IFS approximations of distribution functions and related optimization problems

Stefano M. Iacus and Davide La Torre

Abstract

In this note an iterated function system (IFS) on the space of distribution functions is built with the aim of proposing a new class of distribution function estimators. One IFS estimator is proposed and its properties are studied in details. Relative efficiency of the estimator, for small and moderate sample sizes, are presented via Monte Carlo analysis. It turns out that the IFS distribution function estimator is, in several cases, more accurate than the celebrated empirical distribution function.

M.S.C. 2000: 62E17, 62H10, 37H.
Key words: iterated function systems, optimization, nonparametric statistics.

§1. Introduction

The Iterated Function Systems (IFSs) were born in mid eighties [2, 9] as applications of the theory of discrete dynamical systems and as useful tools to build fractals and other similar sets. Some possible applications of IFSs can be found in image processing theory [6], in the theory of stochastic growth models [17] and in the theory of random dynamical systems [1, 4, 12]. Here we try to apply this methodology to estimation.

The fundamental result [2] on which the IFS method is based is Banach theorem. In practical applications the crucial problem, usually called the *inverse problem* in the IFS literature, is formulated as follows: given f in some metric space (S, d), find a contraction $T : S \to S$ that admits a unique fixed point $\tilde{f} \in S$ such that $d(f, \tilde{f})$ is small enough. In fact if one is able to solve the inverse problem with arbitrary precision, it is possible to identify f with the operator T which has it as fixed point. The paper is organized as follows: Section is devoted to introduce a contractive operator T on the space of distribution functions and to the definition of the inverse problem for T. Section is devoted to estimation. We propose an IFS distribution function estimator and we study its properties. We will also study if there is any advantage of using IFS estimator instead of the celebrated empirical distribution function (e.d.f) estimator when the sample size is small (n from 10 to 30) or moderate (n from 50 to 100). Monte Carlo analysis seems to show some gain of the IFS over the e.d.f.

Applied Sciences, Vol.6, 2004, pp. 27-38.

§2. A contraction on the space of distribution functions

Let us denote by $\mathcal{F}([0,1])$ the space of distribution functions F on $[0,1]$ and by $\mathcal{B}([0,1])$ the space of real bounded functions on $[0,1]$. Let us further define, for $F, G \in \mathcal{B}([0,1])$, $d_\infty(F,G) = \sup_{x\in[0,1]} |F(x) - G(x)|$. So that $(\mathcal{F}([0,1]), d_\infty)$ is a complete metric space.

Let $N \in \mathbb{N}$ be fixed and let:

i) $w_i : [a_i, b_i) \to [c_i, d_i) = w_i([a_i, b_i))$, $i = 1, \ldots, N-1$, $w_N : [a_N, b_N] \to [c_N, d_N]$, with $a_1 = c_1 = 0$ and $b_N = d_N = 1$;

ii) w_i, $i = 1 \ldots N$, are increasing and continuous;

iii) $\bigcup_{i=1}^{N-1} [c_i, d_i) \cup [c_N, d_N] = [0,1]$;

iv) if $i \neq j$ then $[c_i, d_i) \cap [c_j, d_j) = \varnothing$.

v) $p_i \geq 0$, $i = 1, \ldots, N$, $\delta_i \geq 0$, $i = 1 \ldots N-1$, $\sum_{i=1}^{N} p_i + \sum_{i=1}^{N-1} \delta_i = 1$.

On $(\mathcal{F}([0,1], d_\infty)$ we define an operator in the following way:

$$
(2.1) \quad TF(x) = \begin{cases} p_1 F(w_1^{-1}(x)), & x \in [c_1, d_1) \\ p_i F(w_i^{-1}(x)) + \sum_{j=1}^{i-1} p_j + \sum_{j=1}^{i-1} \delta_j\,, & x \in [c_i, d_i)\,,\ i = 2, \ldots, N-1 \\ p_N F(w_N^{-1}(x)) + \sum_{j=1}^{N-1} p_j + \sum_{j=1}^{N-1} \delta_j\,, & x \in [c_N, d_N] \end{cases}
$$

where $F \in \mathcal{F}([0,1])$. In many pratical cases w_i are affine maps. The new distribution function TF is union of distorted copies of F; this is the fractal nature of the operator (see Figure 1). A similar approach has been discussed in [13] but here a more general operator is defined. We stress here that in the following we will think to the maps w_i and to the parameters δ_j as fixed whistle the parameters p_i have to be chosen. To put in evidence the dependence of the operator T on the vector $p = (p_1, \ldots, p_N)$ we will write T_p instead of T.

