



Rate of complete second-order moment convergence and theoretical applications

I. Laroussi & M. Madi

To cite this article: I. Laroussi & M. Madi (2022) Rate of complete second-order moment convergence and theoretical applications, Journal of Taibah University for Science, 16:1, 566-574, DOI: [10.1080/16583655.2022.2082179](https://doi.org/10.1080/16583655.2022.2082179)

To link to this article: <https://doi.org/10.1080/16583655.2022.2082179>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 06 Jun 2022.



Submit your article to this journal [↗](#)



Article views: 48



View related articles [↗](#)



View Crossmark data [↗](#)

Rate of complete second-order moment convergence and theoretical applications

I. Laroussi and M. Madi

Department of Mathematics, Frères Mentouri University, Constantine, Algeria

ABSTRACT

The purpose of this work is to present a novel mode of convergence, complete second-order moment convergence with rate, which implies almost complete convergence and gives a smaller rate of convergence. Indeed, this mode is easier to obtain and gives better performances than those of the almost complete convergence in the case of the nonparametric estimators with kernels of the density function, of the distribution function and of the quantile function. A great advantage of the proposed approach is that less conditions are imposed to the kernel function thanks to the use of the mean squared error expression.

ARTICLE HISTORY

Received 20 September 2021
Revised 18 May 2022
Accepted 20 May 2022

KEYWORDS

Kernel density estimator;
kernel distribution estimator;
kernel quantile estimator;
complete moment
convergence; nonparametric
estimation; rate of
convergence; complete
convergence

1. Introduction

The rate of convergence is an important concept that describes the speed with which a sequence converges to the limit. In this setting, a fundamental question is how fast the convergence is illuminating theoretical studies on the subject has been carried out proposing important new results, algorithms and applications to address this issue. The purpose of this work is to present a novel mode of convergence, namely the complete secondorder convergence with a rate. The novelty here is in the introduction of the complete secondorder convergence rate. Hence, we give proofs of different proprieties of this kind of convergence. The almost complete convergence is induced by second-order convergence in different context (see [1–3] and recently see Yu et al. [29]). The complete convergence concept was introduced by Hsu and Robbins [4]. Then, it has been used by several authors such as Gut and Stadtmüller [5], Gut [6,7], Li et al. [8], Sung [9], Sung and Volodin [10]. The interest of such a notion lies in the fact that almost complete convergence (a.c.) implies almost sure convergence (a.s.) due to the Borel–Cantelli lemma. However in some situations at least, it is much easier to obtain complete second-order moment convergence (c.s.m.) instead of almost complete convergence.

As a practical framework, rates of complete second-order moment convergence for the probability density, the distribution function and the quantile function kernel estimators are established. We also discuss about the speed at which these estimators converge. First, in

the context of estimating probability density function, many studies using different methods have been proposed. The Kernel method is one of the best of these methods which seems to be convenient and does not require a multiple choice of parameters. Rosenblatt [11] was the earliest pioneer of the class of kernel density estimators, using two parameters, namely the kernel and the bandwidth. For this estimator, convergence in probability was established by Parzen [12]. Habbema et al. [13], Hall and Kang [14], Hall and Wand [15], Gosh and Chaudhury [16] and Gosh and Hall [17] can be consulted for various works on the subject, in particular the estimation by classical kernels of the densities. In the case of independent observations, optimal rates of convergence to zero for mean square error and the bias of the kernel estimators have been addressed by several authors under varying conditions on the kernel (K) and the density (f). As a contribution, a new complete second-order moment convergence with a rate is introduced for the first time to improve the rates of convergence for the bias and the mean square error (MSE) of kernel density estimators. Then the complete convergence of the density kernel estimator, under weaker conditions on the density function than those proposed in the literature, is achieved. As a consequence of the complete second-order moment convergence, the almost complete convergence is obtained with a better rate.

Second, the proposed mode of convergence is applied to the distribution function and quantiles kernel

estimators. Notice that for the distribution function, Nadaraya [18] proposed its kernel estimator. While, Parzen [19] retraced the context of Nadaraya [18] and constructed the kernel quantile estimators for which we established the rate of almost complete convergence. To the best of our knowledge, this is a new result obtained from the proposed convergence mode. A great advantage of the proposed approach is that less conditions are imposed to the kernel function thanks to the use of the mean squared error expression.

2. Complete second-order moment convergence with a rate

Throughout this paper, real-valued random variables are defined on a fixed probability space (Ω, \mathcal{A}, P) .

