Mathematics and Statistics 12(2): 142-156, 2024 http://www.hrpub.org
DOI: 10.13189/ms.2024.120204

Recursive Estimation of the Multidimensional Distribution Function Using Bernstein Polynomial

D. A. N. Njamen[1,*], **B. Baldagaï**[1], **G. T. Nguefack**[2], **A. Y. Nana**[3]

[1]Department of Mathematics and Computer Sciences, Faculty of Sciences, University of Maroua, Cameroon

[2]Department of Public Health, Faculty of Medicine and Biomedical Sciences, University of Yaoundé I, Cameroon

[3]Laboratoire de Mathématiques et Informatique, Université de Douala, Cameroon

Received November 28, 2022; Revised January 14, 2024; Accepted February 22, 2024

Cite This Paper in the following Citation Styles

***(a)**: [1] D. A. N. Njamen, B. Baldagaï, G. T. Nguefack, A. Y. Nana, "Recursive Estimation of the Multidimensional Distribution Function Using Bernstein Polynomial," Mathematics and Statistics, Vol.12, No.2, pp. 142-156, 2024. DOI: 10.13189/ms.2024.120204*

***(b)**: D. A. N. Njamen, B. Baldagaï, G. T. Nguefack, A. Y. Nana (2024). Recursive Estimation of the Multidimensional Distribution Function Using Bernstein Polynomial. Mathematics and Statistics, 12(2), 142-156. DOI: 10.13189/ms.2024.120204*

Abstract The recursive method known as the stochastic approximation method, can be used among other things, for constructing recursive nonparametric estimators. Its aim is to ease the updating of the estimator when moving from a sample of size n to $n+1$. Some authors have used it to estimate the density and distribution functions, as well as univariate regression using Bernstein's polynomials. In this paper, we propose a nonparametric approach to the multidimensional recursive estimators of the distribution function using Bernstein's polynomial by the stochastic approximation method. We determine an asymptotic expression for the first two moments of our estimator of the distribution function, and then give some of its properties, such as first- and second-order moments, the bias, the mean square error (MSE), and the integrated mean square error ($IMSE$). We also determine the optimal choice of parameters for which the MSE is minimal. Numerical simulations are carried out and show that, under certain conditions, the estimator obtained converges to the usual laws and is faster than other methods in the case of distribution function. However, there is still a lot of work to be done on this issue. These include the studies of the convergence properties of the proposed estimator and also the estimation of the recursive regression function; the developments of a new estimator based on Bernstein polynomials of a regression function using the semi-recursive estimation method; and also a new recursive estimator of the distribution function, density and regression functions; when the variables are dependent.

Keywords Nonparametric Estimation, Stochastic Approximation Method, Multidimensional Distribution Function, Recursive Estimator, Multidimensional Bernstein Polynomial

MSC 2020 : 62G05, 62L20, 62H10, 03D20, 65D15.

1 Introduction

Estimation theory is one of the most basic branches of statistics. Indeed, in various fields such as medicine, economics, finance, marketing, etc., complex data arise and their analysis must be modeled to better understand them. These data are often described by unknown real random variables. Starting from the observation of such a phenomenon, one would like to have a precise idea of its magnitude. Distribution and density functions are essential concepts in statistics that can better answer this concern, because they allow to describe and characterize the distribution of a random variable.

This theory is usually divided into two main components, namely, parametric and non-parametric estimation. Non-parametric estimation consists, in most cases, in estimating an unknown function from the observations, element of a certain functional class. A non-parametric procedure is defined independently of the distribution or law of the sample of observations. More particularly, we speak of a non-parametric estimation method when it does not boil down to the estimation of a finite number of real parameters associated with the law of the sample.

In fact, for X_1, $X_2, \ldots$ a sequence of random variables independent and identically distributed (i.i.d) with a common distribution function F with density f having a bounded support; without loss of generality, let's consider this as $[0;1]^d$. Knowing that F is continuous, it is natural to estimate F using smooth functions instead of the empirical distribution function which does not continue. Several questions are worth raising: what methods should be used to best estimate the data in a practical context ? For a given method, how to choose different parameters so that the estimator derived from them best describes the data to be modeled ? One way to do this is to use Bernstein's polynomial approximations [1]. This method is particularly interesting because Bernstein polynomials [1] are known to give very smooth estimates that generally have acceptable behavior at the boundaries. This approach using Bernstein's polynomials has been widely discussed in several works such as Tenbusch [2], Leblanc ([3], [4]), Babu et al. [5], Babu and Chaubey [6].

In this paper, we use the stochastic approximation method for finding the zero of an unknown function. The algorithm used is the one introduced by Robbins and Monro [7] and modified by Mokkadem et al. [8] where we integrate the generalized Bernstein polynomial in order to obtain a recursive estimator of the multidimensional distribution function to better address the problem of estimation in the neighborhood of the boundary when the estimation interval is bounded on at least one side. Indeed, the idea is to use the computation of estimates on the basis of the initial data and to update them by taking into account only the new data entering the sample database. In this case, the benefit in terms of calculation can be very interesting. The advantages of recursive methods are numerous (see Nguyen and Saracco, [9]) :

- Taking into account the temporal arrival of information and thus refining the estimation algorithms implemented over time ;
- The benefit in terms of calculation can be very interesting and its applications are numerous;
- It is not necessary to rerun all the calculations of the model's parameter estimates each time the database is completed with new observations.

Recursive estimation methods have never been developed in the framework of the non-parametric multidimensional model. This article is divided as follows: a review of the basic notions of the Bernstein polynomial followed by the presentation of the distribution function using the Bernstein polynomial and their properties. Then, the recursive estimator of the multidimensional distribution function using the generalized Bernstein polynomial is presented as well as some of its properties. Finally, numerical simulations are performed to compare our estimator with that of Vitale [10].

2 Review of the basic concepts

In 1912, while searching for a constructive and probabilistic proof of the classical Weierstrass theorem for the approximation of continuous functions on intervals of type $[a, b]$, Bernstein established a family of polynomials that bears his name:

2.1 Bernstein's polynomial

Bernstein [10] introduced a family of polynomials, which would later bear his name, whose definition is as follows.

Definition 2.1.1 *(Bernstein's polynomial, [1])*
For all $m \in \mathbb{N}$ and $0 \leq k \leq m$, a Bernstein's polynomial is defined as :

$$b_k(m,x) = C_m^k x^k (1-x)^{m-k}.$$

These polynomials have several interesting properties in the field of probability and statistics.

Proposition 2.1.2 *[1]*
Bernstein polynomials have the following properties :

1. *Unit partition :*

$$\sum_{k=0}^{m} b_k(m,x) = 1,\ \forall x \in [0,1].$$

2. *Positivity :*

$$\forall k \in \{0,...,m\},\ b_k(m,x) \geq 0,$$

and $b_k(m,x)$ becomes null only in 0 and 1.

3. *Symetry :*

$$\forall k \in \{0,...,m\},\ b_k(m,x) = b_{m-k}(m,1-x).$$

4. *Recurrence formula : For $m > 0$,*

$$b_k(m,x) = \begin{cases} (1-x)b_k(m-1,x) & if \ \ k=0 \\ (1-x)b_k(m-1,x) + xb_{k-1}(m-1,x) \\ \forall k \in \{1,...,m-1\} \\ xb_{k-1}(m-1,x) & if \ k=m. \end{cases}$$

From a probabilistic point of view, one can view the polynomial $b_k(m,p)$ as the probability $\mathbb{P}(X=k)$, where X is a random variable following a binomial distribution with parameters (m,p).

Theorem 2.1.3 *(Bernstein, [1])*
Let $f : [0,1] \to \mathbb{R}$ be a continuous function. We define the Bernstein's polynomial associated to f with order $m \in \mathbb{N}$:

$$\forall x \in [0,1],\ B_m(f)(x) = \sum_{k=0}^{m} f\left(\frac{k}{m}\right) C_m^k x^k (1-x)^{m-k}.$$

Then

$$\lim_{m\longrightarrow+\infty} ||f - B_m(f)|| = \lim_{m\longrightarrow+\infty} \sup_{x\in[0,1]} ||f(x) - B_m(f)(x)|| = 0.$$

In particular, any continuous function on $[0,1]$ is the uniform limit of a sequence of Bernstein polynomials.

3 Distribution function using the Bernstein polynomial

Let $X_1, ..., X_n$ be iid random variables with unknown distribution function F and associated density f with support $[0,1]$.

Definition 3.0.1 *[10]*
The estimator of F is given by:

$$\hat{F}_{n,m}(x) = \sum_{k=0}^{m} F_n\left(\frac{k}{m}\right) b_k(m,x), \tag{1}$$

where $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i \leq t\}}$ is the empirical distribution function.

Remark 3.0.2 *We note that: $\hat{F}_{n,m}(0) = 0 = F(0)$ and $\hat{F}_{n,m}(1) = 1 = F(1)$.*

This estimator defined in (1) satisfies the properties of a distribution function.

Proposition 3.0.3 *[11]*

i) $\hat{F}_{n,m}$ is continuous on $[0,1]$;

ii) $0 \leq \hat{F}_{n,m}(x) \leq 1$ for all $x \in [0,1]$;

iii) $\hat{F}_{n,m}$ is increasing.

The following theorem provides the expectation and variance of the estimator $\hat{F}_{n,m}$.

Theorem 3.0.4 *[11]*
Let assume that F is a continuous function and has two bounded derivatives on $[0,1]$. For any $x \in [0,1]$, we have:

$$\mathbb{E}[\hat{F}_{n,m}(x)] = F(x) + \frac{1}{m}b(x) + o\left(\frac{1}{m}\right), \tag{2}$$

with $b(x) = \frac{x(1-x)f'(x)}{2}$;
and

$$\mathbb{V}ar[\hat{F}_{n,m}(x)] = \frac{1}{n}\sigma^2(x) - \frac{1}{n\sqrt{m}}\mathbb{V}(x) + o_x\left(\frac{1}{n\sqrt{m}}\right), \tag{3}$$

with $\sigma^2(x) = F(x)[1-F(x)]$ and $\mathbb{V}(x) = f(x)\left[\frac{2x(1-x)}{\pi}\right]$.

Remark 3.0.5 *[9]*
$\hat{F}_{n,m}$ asymptotically dominates the empirical distribution function F_n at each $x \in]0,1[$ in terms of mean square error (MSE).
In fact, we have :

$$\begin{aligned} MSE[\hat{F}_{n,m}(x)] &= \frac{1}{n}\sigma^2(x) - \frac{1}{n\sqrt{m}}\mathbb{V}(x) + \frac{1}{m^2}b(x) \\ &+ o_x\left(\frac{1}{n\sqrt{m}}\right) + o\left(\frac{1}{m^2}\right). \end{aligned} \tag{4}$$

By contrast, it is well known that

$$MSE[F_n(x)] = \mathbb{V}ar[F_n(x)] = \frac{1}{n}\sigma^2(x),$$

so that $\hat{F}_{n,m}$ and F_n are equivalent in terms of MSE till the first order.

