding if compared to the assumptions usually made in alternative settings. So under the assumptions provided above, we will represent the distribution of the data collected as in (4.2) specifying a Dirichlet conjugate prior density for $Q(w)$. We assume that conditions of representation theorem under finite exchangeability hold, i.e. our Dirichlet prior approximates the probability of the long run proportions of data points belonging to the k distinct categories. The choice of the Dirichlet prior process here is shared with the Bayesian Bootstrap, however our choice is made because it will allow for a construction of a prior which has a nice interpretation.

5 Multinomial representation

We assume that k-distinct data points arise from a multinomial distribution with probability mass function given by Multinomial$_k(d_n|w,n)$, with w distributed according to $Q(w)$ which is our prior distribution for relative frequencies. We will derive a Dirichlet prior for w, i.e. we will assume that the probability mass function of the relative frequencies of the k distinct values of the data is Dirichlet$_k(w|\alpha)$, where $w = (w_1, w_2, \ldots, w_k)$ and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{k+1})$ is the Dirichlet parameter vector, that is

$$p(d_n|w) = \frac{n!}{\prod_{j=1}^{k} d_j! \left(n - \sum_{j=1}^{k} d_j\right)!} \prod_{j=1}^{k} w_j^{d_j} \left(1 - \sum_{j=1}^{k} w_j\right)^{n - \sum_{j=1}^{k} d_j}, \tag{5.1}$$

where n and k are given, and

$$p(w) = \frac{\Gamma(\sum_{j=1}^{k+1} \alpha_j)}{\prod_{j=1}^{k+1} \Gamma(\alpha_j)} \prod_{j=1}^{k} w_j^{\alpha_j - 1} \left(1 - \sum_{j=1}^{k} w_j\right)^{\alpha_{k+1} - 1}, \tag{5.2}$$

where the vector α is then specified. Our posterior distribution is computed via Bayes' law as $p(w|d_n,) \propto p(d_n|w)p(w)$, and it is easy to verify that $p(w|d_n) =$ Dirichlet$_k(w|\alpha')$ where

$$\alpha' = \left(d_1 + \alpha_1, d_2 + \alpha_2, \ldots, n - \sum_{j=1}^{k} d_j + \alpha_{k+1}\right).$$

6 A minimum discrepancy objective prior

How to specify a prior distribution for w? It should be specified so that it will fully reflect the information we have, that is observed data and moment conditions. Here, we propose a method that uses the main idea of the

empirical likelihood. Let us to reconsider (2.2), and we modify it slightly as

$$\begin{aligned} \max_{\{w_1,w_2,\ldots,w_n\}} \quad & \sum_{i=1}^{n} -w_i \log \frac{w_i}{1/n}, \\ \text{s.t.} \quad & \sum_{i=1}^{n} w_i g(d_i, \beta) = 0, \\ & \sum_{i=1}^{n} w_i = 1, \quad w_i \in (0,1). \end{aligned} \tag{6.1}$$

The objective function of the problem above is the Kullback-Liebler distance function

$$KL\left[\left\{n^{-1}, n^{-1}, \ldots, n^{-1}\right\} || \left\{w_1, w_2, \ldots, w_n\right\}\right]; \tag{6.2}$$

taken with a minus sign. Despite its name, the Kullback-Liebler distance is not a proper distance since it does not satisfy the triangular inequality, but it can be interpreted as a measure of the discrepancy between two random variables. From a geometrical point of view, (6.2) is a measure of the distance between the probability mass function which assigns probabilities $\{w_1, w_2, \ldots, w_n\}$, and the uniform probability mass function which assigns probabilities $\left\{n^{-1}, n^{-1}, \ldots, n^{-1}\right\}$. In problem (6.1) such a discrepancy is minimized, so that our estimator for β is the value of β which solves the moment condition taken with respect to a distribution of the data which is the nearest to the uniform one. Under ignorance about the true distribution of the data, we should believe that the data come from a uniform distribution, on the other hand if we observe ties in the data this would be not reasonable so that we look for the distribution which is the nearest possibile to the uniform one. In terms of information-theoretic quantities, (6.2) is the expected regret under the logarithm score loss function associated with using the probability mass function which assigns uniform probabilities when instead the data arise from a probability mass function which assigns probabilities $\{w_1, w_2, \ldots, w_n\}$. Again, under ignorance about the true distribution of the data, we should believe that the data come from a uniform distribution, but if this is not the case we would face a regret measured by (6.2), so that we look for all those values of β which solve moment equations

under that distribution which minimizes our regret.