Let $(U_n)_{n \in \mathbb{N}}$ and $(V_n)_{n \in \mathbb{N}}$ be two sequences of real numbers. We assume $(V_n)_{n \in \mathbb{N}}$ does not vanish from a certain rank. We say that $(U_n)_{n \in \mathbb{N}}$ is dominated by $(V_n)_{n \in \mathbb{N}}$ if there exist a real number M and an integer n_0 , such that, for all $n \geq n_0$, we have

$$|U_n| \leq M|V_n|,$$

and we note $U_n = O(V_n)$.

Definition 2.1: A sequence $(X_n)_{n \in \mathbb{N}}$ of random variables is said to be almost complete second-order moment convergent (c.s.m) to the random variable X , with the convergence rate $\frac{1}{\sqrt{U_n}}$, if

$$\sum_{n \in \mathbb{N}} U_n E(X_n - X)^2 < \infty,$$

where (U_n) is a sequence of positive numbers. Note this mode

$$X_n - X = O_{c.s.m} \left(\frac{1}{\sqrt{U_n}} \right).$$

The following theorem shows that if $(X_n)_{n \in \mathbb{N}}$ converges in complete second-order moment to X with a rate $\frac{1}{\sqrt{U_n}}$, then it is almost completely convergent with a rate of $\frac{1}{\sqrt{U_n}}$.

Theorem 2.1: If a sequence $(X_n)_{n \in \mathbb{N}}$ of random variables verifies

$$X_n - X = O_{c.s.m} \left(\frac{1}{\sqrt{U_n}} \right),$$

then

$$X_n - X = O_{ac} \left(\frac{1}{\sqrt{U_n}} \right).$$

And for $\beta > -1$, $\varepsilon > 0$ and the sequence $(V_n)_{n \in \mathbb{N}} = (n^\beta U_n)_{n \in \mathbb{N}}$ we obtain

$$\sum_{n \in \mathbb{N}} n^\beta P(|X_n - X| > \varepsilon) \leq \sum_{n \in \mathbb{N}} n^\beta U_n E(X_n - X)^2 < \infty.$$

Proof: Suppose that $\sum_{n \in \mathbb{N}} U_n E(X_n - X)^2 < \infty$, then from the Markov inequality we obtain

$$P\left(|X_n - X| \geq \frac{1}{\sqrt{U_n}}\right) \leq U_n E(X_n - X)^2,$$

so

$$\sum_{n \in \mathbb{N}} P(|X_n - X| \geq \frac{1}{\sqrt{U_n}}) \leq \sum_{n \in \mathbb{N}} U_n E(X_n - X)^2 < \infty.$$

Now, $\frac{1}{\sqrt{U_n}}$ is a rate of convergence, so $\frac{1}{\sqrt{U_n}} \rightarrow 0$ which equivalent to, for each $\varepsilon > 0$, there exist $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have $\frac{1}{\sqrt{U_n}} < \varepsilon$. So

$$P(|X_n - X| > \varepsilon) \leq P\left(|X_n - X| \geq \frac{1}{\sqrt{U_n}}\right)$$

which implies

$$\sum_{n \in \mathbb{N}} n^\beta P(|X_n - X| > \varepsilon) \leq \sum_{n \in \mathbb{N}} n^\beta U_n E(X_n - X)^2 < \infty.$$

■

The following proposition gives some elementary calculus rules.

Proposition 2.1: Assume $\lim_{n \rightarrow \infty} U_n = +\infty$, $X_n - I_X = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$ and $Y_n - I_Y = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$, where I_X, I_Y are real numbers. We have

- (1) $(X_n + Y_n) - (I_X + I_Y) = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$;
- (2) $X_n \cdot Y_n - I_X \cdot I_Y = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$;
- (3) $\frac{1}{X_n} - \frac{1}{I_X} = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$, with $I_X \neq 0$.