However, considering also higher order terms, it turns out that $\hat{F}_{n,m}$ has smaller MSE than that of F_n when the order m is well chosen. Both estimators achieve a MSE equal to zero for $x = 0, 1$.

By integrating (4), we can obtain the mean integrated square error ($MISE$) of the estimator $\hat{F}_{n,m}$. We then have :

$$\begin{aligned} MISE[\hat{F}_{n,m}(x)] &= \frac{1}{n}C_1 + -\frac{1}{n\sqrt{m}}C_2 + \frac{1}{m^2}C_3 \\ &+ o\left(\frac{1}{n\sqrt{m}}\right) + o\left(\frac{1}{m^2}\right), \end{aligned}$$

with $C_1 = \int_0^1 \sigma^2(x)dx$, $C_2 = \int_0^1 \mathbb{V}(x)dx$, and $C_3 = \int_0^1 b^2(x)dx$.

Remark 3.0.6 *[9]*
The constants C_1, C_2 and C_3 are all strictly positive, except in the trivial case where f is the uniform density, in which case $C_3 = 0$. The optimal choice of the order m that minimizes the $MISE$ is given by

$$m_{opt} = \left[\frac{4C_3}{C_2}\right]^{\frac{2}{3}} n^{\frac{2}{3}};$$

in this case, we have

$$MISE[\hat{F}_{n,m_{opt}}(x)] = \frac{C_1}{n} - \frac{3}{4\sqrt[3]{n^4}}\left[\frac{C_2^{\frac{4}{3}}}{\sqrt[3]{4C_3}}\right] + o\left(\frac{1}{\sqrt[3]{n^4}}\right).$$

Estimator $\hat{F}_{n,m}$ is strongly consistent. Indeed, we have the following result:

Theorem 3.0.7 *[5]*
Let F be a continuous distribution function on $[0,1]$.
If $m, n \to +\infty$, then

$$||\hat{F}_{n,m} - F|| = \sup_{x\in[0,1]} |\hat{F}_{n,m}(x) - F(x)| \longrightarrow 0, \ p.s.$$

The following theorem gives the distance between the estimator $\hat{F}_{n,m}$ and the empirical distribution function F_n.

Theorem 3.0.8 *[5]*
Let F be a continuous distribution function, differentiable on the interval $[0,1]$ and of density f. If f is lipschitzian of order 1, then for $\sqrt[3]{n^2} \leq m \leq (\frac{n}{\ln n})^2$, we have:

$$||\hat{F}_{n,m} - F|| = o\left[\left(\sqrt{\frac{\ln n}{n}}\right)\left(\frac{\ln m}{m}\right)^{\frac{1}{4}}\right].$$

Remark 3.0.9 *This result essentially establishes that when the order m of the Vitale estimator is chosen large enough (so that the bias becomes negligible), the asymptotic distribution of $\hat{F}_{n,m}$ is the same as that of the empirical distribution F_n.*

Definition 3.0.10 *Let $\gamma \in \mathbb{R}$ and $(v_n)_{n\geq 1}$ a positive real sequence. We say that $(v_n) \in \mathcal{GS}(\gamma)$ if*

$$\lim_{n\longrightarrow+\infty} n\left(1 - \frac{v_{n-1}}{v_n}\right) = \gamma. \tag{5}$$

Condition (5) has been introduced by Galambos and Seneta [12] to define sequences with regular variations.

Lemma 3.0.11 *[8]*
Let $(v_n) \in \mathcal{GS}(v^*)$, $(\gamma_n) \in \mathcal{GS}(-\alpha)$, *and* $m > 0$ *such that* $m - v^*\xi > 0$ *where* $\xi = \lim_{n \longrightarrow +\infty}(n\gamma_n)^{-1}$. *We have:*

$$\lim_{n \longrightarrow +\infty} v_n \Pi_n^m \sum_{k=1}^{n} \Pi_k^{-m} \frac{\gamma_k}{v_k} = \frac{1}{m - v^*\xi}. \tag{6}$$

Moreover, for any sequence (α_n) *such that* $\lim_{n \longrightarrow +\infty} \alpha_n = 0$, *and any* $\delta \in \mathbb{R}$,

$$\lim_{n \longrightarrow +\infty} v_n \Pi_n^m \left[\sum_{k=1}^{n} \Pi_k^{-m} \frac{\gamma_k}{v_k} \alpha_k + \delta \right] = 0. \tag{7}$$

4 Recursive estimation of the multidimensional distribution function using the generalized Bernstein polynomial

In this section, we propose a recursive estimator of the multidimensional distribution function using Bernstein polynomial. To achieve this, we use the stochastic approximation method. Indeed, the most famous use of stochastic algorithms in the framework of non-parametric statistics is the work of Kiefer and Wolfowitz [13]. Their work has been improved by several authors (see Blum [14], Chen [15], Dippon [16], Mokkadem and Pelletier [17], Robbins and Monro [7]).
In the following subsection, we present the stochastic approximation algorithm of Robbins and Monro [7].

4.1 Stochastic approximation method

The stochastic approximation method allows the search for the zero θ^* of an unknown function $h : \mathbb{R} \longrightarrow \mathbb{R}$ which can hardly be calculated directly. The best known algorithm is that of Robbins and Monro [7]. We proceed as follows :

i) We choose $\theta_0 \in \mathbb{R}$.

ii) For $n \geq 1$, we construct the sequence (θ_n) by the recursive relation
$$\theta_n = \theta_{n-1} + \gamma_n W_n,$$
where W_n is an observation of the function h at θ_{n-1} and (γ_n) is a sequence of positive real numbers tending to zero, called step of the algorithm.

4.2 Generalization of the Bernstein polynomial in d dimension

Let $m_n \in \mathbb{N}$; $K = (k_1, ..., k_d) \in \mathbb{N}^d$ with $0 \leq k_i \leq m_n$ and $x = (x_1, ..., x_d) \in [0,1]^d$. We can define the Bernstein polynomial of degree m_n by:

$$B_K(m_n, x) = \prod_{i=1}^{d} b_{k_i}(m_n, x_i), \tag{8}$$

with $b_{k_i}(m_n, x_i) = C_{m_n}^{k_i} x_i^{k_i}(1 - x_i)^{m_n - k_i}$.
In the field of probability, $B_K(m_n, x)$ is a binomial distribution with parameters (m_n, x).

4.3 Recursive method in the multidimensional framework

Let $X_1, ..., X_n$ (with $n \in \mathbb{N}$) be i.i.d random variables taking values in $\mathbb{R}^d$ with unknown distribution function F associated to a density function f having support in $[0,1]^d$ with $d \in \mathbb{N}$.
To construct an estimator of F at the point $x \in [0,1]^d$, we will use the stochastic approximation algorithm of Robbins and Monro [7].
We proceed as follows to define the algorithm for finding the zero of the function $h : y \longmapsto F(x) - y$:

i) $\hat{F}_{0,m_0}(x) \in \mathbb{R}$ fixed;

ii) $\forall n \geq 1$, we have:
$$\hat{F}_{n,m_n}(x) = \hat{F}_{n-1,m_{n-1}}(x) + \gamma_n W_{n,m_n}(x),$$
where
- (γ_n) is a numerical sequence designating the step of the algorithm such that $\lim_{n \longrightarrow +\infty} \gamma_n = 0$;
- $W_{n,m_n}(x)$ is the image by h of the point $\hat{F}_{n-1,m_{n-1}}(x)$;
- m_n is the width of the band.

4.4 Application of the recursive stochastic algorithm method to the estimation of the multidimensional distribution function using the Bernstein polynomial

Considering all the above, we can estimate $F(x)$ by $\forall x \in [0,1]^d$,

$$Z_{n,m_n}(x) = \sum_{K=0}^{m_n} F_n\left(\frac{K}{m_n}\right) B_K(m_n, x), \tag{9}$$

with $\sum_{K=0}^{m_n} = \sum_{k_1=0}^{m_n} \sum_{k_2=0}^{m_n} \cdots \sum_{k_d=0}^{m_n}$; $F_n\left(\frac{K}{m_n}\right) = \prod_{i=1}^{d} F_n\left(\frac{k_i}{m_n}\right)$ and where F_n is the empirical distribution function and $B_K(m_n, x)$ is the Bernstein polynomial of degree m_n defined in (8).
We can make the following observation:

Remark 4.4.1 $Z_{n,m_n}(x)$ *defined in (9) can be re-written as follows:*

$$Z_{n,m_n}(x) = \prod_{i=1}^{d} \left[\sum_{k_i=0}^{m_n} F_n\left(\frac{k_i}{m_n}\right) b_{k_i}(m_n, x_i) \right]. \tag{10}$$

In fact,

$$\begin{aligned}
Z_{n,m_n}(x) &= \sum_{k_1=0}^{m_n} \cdots \sum_{k_d=0}^{m_n} \left[\prod_{i=1}^{d} F_n\left(\frac{k_i}{m_n}\right) \right] \left[\prod_{i=1}^{d} b_{k_i}(m_n, x_i) \right] \\
&= \sum_{k_1=0}^{m_n} \cdots \sum_{k_d=0}^{m_n} \left[\prod_{i=1}^{d} F_n\left(\frac{k_i}{m_n}\right) b_{k_i}(m_n, x_i) \right] \\
&= \left(\sum_{k_1=0}^{m_n} F_n\left(\frac{k_1}{m_n}\right) b_{k_1}(m_n, x_1) \right) \times \cdots \times \\
&\quad \left(\sum_{k_d=0}^{m_n} F_n\left(\frac{k_d}{m_n}\right) b_{k_d}(m_n, x_d) \right) \\
&= \prod_{i=1}^{d} \left[\sum_{k_i=0}^{m_n} F_n\left(\frac{k_i}{m_n}\right) b_{k_i}(m_n, x_i) \right].
\end{aligned}$$

We can therefore deduce that

$$W_{n,m_n}(x) = Z_{n,m_n}(x) - \hat{F}_{n-1,m_{n-1}}(x).$$

Thus, in a recursive manner, the estimator of F at point x can be defined by:

$$\hat{F}_{n,m_n}(x) = (1-\gamma_n)\hat{F}_{n-1,m_{n-1}}(x) + \gamma_n Z_{n,m_n}(x). \quad (11)$$

In the following section, we give some properties of the Bernstein polynomial as well as the resulting estimator.

