Laws of the iterated logarithm for nonparametric sequential density estimators

Karima Lagha [1] **and Smail Adjabi** [2]

[1] *Laboratory LAMOS, University of Bejaia, Targa Ouzemour, 06000 Bejaia, Algeria*
[2] *Laboratory LAMOS, University of Bejaia, Targa Ouzemour, 06000 Bejaia, Algeria*
karima_lagha@yahoo.com, adjabi@hotmail.com

Received 6 June 2012
Accepted 5 February 2013

Abstract

In this note, we establish a law of iterated logarithm for a triangular array of a random number of independent random variables and apply it to obtain laws of iterated logarithm for the sequential nonparametric density estimators. We consider the case of Rosenblatt-Parzen kernel estimators and orthogonal polynomial estimators. We point out that we obtain in the present paper sharp pointwise rates of the consistency.

Keywords: Nonparametric density estimator, iterated logarithm, Parzen-Rosenblatt estimator, orthogonal polynomial estimator.

1. Introduction

In this paper, we investigate the problem related to the iterated logarithm law pertaining with the sequential density estimation. Towards this objective, we establish in the first place a result for a statistic built upon a triangular array which allows to deduce results for a number of estimators. In order to be more precise on the matter, let consider, for any $t \in \mathbb{R}^+$, a positive integer random variable N_t representing the number of observations which we may record in time $(0,t]$ as well as a sample $X_1, X_2, \ldots, X_{N_t}$ drown from a real random variable X. Denote by F the distribution function of X and by f its density function with respect to the Lebesgue measure. Suppose that N_t is independent of observations.
The estimators we shall consider are of the type:

$$\widehat{f}_t(x) = N_t^{-1} \sum_{i=1}^{N_t} K_{r(N_t)}(x; X_i),$$

where $\{K_r, r \in I\}$ is an $I-$indexed set of "kernel" functions and I an arbitrary real function index set. This problem has been widely investigated and a number of properties have been studied. We refer to Deheuvels [1], [2] and Prakasa Rao [3], for an overview of results of the subject. The asymptotic properties of the estimate of density function and probability distribution function, are investigated by Srivastava [4], the estimate of density function is asymptotically unbiased, consistent and uniformly consistent. Carroll [5] obtains the asymptotic normality of this estimate. In this note we investigate the problem related to the iterated logarithm law pertaining with the sequential density estimation.
In the non sequential estimation case, iterated logarithm results have been obtained by Hall [6], Stute [7] and

Deheuvels [8]. Hall obtains rates of strong consistency by establishing laws of iterated logarithm for: Rosenblatt-Parzen kernel estimators, trigonometric series estimators and orthogonal polynomial estimators, by making use of the invariance principle due to Komlós et al. [9].

We begin by establishing a law of iterated logarithm for a general class of triangular arrays of independent variables. This result is presented in section 2, we apply it to the Rosenblatt-Parzen kernel estimators in section 3 and to the orthogonal polynomial estimators in the section 4. These results are generalization of that given by Hall by considering the random case.

2. A Law of the Iterated Logarithm for Triangular Arrays

Let X be a random variable, with distribution function F confined to the interval (a,b), where $-\infty \leqslant a < b \leqslant \infty$ and probability density function f. Also, let $X_1, X_2, \dots, X_{N_t}$ be independent observations on X, where N_t, for any $t > 0$, be a non-negative integer-valued random variable, supposed to be independent of the observations.
Let $\{K_r, r \in I\}$ be a sequence of univariate functions each of bounded variation on (a,b). And define, for any $x \in \mathbb{R}$, the following process

$$S_t(x) = \sum_{i=1}^{N_t} [K_{r(N_t)}(x, X_i) - EK_{r(N_t)}(x, X_i)], \tag{1}$$

where $E(X)$ stands as the mathematical expectation of X,

$$\sigma_{rs} = cov[K_r(X_1), K_s(X_1)], \qquad \sigma_r^2 = \sigma_{rr}$$

and

$$g(n) = \left(2n\sigma_{r(n)}^2 \log\log n\right)^{1/2}, \quad n \in N^*. \tag{2}$$

The following theorem involves the triangular array result,

Theorem 1. *Suppose :*

$$(H_1) \quad N_t \xrightarrow{P} \infty, \;\; as \quad t \to \infty,$$

$$(H_2) \quad (\log n)^2 \left[\int |dK_{r(n)}(x)|\right]^2 / n\sigma_{r(n)}^2 \log\log n \quad \searrow 0, \qquad as \qquad n \to \infty,$$

and

$$(H_3) \quad \lim_{\varepsilon \to \infty} \limsup_{n \to \infty} \sup_{m} \left| \frac{r(m)}{r(n)} - 1 \right| = 0,$$

where the inner supremum is taken over values of m with $|m-n| \leqslant \varepsilon n$. Then

$$\limsup_{t \to \infty} \pm g^{-1}(N_t) S_t(x) = 1, \qquad \textit{in probability.}$$

In the proof of theorem (1) we use the following Lemma due to Srivastava [4].