Proof: (1) Immediately from the following inequality:

$$\begin{aligned} \sum_{n \in \mathbb{N}} U_n E((X_n + Y_n) - (I_X + I_Y))^2 \\ \leq 2^2 \left(\sum_{n \in \mathbb{N}} U_n E(X_n - I_X)^2 + \sum_{n \in \mathbb{N}} U_n E(Y_n - I_Y)^2 \right). \end{aligned}$$

(2) We have

$$\begin{aligned} E(X_n \cdot Y_n - I_X \cdot I_Y)^2 \\ = \int_0^{+\infty} P(|X_n \cdot Y_n - I_X \cdot I_Y| > \sqrt{t}) dt \\ = \int_0^{+\infty} P(|(X_n - I_X)(Y_n - I_Y) + I_Y(X_n - I_X) \\ + I_X(Y_n - I_Y)| > \sqrt{t}) dt \\ \leq E|X_n - I_X|^2 + E|Y_n - I_Y|^2 \\ + I_Y^2 E|X_n - I_X|^2 + I_X^2 E|Y_n - I_Y|^2, \end{aligned}$$

then

$$\sum_{n \in \mathbb{N}} U_n E(X_n \cdot Y_n - I_X \cdot I_Y)^2$$

$$\begin{aligned}
&\leq \sum_{n \in \mathbb{N}} U_n E|X_n - I_X|^2 + \sum_{n \in \mathbb{N}} U_n E|Y_n - I_Y|^2 \\
&\quad + \sum_{n \in \mathbb{N}} U_n I_Y^2 E|X_n - I_X|^2 + \sum_{n \in \mathbb{N}} U_n I_X^2 E|Y_n - I_Y|^2 \\
&< \infty.
\end{aligned}$$

(3) For $I_X \neq 0$

$$\begin{aligned}
E\left(\frac{1}{X_n} - \frac{1}{I_X}\right)^2 &= \int_0^{+\infty} P\left(\left|\frac{1}{X_n} - \frac{1}{I_X}\right|^2 > t\right) dt \\
&= \int_0^{+\infty} P(|X_n - I_X|^2 > t.(X_n.I_X)^2) dt \\
&\leq \int_0^{+\infty} P(|X_n - I_X|^2 > t) dt \\
&\leq E|X_n - I_X|^2,
\end{aligned}$$

so

$$\sum_{n \in \mathbb{N}} U_n E\left(\frac{1}{X_n} - \frac{1}{I_X}\right)^2 \leq \sum_{n \in \mathbb{N}} U_n E|X_n - I_X|^2 < \infty.$$

■

Now we have two properties which are consequences of the previous calculus rules.

Corollary 2.1: Assume $\lim_{n \rightarrow \infty} U_n = +\infty$, $X_n = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$ and $Y_n = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$, where I_Y is a reel number. We have

- (1) $(X_n.Y_n) = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$;
- (2) $\frac{X_n}{Y_n} = O_{c.s.m}(\frac{1}{\sqrt{U_n}})$, with $I_X \neq 0$.

Proof: Follows directly from Proposition 2.1. ■

3. Theoretical applications

In this section, rates of complete second-order moment convergence for the probability density, the distribution function and the quantile function kernel estimators are established. The following remark is very important to establish these rates.

Remark 3.1: In all that follows, the supposition of the convergence of $\sum_{n \geq 1} U_n h_n^4$ or $\sum_{n \geq 1} \frac{U_n}{h_n^2}$ follows from the fact that, if $h_n < 1$ we have $U_n h_n^4 \leq \frac{U_n}{h_n^2}$, so the convergence of $\sum_{n \geq 1} \frac{U_n}{h_n^2}$ implies the convergence of $\sum_{n \geq 1} U_n h_n^4$. If $h_n > 1$ and $h_n \rightarrow 0$ while $n \rightarrow \infty$, the convergence is obtained.

3.1. Kernel density estimator

Let X_1, X_2, \dots, X_n be independent and identically distributed copies of a random variable X , which has unknown continuous probability density function f . The

kernel density estimator, noted \hat{f}_n , of the unknown density f defined by Parsen [12]; Rosenblatt [11], is given by

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right),$$

where $(h_n)_{n \geq 1}$ is a sequence of positive numbers, usually called a bandwidth or smoothing parameter, and K is an integrable Borel measurable function satisfying $K \geq 0$ and $\int_{\mathbb{R}} K(x) dx = 1$, called kernel. Assume that the kernel K and the density f functions verify the following conditions:

- (H1) $\forall x \in \mathbb{R}, K(x) = K(-x)$;
- (H2) $\sup_{x \in \mathbb{R}} |K(x)| \leq M < \infty$;
- (H3) $\int_{-\infty}^{+\infty} x^2 K(x) dx < \infty$;
- (H4) $f \in C^2$ and $f^{(2)}$ is bounded.