Based on this, let $\hat{w} = (\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_k)$ be the optimal estimated relative frequencies which solve (6.1). Then we construct the Dirichlet prior $p(w)$ imposing that Mode$[w_j] = \hat{w}_j$. This specification has two nice interpretations. That is, our prior knowledge based on data and moment conditions imposes that the highest probability for each of the distinct data points is the value which minimizes the Kullback-Liebler distance between the distribution of the data and the uniform distribution. Or alternatively, we start being strongly uncertain about the distribution of the data thinking that they are uniformly distributed, and then, after seeing ties in the data, we seek the distribution which — according to the moment conditions — minimizes the regret we would face if the true distribution was not uniform one. The compelling features of such a machinery is that the prior specification fully reflects our prior moment conditions, since optimal relative frequencies are computed taking into account the moments as constraints. Under the Dirichelet prior we impose,

$$\text{Mode}[w_j] = \hat{w}_j = \frac{\alpha_j - 1}{\sum_{j=1}^{k+1} \alpha_j - k - 1}, \qquad \text{for} \quad j = 1, 2, \ldots, k; \quad (6.3)$$

these are k equations but we have $k+1$ unknowns, fixing a prior value for α_{k+1} is always problematic. Let us rewrite the kernel of the Dirichlet prior, which is

$$p(w|\alpha) \propto \prod_{j=1}^{k} w_j^{\alpha_j - 1} \left(1 - \sum_{j=1}^{k} w_j\right)^{\alpha_{k+1} - 1}, \quad (6.4)$$

Observe that $\alpha_{k+1} - 1$ is the prior weight we attach to $1 - \sum_{j=1}^{k} w_j$, in our problem this is the probability that none of the k distinct values of the sample occurs, since in (6.1) we imposed that $\sum_{j=1}^{k} w_j = 1$, our prior weight for this occurrence should be zero, that is we fix $\alpha_{k+1} = 1$, then it is to see that a solution of the k equations in (6.4) is takeing $\alpha_j = \hat{w}_j + 1$ for $j = 1, 2, \ldots, k$. Our final prior for w is $p(w) = \text{Dirichlet}_k(w|\hat{\alpha})$ where $\hat{\alpha} = (\hat{w}_1 + 1, \hat{w}_2 + 1, \ldots, \hat{w}_k + 1, 1)$. Under the specified prior its kernel

becomes

$$p(w) \propto \prod_{j=1}^{k} w_j^{\hat{w}_j}, \tag{6.5}$$

the latter has a nice interpretation since it is the same as saying that we weight each probability w_j with its minimum discrepancy value as derived in (6.1). The distribution on the data is then $p(w|d_n) = \text{Dirichlet}_k(w|\hat{\alpha'})$ where

$$\hat{\alpha'} = (d_1 + \hat{w}_1 + 1, d_2 + \hat{w}_2 + 1, \ldots, \left(n - \sum_{j=1}^{k} d_j\right) + 1).$$

Also the the posterior $p(w|d_n)$ has a nice interpretation, in fact its kernel is

$$p(w|d_n) \propto \prod_{j=1}^{k} w_j^{d_j + \hat{w}_j} \left(1 - \sum_{j=1}^{k} w_j\right)^{n - \sum_{j=1}^{k} d_j}, \tag{6.6}$$

that is the posterior probability, of each w_j conditional on observed data, is weighted considering our prior weights $\hat{w}_j$ we computed with (6.1) according to moment equations, and the actual number of occurrences in each of the k distinct categories. Notice that under our assumptions the last component of the kernel of the posterior will be one in computation. Notice that we now have a continuum of posterior possible values for each marginal proportion.

7 A post-data density for parameters

As well as we now have a posterior distribution for the proportions of observed data, i.e. we have a sort of "posterior histogram", derived under some objective criterion, we still need a distribution for the parameter of interest β. Remember that our goal is to make inference on the parameter of interest, i.e. we want to make probabilistic statements as well as computing confidence sets and other summary statistics.

However, we do not have any knowledge of the distribution of β, and we are not willing to assign to it any prior distribution. In fact, as already pointed out, in practical social sciences empirical analysis problems the parameters do not have any physical meaning so that it would be extremely questionable to impose prior distribution on them, on the other hand we want to get a distribution for the parameters of interest that is conditional on observables. The derivation of such a probability distribution $p(\beta|d_n)$ would be impossible in principle. Here we propose abootstrap algorithm to derive the distribution of β which does not make any use of the Bayes' Law. We used the wording "post-data" to mean that derived distribution make is based on data, however it can not be interpreted as in the standard bayesian analysis.

The only bit of information about the parameter β is tat it should obey to the moment condition $\mathsf{E}[g(d_n, \beta)]$ where this expectation is taken with the respect of the distribution of the data. In the previous section we run a non parametric analysis deriving the posterior distribution of the data distribution, in fact, as we said before we derived a poster histogram for the observed data which we based on an objective prior which agreed with the given moment conditions. Here we make use of this posterior knowledge together with moments to derive post-data distribution for the unknown parameter.