4.5 Fundamental results

The polynomial defined in (8) satisfies the following properties and is the first result of this paper.

Proposition 4.5.1 *Bernstein polynomials have the following properties:*

1. *Unit partition*

$$\sum_{K=0}^{m_n} B_K(m_n,x) = 1, \ \forall\, x \in [0,1]^d,$$

where $\sum_{K=0}^{m_n} = \sum_{k_1=0}^{m_n}\sum_{k_2=0}^{m_n} \cdots \sum_{k_d=0}^{m_n}$.

2. *Positivity*

$$B_K(m_n,x) \geq 0,$$

with $K = (k_i)_{i\in\{1,\,\cdots,\,d\}}$ *such that* $k_i \in \{0, \cdots, m_n\}$, *and* $B_K(m_n,x)$ *is only nullified in* 0 *and* 1.

3. *Symetric*

$$B_K(m_n,x) = B_{m_n-K}(m_n,1-x),$$

with $K = (k_i)_{i\in\{1,\cdots,d\}}$ *such that* $k_i \in \{0, \cdots, m_n\}$.

4. *Recurrence formula : For* $m_n > 0$ *and* $K = (k_i)_{i\in\{1,\cdots,d\}}$ *such that* $k_i \in \{0, \cdots, m_n\}$, *we have :*

$$B_K(m_n,x) = \begin{cases} \prod_{i=1}^{d}(1-x_i)B_K(m_n-1,x_i) \ if \ K=0 \\ \prod_{i=1}^{d}(1-x_i)B_{k_i}(m_n-1,x_i)+ \\ x_iB_{k_i-1}(m_n-1,x_i) \ if \ K\in\{1, \cdots, m_n-1\} \\ \prod_{i=1}^{d} x_iB_{K-1}(m_n-1,x_i) \ if \ K=m_n. \end{cases}$$

Proof

1. Let's show that $\sum_{K=0}^{m_n} B_K(m_n,x) = 1$.
Let's compute the quantity $\sum_{K=0}^{m_n} B_K(m_n,x)$. We have:

$$\begin{aligned} \Sigma_{K=0}^{m_n} B_K(m_n,x) &= \Sigma_{k_1=0}^{m_n}\Sigma_{k_2=0}^{m_n}\cdots\Sigma_{k_d=0}^{m_n}\left(\Pi_{i=1}^{d} b_{k_i}(m_n,x_i)\right) \\ &= \Sigma_{k_1=0}^{m_n}\Sigma_{k_2=0}^{m_n}\cdots\Sigma_{k_d=0}^{m_n}\left(b_{k_1}(m_n,x_1)\times\cdots\times b_{k_d}(m_n,x_d)\right) \\ &= \left(\Sigma_{k_1=0}^{m_n} b_{k_1}(m_n,x_1)\right)\times\cdots\times\left(\Sigma_{k_d=0}^{m_n} b_{k_d}(m_n,x_d)\right) \\ &= 1\times\cdots\times 1 \\ &= 1. \end{aligned}$$

2. Let's show that $B_K(m_n,x) \geq 0$ and deduce that $B_K(m_n,x) = 0$ for $x=0$ and $x=1$.

By definition, $B_K(m_n,x) = \prod_{i=1}^{d} b_{k_i}(m_n,x_i)$; however for any $k_i \in \{0, \cdots, m_n\}$, we have $b_{k_i}(m_n,x_i) \geq 0$; thus $\prod_{i=1}^{d} b_{k_i}(m_n,x_i) \geq 0$; therefore $B_K(m_n,x) \geq 0$.
In addition:

− for $x = 1 = (1, \cdots 1)$, we have $B_K(m_n,x) = b_{k_1}(m_n,1) \times ... \times b_{k_d}(m_n,1) = 0$, because for any $i \in \{1,...,d\},\ b_{k_1}(m_n,1) = 0$.

− for $x = 0 = (0,...,0)$, we have $B_K(m_n,x) = 0$, since $\forall i \in \{1,...,d\},\ b_{k_1}(m_n,0) = 0$.

3. Let's show $B_K(m_n,x) = B_{m_n-K}(m_n,1-x)$.
We have

$$\begin{aligned} B_K(m_n,x) &= \prod_{i=1}^{d} b_{k_i}(m_n,x_i) \\ &= b_{k_1}(m_n,x_1)\times ...\times b_{k_d}(m_n,x_d). \end{aligned}$$

However, according to point 3 of Proposition 2.1.2,

$$\forall\, k_i \in \{1,...,m_n\},\ b_{k_i}(m_n,x_i) = b_{m_n-k_i}(m_n,1-x_i);$$

therefore

$$\begin{aligned} B_K(m_n,x) &= b_{m_n-k_1}(m_n,1-x_1)\times ...\times b_{m_n-k_d}(m_n,1-x_d) \\ &= \prod_{i=1}^{d} b_{m_n-k_i}(m_n,1-x_i) \\ &= B_{m_n-K}(m_n,1-x). \end{aligned}$$

4. Let's show that for $m_n > 0$ and $K = (k_i)_{i\in\{1,...,d\}}$ such that $k_i \in \{0,...,m_n\}$;

$$B_K(m_n,x) = \begin{cases} \prod_{i=1}^{d}(1-x_i)B_K(m_n-1,x) \ if \ K=0; \\ \prod_{i=1}^{d}\Big[(1-x_i)B_{k_i}(m_n-1,x_i) \\ +x_iB_{k_i-1}(m_n-1,x_i)\Big] \ if \ K\in\{1,...,m_n-1\}; \\ \prod_{i=1}^{d} x_iB_{K-1}(m_n-1,x) \ if \ K=m_n \end{cases}.$$

$(*)$ By definition, $B_K(m_n,x) = \prod_{i=1}^{d} b_{k_i}(m_n,x_i)$.
However for $k_i = 0$, according to Proposition 1.2.1,

$$b_{k_i}(m_n,x_i) = (1-x_i)b_{k_i}(m_n-1,x_i);$$

thus,

$$\begin{aligned} B_K(m_n,x) &= \prod_{i=1}^{d} b_{k_i}(m_n,x_i) \\ &= \prod_{i=1}^{d}(1-x_i)b_{k_i}(m_n-1,x_i) \\ &= (1-x_1)b_{k_1}(m_n-1,x_1)\times ...\times (1-x_d)b_{k_d}(m_n-1,x_d) \\ &= \prod_{i=1}^{d}(1-x_i)\prod_{i=1}^{d} b_{k_i}(m_n-1,x_i) \\ &= \prod_{i=1}^{d}(1-x_i)B_K(m_n-1,x); \end{aligned}$$

therefore $B_K(m_n,x) = \prod_{i=1}^{d}(1-x_i)B_K(m_n-1,x)$ for $K=0$.
$(**)$ By definition, $B_K(m_n,x) = \prod_{i=1}^{d} b_{k_i}(m_n,x_i)$.
However, according to Proposition 1.2.1, and for all $i \in \{1, \cdots, d\}$,

$$For\ k_i = 1,\ \ b_{k_i}(m_n, x_i) = (1 - x_i)B_{k_i}(m_n - 1, x_i) + x_i B_{k_i - 1}(m_n - 1, x_i)$$
$$For\ k_i = 2,\ \ b_{k_i}(m_n, x_i) = (1 - x_i)B_{k_i}(m_n - 1, x_i) + x_i B_{k_i - 1}(m_n - 1, x_i)$$
$$\vdots$$
$$For\ k_i = m_n - 1,\ \ b_{k_i}(m_n, x_i) = (1 - x_i)B_{k_i}(m_n - 1, x_i) + x_i B_{k_i - 1}(m_n - 1, x_i).$$

We obtain the following iteration by doing the member-by-member product :

$$B_K(m_n, x) = \prod_{i=1}^{d} \Big[(1 - x_i)B_{k_i}(m_n - 1, x_i) + x_i B_{k_i - 1}(m_n - 1, x_i)\Big].$$

$(***)$ By definition, $B_K(m_n, x) = \prod_{i=1}^{d} b_{k_i}(m_n, x_i)$. However for $k_i = m_n$, according to Proposition 1.2.1

$$b_{k_i}(m_n, x_i) = x_i b_{k_i - 1}(m_n - 1, x_i);$$

thus,

$$\begin{aligned} B_K(m_n, x) &= \prod_{i=1}^{d} b_{k_i}(m_n, x_i) \\ &= \prod_{i=1}^{d} x_i b_{k_i}(m_n - 1, x_i) \\ &= x_1 b_{k_1 - 1}(m_n - 1, x_1) \times ... \times x_d b_{k_d - 1}(m_n - 1, x_d) \\ &= \left(\prod_{i=1}^{d} x_i\right) \prod_{i=1}^{d} b_{k_i - 1}(m_n - 1, x_i) \\ &= \prod_{i=1}^{d} x_i B_{K-1}(m_n - 1, x); \end{aligned}$$

therefore $B_K(m_n, x) = \prod_{i=1}^{d} x_i B_{K-1}(m_n - 1, x)$ for $K = m_n$.

■

For the development of our results, we will need the following assumptions:

A1) $\forall i, j \in \{1, ..., d\}, F_{i,j} = \frac{\partial^2 F}{\partial_i \partial_j}(x)$ exists and is finite;

A2) $(\gamma_n) \in \mathcal{GS}(-\alpha)$ with $\alpha \in \left]\frac{1}{2}, 1\right]$;

A3) $(m_n) \in \mathcal{GS}(a)$ with $a \in]0, \frac{\alpha}{d}[$;

A4) $\lim_{n \longrightarrow +\infty}(n\gamma_n) \in \left]\min\left\{ad, \frac{2\alpha + ad}{4}\right\}, +\infty\right[$.

Remark 4.5.2 *a) Assumption $(A1)$ is standard for the bias and variance of the estimator of the distribution using the Bernstein polynomial [9].*

b) Assumption $(A2)$ on (γ_n) is used in a similar way to that of Mokkadem et al. [8].

c) Assumption $(A3)$ on (m_n) is introduced in a similar way to the assumption on the smooth parameter width used for the recursive estimator of the kernel distribution (See Slaoui [18, 19]).

d) Assumption $(A4)$ on the limit of $(n\gamma_n)_{n \in \mathbb{N}}$ when n tends to infinity is usually used in stochastic approximation algorithms.