Theorem 3.1: Under (H1)–(H4) and supposing that $\sum_{n \geq 1} U_n h_n^4$ or $\sum_{n \geq 1} \frac{U_n}{h_n^2}$ converges (see Remark 3.1), we have

$$\hat{f}_n(x) - f(x) = O_{c.s.m}\left(\frac{1}{\sqrt{U_n}}\right), \quad (1)$$

Proof: To prove (1), notice first that the mean squared error MSE of \hat{f}_n , defined by

$$MSE(\hat{f}_n(x)) = E(\hat{f}_n(x) - f(x))^2,$$

can be written as

$$MSE(\hat{f}_n(x)) = Bias(\hat{f}_n(x))^2 + Var(\hat{f}_n(x)). \quad (2)$$

Hence to prove (1), it suffices to show that

$$\sum_{n \geq 1} U_n Var[\hat{f}_n(x)] < \infty \quad (3)$$

and

$$\sum_{n \geq 1} U_n Bias(\hat{f}_n(x))^2 < \infty. \quad (4)$$

For the inequality (3)

$$\begin{aligned}
Var[\hat{f}_n(x)] &= \frac{1}{n^2 h_n^2} Var \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \\
&\leq \frac{1}{nh_n^2} E\left[K^2\left(\frac{x - X}{h_n}\right)\right].
\end{aligned}$$

Using the condition (H2) of the kernel K , it follows that

$$Var[\hat{f}_n(x)] \leq \frac{M^2}{nh_n^2}.$$

Therefore

$$\sum_{n \geq 1} U_n Var[\hat{f}_n(x)] \leq \sum_{n \geq 1} \frac{U_n M^2}{h_n^2}.$$

Since $\sum_{n \geq 1} \frac{U_n}{h_n^2}$ converges, so $\sum_{n \geq 1} U_n Var[\hat{f}_n(x)] < \infty$.

On the other hand for the inequality (4), we have

$$\begin{aligned} E[\hat{f}_n(x)] &= E\left[\frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)\right] \\ &= \int_{\mathbb{R}} K(z) f(x - zh_n) dz. \end{aligned}$$

Using Taylor's series expansion of the function f about a point x up to order 3, (H3), (H1) and (H4), one obtain

$$\begin{aligned} E[\hat{f}_n(x)] &= f(x) \int_{\mathbb{R}} K(z) dz - h_n f'(x) \int_{\mathbb{R}} z K(z) dz \\ &\quad + \frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz - \frac{h_n^3}{6} f^{(3)}(\theta) \\ &\quad \times \int_{\mathbb{R}} z^3 K(z) dz, \end{aligned}$$

where θ is a real number between x and zh_n . Hence

$$E[\hat{f}_n(x)] \leq f(x) + \frac{h_n^2}{2} f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz,$$

Thus

$$\begin{aligned} \sum_{n \geq 1} U_n \text{Bias}(\hat{f}_n(x))^2 &\leq \sum_{n \geq 1} \frac{U_n h_n^4}{4} \left[f^{(2)}(x) \int_{\mathbb{R}} z^2 K(z) dz \right]^2 \\ &= C \sum_{n \geq 1} U_n h_n^4 < \infty. \end{aligned}$$

The proof is completed. ■

Remark 3.2: If K is a symmetric compactly supported kernel, we obtain the complete second-order moment convergence of \hat{f}_n under H4.

The choices of U_n and h_n are not arbitrary because their expressions must be selected so that the convergence of the obtained series is ensured.

Example 3.1: By choosing $h_n = n^{-s}$ and $U_n = n^{-\alpha}$, where $s \in]0, 1[$ and $\alpha > 1$, we obtain

$$\sum_{n \geq 1} U_n \text{Var}[\hat{f}_n(x)] \leq \sum_{n \geq 1} \frac{M^2}{n^{1+\alpha-2s}},$$

This Riemann's serie converges if $1 + \alpha - 2s > 1$, thus if $\alpha > 2s$. And

$$\sum_{n \geq 1} U_n (E[\hat{f}_n(x)] - f(x))^2 \leq C \sum_{n \geq 1} \frac{1}{n^{4s+\alpha}},$$

the right-hand side converge if $\alpha > 1 - 4s$.

Consequently, combining the two conditions of right-hand side series, one can obtain (1) for $\alpha > 2$. Indeed, it can be checked that

- (1) if $(1 - 4s) \in]-3, 0]$, then $2s \in]0.5, 1]$;
- (2) if $(1 - 4s) \in]0, 1]$, then $2s \in]0, 0.5]$;
- (3) if $2s \in]1, 2]$, then $1 - 4s \in]-3, -1]$;

(4) if $2s = 1 - 4s$, then $s = \frac{1}{6}$.