The Monte Carlo integration algorithm is as follows:

1. we use the posterior $p(w|d_n)$ to draw M i.i.d. samples for w, where M is some large integer. At each iteration $s = 1, 2, \ldots, M$ we store a vector of k posterior weights $w^{(s)} = (w_1^{(s)}, w_2^{(s)}, \ldots, w_k^{(s)})$ where w_k^s is the posterior proportion for the kth data cell at the sth drawn.
2. At each iteration we compute $\beta^{(s)}$ solving the moment conditions for β, i.e. $\beta^{(s)}$ is such that $\sum_{j=1}^{k} g(d_j, \beta^{(s)}) = 0$.
3. We now have an M-dimensional vector $\beta^{(M)} = (\beta^{(1)}, \beta^{(2)}, \ldots, \beta^{(M)})$; then we use Monte Carlo integration to compute confidence intervals, quantiles and so on.

Even if the model proposed implies a huge amount of computations, the increased computational capability makes it viable to be run. In the next

section we propose a simple toy and extreme example showing that in small samples the model proposed exhibits better behaviour if compared with non parametric classical inference methods such as GMM or EL methods.

8 A toy example and conclusions

In this example we propose an extreme situation and we compare our model with GMM and EL methods. In the future we will also explore comparisons with Bayesian Bootstrap. Suppose we observe a sample of $\{y_i, x_i\}_{i=1}^{14}$ and that based on some theoretical result these observations should obey to the following model $y_i = \gamma x_i + \varepsilon_i$, where $\{\varepsilon_i\}_{i=1}^{14}$ is a sequence of unobservable i.i.d. shocks for which our theory states that E$[\varepsilon_i] = 0$ for each i. This model implicitly assumes the following moment condition E$[y_i - \gamma x_i] = 0$ for each i.

In classical inference setting we would look at the parameter γ as fixed and, assuming random sampling we derive the asymptotic distribution for a sufficient statistic for γ or the asymptotic distribution for some test statistics. By using GMM or EL methods this can be easily done. However in this case we only have fourteen points observations, so that any asymptotic result for GMM or EL methods is not valid and the derived inference would be doubtful. There is a researcher who refuses to look at γ as an unknown random quantity, and as result she or he has two alternatives: he or she can choose to not make inference on the parameter of interest, or she or he will use asymptotic results to make inference knowing in principle that the conclusions he or she will derive are subject to possibly a tremendous bias.

On the other hand, γ is unknown and unobservable as the sequence $\{\varepsilon_i\}_{i=1}^{14}$ does. Another researcher starting from this consideration looks at γ as a random variable, in principle he or she cannot say anything about γ, but he or she wants to learn about it looking at the data. So she or he decides to derive its post-data distribution as described in the previous section.

We generated a $\{x_i\}_{i=1}^{14}$ of i.d. drawn from a discrete uniform distrib-

ution with support $[1, 14]$, then we generated the sequence $\{\varepsilon_i\}_{i=1}^{14}$ of i.d. from a continuous uniform distribution with support $[-5, 5]$ which has zero expectation as our theory states. Finally we fixed $\gamma = 2$ and we generated the sequence $\{y_i\}_{i=1}^{14}$ where $y_i = 2x_i + \varepsilon_i$.

The first researcher using GMM would estimate $\hat{\gamma} = 1.83474$ and the 95%-confidence interval derived for $\hat{\gamma}$ would be $[1.637645, 2.031838]$, also he or she knows that this confidence interval is derived under the asymptotic normality of $\hat{\gamma}$. He or she also tries an EL estimation, and the estimates of γ is now better than before as it results to be $\hat{\gamma} = 1.88273$ with the 95%-confidence being $[1.843363, 2.029838]$. The second researcher starts running the algorithm proposed in the previous section and using Monte Carlo integration over 100 i.d. drawn from the posterior density of proportions for the data he or she gets summary statistics for the post-data density of γ. She or he computes Monte Carlo mean for γ which is 1.89702, and the median is 1.915013 and the 95%-confidence interval is $[1.879363, 2.021638]$.

Even though this is a simple example, it shows that the method proposed in the previous section offers better performances if compared with other well known nonparametric non-bayesian procedures. Future analysis should be devoted to assess the behaviour of this estimators in more complicated estimating problems, also suitable computational methods have to be studied to make computation viable in more complicated real empirical problems. A fundamental issue is to compare our result with those of the Bayesian Bootstrap.

References

CHOW, Y. S., AND H. TEICHER (1978): *Probability Theory*. Springer-Verlag, New York.

CORCORAN, S. A. (1998): "Bartlett adjustment of empirical discrepancy statistics," *Biometrika*, 85, 967–972.

DE FINETTI, B. (1930a): "Funzione Caratteristica di un Fenomeno Aleatorio," *Mem. Accad. Naz. Lincei*, 4, 86–133.