We will use the following notations throughout the rest of the work:

$$\Pi_n = \prod_{j=1}^{n}(1 - \gamma_j),\ \xi = \lim_{n \longrightarrow +\infty}(n\gamma_n)^{-1}, \qquad (i)$$

$$B(x) = \prod_{i=1}^{d} b(x_i) = \prod_{i=1}^{d} \frac{x_i(1 - x_i)F_{(1)i,i}(x)}{2}, \qquad (ii)$$

$$\sigma^2(x) = \prod_{i=1}^{d} \sigma^2(x_i) = \prod_{i=1}^{d} F_{(1)}(x_i)[1 - F_{(1)}(x_i)], \quad (iii)$$

$$\mathbb{V}(x) = \prod_{i=1}^{d} v(x_i) = \prod_{i=1}^{d} f_{(1)}(x_i)\left[\frac{2x_i(1 - x_i)}{\pi}\right]^{\frac{1}{2}}, \quad (iv)$$

$$C_1 = \int_{[0,1]^d} \sigma^2(x)dx,\ \ C_2 = \int_{[0,1]^d} \mathbb{V}(x)dx,\ \ C_3 = \int_{[0,1]^d} B^2(x)dx. \ (v)$$

To provide the characteristics of the estimator $\hat{F}_{n,m_n}(x)$, we consider the class of sequences with regular variations given by Galambos and Seneta [12]. The following result gives the bias and variance of the estimator defined in (11) and is the second result of this paper.

Theorem 4.5.3 *Under Assumptions $(A1)-(A4)$, the bias and variance of the estimator (11) are given respectively by :*

1.

$$\mathbb{E}\left[\hat{F}_{n,m_n}(x)\right] - F(x) = \begin{cases} m_n^{-d}\frac{1}{(1 - ad\xi)}B(x) + o\left(m_n^{-d}\right) & \text{if } 0 < a \leq \frac{2}{3d}\alpha \\ o\left(\sqrt{\frac{\gamma_n}{\sqrt{m_n^d}}}\right) & \text{if } \frac{2}{3d}\alpha < a \leq \frac{\alpha}{d}, \end{cases} \qquad (12)$$

with $B(x)$ defined in (ii).

2.

$$\mathbb{V}ar\left[\hat{F}_{n,m_n}(x)\right] = \begin{cases} n^{-d}\gamma_n\frac{\sigma^2(x)}{2 - (\alpha + d)\xi} - \frac{\gamma_n}{n^d\sqrt{m_n^d}}\frac{V(x)}{2 - (d + \alpha + a)\xi} + o\left(\frac{\gamma_n}{n^d\sqrt{m_n^d}}\right) & \text{if } \frac{2}{3d}\alpha \leq a \leq \frac{\alpha}{d} \\ n^{-d}\gamma_n\frac{1}{2 - (\alpha + d)\xi}\sigma^2(x) + o\left(n^{-d}\gamma_n\right) & \text{if } \frac{1}{2d}\alpha \leq a < \frac{2}{3d}\alpha \\ o\left(\frac{1}{m_n^d}\right) & \text{if } 0 < a \leq \frac{1}{2d}\alpha, \end{cases} \qquad (13)$$

with $\sigma^2(x)$ and $\mathbb{V}(x)$ defined respectively in (iii) and (iv).

3. If $\lim_{n \longrightarrow \infty}(n\gamma_n) > \max\left(ad, \frac{2\alpha + ad}{4}\right)$, then the first condition of the bias (12) and the first condition of the variance (13) are simultaneously verified.

Proof - Let's compute the quantity $\hat{F}_{n,m_n}(x) - F(x)$. We have :

$$\begin{aligned}\hat{F}_{n,m_n}(x) - F(x) &= (1-\gamma_n)\hat{F}_{n-1,m_{n-1}}(x) + \\ &\quad \gamma_n Z_{n,m_n}(x) - F(x) \\ &= (1-\gamma_n)\hat{F}_{n-1,m_{n-1}}(x) + \gamma_n Z_{n,m_n}(x) - \\ &\quad F(x) + \gamma_n F(x) - \gamma_n F(x) \\ &= (1-\gamma_n)(\hat{F}_{n-1,m_{n-1}}(x) - F(x)) + \\ &\quad \gamma_n (Z_{n,m_n}(x) - F(x)).\end{aligned}$$

Likewise,

$$\begin{aligned}\hat{F}_{n-1,m_{n-1}}(x) - F(x) &= (1-\gamma_{n-1})(\hat{F}_{n-2,m_{n-2}}(x) - F(x)) + \\ &\quad \gamma_{n-1}(Z_{n-1,m_{n-1}}(x) - F(x)). \qquad (14)\end{aligned}$$

In addition,

$$\begin{aligned}\hat{F}_{n-2,m_{n-2}}(x) - F(x) &= (1-\gamma_{n-2})(\hat{F}_{n-3,m_{n-3}}(x) - F(x)) + \\ &\quad \gamma_{n-2}(Z_{n-2,m_{n-2}}(x) - F(x)). \qquad (15)\end{aligned}$$

Replacing (15) in (14), we have :

$$\begin{aligned}\hat{F}_{n,m_n}(x) - F(x) &= \prod_{j=n-2}^{n}(1-\gamma_j)(\hat{F}_{n-3,m_{n-3}}(x) - F(x)) \\ &+ \prod_{j=n-2}^{n}(1-\gamma_j)\sum_{l=n-2}^{n}\Pi_l^{-1}\gamma_l(Z_{l,m_l}(x) - F(x)),\end{aligned}$$

with $\Pi_l = \prod_{j=n-2}^{l}(1-\gamma_j)$.
From one to the next, we can generalize as follows:

$$\begin{aligned}\hat{F}_{n,m_n}(x) - F(x) &= \prod_{j=1}^{n}(1-\gamma_j)\sum_{l=1}^{n}\Pi_l^{-1}\gamma_l(Z_{l,m_l}(x) - F(x)) \\ &\quad + \prod_{j=1}^{n}(1-\gamma_j)(\hat{F}_{0,m_0}(x) - F(x)) \\ &= \Pi_n\sum_{l=1}^{n}\Pi_l^{-1}\gamma_l(Z_{l,m_l}(x) - F(x)) + \\ &\quad \Pi_n(\hat{F}_{0,m_0}(x) - F(x)).\end{aligned}$$

1. Let's show the results on the bias of our estimator defined in (11).
   Applying the expectation to $\hat{F}_{n,m_n}(x) - F(x)$, we have

$$\begin{aligned}\mathbb{E}[\hat{F}_{n,m_n}(x) - F(x)] &= \mathbb{E}[\hat{F}_{n,m_n}(x)] - F(x) \\ &= \Pi_n\sum_{l=1}^{n}\Pi_l^{-1}\gamma_l(\mathbb{E}[Z_{l,m_l}(x)] - F(x)) + \\ &\quad \Pi_n(\hat{F}_{0,m_0}(x) - F(x)).\end{aligned}$$

   However

$$\begin{aligned}\mathbb{E}[Z_{l,m_l}(x)] - F(x) &= \prod_{i=1}^{d}\left[\sum_{k_i=0}^{m_l}F_{(1)}\left(\frac{k_i}{m_l}\right)b_{k_i}(m_l,x_i)\right] - \\ &\quad \prod_{i=1}^{d}F_{(1)}(x_i) \\ &= \prod_{i=1}^{d}\left[\sum_{k_i=0}^{m_l}F_{(1)}\left(\frac{k_i}{m_l}\right)b_{k_i}(m_l,x_i) - \right. \\ &\quad \left. F_{(1)}(x_i)\right].\end{aligned}$$

   It follows from Theorem 3.0.4 that

$$\sum_{k_i=0}^{m_l}F_{(1)}\left(\frac{k_i}{m_l}\right)b_{k_i}(m_l,x_i) = F_{(1)}(x_i) + \frac{1}{m_l}b(x_i) + o\left(\frac{1}{m_l}\right);$$

   therefore

$$\begin{aligned}\mathbb{E}[Z_{l,m_l}(x)] - F(x) &= \prod_{i=1}^{d}\left[\frac{1}{m_l}b(x_i) + o\left(\frac{1}{m_l}\right)\right] \\ &= \frac{\prod_{i=1}^{d}b(x_i)}{m_l^d} + o\left(\frac{1}{m_l^d}\right) \\ &= \frac{1}{m_l^d}B(x) + o\left(\frac{1}{m_l^d}\right);\end{aligned}$$

hence,

$$\begin{aligned}\mathbb{E}[\hat{F}_{n,m_n}(x)] - F(x) &= \Pi_n\sum_{l=1}^{n}\Pi_l^{-1}\gamma_l\left[\frac{1}{m_l^d}B(x) + o\left(\frac{1}{m_l^d}\right)\right] \\ &\quad + \Pi_n(\hat{F}_{0,m_0}(x) - F(x)) \\ &= B(x)\Pi_n\sum_{l=1}^{n}\Pi_l^{-1}\frac{\gamma_l}{m_l^d} \\ &\quad + \Pi_n\sum_{l=1}^{n}\Pi_l^{-1}\gamma_l\left[o\left(\frac{1}{m_l^d}\right)\right. \\ &\quad \left. + (\hat{F}_{0,m_0}(x) - F(x))\right].\end{aligned}$$

We have :
$-$ if $0 < a \leq \frac{2}{3d}\alpha$, then $\min\left\{ad, \frac{2\alpha+da}{4}\right\} = ad$.
According to Assumption $(A4)$, we have $\lim_{n\longrightarrow+\infty}(n\gamma_n) > ad$.
Applying Lemma 3.0.11 with $m = 1,\ v_n = m_n^d$, we can deduced that $(v_n) \in \mathcal{GS}(ad)$ since Assumption $(A2)$ is verified; thus, $v^* = ad$, hence $m - v^*\xi = 1 - ad\xi > 0$; therefore

$$\mathbb{E}\left[\hat{F}_{n,m_n}(x)\right] - F(x) = m_n^{-d}\frac{1}{(1-ad\xi)}B(x) + o(m_n^{-d}).$$

$-$ If $\frac{2}{3d}\alpha < a \leq \frac{\alpha}{d}$, then $\min\left\{ad, \frac{2\alpha+da}{4}\right\} = \frac{2\alpha+da}{4}$.
According to Assumption $(A4)$, we have $\lim_{n\longrightarrow+\infty}(n\gamma_n) > \frac{2\alpha+da}{4}$.
In this case, we can set $m_n^{-d} = o\left(\sqrt{\frac{\gamma_n}{m_n^{d/2}}}\right)$ because both expressions belong to the same class of sequences with regular variations;
therefore

$$\begin{aligned}\mathbb{E}[\hat{F}_{n,m_n}(x)] - F(x) &= B(x)\Pi_n\sum_{l=1}^{n}\Pi_l^{-1}\gamma_l o\left(\sqrt{\frac{\gamma_n}{m_n^{d/2}}}\right) \\ &\quad + o\left(\sqrt{\frac{\gamma_n}{m_n^{d/2}}}\right) \\ &= o\left(\sqrt{\frac{\gamma_n}{m_n^{d/2}}}\right).\end{aligned}$$