In Remark 2.1 the hypotheses ii) and v) will be weakened to allow more general functionals.

Theorem 2.1. *T_p is an operator from $\mathcal{F}([0,1])$ to itself.*

Proof. It is trivial that $T_pF(0) = 0$ and $T_pF(1) = 1$. Furthermore if $x_1 > x_2$, without loss of generality, we can consider the two cases:

i) $x_1, x_2 \in w_i([a_i, b_i))$;

ii) $x_1 \in w_{i+1}([a_{i+1}, b_{i+1}))$ and $x_2 \in w_i([a_i, b_i))$.

In case i), recalling that w_i are increasing maps, we have:

$$\begin{aligned} T_pF(x_1) &= p_i F(w_i^{-1}(x_1)) + \sum_{j=1}^{i-1} p_j + \sum_{j=1}^{i-1} \delta_j \\ &\geq p_i F(w_i^{-1}(x_2)) + \sum_{j=1}^{i-1} p_j + \sum_{j=1}^{i-1} \delta_j = T_pF(x_2) \end{aligned}$$

In case ii) we obtain:

$$\begin{aligned} T_pF(x_1) - T_pF(x_2) &= p_i + \delta_i + p_{i+1}F(w_{i+1}^{-1}(x_1)) - p_i F(w_i^{-1}(x_2)) \\ &= p_i(1 - F(w_i^{-1}(x_2))) + p_{i+1}F(w_{i+1}^{-1}(x_1)) + \delta_i \geq 0 \end{aligned}$$

since $p_i \geq 0$, $\delta_i \geq 0$ and $0 \leq F(y) \leq 1$. Finally, one can prove without difficulties the right continuity of T_pf. □

The following remark will be useful for the applications in Section .

Remark 2.1. *If hypotheses i), ii) and v) in the definition of T_p are replaced by the following*

i'+ii') $w_i(x) = x$, $a_i = c_i$, $b_i = d_i$, $i = 1, \dots, N$,

v') $p_i = p$, $\delta_i \geq -p$, $Np + \sum_{i=1}^{N-1} \delta_i = 1$,

then $T_p : \mathcal{F}([0,1]) \to \mathcal{F}([0,1])$.

Theorem 2.2. *If* $c = \max_{i=1,\dots,N} p_i < 1$, *then T_p is a contraction on $(\mathcal{F}([0,1]), d_\infty)$ with contractivity constant c.*

Proof. Let $F, G \in (\mathcal{F}([0,1]), d_\infty)$ and let it be $x \in w_i([a_i, b_i))$. We have

$$|T_pF(x) - T_pG(x)| = p_i \left| F(w_i^{-1}(x)) - G(w_i^{-1}(x)) \right| \leq c\, d_\infty(F, G)\,.$$

This implies $d_\infty(T_pF, T_pG) \leq c\, d_\infty(F, G)$. □

The following theorem states that small perturbations of the parameters p_i produce small variations on the fixed point of the operator.

Theorem 2.3. *Let p, $p^* \in \mathbb{R}^N$ such that $T_pF_1 = F_1$ and $T_{p^*}F_2 = F_2$. Then*

$$d_\infty(F_1, F_2) \leq \frac{1}{1-c} \sum_{j=1}^{N} \left| p_j - p_j^* \right|$$

where c is the contractivity constant of T_p.