Lemma

Let $\{Y_n\}$ be a sequence of random variables that is independent of the random variable N_t, for any t with $N_t \xrightarrow{P} \infty$ as $t \to \infty$. If $Y_n \xrightarrow{P} \theta$ as $n \to \infty$, then

$$Y_{N_t} \xrightarrow{P} \theta \;\text{ as }\; t \to \infty.$$

Here, θ is a constant.

Proof. We see that it suffices to consider the case where each X_i is uniform on $(0,1)$. In this case $a = 0$ and

$b=1$. Let F_t denote the empiric distribution function of $X_1, X_2, \ldots, X_{N_t}$. So, $F_t(x) = N_t^{-1}\sum_{i=1}^{N_t} \mathbf{1}_{\{X_i \leqslant x\}}$.
Using the result given in theorem 4 of [9] on a rich enough probability space and the relation (H1) it is easy to obtain that:

$$N_t[F_t(x)-x] = \sum_{i=1}^{N_t} W_i(x) + e_t(x), \;\; 0 \leqslant x \leqslant 1, t \in R \quad \text{in probability.}$$

Where $W_i, i \geqslant 1$ are independent Brownian bridges and there exist positive absolute constants C_1, C_2 and λ such that the error $e_n(x)$ verify :

$$P\left(\sup_{0\leqslant x\leqslant 1} |e_n(x)| > (C_1 \log n + x)\log n\right) < C_2 e^{-\lambda x}, \tag{3}$$

for all x and n. Therefore, the process defined in (1) can be written as:

$$\begin{aligned} S_t(x) &= \sum_{i=1}^{N_t}[K_{r(N_t)}(x,X_i) - EK_{r(N_t)}(x,X_i)] \\ &= N_t \int K_{r(N_t)}(x) dF_t(x) - N_t E\left(K_{r(N_t)}(x,X_1)\right) \\ &= -\sum_{i=1}^{N_t} \int W_i(x) dK_{r(N_t)}(x) - \int e_t(x) dK_{r(N_t)}(x) + \\ &\quad N_t \int \left[K_{r(N_t)}(x) - E(K_{r(N_t)}(x))\right] dF(x). \end{aligned}$$

We have, $\forall \varepsilon > 0$,

$$P\left((g(N_t))^{-1}\left|\int e_t(x) dK_{r(N_t)}(x)\right| > \varepsilon\right) = \sum_n P\left(g^{-1}(n)\left|\int e_n(x) dK_{r(n)}(x)\right| > \varepsilon\right) P(N_t = n)$$

and

$$\begin{aligned} P &\quad \left(g^{-1}(n)\left|\int e_n(x) dK_{r(n)}(x)\right| > \varepsilon\right) \leqslant P\left(\sup_{0\leqslant x\leqslant 1} |e_n(x)| > \varepsilon g(n) / \int |dK_{r(n)}(x))|\right) \\ &= P\left(\sup_{0\leqslant x\leqslant 1} |e_n(x)| > \varepsilon \log n [2n\sigma_{r(n)}^2 \log\log n / \log^2 n \{\int |dK_{r(n)}(x)|\}^2]^{1/2}\right) \\ &\leqslant C_2 \exp\left(-\delta [n\sigma_{r(n)}^2 \log\log n / \log^2 n \{\int |dK_{r(n)}(x)|\}^2]^{1/2}\right) \longrightarrow 0 \;\text{ as }\; n \to \infty, \end{aligned}$$

using the relation (3) and (H2), where $\delta > 0$ does not depend on n.