2. Let's show the results on the variance of the estimator defined in (11).

$$\begin{aligned}&\mathbb{V}ar\left(\hat{F}_{n,m_n}(x)\right) = \Pi_n^2\sum_{l=1}^{n}\Pi_l^{-2}\gamma_l^2\left[\mathbb{V}ar(Z_{l,m_l}(x))\right] = \\ &\Pi_n^2\sum_{l=1}^{n}\Pi_l^{-2}\gamma_l^2\mathbb{V}ar\left[\prod_{i=1}^{d}\left(\sum_{k_i=0}^{m_l}F_l\left(\frac{k_i}{m_l}\right)b_{k_i}(m_l,x_i)\right)\right] \\ &= \Pi_n^2\sum_{l=1}^{n}\Pi_l^{-2}\gamma_l^2\prod_{i=1}^{d}\mathbb{V}ar\left[\sum_{k_i=0}^{m_l}F_l\left(\frac{k_i}{m_l}\right)b_{k_i}(m_l,x_i)\right].\end{aligned}$$

According to Theorem 3.0.4, we have :

$$\begin{aligned}\mathbb{V}ar\left[\sum_{k_i=0}^{m_l}F_l\left(\frac{k_i}{m_l}\right)b_{k_i}(m_l,x_i)\right] &= \frac{v(x_i)}{l} - \frac{\sigma^2(x_i)}{l\sqrt{m_l}} + \\ &\quad o\left(\frac{1}{l\sqrt{m_l}}\right);\end{aligned}$$

then

$$\begin{aligned}\mathbb{V}ar[Z_{l,m_l}(x)] &= \frac{\prod_{i=1}^d v(x_i)}{l^d} - \frac{\prod_{i=1}^d \sigma^2(x_i)}{l^d m_l^{d/2}} + o\left(\frac{1}{l^d m_l^{d/2}}\right), \\ &= \frac{V(x)}{l^d} - \frac{\sigma^2(x)}{l^d m_l^{d/2}} + o\left(\frac{1}{l^d m_l^{d/2}}\right).\end{aligned}$$

Applying Lemma 3.0.11, we have :

$$\begin{aligned}\mathbb{V}ar[\hat{F}_{n,m_n}(x)] &= \mathbb{V}(x)\Pi_n^2 \sum_{l=1}^n \Pi_l^{-2}\frac{\gamma_l^2}{l^d} - \sigma^2(x)\Pi_n^2 \sum_{l=1}^n \Pi_l^{-2}\frac{\gamma_l^2}{l^d m_l^{d/2}} \\ &\quad + o\left(\Pi_n^2 \sum_{l=1}^n \Pi_l^{-2}\frac{\gamma_l^2}{l^d m_l^{d/2}}\right). \end{aligned} \tag{16}$$

Therefore in the first component of (16), take $m = 2$ and $v_n = n^d\gamma_n^{-1}$ and this case $v^* = \alpha + d$.
For the second component of (16), take $m = 2$ and $v_n = n^d m_n^{d/2}\gamma_n^{-1}$, thus $v^* = a + \alpha + d$.
We have the following cases :

− if $\frac{2}{3d}\alpha \le a \le 1$, then $\min\left\{ad, \frac{2\alpha+da}{4}\right\} = \frac{2\alpha+da}{4}$.
From Assumption $(A4)$, we have : $\lim_{n\longrightarrow+\infty}(n\gamma_n) > \frac{2\alpha+da}{4}$, thus according to Lemma 3.0.11, we have :

$$\begin{aligned}\mathbb{V}ar[\hat{F}_{n,m_n}(x)] &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad - \frac{\gamma_n}{n^d\sqrt{m_n^d}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad + o\left(\frac{\gamma_n}{n^d\sqrt{m_n^d}}\right).\end{aligned}$$

− In case $\frac{1}{2d}\alpha \le a \le \frac{2}{3d}\alpha$, we have $\gamma_n n^{-d} m_n^{-d/2} = o(m_n^{-d/2})$ and $\lim_{n\longrightarrow+\infty}(n\gamma_n) > \frac{1}{2d}\alpha$ and according to Lemma 3.0.11, we have :

$$\begin{aligned}\mathbb{V}ar[\hat{F}_{n,m_n}(x)] &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} - \Pi_n^2\sum_{l=1}^n \Pi_l^{-2}\gamma_l^2 o_x(m_l^{d/2}) \\ &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} + o_x(n^{-d}\gamma_n).\end{aligned}$$

− When $0 < a < \frac{1}{2d}\alpha$, we have $n^{-d}\gamma_n = o(m_n^{-d/2})$ and $\lim_{n\longrightarrow+\infty}(n\gamma_n) > ad$ and according to Lemma 3.0.11, we have

$$\mathbb{V}ar[\hat{F}_{n,m_n}(x)] = \Pi_n^2\sum_{l=1}^n \Pi_l^{-2}\gamma_l o(m_l^{-d/2}) = o(m_n^{-d/2}).$$

■

The following proposition shows the mean square error of the estimator $\hat{F}_{n,m_n}$ defined in (11) and is the third result of this paper.

Proposition 4.5.4 *Under Assumptions (A1)-(A4), the mean square error of the estimator defined in (11) is given by :*

$$MSE\left[\hat{F}_{n,m_n}(x)\right] = \begin{cases} m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} + o(m_n^{-2d}) & \text{if } 0 < a < \frac{1}{2d}\alpha \\ n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} + m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} + o\left(m_n^{-2d} + n^{-d}\gamma_n\right) & \text{if } a \in [\frac{1}{2d}\alpha, \frac{2\alpha}{3d}[\\ n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} - \frac{\gamma_n}{n^d m_n^{d/2}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} + m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} + o\left(\frac{\gamma_n}{n^d m_n^{d/2}} + m_n^{-2d}\right) & \text{if } a = \frac{2\alpha}{3d} \\ n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} - \frac{\gamma_n}{n^d m_n^{d/2}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} + o\left(\frac{\gamma_n}{n^d m_n^{d/2}}\right) & \text{if } a \in]\frac{2\alpha}{3d}, \frac{\alpha}{d}[, \end{cases} \tag{17}$$

with $B(x)$, $\sigma^2(x)$ and $\mathbb{V}(x)$ defined respectively in (ii), (iii) and (iv).

Proof By definition we have

$$MSE[\hat{F}_{n,m_n}(x)] = \mathbb{V}ar[\hat{F}_{n,m_n}(x)] + Bias^2[\hat{F}_{n,m_n}(x)]. \tag{18}$$

1. if $0 < a < \frac{1}{2d}\alpha$, according to Theorem 4.5.3, the bias and variance of the estimator defined in (11) are given respectively by :

$$Bias[\hat{F}_{n,m_n}(x)] = m_n^{-d}\frac{1}{(1-ad\xi)}B(x) + o(m_n^{-d}),$$

and

$$\mathbb{V}ar[\hat{F}_{n,m_n}(x)] = o\left(\frac{1}{m_n^d}\right). \tag{19}$$

However,

$$\begin{aligned}Bias^2[\hat{F}_{n,m_n}(x)] &= \left[m_n^{-d}\frac{1}{(1-ad\xi)}B(x) + o(m_n^{-d})\right]^2 \\ &= m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} + o(m_n^{-2d}).\end{aligned} \tag{20}$$

Putting (19) and (20) into (18), we have

$$\begin{aligned}MSE[\hat{F}_{n,m_n}(x)] &= o(\frac{1}{m_n^d}) + m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} + o(m_n^{-2d}) \\ &= m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} + o(m_n^{-2d}).\end{aligned}$$

2. If $\frac{1}{2d}\alpha \le a < \frac{2}{3d}\alpha$, according to Theorem 4.5.3, the bias and variance of the estimator defined in (11) are respectively given by :

$$Bias[\hat{F}_{n,m_n}(x)] = m_n^{-d}\frac{1}{(1-ad\xi)}B(x) + o\left(m_n^{-d}\right),$$

and

$$\mathbb{V}ar[\hat{F}_{n,m_n}(x)] = n^{-d}\gamma_n\frac{1}{2-(\alpha+d)\xi}\sigma^2(x) + o\left(n^{-d}\gamma_n\right). \tag{21}$$

However,

$$\begin{aligned} Bias^2[\hat{F}_{n,m_n}(x)] &= \left[m_n^{-d}\frac{1}{(1-ad\xi)}B(x)+o\left(m_n^{-d}\right)\right]^2 \\ &= m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2}+o\left(m_n^{-2d}\right). \end{aligned} \tag{22}$$

Putting (22) and (21) into (18), we have :

$$\begin{aligned} MSE[\hat{F}_{n,m_n}(x)] &= n^{-d}\gamma_n\frac{1}{2-(\alpha+d)\xi}\sigma^2(x) \\ &\quad +o\left(n^{-d}\gamma_n\right)+m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} \\ &\quad +o\left(m_n^{-2d}\right) \\ &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi}+m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} \\ &\quad +o\left(m_n^{-2d}+n^{-d}\gamma_n\right). \end{aligned}$$

3. If $a=\frac{2}{3d}\alpha$, according to Theorem 4.5.3, the bias and variance of the estimator defined in (11) are respectively given by :

$$Bias[\hat{F}_{n,m_n}(x)]=m_n^{-d}\frac{1}{(1-ad\xi)}B(x)+o\left(m_n^{-d}\right)$$

and

$$\begin{aligned} \mathbb{V}ar[\hat{F}_{n,m_n}(x)] &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad -\frac{\gamma_n}{n^d\sqrt{m_n^d}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +o\left(\frac{\gamma_n}{n^d\sqrt{m_n^d}}\right). \end{aligned} \tag{23}$$

However,

$$\begin{aligned} Bias^2[\hat{F}_{n,m_n}(x)] &= \left[m_n^{-d}\frac{1}{(1-ad\xi)}B(x)+o\left(m_n^{-d}\right)\right]^2 \\ &= m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2}+o\left(m_n^{-2d}\right). \end{aligned} \tag{24}$$

Putting (23) and (24) into (18), we have :

$$\begin{aligned} MSE[\hat{F}_{n,m_n}(x)] &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi}-\frac{\gamma_n}{n^d\sqrt{m_n^d}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +o\left(\frac{\gamma_n}{n^d\sqrt{m_n^d}}\right)+m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} \\ &\quad +o\left(m_n^{-2d}\right) \\ &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi}-\frac{\gamma_n}{n^d m_n^{d/2}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2}+o\left(\frac{\gamma_n}{n^d m_n^{d/2}}+m_n^{-2d}\right). \end{aligned}$$

4. If $\frac{2}{3d}\alpha<a<\frac{\alpha}{d}$, according to Theorem 4.5.3, the bias and variance of the estimator defined in (11) are respectively given by:

$$Bias[\hat{F}_{n,m_n}(x)]=o\left(\sqrt{\frac{\gamma_n}{\sqrt{m_n^d}}}\right),$$

and

$$\begin{aligned} \mathbb{V}ar\left[\hat{F}_{n,m_n}(x)\right] &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad -\frac{\gamma_n}{n^d\sqrt{m_n^d}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +o\left(\frac{\gamma_n}{n^d\sqrt{m_n^d}}\right). \end{aligned} \tag{25}$$

However,

$$Bias^2[\hat{F}_{n,m_n}(x)]=o\left(\frac{\gamma_n}{\sqrt{m_n^d}}\right). \tag{26}$$

Putting (25) and (26) into (18), we have :

$$\begin{aligned} MSE[\hat{F}_{n,m_n}(x)] &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad -\frac{\gamma_n}{n^d\sqrt{m_n^d}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +o\left(\frac{\gamma_n}{n^d\sqrt{m_n^d}}\right)+o\left(\frac{\gamma_n}{\sqrt{m_n^d}}\right) \\ &= n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad -\frac{\gamma_n}{n^d m_n^{d/2}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +o\left(\frac{\gamma_n}{n^d m_n^{d/2}}\right). \end{aligned}$$

■

The following proposition gives the integrated mean square error and is the fourth result of this paper.