Proof. In fact, recalling that w_i and δ_i are fixed, we have

$$\begin{aligned}
d_\infty(F_1, F_2) = d_\infty(T_pF_1, T_pF_2) \\
&= \max_{i=1,\dots,N} \sup_{x\in[c_i,d_i)} \left\{ \left| p_i F_1(w_i^{-1}(x)) + \sum_{j=1}^{i-1} p_j \right.\right. \\
&\qquad \left.\left. - p_i^* F_2(w_i^{-1}(x)) - \sum_{j=1}^{i-1} p_j^* \right| \right\} \\
&\le \sum_{i=1}^{N} |p_i - p_i^*| + c\, d_\infty(F_1, F_2)\,,
\end{aligned}$$

since

$$\begin{aligned}
&\left| p_i F_1(w_i^{-1}(x)) + \sum_{j=1}^{i-1} p_j - p_i^* F_2(w_i^{-1}(x)) - \sum_{j=1}^{i-1} p_j^* \right| \\
&\le \sum_{j=1}^{i-1} |p_j - p_j^*| + |p_i F_1(w_i^{-1}(x)) - p_i F_2(w_i^{-1}(x))| \\
&\quad + |p_i F_2(w_i^{-1}(x)) - p_i^* F_2(w_i^{-1}(x))| \\
&\le \sum_{j=1}^{i-1} |p_j - p_j^*| + p_i d_\infty(F_1, F_2) + |p_i - p_i^*| \\
&\le c\, d_\infty(F_1, F_2) + \sum_{j=1}^{N} |p_j - p_j^*|\,.
\end{aligned}$$

□

Choose $F \in (\mathcal{F}([0,1]), d_\infty)$. The goal now is to find a contractive map $T : \mathcal{F}([0,1]) \to \mathcal{F}([0,1])$ which has a fixed point "near" to F. In fact if it is possible to solve the inverse problem with an arbitrary precision one can identify the operator T with its fixed point. With a fixed system of maps w_i and parameters δ_j, the inverse problem can be solved, if it is possible, by using the parameters p_i. These have to be choosen in the following convex set:

$$C = \left\{ p \in \mathbb{R}^N : p_i \ge 0, i = 1, \dots, N, \sum_{i=1}^{N} p_i = 1 - \sum_{i=1}^{N-1} \delta_i \right\},$$

We have the following result that is trivial to prove.

Proposition 2.1. *Choose $\epsilon > 0$ and $p \in C$ such that $p_i \cdot p_j > 0$ for some $i \neq j$. If $d_\infty(T_pF, F) \le \epsilon$, then:*

$$d_\infty(F, \tilde{F}) \le \frac{\epsilon}{1-c},$$

where $\tilde{F}$ is the fixed point of T_p on $\mathcal{F}([0,1])$ and $c = \max_{i=1,\dots,N} p_i$ is the contractivity constant of T_p.

If we wish to find an approximated solution of the inverse problem, we have to solve the following constrained optimization problem:

$$(P) \qquad \min_{p \in C} d_\infty(T_p F, F)$$

It is clear that the ideal solution of (**P**) consists of finding a $p^* \in C$ such that $d_\infty(T_{p^*}F, F) = 0$. In fact this means that, given a distribution function F, we have found a contractive map T_p which has exactly F as fixed point. Indeed the use of Proposition 2.1 gives us only an approximation of F. This can be improved by increasing the number of parameters p_i (and maps w_i).

The following result proves the convexity of the function $D(p){=}d_\infty(T_pF,\ F)$, $p \in \mathbb{R}^N$.

Theorem 2.4. *The function $D(p) : \mathbb{R}^N \to \mathbb{R}$ is convex.*

Proof. If we choose $p_1, p_2 \in \mathbb{R}^N$ and $\lambda \in [0, 1]$ then:

$$\begin{aligned} D(\lambda p_1 + (1-\lambda)p_2) &= \sup_{x \in [0,1]} |T_{\lambda p_1 + (1-\lambda)p_2} F(x) - F(x)| \\ &\le \lambda \sup_{x \in [0,1]} |T_{p_1} F(x) - F(x)| \\ &\quad + (1-\lambda) \sup_{x \in [0,1]} |T_{p_2} F(x) - F(x)| \\ &= \lambda D(p_1) + (1-\lambda) D(p_2). \end{aligned}$$

□

Hence for solving problem (**P**) one can recall classical results about convex programming problems (see for instance [18]). A necessary and sufficient condition for $p^* \in C$ to be a solution of (**P**) can be given by Kuhn-Tucker conditions.