——— (1930b): "La prévision : ses lois logiques, ses sources subjectives,"

Annales de l'institut Henri Poincaré, 7, 1–68.

DIACONIS, P., AND D. FREEDMAN (1980): "Finite Exchangeable Sequences," *Annals of Probability*, 8, 745–764.

DICICCIO, T. J., P. HALL, AND J. P. ROMANO (1991): "Empirical likelihood is Bartlett correctable," *Annals of Statististic*, 19, 1053–1061.

EFRON, B. (1981): "Nonparametric standard errors and confidence intervals," *Can. J. Statist.*, 9, 139–172.

HANSEN, L. P. (1982): "Large sample properties of generalized method of moment estimators," *Econometrica*, 50, 1029–1054.

HANSEN, L. P., J. HEATON, AND A. YARON (1996): "Finite sample properties of some alternative gmm estimators," *J. Bus. Econ. Statist.*, 14, 262280.

IMBENS, G. W. (1997): "One-step estimators for over-identified generalized method of moments models," *Rev. Econ. Studies*, 64, 359–383.

IMBENS, G. W., R. H. SPADY, AND P. JOHNSON (1998): "Information theoretic approaches to inference in moment condition models," *Econometrica*, 66, 333–357.

JAYNES, E. T. (1986): *Some Applications and Extensions of the de Finetti Representation Theorem*, vol. Bayesian Inference and Decision Techniques. Elsevier, Amsterdam.

JING, B. Y., AND A. T. A. WOOD (1996): "Exponential empirical likelihood is not Bartlett correctable," *Annals of Statistics*, 24, 365–369.

KITAMURA, Y., AND M. STUTZER (1997): "An information-theoretic alternative to generalized method of moment estimation," *Econometrica*, 65, 861–874.

LAZAR, N. (2003): "Bayesian Empirical Likelihood," *Biometrika*, 90.

LAZAR, N. A., AND P. MYKLAND (1999): "Empirical likelihood in the presence of nuisance parameters," *Biometrika*, 86, 203211.

LEE, S. M. S., AND G. A. YOUNG (1999): "Nonparametric likelihood ratio confidence intervals," *Biometrika*, 86, 107–118.

NEWEY, W., AND R. J. SMITH (2004): "Higher-order properties of GMM and generalized empirical likelihood estimators," *Econometrica*, 72, 219255.

OWEN, A. B. (1988): "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, 75, 237–249.

OWEN, A. B. (2001): *Empirical Likelihood*. Chapman&Hall/CRC, New York.

QIN, J., AND J. LAWLESS (1994): "Empirical likelihood and general estimating equations," *Annals of Statistics*, 22, 300–325.

RUBIN, D. B. (1981): "The Bayesian bootstrap," *Annals of Statistics*, 9, 130134.

SCHENACH, S. M. (2005): "Bayesian Exponentially Tilted Empirical Likelihood," forthcoming on Biometrika.

SMITH, R. J. (1997): "Alternative Semi-Parametric Likelihood Approaches to Generalized Method of Moments Estimation," *Economic Journal*, 107, 503–519.

ZELLNER, A. (1997): "The Bayesian Method of Moments (BMOM): Theory and Applications," *Advances in Econometrics*, 12, 85–105.

Dipartimento di Scienze Economiche e Statistiche
Università degli Studi di Salerno

*W*ORKING *P*APERS DEL *D*IPARTIMENTO

1988, 3.1 Guido CELLA
Linkages e moltiplicatori input-output.

1989, 3.2 Marco MUSELLA
La moneta nei modelli di inflazione da conflitto.

1989, 3.3 Floro E. CAROLEO
Le cause economiche nei differenziali regionali del tasso di disoccupazione.

1989, 3.4 Luigi ACCARINO
Attualità delle illusioni finanziarie nella moderna società.

1989, 3.5 Sergio CESARATTO
La misurazione delle risorse e dei risultati delle attività innovative: una valutazione dei risultati dell'indagine CNR- ISTAT sull'innovazione tecnologica.

1990, 3.6 Luigi ESPOSITO - Pasquale PERSICO
Sviluppo tecnologico ed occupazionale: il caso Italia negli anni '80.

1990, 3.7 Guido CELLA
Matrici di contabilità sociale ed analisi ambientale.

1990, 3.8 Guido CELLA
Linkages e input-output: una nota su alcune recenti critiche.

1990, 3.9 Concetto Paolo VINCI
I modelli econometrici sul mercato del lavoro in Italia.

1990, 3.10 Concetto Paolo VINCI
Il dibattito sul tasso di partecipazione in Italia: una rivisitazione a 20 anni di distanza.

1990, 3.11 Giuseppina AUTIERO
Limiti della coerenza interna ai modelli con la R.E.H..