Proposition 4.5.5 *Under Assumptions (A1)-(A4), the integrated mean square error of the estimator defined in (11) is given by :*

$$MISE\left[\hat{F}_{n,m_n(x)}\right]=\begin{cases} K_3(\xi)m_n^{-2d}\left[1\ +\ o(1)\right] & \text{if } 0<a<\frac{1}{2d}\alpha \\ K_1(\xi)n^{-d}\gamma_n+K_3(\xi)m_n^{-2d} & \\ +o\left(m_n^{-2d}+n^{-d}\gamma_n\right) & \text{if } a\in[\frac{1}{2d}\alpha,\frac{2\alpha}{3d}[\\ K_1(\xi)n^{-d}\gamma_n+K_3(\xi)m_n^{-2d}- & \\ K_2(\xi)\frac{\gamma_n}{n^d m_n^{d/2}}+o(\frac{\gamma_n}{n^d m_n^{d/2}}+m_n^{-2d}) & \text{if } a=\frac{2\alpha}{3d} \\ K_1(\xi)n^{-d}\gamma_n-K_2(\xi)n^{-d}\gamma_n m_n^{-d/2} & \\ +o(n^{-d}\gamma_n m_n^{-3d/2}) & \text{if } a\in]\frac{2\alpha}{3d},\frac{\alpha}{d}[. \end{cases}$$

where

$$K_1(\xi)=\frac{C_1}{2-(\alpha+d)\xi},\qquad K_2(\xi)=\frac{C_2}{2-(\alpha+a+d)\xi}$$
$$and\qquad K_3(\xi)=\frac{C_3}{(1-da\xi)^2}. \tag{27}$$

Proof By definition, we have

$$MISE[\hat{F}_{n,m_n}(x)]=\int_{[0,1]^d}MSE[\hat{F}_{n,m_n}(x)]dx.$$

1. If $0 < a < \frac{1}{2d}\alpha$, then we have

$$\begin{aligned} MISE[\hat{F}_{n,m_n}(x)] &= \int_{[0,1]^d}\left[m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2}+o\left(m_n^{-2d}\right)\right]dx \\ &\quad \text{from Proposition 4.5.4} \\ &= m_n^{-2d}\frac{1}{(1-da\xi)^2}\int_{[0,1]^d}B^2(x)dx \\ &\quad +o\left(m_n^{-2d}\right) \\ &= m_n^{-2d}\frac{1}{(1-da\xi)^2}C_3 \\ &\quad +o\left(m_n^{-2d}\right) \text{ from } (v) \\ &= K_3(\xi)m_n^{-2d}\left[1 + o(1)\right], \end{aligned}$$

with $K_3(\xi) = \frac{C_3}{(1-da\xi)^2}$.

2. If $\frac{1}{2d}\alpha \le a < \frac{2}{3d}\alpha$, then we have

$$\begin{aligned} MISE[\hat{F}_{n,m_n}(x)] &= \int_{[0,1]^d}\Bigg(n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad +m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} \\ &\quad + o(m_n^{-2d}+n^{-d}\gamma_n)\Bigg)dx \\ &\quad \text{from, Proposition 4.5.4} \\ &= n^{-d}\gamma_n\frac{1}{2-(\alpha+d)\xi}\int_{[0,1]^d}\sigma^2(x)dx \\ &\quad + m_n^{-2d}\frac{1}{(1-da\xi)^2}\int_{[0,1]^d}B^2(x)dx \\ &\quad +o\left(m_n^{-2d}+n^{-d}\gamma_n\right) \\ &= n^{-d}\gamma_n\frac{C_1}{2-(\alpha+d)\xi}+m_n^{-2d}\frac{C_3}{(1-da\xi)^2} \\ &\quad +o(m_n^{-2d}+n^{-d}\gamma_n) \text{ from } (v) \\ &= K_1(\xi)n^{-d}\gamma_n+K_3(\xi)m_n^{-2d} \\ &\quad +o\left(m_n^{-2d}+n^{-d}\gamma_n\right), \end{aligned}$$

with $K_1(\xi) = \frac{C_1}{2-(\alpha+d)\xi}$ and $K_3(\xi) = \frac{C_3}{(1-da\xi)^2}$.

3. If $a = \frac{2}{3d}\alpha$, then we have

$$\begin{aligned} MISE[\hat{F}_{n,m_n}(x)] &= \int_{[0,1]^d}\Bigg(n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad -\frac{\gamma_n}{n^d m_n^{d/2}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +m_n^{-2d}\frac{B^2(x)}{(1-da\xi)^2} \\ &\quad +o\left(\frac{\gamma_n}{n^d m_n^{d/2}}+m_n^{-2d}\right)\Bigg)dx \\ &\quad \text{from Proposition 4.5.4} \\ &= n^{-d}\gamma_n\frac{1}{2-(\alpha+d)\xi}\int_{[0,1]^d}\sigma^2(x)dx \\ &\quad -\frac{\gamma_n}{n^d m_n^{d/2}}\frac{1}{2-(d+\alpha+a)\xi}\int_{[0,1]^d}\mathbb{V}(x)dx \\ &\quad +m_n^{-2d}\frac{1}{(1-da\xi)^2}\int_{[0,1]^d}B^2(x)dx \\ &\quad +o\left(\frac{\gamma_n}{n^d m_n^{d/2}}+m_n^{-2d}\right) \\ &= n^{-d}\gamma_n\frac{C_1}{2-(\alpha+d)\xi}+m_n^{-2d}\frac{C_3}{(1-da\xi)^2} \\ &\quad -\frac{\gamma_n}{n^d m_n^{d/2}}\frac{C_2}{2-(\alpha+a+d)\xi} \\ &\quad +o(\frac{\gamma_n}{n^d m_n^{d/2}}+m_n^{-2d}) \text{ from } (v) \\ &= K_1(\xi)n^{-d}\gamma_n+K_3(\xi)m_n^{-2d} \\ &\quad -K_2(\xi)\frac{\gamma_n}{n^d m_n^{d/2}}+o\left(\frac{\gamma_n}{n^d m_n^{d/2}}+m_n^{-2d}\right), \end{aligned}$$

where $K_1(\xi) = \frac{C_1}{2-(\alpha+d)\xi}$, $K_2(\xi) = \frac{C_2}{2-(\alpha+a+d)\xi}$ and $K_3(\xi) = \frac{C_3}{(1-da\xi)^2}$.

4. If $\frac{2}{3d}\alpha < a < \frac{\alpha}{d}$, then we have

$$\begin{aligned} MISE[\hat{F}_{n,m_n}(x)] &= \int_{[0,1]^d}\Bigg[n^{-d}\gamma_n\frac{\sigma^2(x)}{2-(\alpha+d)\xi} \\ &\quad -\frac{\gamma_n}{n^d m_n^{d/2}}\frac{\mathbb{V}(x)}{2-(d+\alpha+a)\xi} \\ &\quad +o\left(\frac{\gamma_n}{n^d m_n^{d/2}}\right)\Bigg]dx \\ &\quad \text{from Proposition 4.5.4} \\ &= n^{-d}\gamma_n\frac{1}{2-(\alpha+d)\xi}\int_{[0,1]^d}\sigma^2(x)dx \\ &\quad -\frac{\gamma_n}{n^d m_n^{d/2}}\frac{1}{2-(d+\alpha+a)\xi}\int_{[0,1]^d}\mathbb{V}(x)dx \\ &\quad +o\left(\frac{\gamma_n}{n^d m_n^{d/2}}\right) \\ &= n^{-d}\gamma_n\frac{1}{2-(\alpha+d)\xi}C_1 \\ &\quad -\frac{\gamma_n}{n^d m_n^{d/2}}\frac{1}{2-(d+\alpha+a)\xi}C_2 \\ &\quad +o\left(\frac{\gamma_n}{n^d m_n^{d/2}}\right) \quad \text{from } (v) \\ &= K_1(\xi)n^{-d}\gamma_n - K_2(\xi)n^{-d}\gamma_n m_n^{-d/2} \\ &\quad +o\left(n^{-d}\gamma_n m_n^{-3d/2}\right), \end{aligned}$$

with $K_1(\xi) = \frac{C_1}{2-(\alpha+d)\xi}$ and $K_2(\xi) = \frac{C_2}{2-(\alpha+a+d)\xi}$.

■

The previous proposition allows us to directly obtain the optimal choice of m_n for which the integrated mean square error is minimal for the estimator defined in (11) and the corresponding $MISE$ in the case $(\gamma_n) \in \mathcal{GS}(-1)$. It is thus the fifth result of this paper.

Corollary 4.5.6 *Under Assumptions (A1)-(A4), the $MISE$ of the estimator defined in (11) is minimal for $(\gamma_n) \in \mathcal{GS}(-1)$ and $\lim_{n\longrightarrow+\infty}(n\gamma_n) = 1$ if m_n is defined by:*

$$m_{n_{opt}} = \left(\frac{4K_3(\xi)}{n^d\gamma_n K_2(\xi)}\right)^{\frac{2}{3d}}.$$

Then we have

$$\begin{aligned} MISE\left[\hat{F}_{n,m_{n_{opt}}}(x)\right] &= \frac{n^{-d-1}C_1}{1-d} \\ &= +\frac{3^{2/3}4^{-4/3}C_3^{1/3}D^{-2/3}}{n^{\frac{-4(d-1)}{3}}C_2^{-4/3}}\times \\ &\quad \left(D^{-2/3}-4n^{-d}3^{-2/3}C_2^{-2/3}\right), \end{aligned}$$

with $D = -3d^2+3d-2$.