§3. Distribution function estimation and applications

In this section we focus the attention on some estimation problems. Instead of trying to solve exactly the problem in (**P**) we will use the properties of distribution functions to obtain a *good* approximator of F that can be directly used in distribution function estimation. Via Monte Carlo analysis, we will also show that, for small sample sizes, this IFS estimator is better than the celebrated empirical distribution function in several situations.

As is usual in statistical applications, given a sample of n independent and identically distributed observations, $(x_1, x_2, \ldots, x_n)$, drawn from an unknown continuous distribution function $F \in \mathcal{F}([0, 1])$, one can easily contruct the empirical distribution function (e.d.f.) $\hat{F}_n$ that reads as

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \chi_{(-\infty, x]}(x_i), \quad x \in \mathbb{R}\,,$$

where χ_A is the indicator function of the set A. Asymptotic properties of optimality of $\hat{F}_n$ as estimator of the unknown F when n goes to infinity are well known and studied [15, 16].

Remark 3.2. *This function has an IFS representation that is exact and can be found without solving any optimization problem. We assume that the x_i in the sample are all different (this assumption is natural if F is a continuous distribution function). Let $w_i(x) : [x_{i-1}, x_i) \to [x_{i-1}, x_i)$, when $i = 1 \dots n$ and $w_1(x) : [0, x_1) \to [0, x_1)$, $w_{n+1}(x) : [x_n, x_{n+1}] \to [x_n, x_{n+1}]$, with $x_0 = 0$ and $x_{n+1} = 1$. Assume also that every map is of the form $w_i(x) = x$. If we choose $p_i = \frac{1}{n}$, $i = 2 \dots n+1$, $p_1 = 0$ and*

$$\delta_1 = \frac{n-1}{n^2}, \quad \delta_i = -\frac{1}{n^2}$$

then the following representation holds:

$$T_p \hat{F}_n(x) = \begin{cases} 0, & i = 1 \\ \frac{1}{n} \hat{F}_n(x) + \frac{n-1}{n^2}, & i = 2 \\ \frac{1}{n} \hat{F}_n(x) + \frac{i-1}{n} + \frac{n-i+1}{n^2}, & i = 3, \dots, n+1. \end{cases}$$

for $x \in [x_{i-1}, x_i)$. It should be clear that any discrete distribution function can be represented exactly with an IFS by similar arguments.

From now on we assume that $\delta_i = 0$, $\forall\, i$. To produce an estimator we should first provide a good approximator of F. So fix an $F \in \mathcal{F}([0,1])$ and choose $N+1$ ordered points $(x_1, \dots, x_{N+1})$ such that $x_1 = 0$ and $x_{N+1} = 1$. Define the maps w_i and coefficients p_i as follows

$$p_i(F) = F(x_{i+1}) - F(x_i)\,,$$

$$w_i(x) : [0,1) \to [x_i, x_{i+1}) = (x_{i+1} - x_i) \cdot x + x_i\,, \qquad i = 1, \dots, N.$$

The functional T_p can be denoted as T_N with this given set of maps and coefficients as it depends only on the number of points. For any $u \in \mathcal{F}([0,1])$, T_N can be written as

$$T_N u(x) = \sum_{i=1}^{N} p_i u\left(w_i^{-1}(x)\right) = \sum_{i=1}^{N} \Big(F(x_{i+1}) - F(x_i) \Big) \cdot u\left(\frac{x - x_i}{x_{i+1} - x_i}\right),$$

$x \in \mathbb{R}$, T_N is a contraction on $(\mathcal{F}([0,1]), d_{\sup})$ and it is such that $T_N u(x_i) = F(x_i)$, $\forall\, i$. T_N has been built by first rescaling in abscissa the whole function F from $[0,1]$ to each of the intervals $[x_i, x_{i+1})$ and then copying it in ordinate after a translation equal to $F(x_i)$. The idea is to use the fractal nature of the IFS (see again Figure 1).