1990, 3.12 Gaetano Fausto ESPOSITO
Evoluzione nei distretti industriali e domanda di istituzione.

1990, 3.13 Guido CELLA
Measuring spatial linkages: input-output and shadow prices.

1990, 3.14 Emanuele SALSANO
Seminari di economia.

1990, 3.15 Emanuele SALSANO
Investimenti, valore aggiunto e occupazione in Italia in contesto biregionale: una prima analisi dei dati 1970/1982.

1990, 3.16 Alessandro PETRETTO- Giuseppe PISAURO
Uniformità vs selettività nella teoria della ottima tassazione e dei sistemi tributari ottimali.

1990, 3.17 Adalgiso AMENDOLA
Inflazione, disoccupazione e aspettative. Aspetti teorici dell'introduzione di aspettative endogene nel dibattito sulla curva di Phillips.

1990, 3.18 Pasquale PERSICO
Il Mezzogiorno e le politiche di sviluppo industriale.

1990, 3.19 Pasquale PERSICO
Priorità delle politiche strutturali e strategie di intervento.

1990, 3.20 Adriana BARONE - Concetto Paolo VINCI
La produttività nella curva di Phillips.

1990, 3.21 Emiddio GALLO
Varianze ed invarianze socio-spaziali nella transizione demografica dell'Italia post-industriale.

1991, 3.22 Alfonso GAMBARDELLA
I gruppi etnici in Nicaragua. Autonomia politica ed economica.

1991, 3.23 Maria SCATTAGLIA
La stima empirica dell'offerta di lavoro in Italia: una rassegna.

1991, 3.24 Giuseppe CELI
La teoria delle aree valutarie: una rassegna.

1991, 3.25 Paola ADINOLFI
Relazioni industriali e gestione delle risorse umane nelle imprese italiane.

1991, 3.26 Antonio e Bruno PELOSI
Sviluppo locale ed occupazione giovanile: nuovi bisogni formativi.

1991, 3.27 Giuseppe MARIGLIANO
La formazione del prezzo nel settore dell'intermediazione commerciale.

1991, 3.28 Maria PROTO
Risorse naturali, merci e ambiente: il caso dello zolfo.

1991, 3.29 Salvatore GIORDANO
Ricerca sullo stato dei servizi nelle industrie del salernitano.

1992, 3.30 Antonio LOPES
Crisi debitoria e politiche macroeconomiche nei paesi in via di sviluppo negli anni 80.

1992, 3.31 Antonio VASSILLO
Circuiti economici semplici, complessi, ed integrati.

1992, 3.32 Gaetano Fausto ESPOSITO
Imprese ed istituzioni nel Mezzogiorno: spunti analitici e modalità di relazione.

1992, 3.33 Paolo COCCORESE
Un modello per l'analisi del sistema pensionistico.

1994, 3.34 Aurelio IORI
Il comparto dei succhi di agrumi: un caso di analisi interorganizzativa.

1994, 3.35 Nicola POSTIGLIONE
Analisi multicriterio e scelte pubbliche.

1994, 3.36 Adriana BARONE
Cooperazione nel dilemma del prigioniero ripetuto e disoccupazione involontaria.

1994, 3.37 Adriana BARONE
Le istituzioni come regolarità di comportamento.

1994, 3.38 Maria Giuseppina LUCIA
Lo sfruttamento degli idrocarburi offshore tra sviluppo economico e tutela dell'ambiente.

1994, 3.39 Giuseppina AUTIERO
Un'analisi di alcuni dei limiti strutturali alle politiche di stabilizzazione nei LCDs.

1994, 3.40 Bruna BRUNO
Modelli di contrattazione salariale e ruolo del sindacato.

1994, 3.41 Giuseppe CELI
Cambi reali e commercio estero: una riflessione sulle recenti interpretazioni teoriche.

1995, 3.42 Alessandra AMENDOLA, M. Simona ANDREANO
The TAR models: an application on italian financial time series.

1995, 3.43 Leopoldo VARRIALE
Ambiente e turismo: Parco dell'Iguazù - Argentina.

1995, 3.44 A. PELOSI, R. LOMBARDI
Fondi pensione: equilibrio economico-finanziario delle imprese.

1995, 3.45 Emanuele SALSANO, Domenico IANNONE
Economia e struttura produttiva nel salernitano dal secondo dopoguerra ad oggi.

1995, 3.46 Michele LA ROCCA
Empirical likelihood and linear combinations of functions of order statistics.

1995, 3.47 Michele LA ROCCA
L'uso del bootstrap nella verosimiglianza empirica.

1996, 3.48 Domenico RANESI
Le politiche CEE per lo sviluppo dei sistemi locali: esame delle diverse tipologie di intervento e tentativo di specificazione tassonomica.