So, by lemma, we have,

$$\lim_{t\to\infty} g^{-1}(N_t) \int e_t(x)\, dK_{r(N_t)}(x) = 0, \qquad \text{in probability.}$$

In other side, for all $\varepsilon' > 0$ we have

$$\begin{aligned} &P\left(N_t (g(N_t))^{-1}\left|\int \left(K_{r(N_t)}(x) - EK_{r(N_t)}(x)\right) dF(x)\right| > \varepsilon'\right) = \\ &P\left(N_t (g(N_t))^{-1}\left|\sum_n \int \left(K_{r(N_t)}(x) - K_{r(n)}(x)\right) dF(x) P(N_t = n)\right| > \varepsilon'\right) = \\ &P\left(N_t (g(N_t))^{-1}\left|\sum_n \int \left(K_{r(N_t)}(x) - EK_{r(n)}(x,X_1)\right) dF(x) P(N_t = n)\right| > \varepsilon'\right) \end{aligned} \tag{4}$$

By Markov inequality we have,

$$\begin{aligned}(4) \quad &\leqslant \frac{1}{\varepsilon'} E\left[N_t (g(N_t))^{-1} \left| \sum_n \int \left(K_{r(N_t)}(x) - EK_{r(n)}(x,X_1)\right) dF(x) P(N_t = n)\right|\right] \\ &\leqslant \frac{1}{\varepsilon'} \sum_m \sum_n m g^{-1}(m) \left| \int \left(K_{r(m)}(x) - EK_{r(n)}(x,X_1)\right) dF(x)\right| P(N_t = n) P(N_t = m) \\ &= \frac{1}{\varepsilon'} \sum_m \sum_n m g^{-1}(m) \left| E\left(K_{r(m)} - K_{r(n)}\right)(x,X_1)\right| P(N_t = n) P(N_t = m) \qquad (5)\end{aligned}$$

using (H3) we obtain that for $t \to \infty$ we have $m,n \to \infty$ and (5) tends to zero. So,

$$\lim_{t\to\infty} N_t g^{-1}(N_t) \int \left(K_{r(N_t)}(x) - EK_{r(N_t)}(x)\right) dF(x) = 0, \text{ in probability.}$$

The law of iterated logarithm will be then obtained on the following process

$$T_t(x) = \sum_{i=1}^{N_t} \int W_i(x) dK_{r(N_t)}(x).$$

When N_t take a value n, we can consider $T_n(x) = n^{1/2} \int W(x) dK_{r(n)}(x)$, for a Brownian bridge W. So, $T_n(x) \curvearrowright \mathcal{N}(0, n\sigma^2_{r(n)})$.

In view of the lemma and the usual approximation to the tail of the normal distribution, for all $\varepsilon > 0$

$$P\left(T_n(x) > (1+\varepsilon)g(n)\right) \leqslant (2\pi)^{-1/2} exp\{-(1+\varepsilon)^2 \log\log n\}, (\longrightarrow 0 \text{ as } n \to \infty).$$

Therefore,

$$P\left(g^{-1}(N_t)T_t(x) > (1+\varepsilon)\right) = \sum_n P\left(T_n(x) > (1+\varepsilon)g(n)\right) P(N_t = n) \longrightarrow 0 \text{ as } t \to \infty.$$

In other side, $P\left(T_n(x) < (1-\varepsilon)g(n)\right) \longrightarrow 0$ as $n \to \infty$ see theorem 1 in [6], which implies that,

$$P\left(g^{-1}(N_t)T_t(x) < (1-\varepsilon)\right) = \sum_n P\left(T_n(x) < (1-\varepsilon)g(n)\right) P(N_t = n) \longrightarrow 0 \text{ as } t \to \infty.$$

So,

$$\limsup_{t\to\infty} \pm g^{-1}(N_t) T_t(x) = 1 \quad \text{in probability}$$

□

and theorem (1) holds.

3. Rosenblatt-Parzen kernel Estimators

Let K be a function of bounded variation on $(-\infty, +\infty)$ satisfying

$$zK(z) \longrightarrow 0 \quad \text{as } |z| \longrightarrow \infty \text{ and } \int_R K^2(z)\, dz < \infty.$$

Let $X_1, X_2, \ldots, X_{N_t}$ be independent random variables whose common distribution function F has a derivative $F'(x) = f(x) \neq 0$ at x. Where N_t, for any $t > 0$, be a non-negative integer valued random variable, independent

of the observations.
A kernel estimators of $f(x)$ is defined , in the random case following Parzen [10], by :

$$\widehat{f}_t(x) = (N_t h(N_t))^{-1} \sum_{i=1}^{N_t} K_{h_t}(X_i),$$

where $K_{h_t}(X_i) = K\left(\frac{x-X_i}{h(N_t)}\right)$.
We shall assume in addition that F satisfies a Lipshitz condition of order one in a neighborhood of x; that