Before proceeding to the proof of this corollary 4.5.6, we will state a technical result which will be useful to us:

Lemma 4.5.7 *Considering $K_2(\xi)$ and $K_3(\xi)$ defined in (27) and the sequences with regular variations defined by Galambos and Seneta [12], we have*

1. $[K_3(\xi)m_n^{-2d}] \in \mathcal{GS}(-2ad)$.
2. $[K_2(\xi)n^{-d}\gamma_n m_n^{-d/2}] \in \mathcal{GS}(-(d+\alpha+\frac{ad}{2}))$.

We now prove the Corollary 4.5.6:

Proof According to Proposition 4.5.7, the integrated mean square error is given by

$$MISE[\hat{F}_{n,m_n}(x)] = \begin{cases} K_3(\xi)m_n^{-2d}\,[1 \;+\; o(1)] & \text{if } 0 < a < \frac{1}{2d}\alpha \\ K_1(\xi)n^{-d}\gamma_n + K_3(\xi)m_n^{-2d} & \\ +o\left(m_n^{-2d} + n^{-d}\gamma_n\right) & \text{if } a \in [\frac{1}{2d}\alpha,\, \frac{2\alpha}{3d}[\\ K_1(\xi)n^{-d}\gamma_n + K_3(\xi)m_n^{-2d} - K_2(\xi)\frac{\gamma_n}{n^d m_n^{d/2}} & \\ +o\left(\frac{\gamma_n}{n^d m_n^{d/2}} + m_n^{-2d}\right) & \text{if } a = \frac{2\alpha}{3d} \\ K_1(\xi)n^{-d}\gamma_n - K_2(\xi)n^{-d}\gamma_n m_n^{-d/2} & \\ +o\left(n^{-d}\gamma_n m_n^{-3d/2}\right) & \text{if } a \in]\frac{2\alpha}{3d},\, \frac{\alpha}{d}[; \end{cases}$$

which can be written as follows:

$$MISE[\hat{F}_{n,m_n}(x)] = \begin{cases} K_3(\xi)m_n^{-2d}[1 \;+\; o(1)] & \text{if } 0 < a < \frac{1}{2d}\alpha, \\ K_1(\xi)n^{-d}\gamma_n + m_n^{-2d}K_3(\xi)[1 + o(1)] & \\ & \text{if } a \in [\frac{1}{2d}\alpha,\, \frac{2\alpha}{3d}[, \\ K_1(\xi)n^{-d}\gamma_n + m_n^{-2d}[K_3(\xi) - & \\ K_2(\xi)\frac{\gamma_n}{n^d m_n^{-3d/2}} + o(1)] & \text{if } a = \frac{2\alpha}{3d}, \\ K_1(\xi)n^{-d}\gamma_n - n^{-d}\gamma_n m_n^{-d/2}[K_2(\xi) & \\ +o(1)] & \text{if } a \in]\frac{2\alpha}{3d}\, \frac{\alpha}{d}[. \end{cases}$$

We distinguish the following cases :

- If $0 < a < \frac{1}{2d}\alpha$, then we have $K_3(\xi)m_n^{-2d} \in \mathcal{GS}(-2ad)$, according to 1) of Lemma 4.5.6; and since $-2ad > -\alpha$, we conclude that

$$MISE(\hat{F}_{n,m_n}) \in \mathcal{GS}(-\alpha).$$

- In case $\frac{1}{2d}\alpha \leq a < \frac{\alpha}{d}$, we make the comparison on

$$MISE(\hat{F}_{n,m_n}) \in \mathcal{GS}(-\alpha) - K_1(\xi)n^{-d}\gamma_n.$$

We have

$$MISE[\hat{F}_{n,m_n}(x)] - K_1(\xi)\frac{n^{-d}}{\gamma_n} = \begin{cases} K_3(\xi)m_n^{-2d} + o(m_n^{-2d}) \\ \text{if } a \in [\frac{\alpha}{2d},\, \frac{2\alpha}{3d}[\\ K_3(\xi)m_n^{-2d} - K_2(\xi)\frac{\gamma_n}{n^d m_n^{d/2}} \\ +o(m_n^{-2d}) \quad \text{if } a = \frac{2\alpha}{3d} \\ -K_2(\xi)\gamma_n n^{-d}m_n^{-d/2} \\ +o(\gamma_n n^{-d}m_n^{-d/2}) \quad \text{if } a \in]\frac{2\alpha}{3d},\, \frac{\alpha}{d}[. \end{cases}$$

We distinguish several situations :

- • If $a \in [\frac{\alpha}{2d}, \frac{2\alpha}{3d}[$, and since $[K_3(\xi)m_n^{-2d}] \in \mathcal{GS}(-2ad)$, according to 1) of Lemma 4.5.7, then :

$$MISE(\hat{F}_{n,m_n}) - K_1(\xi)\frac{n^{-d}}{\gamma_n} \in \mathcal{GS}(-2ad),$$

  and $-2ad > -\frac{4\alpha}{3d}$.

- • If $a = \frac{2}{3d}\alpha$, and because

$$[K_3(\xi)m_n^{-2d}] \in \mathcal{GS}(-2ad)$$

  and

$$[-K_2(\xi)\gamma_n n^{-d}m_n^{-d/2}] \in \mathcal{GS}\left[-\left(d + \alpha + \frac{ad}{2}\right)\right],$$

  according to 1) and 2) of Lemma 4.5.7, we have

$$MISE(\hat{F}_{n,m_n}) - K_1(\xi)\frac{n^{-d}}{\gamma_n} \in \mathcal{GS}\left[-\left(\frac{4\alpha}{3d}\right)\right],$$

  with $\frac{4\alpha}{3d} = 2a$.

- • If $\frac{2}{3d}\alpha \leq a < \frac{\alpha}{d}$, and since

$$\left[-K_2(\xi)\gamma_n n^{-d}m_n^{-d/2}\right] \in \mathcal{GS}\left[-(d + \alpha + \frac{ad}{2})\right],$$

  according to 2) of Lemma 4.5.7 and because $-(d + \alpha + \frac{ad}{2}) > -2ad$, we have

$$MISE(\hat{F}_{n,m_n}) - K_1(\xi)\frac{n^{-d}}{\gamma_n} \in \mathcal{GS}(-2ad).$$

  It follows that, for a given α, in order to minimize

$$MISE(\hat{F}_{n,m_n}) - K_1(\xi)\frac{n^{-d}}{\gamma_n},$$

  the parameter "a" should be chosen equal to $\frac{2\alpha}{3d}$, and we choose $\alpha = 1$.

We can conclude that to minimise $MISE(\hat{F}_{n,m_n})$, the step must be chosen in $\mathcal{GS}(-1)$ and the order $(m_n) \in \mathcal{GS}\left(\frac{2}{3d}\right)$.

Consider the following function:

$$g : x \longmapsto x^{-2d}n^{-d}K_3(\xi) - \gamma_n x^{-d/2}K_2(\xi).$$

The function $g : x \longmapsto x^{-2d}n^{-d}K_3(\xi) - \gamma_n x^{-d/2}K_2(\xi)$ is continuous and derivable on $\mathbb{R}^*$ and its derivative is :

$$g'(x) = -2dx^{-2d-1}n^{-d}K_3(\xi) + \frac{d}{2}\gamma_n x^{-d/2-1}K_2(\xi).$$

We have :

$$\begin{aligned} g'(x) = 0 \quad &\Longleftrightarrow \quad -2dx^{-2d-1}n^{-d}K_3(\xi) \\ &\qquad\quad + \tfrac{d}{2}\gamma_n x^{-d/2-1}K_2(\xi) = 0 \\ &\Longleftrightarrow \quad x = \left(\frac{4K_3(\xi)}{n^d\gamma_n K_2(\xi)}\right)^{\frac{2}{3d}}. \end{aligned}$$

It follows that g reaches its minimum at

$$x = \left(\frac{4K_3(\xi)}{n^d\gamma_n K_2(\xi)}\right).$$

Therefore $m_{n_{opt}} = \left(\frac{4K_3(\xi)}{n^d\gamma_n K_2(\xi)}\right)^{\frac{2}{3d}}$.

By replacing m_n by $m_{n_{opt}}$ in the following quantity:

$$\begin{aligned} MISE(\hat{F}_{n,m_n}) \quad &= \quad K_1(\xi)\frac{n^{-d}}{\gamma_n} + K_3(\xi)m_n^{-2d} \\ &\qquad - K_2(\xi)\frac{\gamma_n}{n^d m_n^{d/2}} + o(m_n^{-2d}), \end{aligned}$$

we have

$$MISE[\hat{F}_{n,m_{n_{opt}}}(x)] = \frac{n^{-d-1}C_1}{1-d} + \frac{3^{2/3}4^{-4/3}C_3^{1/3}D^{-2/3}}{n^{\frac{-4(d-1)}{3}}C_2^{-4/3}} \times \left(D^{-2/3} - 4n^{-d}3^{-2/3}C_2^{-2/3}\right) + o\left(\frac{1}{n^{d+1}}\right),$$

with $D = -3d^2 + 3d - 2$.

Hence the result. ■

5 Numerical studies

5.1 Simulation studies

Simulation studies

In this section, we perform simulation studies to compare the estimator

$$\hat{F}_{n,m_n}(x) = (1-\gamma_n)\hat{F}_{n-1,m_{n-1}}(x) + \gamma_n Z_{n,m_n}(x), \quad (28)$$

with that of Vitale [10]

$$F_{n,m_n}(x) = \prod_{i=1}^{d}\left[\sum_{k_i=0}^{m_n} F_n\left(\frac{k_i}{m_n}\right) b_{k_i}(m_n, x_i)\right], \quad (29)$$

for $d = 1$.

For estimator (28), two quantities must be chosen:

1. The step size of the algorithm $(\gamma_n) = (\gamma_0 n^{-1})$, where $\gamma_0 = 2/3 + c$ with $c \in]0, 1/3]$.
2. The order m_n is relative to the optimal choice.

In our simulation study, we consider:

$n = 50, \ n = 150 \ and \ n = 1000$ and the following three distribution functions :

- Distribution beta $\mathcal{B}(2,1)$ with an approximately symmetric distribution function.

- Distribution $1/2\mathcal{B}eta(12,6) + 1/2\mathcal{B}eta(9,1)$ with a completely asymmetric distribution function.