So, the condition $\alpha > 2$ is always verified.

Corollary 3.1: Under (H1)–(H4), we have $\hat{f}_n - f = O_{a.c}(\frac{nh_n}{n^\alpha(\log n)^2})$, $\alpha > 1$.

Proof: For the optimal bandwidth $h_n = n^{-\frac{1}{5}}$ and a rate of convergence $U_n = \frac{nh_n}{n^\alpha(\log n)^2}$, which satisfies the inequality $U_n < \frac{nh_n}{\log n}$, where $\frac{nh_n}{\log n}$ is the rate of almost complete convergence of the density kernel estimator, one obtain

$$\sum_{n \geq 1} U_n \text{Var}[\hat{f}_n(x)] \leq \sum_{n \geq 1} \frac{M^2}{n^{\alpha-\frac{1}{5}}(\log n)^2}$$

and

$$\sum_{n \geq 1} U_n (\text{Bias}[\hat{f}_n(x)])^2 \leq C \sum_{n \geq 1} \frac{1}{n^\alpha(\log n)^2}.$$

Combining the convergence conditions of the two series in the right hand side, we obtain (1) for $\alpha > 1$. ■

3.2. Kernel distribution function estimator

Let X_1, X_2, \dots, X_n be independent and identically distributed copies of a random variable X , which has unknown continuous probability density f and distribution F functions. The kernel distribution function estimator \hat{F}_n , that was proposed by Nadaraya [18], can be obtained by integrating the kernel density estimator \hat{f}_n , as follows:

$$\hat{F}_n(x) = \int_{-\infty}^x \hat{f}_n(t) dt = \frac{1}{n} \sum_{i=1}^n H\left(\frac{x - X_i}{h_n}\right),$$

where the function H is defined from the kernel K as

$$H(x) = \int_{-\infty}^x K(t) dt.$$

Function H is a cumulative distribution function because K is a probability density function.

Assume that the following hypothesis are satisfied:

- (H5) $\int_{-\infty}^{+\infty} K(x)H(x) dx < \infty$;
- (H6) $\int_{-\infty}^{+\infty} x^2 K(x)H(x) dx < \infty$;
- (H7) $\int_{-\infty}^{+\infty} x K(x)H(x) dx \geq 0$;
- (H8) $F \in C^2$ and F' is bounded.

Then Theorem 3.2 states the complete second-order moment convergence of \hat{F}_n to F .

Theorem 3.2: Under (H1), (H3), (H5)–(H8) and supposing that $\sum_{n \geq 1} U_n h_n^4$ or $\sum_{n \geq 1} \frac{U_n}{h_n^2}$ converges (see Remark 3.1), we have

$$\hat{F}_n(x) - F(x) = O_{c.s.m}\left(\frac{1}{\sqrt{U_n}}\right). \quad (5)$$

Proof: To prove (5), we use the same argument used to check (1). First for the bias, we have

$$\begin{aligned} E[\hat{F}_n(x)] &= E\left[\frac{1}{n} \sum_{i=1}^n H\left(\frac{x - X_i}{h_n}\right)\right] \\ &= \int_{\mathbb{R}} H\left(\frac{x - z}{h_n}\right) f(z) dz. \end{aligned}$$

Using integration by part, substitution $\frac{x-z}{h_n} = y$, a Taylor series expansion of the function F about the point x up to order 2, (H1), (H3) and (H8), we obtain

$$\sum_{n \geq 1} U_n \text{Bias}(\hat{F}_n(x))^2 \leq C \sum_{n \geq 1} U_n h_n^4 < \infty.$$

Now for the variance

$$\begin{aligned} \text{Var}[\hat{F}_n(x)] &= \frac{1}{n} \text{Var} \sum_{i=1}^n H\left(\frac{x - X_i}{h_n}\right) \\ &\leq E\left[H^2\left(\frac{x - X}{h_n}\right)\right]. \end{aligned}$$