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \sup_{m:|m-n| \leqslant n\varepsilon} \left| \frac{h(n)}{h(m)} - 1 \right| = 0, \tag{6}$$

where $h = h(n)$ is a sequence of positive constants converging to zero; and that

$$\log^2 n / nh \log\log n \longrightarrow 0. \tag{7}$$

Conditions (6) and (7) would be satisfied in practice, since it is usual to take $h(n) \sim an^{-b}$ for positive numbers a and b with $b < 1$.
From theorem (1) we obtain

Theorem 2.
Suppose $N_t \xrightarrow{P} \infty$ as $t \to \infty$. Under the conditions above,

$$\limsup_{t \to \infty} \pm \left\{ \frac{N_t h^2(N_t)}{2Eh(N_t) \log\log N_t} \right\}^{1/2} \left[\widehat{f}_t(x) - E\widehat{f}_t(x) \right] = \left[f(x) \int K^2(z)\, dz \right]^{1/2} \quad \text{in probability.}$$

Proof. Let $S_t(x) = \sum_{i=1}^{N_t} \left[K_{h_t}(x, X_i) - EK_{h_t}(x, X_i) \right]$ and $A_{h_t}(x) = N_t \left\{ EK_{h_t}(x, X_1) - h(N_t) E\widehat{f}_t(x) \right\}$. We can write,

$$N_t h(N_t) \left[\widehat{f}_t(x) - E\widehat{f}_t(x) \right] = S_t(x) + A_{h_t}(x)$$

Also we have, using Markov inequality, for all $\varepsilon > 0$,

$$\begin{aligned} P\left((g(N_t))^{-1} |A_{h_t}(x)| > \varepsilon \right) &\leqslant \frac{1}{\varepsilon} E\left[(g(N_t))^{-1} |A_{h_t}(x)| \right] \\ &\leqslant \frac{1}{\varepsilon} \sum_m \sum_n m g^{-1}(m) \left| \left(1 - \frac{h(m)}{h(n)} \right) E\left[K_{h(n)}(x, X_1) \right] \right| \times \\ &\quad P(N_t = n) P(N_t = m). \end{aligned}$$

Where g is the function defined in (2). So that under condition (6),

$$g^{-1}(N_t) A_{h_t}(x) \longrightarrow 0, \quad \text{as } t \to \infty \quad \text{in probability.}$$

Therefore, the law of the Iterated Logarithm is established on $S_t(x)$ process.

Let $\sigma_t^2 = Var\left(K_{h_t}(x, X_1)\right)$. We have (see [6]), that
$EK_{h(n)}(x, X_1) \simeq o(h^{1/2}(n)) \longrightarrow 0$, as $n \to \infty$ and $EK^2_{h(n)}(X) \simeq f(x) h(n) \int K^2(z)\, dz$. It follows from lemma that,

$$EK_{h_t}(x, X_1) = \sum_n E\left[K_{h(n)}(x, X_1) \right] P(N_t = n) \longrightarrow 0, \text{ as } t \to \infty$$

and

$$\begin{aligned}\sigma_t^2 &\simeq EK_{h_t}^2(x,X_1) = \sum_n f(x)h(n)\int K^2(z)\,dz\,P(N_t = n)\\ &= f(x)E[h(N_t)]\int K^2(z)\,dz.\end{aligned}$$

Condition (H2) now follow from (7), condition (H3) follow from (6) and theorem (2) is then established from theorem (1). □

4. Orthogonal Polynomial Estimators

We consider only the case of an estimate based on the Legendre polynomials.
Let $X_1, X_2, \ldots, X_{N_t}$ be independent random variables of distribution with unknown density f having its support confined to $(-1,1)$, and N_t, for any $t > 0$ a non-negative integer-valued random variable, supposed to be independent of observations. Suppose that f is continuous at x, $x \in (-1,1)$ and of bounded variation in a neighborhood of x, and $f(x) \neq 0$. We assume in addition that $(1-y^2)^{-1/4} f(y)$ is integrable on $(-1,1)$.
The orthogonal Legendre system is defined by

$$p_i(z) = \left[\frac{1}{2}(2i+1)\right]^{1/2} q_i(z), \qquad i \geqslant 0, \tag{8}$$

where the functions $q_i(z)$ are the Legendre polynomials. An estimator of $f(x)$ is given by

$$\widehat{f}_t(x) = \sum_{i=0}^{m(t)} \widehat{a}_{i,t}(x) p_i(x),$$

where $m(t) = m(N_t)$ is a sequence of integers tending to infinity, when N_t take a value n, and