This functional is indeed a function of F and can't be used directly in statistical applications as F is unknown. To this end, just think at the point x_i as the quantiles q_i of F, i.e. choose $N+1$ points $u_1 = 0 < u_2 < \cdots u_n < u_{N+1} = 1$ equally spaced on [0,1] and set $q_i = F^{-1}(u_i)$. The function T_N becomes

$$T_N u(x) = \sum_{i=1}^{N} \frac{1}{N} u\left(\frac{x - q_i}{q_{i+1} - q_i}\right), \quad x \in \mathbb{R}$$

and T_N depends on F only through the quantiles q_i, moreover in this way, it is assured that the profile of F is followed as smooth as possible. In fact, if two quantiles q_i and q_{i+1} are relatively distant each other, then F is slowly increasing in the interval

(q_i, q_{i+1}) and viceversa. As the quantiles can be easily estimate from a sample we have now a possible candidate for an IFS distribution function estimator. Thus, let now $x_1, x_2, \ldots, x_n$ be a sample drawn from F and let $\hat{q}_i$, $i = 1, \ldots, N+1$ the empirical quantiles of order $1/N$ such that $\hat{q}_1 = 0$ and $\hat{q}_{N+1} = 1$, then we propose as IFS distribution function estimator the fixed point of the following IFS

$$\hat{T}_N u(x) = \sum_{i=1}^{N} \frac{1}{N} u\left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i}\right), \quad x \in \mathbb{R},$$

with $u \in \mathcal{F}([0,1])$.

Let $N = N_n$ be a sequence depending on the sample size n. Denote the fixed point of $\hat{T}_{N_n}$ by $\hat{T}^*_{N_n}$. Then $\hat{T}^*_{N_n}$ satifies

$$\hat{T}^*_{N_n}(x) = \sum_{i=1}^{N} \frac{1}{N} \hat{T}^*_{N_n}\left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i}\right), \quad x \in \mathbb{R}.$$

Theorem 3.5 (Glivenko-Cantelli). *Let $N_n \to \infty$ as $n \to \infty$, then for any fixed F*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} \left| \hat{T}^*_{N_n}(x) - F(x) \right| \overset{a.s.}{=} 0\,.$$

Proof. We can write

$$\left| \hat{T}^*_{N_n}(x) - F(x) \right| \leq \left| \hat{T}^*_{N_n}(x) - \hat{F}_n(x) \right| + \left| \hat{F}_n(x) - F(x) \right|$$

and the first term can be estimated by $1/N_n$ while the second one converges to 0 almost surely by the Glivenko-Cantelli theorem for the e.d.f. □

3.1 Monte Carlo analysis for small samples

The IFS estimator is as efficient as the e.d.f. in the large sample case; we will give now empirical evidence that it can be even better in some situations for the small sample case. The main difference between the e.d.f and the IFS estimator is that the e.d.f. is a step-wise function whistle the IFS is somewhat "smooth" in the sense that the IFS jumps are several order of magnitude smaller then the ones of the e.d.f. Remember that we assume that the underline distribution function F is a continuous one.

Tables 1 and 2 report the results of Monte Carlo analysis for distribution function estimation. At each Monte Carlo step, we have drawn samples of $n = 10, 20, 30, 50, 75,$ 100 replications for several type of distributions. For each distribution and sample size n we have done 100 Monte Carlo simulations. We have chosen the beta family of distribution functions because they allow very good and well tested random number generators, different kinds of asymmetry (beta(3,5) and beta(5,3)), bell shaped distributions with (beta(5,5)) or without (beta(2,2)) tails, and also U-shaped distributions (beta(.1,.1)). For distribution function estimation we have considered the $d_{\sup}$ (SUP-NORM) distance and the average mean square error (AMSE) both for $\hat{T}_N$ and $\hat{F}_n$, then we have reported in the table only the ratio of the indexes. Thus, each entry in