Using (H1), (H5)–(H8), integration by part, substitution $\frac{x-z}{h_n} = y$, one have

$$\begin{aligned} \text{Var}[\hat{F}_n(x)] &\leq \frac{2h_n^2 F(x) \int_{-\infty}^{+\infty} K(y) H(y) dy + h_n^4 f'(x) \int_{-\infty}^{+\infty} y^2 K(y) H(y) dy}{h_n^2} \\ &\leq \frac{C_1 h_n^2 + C_2 h_n^4}{h_n^2}, \end{aligned}$$

where C_1 and C_2 are two constants. Since (h_n) converges to 0, so for every $\epsilon > 0$ there exists n_0 such that $h_n \leq \epsilon$ for all $n \geq n_0$, then

$$\sum_{n \geq 1} U_n \text{Var}[\hat{F}_n(x)] \leq M \sum_{n \geq 1} \frac{U_n}{h_n^2},$$

and the desired result is obtained \blacksquare

Remark 3.3: When K is symmetric and has a compact support $[-1, 1]$, the proprieties of H are given in Baszczynska [20], then Hypothesis H5 and H7 are verified. We can obtain the MSE of the kernel distribution estimator \hat{F}_n with the assumptions used by Azzalini [21] and the kernel satisfying the above assumptions, we have the squared bias

$$\text{Bias}^2[\hat{F}_n(x)] = \frac{1}{4} h_n^4 f'^2(x) \left(\int_{-1}^1 y^2 K(y) dy \right)^2 + o(h_n^4),$$

and the variance

$$\begin{aligned} \text{Var}[\hat{F}_n(x)] &= \frac{1}{n} F(x)(1 - F(x)) \\ &\quad - 2 \frac{1}{n} h_n f(x) \int_{-1}^1 y K(y) H(y) dy + o\left(\frac{h_n}{n}\right). \end{aligned}$$

Then the MSE is given by

$$\begin{aligned} \text{MSE}[\hat{F}_n(x)] &= \frac{1}{4} h_n^4 f'^2(x) \left(\int_{-1}^1 y^2 K(y) dy \right)^2 \\ &\quad + \frac{1}{n} F(x)(1 - F(x)) - 2 \frac{1}{n} h_n f(x) \\ &\quad \times \int_{-1}^1 y K(y) H(y) dy + o\left(h_n^4 + \frac{h_n}{n}\right). \end{aligned}$$

So under H8 we obtain

$$\sum_{n \geq 1} U_n \text{MSE}(\hat{F}_n(x)) < \infty.$$

The next corollary gives a new rate of the almost complete convergence of \hat{F}_n to F .

Corollary 3.2: Under (H1), (H3), (H5)–(H8), we have

$$\hat{F}_n - F = O_{c.c} \left(\sqrt{\frac{\log n}{n^{2\alpha} h_n^8}} \right), \quad \alpha > 1.$$

Proof: Remarking that $\sqrt{\frac{\log n}{n^{2\alpha} h_n^8}} < \sqrt{\frac{\log n}{n}}$, where $\sqrt{\frac{\log n}{n}}$ is the rate of the almost complete convergence of \hat{F}_n to F , we get

$$\sum_{n \geq 1} U_n E(\hat{F}_n(x) - F(x))^2 \leq \sum_{n \geq 1} \sqrt{\frac{\log n}{n^{2\alpha} h_n^8}} h_n^4 \leq \sum_{n \geq 1} \frac{\sqrt{\log n}}{n^\alpha}.$$

The last series converge for $\alpha > 1$. \blacksquare

Example 3.2: For the optimal bandwidth $h_n = n^{-\frac{1}{3}}$ and $U_n = n^{-\alpha}$ where $\alpha > 1$, one obtain

$$\sum_{n \geq 1} U_n \text{Bias}(\hat{F}_n(x))^2 \leq C \sum_{n \geq 1} n^{-\frac{4}{3}} n^{-\alpha}$$

and

$$\sum_{n \geq 1} U_n \text{Var}[\hat{F}_n(x)] \leq M \sum_{n \geq 1} \frac{n^{-\alpha}}{n^{\frac{2}{3}}}.$$

The two right-hand side series converge simultaneously if $\alpha > \frac{5}{3}$.

3.3. Kernel quantile function estimator

Let X_1, X_2, \dots, X_n be independent and identically distributed copies of a random variable with absolutely continuous distribution function F . Denoting $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ the corresponding order statistics. The quantile function, noted Q , is defined to be the left

continuous inverse of F , given by

$$Q(p) = \inf\{x : F(x) \geq p\}, 0 < p < 1.$$

A kernel quantile estimator, based on the Nadaraya [18] kernel distribution function estimator \hat{Q}_n , is defined as

$$\hat{Q}_n(p) = \inf\{x : \hat{F}_n(x) \geq p\}, \quad 0 < p < 1.$$

and given by

$$\hat{Q}_n(p) = \sum_{i=1}^n X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{h_n} K\left(\frac{x-p}{h_n}\right) dx,$$

where K is a density function, while $h_n \rightarrow 0$ as $n \rightarrow \infty$.