$$\widehat{a}_{i,t}(x) = N_t^{-1} \sum_{j=1}^{N_t} p_i(X_j).$$

Assume that,

$$\lim_{\varepsilon\to 0} \limsup_{n\to\infty} \sup_k |m(k)/m(n) - 1| = 0, \tag{9}$$

where the inner supremum is taken over integers k with $|k-n| \leqslant n\varepsilon$, and

$$m^3(\log n)^2/n \log\log n \longrightarrow 0. \tag{10}$$

Theorem 3.
Suppose $N_t \xrightarrow{P} \infty$ *as* $t \to \infty$. *Under the conditions above,*

$$\limsup_{t\to\infty} \pm \left[\frac{N_t}{2E(m(t))\log\log N_t}\right]^{1/2} \left[\widehat{f}_t(x) - E\widehat{f}_t(x)\right] = \left[\frac{f(x)}{\pi(1-x^2)^{1/2}}\right]^{1/2}, \quad \textit{in probability}.$$

Proof. Using the relation (8), we can write

$$\widehat{f}_t(x) = N_t^{-1} \sum_{j=1}^{N_t} K_{m(t)}(x, X_j),$$

where $K_{m(t)}(x,X_j) = \sum_{i=0}^{m(t)} p_i(X_j)p_i(x)$. So, we can easily obtain that,

$$N_t\{\widehat{f}_t(x) - E\widehat{f}_t(x)\} = \sum_{j=1}^{N_t} \left[K_{m(t)}(x,X_j) - E(K_{m(t)}(x,X_j))\right]$$

and the law of the iterated logarithm is established on the random process $\left(N_t\{\widehat{f}_t(x) - E\widehat{f}_t(x)\}\right)_t$.

Let $\sigma_t^2 = Var(K_{m(t)}(x,X))$. We have that $E[K_{m(n)}(x,X)] \longrightarrow 0,$ as $n \to \infty$ and $E[K^2_{m(n)}(x,X)] \simeq m(n)f(x)/\pi(1-x^2)^{1/2}$ (see [6]).
It follows from lemma and the conditions imposed on f that,

$$EK_{m(t)}(x,X) = \sum_n E[K_{m(n)}(x,X)]P(N_t = n) \longrightarrow 0, \text{ as } t \to \infty$$

and

$$\begin{aligned}\sigma_t^2 &\simeq EK^2_{m(t)}(x,X) = \sum_n \left[m(n)f(x)/\pi(1-x^2)^{1/2}\right] P(N_t = n) \\ &= \frac{f(x)}{\pi(1-x^2)^{1/2}} E(m(N_t)).\end{aligned}$$

Condition (H2) now follow from (10), (H3) follow from (9) and the theorem (3) is then established from theorem (1). □

Acknowledgments

We sincerely thank an Associate Editor and the anonymous referee

References

1. P. Deheuvels, “ Sur l’estimation séquentielle de la densité, ”*C.R. Acad. Sci., Paris* , **76** sér. A-B, 1119–1121 (1973).
2. P. Deheuvels, “ Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité, *C.R. Acad. Sci., Paris*, **278**, sér. A. 1217–1220 (1974).
3. B.L.S. Prakasa Rao, “ Non parametric Functional Estimation, Probability and Mathematical Statistics”. *A serie of Monographs and Textbooks*. **72** PRA 83a. (1983).
4. R.C. Srivastava, “ Estimation of Probability Density Function based on Random Number of Observations with Applications”, *Int. Stat. Rev.*, **41** N1, 77–86 (1973).
5. R. J. Carroll, “ On Sequential Density Estimation”, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **36**, 137–151 (1976).
6. P. Hall, “ Laws of the Iterated Logarithm for Nonparametric Density Estimators”, *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **56**, 47–61 (1981).
7. W. Stute, “ A Law of the logarithm for kernel density estimators”, *The annals of Probability*, **10** N2, 414–422 (1982).
8. P. Deheuvels, “ Laws of iterated logarithm for density estimators”, *Nonparametric Functional Estimation and Related Topics*, G. Roussa Edit., 19–29 (1991).
9. J. Komlós, P. Major, and G. Tusnady, “ An approximation of partial sums of independent rv’s, and the sample” *Wahrscheinlichkeitstheorie verw. Gebiete*, **32**, 111–131 (1975).
10. E. Parzen, “ On estimation of a probability density and mode”, *Ann. Math. Statist.*, **33**, 1065–1076 (1962).