1996, 3.49 Michele LA ROCCA
L'uso della verosimiglianza empirica per il confronto di due parametri di posizione.

1996, 3.50 Massimo SPAGNOLO
La domanda dei prodotti della pesca in Italia.

1996, 3.51 Cesare IMBRIANI, Filippo REGANATI
Macroeconomic stability and economic integration. The case of Italy.

1996, 3.52 Annarita GERMANI
Gli effetti della mobilizzazione della riserva obbligatoria. Analisi sull'efficienza del suo utilizzo.

1996, 3.53 Massimo SPAGNOLO
A model of fish price formation in the north sea and the Mediterranean.

1996, 3.54 Fernanda MAZZOTTA
RTFL: problemi e soluzioni per i dati Panel.

1996, 3.55 Angela SPAGNUOLO
Concentrazione industriale e dimensione del mercato: il ruolo della spesa per pubblicità e R&D.

1996, 3.56 Giuseppina AUTIERO
The economic case for social norms.

1996, 3.57 Francesco GIORDANO
Sulla convergenza degli stimatori Kernel.

1996, 3.58 Tullio JAPPELLI, Marco PAGANO
The determinants of saving: lessons from Italy.

1997, 3.59 Tullio JAPPELLI
The age-wealth profile and the life-cycle hypothesis: a cohort analysis with a time series of cross sections of Italian households.

1997, 3.60 Marco Antonio MONACO
La gestione dei servizi di pubblico interesse.

1997, 3.61 Marcella ANZOLIN
L'albero della qualità dei servizi pubblici locali in Italia: metodologie e risultati conseguiti.

1997, 3.62 Cesare IMBRIANI, Antonio LOPES
Intermediazione finanziaria e sistema produttivo in un'area dualistica. Uno studio di caso.

1997, 3.63 Tullio JAPPELLI
Risparmio e liberalizzazione finanziaria nell'Unione europea.

1997, 3.64 Alessandra AMENDOLA
Analisi dei dati di sopravvivenza.

1997, 3.65 Francesco GIORDANO, Cira PERNA
Gli stimatori Kernel per la stima non parametrica della funzione di regressione.

1997, 3.66 Biagio DI SALVIA
Le relazioni marittimo-commerciali nell'imperiale regio litorale austriaco nella prima metà dell'800.
I. Una riclassificazione delle Tafeln zur Statistik der Öesterreichischen Monarchie.

1997, 3.67 Alessandra AMENDOLA
Modelli non lineari di seconda e terza generazione: aspetti teorici ed evidenze empiriche.

1998, 3.68 Vania SENA
L'analisi econometrica dell'efficienza tecnica. Un'applicazione agli ospedali italiani di zona.

1998, 3.69 Domenico CERBONE
Investimenti irreversibili.

1998, 3.70 Antonio GAROFALO
La riduzione dell'orario di lavoro è una soluzione al problema disoccupazione: un tentativo di analisi empirica.

1998, 3.71 Jacqueline MORGAN, Roberto RAUCCI
New convergence results for Nash equilibria.

1998, 3.72 Rosa FERRENTINO
Niels Henrik Abel e le equazioni algebriche.

1998, 3.73 Marco MICOCCI, Rosa FERRENTINO
Un approccio markoviano al problema della valutazione delle opzioni.

1998, 3.74 Rosa FERRENTINO, Ciro CALABRESE
Rango di una matrice di dimensione K.

1999, 3.75 Patrizia RIGANTI
L'uso della valutazione contingente per la gestione del patrimonio culturale: limiti e potenzialità.

1999, 3.76 Annamaria NESE
Il problema dell'inefficienza nel settore dei musei: tecniche di valutazione.

1999, 3.77 Gianluigi COPPOLA
Disoccupazione e mercato del lavoro: un'analisi su dati provinciali.

1999, 3.78 Alessandra AMENDOLA
Un modello soglia con eteroschedasticità condizionata per tassi di cambio.

1999, 3.79 Rosa FERRENTINO
Su un'applicazione della trasformata di Laplace al calcolo della funzione asintotica di non rovina.

1999, 3.80 Rosa FERRENTINO
Un'applicazione della trasformata di Laplace nel caso di una distribuzione di Erlang.

1999, 3.81 Angela SPAGNUOLO
Efficienza e struttura degli incentivi nell'azienda pubblica: il caso dell'industria sanitaria.

1999, 3.82 Antonio GAROFALO, Cesare IMBRIANI, Concetto Paolo VINCI
Youth unemployment: an insider-outsider dynamic approach.

1999, 3.83 Rosa FERRENTINO
Un modello per la determinazione del tasso di riequilibrio in un progetto di fusione tra banche.

1999, 3.84 DE STEFANIS, PORZIO
Assessing models in frontier analysis through dynamic graphics.