Our result of convergence is based on the expression of the MSE of the kernel quantile estimator, given by Sheather and Marron [22] (Theorem 1, p. 5), when p is in the interior of $(0, 1)$, under conditions that the kernel K is symmetric about 0 with compact support, and $Q^{(2)}$ is continuous in a neighbourhood of p .

Theorem 3.3: Supposing that $\sum_{n \geq 1} U_n h_n^4$ or $\sum_{n \geq 1} \frac{U_n}{h_n^2}$ converges (see Remark 3.1), we have

$$\hat{Q}_n(x) - Q(x) = O_{c.s.m} \left(\frac{1}{\sqrt{U_n}} \right). \quad (6)$$

Proof: Building on Falk [23] and David [24], Sheather and Marron [22] give the expressions of bias and variance of \hat{Q}_n as

$$\begin{aligned} \text{Bias}[\hat{Q}_n(x)] &= \frac{1}{2} h_n^2 Q^{(2)}(p) \int_{-\infty}^{+\infty} x^2 K(x) dx \\ &\quad + o(h_n^2) + o(n^{-1}), \end{aligned}$$

and

$$\begin{aligned} \text{Var}[\hat{Q}_n(x)] &= \frac{1}{n} p(1-p) (Q'(p))^2 \\ &\quad - \frac{1}{n} h_n (Q'(p))^2 \int_{-\infty}^{+\infty} x K(x) H(x) dx \\ &\quad + o\left(\frac{h_n}{n}\right). \end{aligned}$$

So

$$\sum_{n \geq 1} U_n \text{Bias}(\hat{Q}_n(x))^2 \leq C \sum_{n \geq 1} U_n h_n^4 < \infty$$

and

$$\begin{aligned} \sum_{n \geq 1} U_n \text{Var}[\hat{Q}_n(x)] &\leq M \sum_{n \geq 1} \frac{U_n}{n} \\ &\leq M \sum_{n \geq 1} \frac{U_n}{h_n^2} < \infty. \end{aligned}$$

Finally we have obtained the almost complete convergence and the complete second-order moment convergence of the kernel quantile estimator \hat{Q}_n . ■

Example 3.3: In the same way of Example 3.2 and using the same h_n and U_n , we obtain

$$\hat{Q}_n(x) - Q(x) = O_{c.s.m} \left(\frac{1}{\sqrt{U_n}} \right).$$

4. Simulation study

In this section, to present the performance of the new rate of convergence for a finite-size sample, we realize a simulation study. We give a visual impression of the quality of convergence by calculating the correspondent MSE together with the value of the rate of complete second-moment convergence and the value of the rate of almost complete convergence, based on a sample obtained from two theoretical models: Gamma kernel density estimates and the innovation one said Laplace kernel density developed by Khan and Akbar [25], inspired to Chen's idea [26]. Defined by

$$K_{\text{Laplace}(x, h^{1/2})}(u) = \frac{1}{2\sqrt{h}} \exp\left(-\frac{|u-x|}{\sqrt{h}}\right).$$

In the second part, Normal and Epanechnikov kernel distribution estimators are used with optimal bandwidth and different sizes of normal and exponential sample to give a performance of the new rate of convergence. Finally, with the Normal model, one perform the quantile (25%, 50%, 75%) estimates for the same sample size.

4.1. Kernel density MSE

We propose two schemes of Kernel estimations, Laplace and Gamma kernel density's estimates with optimal bandwidth $h = n^{-1/5}$. One conduct simulations of samples data from Exponential and Gamma density with sample sizes $n = 100, n = 150, 200, 250, 300, 500, 800, 1000$. We summarize the numerical calculations in Table 1.

The convergence CSM rate is more efficient in terms of speed towards zero. Even if the rate is almost efficient in the AC convergence, we attend the good behaviour of the new rate for both kernel models.

According to the results obtained in Table 1, we remark that the values of the CSMC rate of the kernel density estimator for Laplace kernel (resp. Gamma kernel) are closer to the Exp MSE (resp. Gamma MSE) values, than the ACC rate one.

4.2. Kernel distribution MSE

In this case, the Normal kernel distribution and Epanechnikov kernel distribution with optimal bandwidth $h = n^{-1/3}$ are compared to Normal sample. The size varies between 100 and 1000. We summarize the numerical calculations in Table 2. Here too, we notice the fast convergence speed of the CSM rate. We also notice its

Table 1. Mean simulated values and rate of convergences for density.