- Troncated normal distribution $\mathcal{N}_{[0,1]}(0.5, 0.25)$ with a symmetric distribution function.

All our figures have been generated using R software version 4.0.3 $(2020-10-10)$. Our estimator (28) is shown in **red** and that of Vitale (29) in **green**.

For distribution beta $\mathcal{B}(2,1)$, we obtain the following figure 1 below.

Through Figure 1, we notice that if the sample size is large $(n = 1000)$, the estimator (28) converges faster to the same direction as that of beta $\mathcal{B}(2,1)$ and during this time, the Vitale estimator (29) overflows the study interval. For small sample sizes $(n = 50)$, the estimator (29) converges better to the beta $(2,1)$ than the estimator (28). However, when the sample size is average $(n = 150)$, we observe a considerable reduction of the mean square errors of the estimator (28) compared to the estimator (29).

For the distribution $1/2\mathcal{B}eta(12,6) + 1/2\mathcal{B}eta(9,1)$, we have the following Figure 2 below.

Through Figure 2, the distribution $1/2Beta(12,6) + 1/2Beta(9,1)$ represented in **black** is asymmetric. Both estimators follow the same pattern as the distribution $1/2Beta(12,6) + 1/2Beta(9,1)$. When the sample size is small $(n = 50)$, over the interval $[0;\ 0.6]$, the estimator (28) has a smaller mean square error than Vitale's (29) ; but on the interval $[0.6;\ 1]$, the Vitale estimator (29) has a reduced mean square error than the estimator (28). When the sample size is average $(n = 150)$, the estimators (28) and (29) are nearly similar. On the other hand, when the sample size is large $(n = 1000)$, the estimator (28) on $[0;\ 0.8]$ is preferable to that of Vitale (29). On the other hand, on the interval $[0.8;\ 1]$, the Vitale estimator (29) is preferable to that of the estimator (28). Thus, globally, the larger the sample size, the larger the interval of preference of the estimator (28).

For the troncated normal distribution $\mathcal{N}_{[0,1]}(0.5, 0.25)$, we have the following figure 3.

We observe that for samples from the normal distribution (see Figure 3) represented in **black** whose data are symmetrical, when the sample size is small $(n = 50)$, Vitale's estimator (29) is closer to the normal distribution than the estimator (28). On the other hand, when the sample size is average $(n = 150)$, we observe a significant improvement with respect to the estimator (28), especially in the neighborhood of the edge, and this behavior improves as the sample size is large $(n = 1000)$.

Figure 3. Bernstein's estimator for a sample of distribution $\mathcal{N}_{[0,1]}(0.5, 0.25)$.

5.2 Examples of multivariate recursive distribution functions

In this section, we give some examples of multivariate recursive distribution function estimators using Bernstein's polynomial and the stochastic approximation method derived from 11.

Example 5.2.1 *Let $X_1; ....; X_n$ sequences of i.i.d random variables with value in $[0; 1]^d$ follow a normal distribution function Φ. The recursive estimator $\hat{\Phi}_{n,m_n}$ of the distribution function Φ by the stochastic approximation method using the Bernstein polynomial is given by : $\forall x \in [0; 1]^d$*

$$\hat{\Phi}_{n,m_n}(x) = (1-\gamma_n)\hat{\Phi}_{n-1,m_{n-1}}(x) + \gamma_n Q_{n,m_n}(x), \quad (30)$$

with

$$Q_{n,m_n}(x) = \prod_{i=1}^{d}\Big(\sum_{k_i=0}^{m_n}\Phi(\frac{k_i}{m_n})b_{k_i}(m_n, x_i)\Big). \quad (31)$$

Example 5.2.2 *Consider $(X_i)_{i\in\{1;...;n\}}$ sequences of i.i.d random variables with value in $[0; 1]^d$ which follow the Beta distribution function β. The recursive estimator $\hat{\beta}_{n,m_n}$ of β by the stochastic approximation method using the Bernstein polynomial is defined by : $\forall x \in [0; 1]^d$*

$$\hat{\beta}_{n,m_n}(x) = (1-\gamma_n)\hat{\beta}_{n-1,m_{n-1}}(x) + \gamma_n G_{n,m_n}(x), \quad (32)$$

with

$$G_{n,m_n}(x) = \prod_{i=1}^{d}\Big(\sum_{k_i=0}^{m_n}\beta(\frac{k_i}{m_n})b_{k_i}(m_n, x_i)\Big). \quad (33)$$

Example 5.2.3 *Let $(X_1; ....; X_n)$ be sequences of random variables i.i.d with value in $[0; 1]^d$ which follow the exponential distribution function ξ. The recursive estimator $\hat{\xi}_{n,m_n}$ of ξ by the stochastic approximation method using the Bernstein polynomial is given by : $\forall x \in [0; 1]^d$*

$$\hat{\xi}_{n,m_n}(x) = (1-\gamma_n)\hat{\xi}_{n-1,m_{n-1}}(x) + \gamma_n E_{n,m_n}(x), \quad (34)$$

with

$$E_{n,m_n}(x) = \prod_{i=1}^{d}\Big(\sum_{k_i=0}^{m_n}\xi(\frac{k_i}{m_n})b_{k_i}(m_n, x_i)\Big). \quad (35)$$

6 Conclusions

In this paper, we have outlined some basic notions about the Bernstein polynomial and the univariate iterative distribution function estimator using the Bernstein polynomial proposed by Vitale [10]. We then presented a recursive method called stochastic approximation method whose application allowed us to construct a recursive multidimensional estimator of the distribution function using the generalized Bernstein polynomial and its consistency properties. We ended with a numerical study highlighting a comparative study between the proposed estimator of the distribution function and that of Vitale. This study has shown us that in addition to the fact that our estimator is recursive and will therefore facilitate the updating of the data in the database, it appears that when the sample size is large our estimator is more efficient than that of Vitale. However, many works are planned for the future. These include the study of the convergence properties of the proposed estimator; the recursive estimation of the multidimensional density function using the generalized Bernstein polynomial, the recursive estimation of the multidimensional regression function using the generalized Bernstein polynomial, the construction of a recursive estimator of the distribution, density and regression function when the variables are dependent, the construction of a recursive estimator based on the Bernstein polynomials of a distribution, density and regression function using the semi-recursive estimation method.

Acknowledgements

We are very grateful to the reviewers for their thoughtful comments which have improved the work.

REFERENCES

[1] Bernstein, S. N., "Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités", Communications of the Kharkov Mathematical Society, vol. 13, no 1, pp. 1-2, 1912. Zbl43.0301.03.

[2] Tenbusch, A., "Two-dimensional Bernstein polynomial density estimation", Metrika, vol. 41, no 1, pp. 233-253, 1994. DOI: 10.1007/BF01895321.

[3] Leblanc, A., "A bias-reduced approach to density estimation using Bernstein polynomials", Journal of Nonparametric Statistics, vol. 22, no 4, pp. 459-475, 2010. DOI: 10.1080/10485250903318107.

[4] Leblanc, A., "On estimating distribution function using Bernstein polynomials", Annals of the Institute of Statistical Mathematics, vol. 64, pp. 919-943, 2012. DOI: 10.1007/s10463-011-0339-4.

[5] Babu, G. J., Canty, A. J., & Chaubey, Y. P., "Application of Bernstein polynomials for smooth estimation of a distribution and density function", Journal of Statistical Planning and Inference, vol. 105, no 2, pp. 377-392, 2002. DOI: 10.1016/S0378-3758(01)00265-8.

[6] Babu, G. J., & Chaubey, Y. P., "Smooth estimation of a distribution function and density function on a hypercube using Bernstein polynomials for dependent random vectors", Statistics and Probability Letters, vol. 76, no 9, pp. 959-969, 2006. DOI: 10.1016/j.spl.2005.10.031.

[7] Robbins, H., & Monro, S., "A Stochastic Approximation Method", Annals of Mathematical Statistics, vol. 22, no 3, pp. 400-407, 1951. DOI: 10.1214/aoms/1177729586.

[8] Mokkadem, A., Pelletier, M., & Slaoui, Y., "The stochastic approximation method for the estimation of a multivariate probability density", Journal of Statistical Planning and Inference, vol. 139, no 7, pp. 2459-2478, 2009. DOI : 10.1016/j.jspi.2008.11.012.

[9] Nguyen, T. M. N., & Saracco, J., "Estimation rcursive en rgression inverse par tranches (sliced inverse regression)", Journal de la socité franaise de statistique, vol. 151, no 2, pp. 19-46, 2010. Zbl: 1316.62071.

[10] Vitale, R. A., "A Bernstein polynomial approach to density function estimation", Statistical Inference and Related Topics, pp. 87-99, 1975. DOI: 10.1016/B978-0-12-568002-8.50011-2.

[11] Lorentz, G. G., "Bernstein polynomials" in American Mathematical Soc., 2^{th} ed, Chelsea Publishing, vol. 323, pp.134, 2012. https://bookstore.ams.org/chel-323.h.

[12] Galambos, J., & Seneta, E., "Regularly varying sequences", Proceedings of the American Mathematical Society, vol. 41, no 1, pp. 110-116, 1973. DOI: 10.2307/2038824.

[13] Kiefer, J., & Wolfowitz, J., "Stochastic approximation of the maximum of a regression functions", Annals of Mathematical Statistics, vol. 23, no 3, pp. 462-466, 1952. DOI: 10.1214/aoms/1177729392.

[14] Blum, J. R., "Multidimensional stochastic approximation methods", Annals of Mathematical Statistics, vol. 25, no 4, pp. 737-744, 1954. DOI: 10.1214/aoms/1177728659.

[15] Chen, H., "Lower rate of convergence for locating a maximum of a function", Annals of Statistics, vol. 16, no 3, pp. 1330-1334, 1988. DOI: 10.1214/aos/1176350965.

[16] Dippon, J., "Accelerated randomized stochastic optimization", Annals of Statistics, vol. 31, no 4, pp. 1260-1281, 2003. DOI: 10.1214/aos/1059655913.

[17] Mokkadem, A., & Pelletier, M., "A companion for the Kiefer-Wolfowitz- Blum stochastic approximation algorithm", vol. 35, no 4, pp. 1749-1772, 2007. DOI: 10.1214/009053606000001451.

[18] Slaoui, Y., "The stochastic approximation method for the estimation of a distribution function", Mathematical Methods of Statistics, vol. 23, no 4, pp. 306-325, 2014. DOI: 10.3103/S1066530714040048.

[19] Slaoui, Y., "Bandwidth selection for recursive kernel density estimators defined by stochastic approximation method", Journal of Probability and Statistics, pp. 1-11, 2014. DOI: 10.1155/2014/739640.