1999, 3.85 Annunziato GESUALDI
Inflazione e analisi delle politiche fiscali nell'U.E..

1999, 3.86 R. RAUCCI, L. TADDEO
Dalle equazioni differenziali alle funzioni e^x, $\log x$, a^x, $\log_a x$, x^α.

1999, 3.87 Rosa FERRENTINO
Sulla determinazione di numeri aleatori generati da equazioni algebriche.

1999, 3.88 C. PALMISANI, R. RAUCCI
Sulle funzioni circolari: una presentazione non classica.

2000, 3.89 Giuseppe STORTI, Pierluigi FURCOLO, Paolo VILLANI
A dynamic generalized linear model for precipitation forecasting.

2000, 3.90 Rosa FERRENTINO
Un procedimento risolutivo per l'equazione di Dickson.

2000, 3.91 Rosa FERRENTINO
Un'applicazione della mistura di esponenziali alla teoria del rischio.

2000, 3.92 Francesco GIORDANO, Michele LA ROCCA, Cira PERNA
Bootstrap variance estimates for neural networks regression models.

2000, 3.93 Alessandra AMENDOLA, Giuseppe STORTI
A non-linear time series approach to modelling asymmetry in stock market indexes.

2000, 3.94 Rosa FERRENTINO
Sopra un'osservazione di De Vylder.

2000, 3.95 Massimo SALZANO
Reti neurali ed efficacia dell'intervento pubblico: previsioni dell'inquinamento da traffico nell'area di Villa S. Giovanni.

2000, 3.96 Angela SPAGNUOLO
Concorrenza e deregolamentazione nel mercato del trasporto aereo in Italia.

2000, 3.97 Roberto RAUCCI, Luigi TADDEO
Teoremi ingannevoli.

2000, 3.98 Francesco GIORDANO
Una procedura per l'inizializzazione dei pesi delle reti neurali per l'analisi del trend.

2001, 3.99 Angela D'ELIA
Some methodological issues on multivariate modelling of rank data.

2001, 3.100 Roberto RAUCCI, Luigi TADDEO
Nuove classi di funzioni scalari quasiconcave generalizzate: caratterizzazioni ed applicazioni a problemi di ottimizzazione.

2001, 3.101 Adriana BARONE, Annamaria NESE
Some insights into night work in Italy.

2001, 3.102 Alessandra AMENDOLA, Marcella NIGLIO

Predictive distributions of nonlinear time series models.

2001, 3.103 Roberto RAUCCI
Sul concetto di certo equivalente nella teoria HSSB.

2001, 3.104 Roberto RAUCCI, Luigi TADDEO
On stackelberg games: a result of unicity.

2001, 3.105 Roberto RAUCCI
Una definizione generale e flessibile di insieme limitato superiormente in $\Re^n$

2001, 3.106 Roberto RAUCCI
Stretta quasiconcavità nelle forme funzionali flessibili.

2001, 3.107 Roberto RAUCCI
Sugli insiemi limitati in $\Re^m$ *rispetto ai coni.*

2001, 3.108 Roberto RAUCCI
Monotonie, isotonie e indecomponibilità deboli per funzioni a valori vettoriali con applicazioni.

2001, 3.109 Roberto RAUCCI
Generalizzazioni del concetto di debole Kuhn-Tucker punto-sella.

2001, 3.110 Antonia Rosa GURRIERI, Marilene LORIZIO
Le determinanti dell'efficienza nel settore sanitario. Uno studio applicato.

2001, 3.111 Gianluigi COPPOLA
Studio di una provincia meridionale attraverso un'analisi dei sistemi locali del lavoro. Il caso di Salerno.

2001, 3.112 Francesco GIORDANO
Reti neurali per l'analisi del trend: un approccio per identificare la topologia della rete.

2001, 3.113 Marcella NIGLIO
Nonlinear time series models with switching structure: a comparison of their forecast performances.

2001, 3.114 Damiano FIORILLO
Capitale sociale e crescita economica. Review dei concetti e dell'evidenza empirica.

2001, 3.115 Roberto RAUCCI, Luigi TADDEO
Generalizzazione del concetto di continuità e di derivabilità.

2001, 3.116 Marcella NIGLIO
Ricostruzione dei dati mancanti in serie storiche climatiche.

2001, 3.117 Vincenzo VECCHIONE
Mutamenti del sistema creditizio in un'area periferica.

2002, 3.118 Francesco GIORDANO, Michele LA ROCCA, Cira PERNA
Bootstrap variable selection in neural network regression models.

2002, 3.119 Roberto RAUCCI, Luigi TADDEO
Insiemi debolmente convessi e concavità in senso generale.

2002, 3.120 Vincenzo VECCHIONE
Know how locali e percorsi di sviluppo in aree e settori marginali.