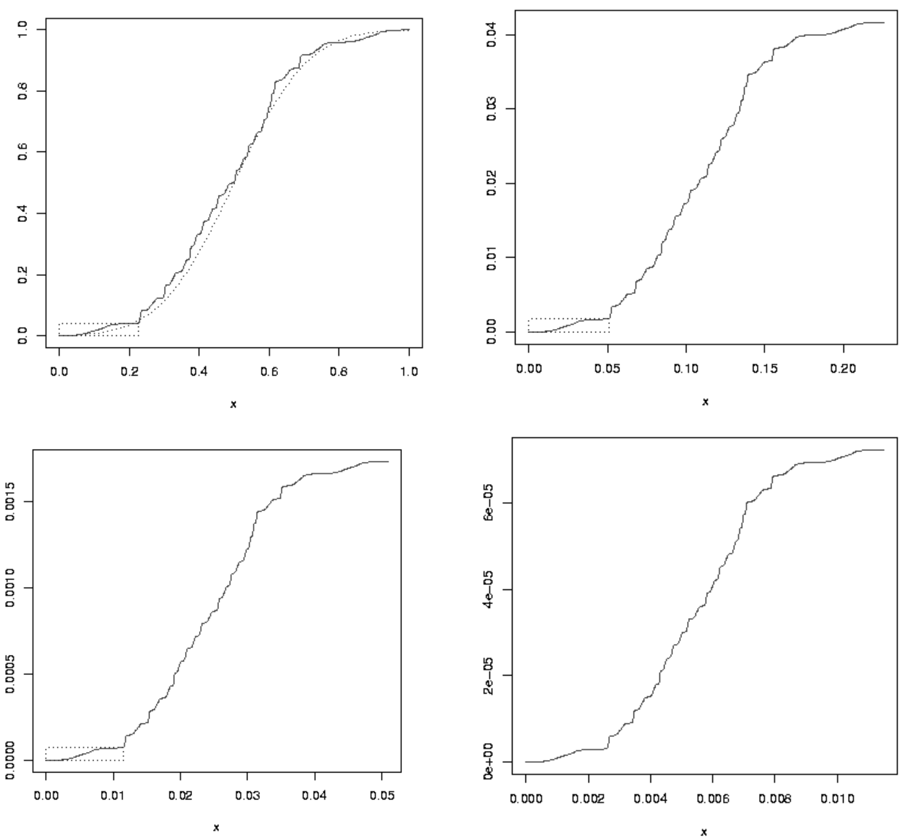

Figure 1: *The fractal nature of the IFS distribution function estimator* $\hat{T}_N$*. The dotted line is the underlying truncated Gaussian distribution. The dotted rectangle is to represent the area zoomed in the next plot (left to right, up to down). The dotted boxes are in the order:* $[0, \hat{q}_2] \times [0, \hat{q}_2]$, $[0, \hat{q}_2^2] \times [0, \hat{q}_2^2]$ *and* $[0, \hat{q}_2^3] \times [0, \hat{q}_2^3]$.

n	law	AMSE $\hat{T}_N$ w.r.t. $\hat{F}_n$	SUP-NORM $\hat{T}_N$ w.r.t. $\hat{F}_n$
10	beta(.9,.1)	94.86	98.10
10	beta(.1,.9)	99.79	100.33
10	beta(.1,.1)	99.26	100.05
10	beta(2, 2)	80.99	82.18
10	beta(5, 5)	89.91	89.74
10	beta(5, 3)	98.57	92.86
10	beta(3, 5)	90.70	91.34
10	beta(1, 1)	81.23	80.04

n	law	AMSE $\hat{T}_N$ w.r.t. $\hat{F}_n$	SUP-NORM $\hat{T}_N$ w.r.t. $\hat{F}_n$
20	beta(.9,.1)	101.56	99.1
20	beta(.1,.9)	113.31	114.50
20	beta(.1,.1)	99.77	98.84
20	beta(2, 2)	86.05	83.34
20	beta(5, 5)	89.27	87.49
20	beta(5, 3)	92.89	89.31
20	beta(3, 5)	88.27	87.29
20	beta(1, 1)	84.89	79.27

n	law	AMSE $\hat{T}_N$ w.r.t. $\hat{F}_n$	SUP-NORM $\hat{T}_N$ w.r.t. $\hat{F}_n$
30	beta(.9,.1)	97.17	92.91
30	beta(.1,.9)	118.80	115.38
30	beta(.1,.1)	99.62	98.57
30	beta(2, 2)	86.23	84.01
30	beta(5, 5)	89.54	89.21
30	beta(5, 3)	91.67	87.90
30	beta(3, 5)	88.31	88.34
30	beta(1, 1)	87.16	81.00