2002, 3.121 Michele LA ROCCA, Cira PERNA
Neural networks with dependent data.

2002, 3.122 Pietro SENESI
Economic dynamics: theory and policy. A stability analysis approach.

2002, 3.123 Gianluigi COPPOLA
Stima di un indicatore di pressione ambientale: un'applicazione ai comuni della Campania.

2002, 3.124 Roberto RAUCCI
Sull'esistenza di autovalori e autovettori positivi anche nel caso non lineare.

2002, 3.125 Maria Carmela MICCOLI
Identikit di giovani lucani.

2002, 3.126 Sergio DESTEFANIS, Giuseppe STORTI
Convexity, productivity change and the economic performance of countries.

2002, 3.127 Giovanni C. PORZIO, Maria Prosperina VITALE
Esplorare la non linearità nei modelli Path.

2002, 3.128 Rosa FERRENTINO
Sulla funzione di Seal.

2003, 3.129 Michele LA ROCCA, Cira PERNA
Identificazione del livello intermedio nelle reti neurali di tipo feedforward.

2003, 3.130 Alessandra AMENDOLA, Marcella NIGLIO, Cosimo VITALE
The exact multi-step ahead predictor of SETARMA models.

2003, 3.131 Mariangela BONASIA
La dimensione ottimale di un sistema pensionistico: means tested vs programma universale.

2003, 3.132 Annamaria NESE
Abitazione e famiglie a basso reddito.

2003, 3.133 Maria Lucia PARRELLA
Le proprietà asintotiche del Local Polynomial Bootstrap.

2003, 3.134 Silvio GIOVE, Maurizio NORDIO, Stefano SILVONI
Stima della prevalenza dell'insufficienza renale cronica con reti bayesiane: analisi costo efficacia delle strategie di prevenzione secondaria.

2003, 3.135 Massimo SALZANO
Globalization, complexity and the holism of the italian school of public finance.

2003, 3.136 Giuseppina AUTIERO
Labour market institutional sistems and unemplyment performance in some Oecd countries.

2003, 3.137 Marisa FAGGINI
Recurrence analysis for detecting non-stationarity and chaos in economic times series.

2003, 3.138 Marisa FAGGINI, Massimo SALZANO
The reverse engineering of economic systems. Tools and methodology.

2003, 3.139 Rosa FERRENTINO
In corso di pubblicazione.

2003, 3.140 Rosa FERRENTINO, Roberto RAUCCI
Sui problemi di ottimizzazione in giochi di Stackelberg ed applicazioni in modelli economici.

2003, 3.141 Carmine SICA
In corso di pubblicazione.

2004, 3.142 Sergio DESTEFANIS, Antonella TADDEO, Maurizio TORNATORE
The stock of human capital in the Italian regions.

2004, 3.143 Elena Laureana DEL MERCATO
Edgeworth equilibria with private provision of public good.

2004, 3.144 Elena Laureana DEL MERCATO
Externalities on consumption sets in general equilibrium.

2004, 3.145 Rosa FERRENTINO, Roberto RAUCCI
Su alcuni criteri delle serie a termini non negativi.

2004, 3.146 Rosa FERRENTINO, Roberto RAUCCI
Legame tra le soluzioni di Minty e di Stempacenhia nelle disequazioni variazionali.

2004, 3.147 Gianluigi COPPOLA
In corso di pubblicazione.

2004, 3.148 Massimo Spagnolo
The Importance of Economic Incentives in Fisheries Management

2004, 3.149 F. Salsano
La politica monetaria in presenza di non perfetta osservabilità degli obiettivi del banchiere centrale.

2004, 3.150 A. Vita
La dinamica del cambiamento nella rappresentazione del territorio. Una mappa per i luoghi della Valle dell'Irno.

2004, 3.151 Celi
Empirical Explanation of vertical and horizontal intra-industry trade in the UK: a comment.

2004, 3.152 Amendola – P. Vitale
Self-Assessment and Career Choices: An On-line resource for the University of Salerno.

2004, 3.153 A. Amendola – R. Troisi
Introduzione all'economia politica dell'organizzazione: nozioni ed applicazioni.

2004, 3.154 A. Amendola – R. Troisi
Strumenti d'incentivo e modelli di gestione del personale volontario nelle organizzazioni non profit.

2004, 3.155 Lavinia Parisi
La gestione del personale nelle imprese manifatturiere della provincia di Salerno.

2004, 3.156 Angela Spagnuolo – Silvia Keller
La rete di accesso all'ultimo miglio: una valutazione sulle tecnologie alternative.

2005, 3.157 Davide Cantarelli
Elasticities of Complementarity and Substitution in Some Functional Forms. A Comparative Review.

2005, 3.158 Pietro Coretto – Giuseppe Storti
Subjective Sxpectations in Economics: a Statistical overview of the main findings
.