Table 1: *Relative efficiency of IFS-based estimator with respect to the empirical distribution function. Small sample sizes. In many situations the $\hat{T}_N$ seems to be better (10 to 20%) then the usual empirical distribution function estimator.*

the table reports the percentage of error of $\hat{T}_N$ with respect to $\hat{F}_n$, the error of $\hat{F}_n$ being 100. The software used is R [10], freely available on `http://cran.R-project.org`, using a beta packages `ifs` available as an additional contributed package.

For the estimator $\hat{T}_N$ we have chosen to take $N_n = n/2$. In the small sample size case $n = 10, 20, 30$ it can be noted that $\hat{T}_N$ is sometimes better (from 10 to 20%) then the empirical distribution function. Our advise is that this happen due to the difference of the jumps (very high) of the e.d.f. and of $\hat{T}_N$ (rather small ones). This allows to follow the profile of a continuous curve better.

For moderate sample sizes ($n = 50, 75, 100$) the distance between $\hat{T}_N$ and $\hat{F}_n$ decreases and consequently the gain of using the IFS estimator is not that evident (on the average 5 to 10%).

Final remarks about the method

There is at least one open issue in this topic as this is a first attempt to introduce IFS in distribution function estimation: are there other maps then the ones used in T_N that can improve the performance of the corresponding IFS estimator? We

n	law	AMSE $\hat{T}_N$ w.r.t. $\hat{F}_n$	SUP-NORM $\hat{T}_N$ w.r.t. $\hat{F}_n$
50	beta(.9,.1)	98.59	95.87
50	beta(.1,.9)	107.85	113.85
50	beta(.1,.1)	97.77	99.87
50	beta(2, 2)	91.55	87.40
50	beta(5, 5)	91.57	89.64
50	beta(5, 3)	97.80	92.73
50	beta(3, 5)	91.90	91.40
50	beta(1, 1)	92.84	86.24

n	law	AMSE $\hat{T}_N$ w.r.t. $\hat{F}_n$	SUP-NORM $\hat{T}_N$ w.r.t. $\hat{F}_n$
75	beta(.9,.1)	98.85	97.66
75	beta(.1,.9)	122.62	122.78
75	beta(.1,.1)	100.37	106.85
75	beta(2, 2)	94.77	90.42
75	beta(5, 5)	97.41	92.39
75	beta(5, 3)	104.31	96.29
75	beta(3, 5)	95.14	91.80
75	beta(1, 1)	98.39	89.05

n	law	AMSE $\hat{T}_N$ w.r.t. $\hat{F}_n$	SUP-NORM $\hat{T}_N$ w.r.t. $\hat{F}_n$
100	beta(.9,.1)	98.22	96.65
100	beta(.1,.9)	114.91	152.41
100	beta(.1,.1)	102.02	110.96
100	beta(2, 2)	97.03	91.68
100	beta(5, 5)	95.08	93.37
100	beta(5, 3)	95.52	93.96
100	beta(3, 5)	97.59	94.23
100	beta(1, 1)	101.58	92.26

Table 2: *Relative efficiency of IFS-based estimator with respect to the empirical distribution function. Moderate sample sizes. In many situations the $\hat{T}_N$ seems to be better (around 10% in average) then the usual empirical distribution function estimator. As $\hat{T}_N$ is asymptotically equivalent to the e.d.f it is clear that the advantage of using $\hat{T}_N$ instead of $\hat{F}_n$ reduces as the sample size increases.*

have suggested a quantile approach but some other good partition of the space, like a dyadic sequence, can be used at the cost of the need to solve some optimization problems. In [6] this problem is touched incidentally but not in a statistical context.

To put in evidence the fractal nature of the IFS we have added a graph (see Figure 1) of the $\hat{T}_N$ estimator on sampled data from a Gaussian distribution rescaled on [0,1]. We have zoomed the graph three times to show the self-similarity of the fixed point.

Acknowledgements

The authors are thanksfull to Anestis Antoniadis for his *conseils* and for having brought to their attention the papers of Winter [20, 21] and Yukich [22].

References

[1] L. Arnold and H. Crauel, Iterated function systems and multiplicative ergodic theory, *Diffusion processes and related problems in analysis* (Pinsky M.A., Vihstutz V. eds.), vol. II, Stochastics flows, Progress in probability 27, Birkäuser, Boston, 283-305, 1992.

[2] M.F. Barnsley and S. Demko, Iterated function systems and the global construction of fractals, *Proc. Roy. Soc. London, Ser A*, **399** (1985), 243-275.

[3] C.A. Cabrelli, B. Forte, U.M. Molter and E.R. Vrscay, Iterated fuzzy set systems: a new approach to the inverse problem for fractals and other sets, *J. Math. Anal. Appl.*, **171** (1992), 79-100.

[4] J.H. Elton and M. Piccioni, Iterated function systems arising from recursive estimation problems, *Probab. Theory Related Fields*, **91** (1992), 103-114.

[5] T.R. Fleming and D.P. Harrington, *Counting processes and survival analysis*, Wiley, New York, 1991.

[6] B. Forte and E.R. Vrscay, Solving the inverse problem for function/image approximation using iterated function systems, I. Theoretical basis, *Fractal*, **2**, 3 (1994), 325-334.

[7] B. Forte and E.R. Vrscay, Inverse problem methods for generalized fractal transforms, in *Fractal Image Encoding and Analysis*, NATO ASI Series F, Vol. 159, ed. Y. Fisher, Springer Verlag, Heidelberg, 1998.

[8] R.D. Gill and B.Y. Levit, Applications of the van Trees inequality: A Bayesian Cramér-Rao bound, *Bernoulli*, 1 (1995), 59-79.

[9] J. Hutchinson, Fractals and self-similarity, *Indiana Univ. J. Math.*, **30**, 5 (1981), 713-747.

[10] R. Ihaka and R. Gentleman, R: A Language for Data Analysis and Graphics, *Journal of Computational and Graphical Statistics*, 5 (1996), 299-314.

[11] E. Kaplan and P. Meyer, Nonparametric estimator from incomplete observations, J. Amer. Statist. Ass, **53** (1958), 457-81.

[12] A.A. Kwiecinska and W. Slomczynski, Random dynamical systems arising from iterated function systems with place-dependent probabilities, *Stat. and Prob. Letters*, **50** (2000), 401-407.

[13] D. La Torre and M. Rocca, Iterated function systems and optimization problems, *Rend. Sem. Mat. Messina SER II*, **6**, 21 (1999), 165-173.

[14] B.Y. Levit, Infinite-dimensional information inequalities, *Theory Probab. Applic.*, 23 (1978), 371-377.

[15] P.W. Millar, Asymptotic minimax theorems for the sample distribution function, *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, **48** (1979), 233-252.

[16] P.W. Millar, The minimax principle in asymptotic statistical theory, *Lecture Notes in Math.*, **976** (1983), 76-265.

[17] L. Montrucchio and F. Privileggi, Fractal steady states in stochastic optimal control models, *Ann. Op. Res.*, **88** (1999), 183-197.

[18] R.T. Rockafellar and R.J-B. Wets, *Variational analysis*, Springer, New York, 1998.

[19] M.E Tarter and M.D. Lock, *Model free curve estimation*, Chapman & Hall, New York, 1993.

[20] B.B. Winter, Strong uniform consistency of integrals of density estimators, *Can. J. Statist.*, **1** (1973), 247–253.

[21] B.B. Winter, Convergence rate of perturbed empirical distribution functions, *J. Appl. Prob.*, **16** (1979), 163–173.

[22] J.E. Yukish, A note on limit theorems for perturbed empirical processes. *Stoch. Proc. Appl.*, **33** (1989), 163–173.

Authors' addresses:

Stefano M. Iacus
Department of Economics, Via Conservatorio 7
I-20122 Milan, Italy
e-mail: stefano.iacus@unimi.it

Davide La Torre
Department of Economics, Via Conservatorio 7
I-20122 Milan, Italy
e-mail: davide.latorre@unimi.it