Kernel density	Sample size	Gamma MSE	Exp MSE	CSMC rate	ACC rate
Laplace	100	0.043479007	0.001242875	0.0018771917	0.013831660
	150	0.033926113	0.003267882	0.0011938521	0.011722195
	200	0.032411091	0.002689498	0.0008729775	0.010465886
	250	0.027020496	0.004629636	0.0006875970	0.009604571
	300	0.019391081	0.005242242	0.0005671386	0.008964638
	500	0.008215233	0.008617148	0.000334113	0.007428690
	800	0.023472428	0.007204569	0.000207819	0.006286744
	1000	0.026401011	0.006968009	0.000166466	0.005818125
Gamma	100	0.4151408	0.026409057	0.001877192	0.013831660
	150	0.2976568	0.050039754	0.001193852	0.011722195
	200	0.2371582	0.040335448	0.000872978	0.010465886
	250	0.2188774	0.002523508	0.000687597	0.009604571
	300	2.6515305	0.031110228	0.000567139	0.008964638
	500	0.4655139	0.028309673	0.000334113	0.007428690
	800	1.0592910	0.048968084	0.000207819	0.006286744
	1000	0.6732155	0.030791537	0.000166466	0.005818125

Table 2. Mean simulated values and rate of convergences for distribution.

Kernel distribution	Sample size	Normal MSE	Exp MSE	CSMC rate	ACC rate
Normal	100	0.001111	0.00106656	0.0001468843	0.13788621
	200	0.000311	0.00028120	0.0047592004	0.16276236
	300	0.000166	0.00017937	0.0030768525	0.13788621
	500	0.000032	0.00003562	0.0017697362	0.11148639
	800	0.000070	0.00645965	0.0010607176	0.09140987
	1000	0.000052	0.00005699	0.0008311291	0.08311291
Epanechnikov	100	0.000282	0.00028770	0.0001468843	0.13788621
	200	0.000188	0.00024757	0.0047592004	0.16276236
	300	0.000087	0.00004802	0.0030768525	0.13788621
	500	0.000410	0.00043860	0.0017697362	0.11148639
	800	0.000191	0.00003853	0.0010607176	0.09140987
	1000	0.000096	0.00007708	0.0008311291	0.08311291

similarity to the values of the MSE in both distributions. We remark that the CSM rate gives a good results for the Normal kernel.

By the results given in Table 2, we notice that, in the both cases normal and Epanechnikov kernels distribution, CSMC rate values of the kernel distribution estimator, are closer to the normal MSE and Exp MSE, respectively values than that of the ACC rate.

4.3. Kernel quantile MSE

Now we calculate the MSE of the normal kernel quantile estimator. The size varies between 100 and 1000. We summarize the numerical calculations in Table 3. We also remark that the CSMC rate values are closer to the Kernel quantile estimators than the ACC rate values.

Eventually, we conclude that, in all cases, the CSMC rate of the kernel density, distribution and quantile estimators, gives better results than the ACC rate of the same estimators.

5. Real data analysis

Female infertility and BMI: This study aims to investigate the body mass index (BMI) of infertile women of child-bearing age. We use data from 200 participants from the Ben Badis University Hospital Centre of Constantine, Algeria, in 2018.

The first step is to establish the conformity test between the sample and a normal distribution. One obtain, $D = 0.079226$ smaller than p -value 0.1623. This implies the use of the density and the distribution of the normal law. The second step is to estimate the density and the distribution functions and to represent them in a graph (Figure 1). For the MSE between the real density and a normal kernel density is equal to 0.003128583 and the rate of CSM convergence is 0.0008729775 and the AC convergence is 0.06541179. We obtain for the distribution, $MSE = 0.003128583$ and the rate of CSM convergence is 0.0047592 and the AC convergence is 0.1627624. For the same rate of convergence, one have

Table 3. Mean simulated values and rate of convergence for quantiles.

Sample size	MSE 25%	MSE 50%	MSE 75%	CSMC rate	ACC rate
100	0.7203108	0.01220052	0.28767206	0.0001468843	0.13788621
200	0.2117512	0.0127754	0.507382	0.0047592004	0.16276236
300	0.449008368	0.000455053	0.372990054	0.0030768525	0.13788621
500	0.21501058	0.0165624	0.07083821	0.0017697362	0.11148639
800	0.17837084	0.06426792	0.92478425	0.0010607176	0.09140987
1000	0.43201245	0.02671836	0.87488135	0.0008311291	0.08311291

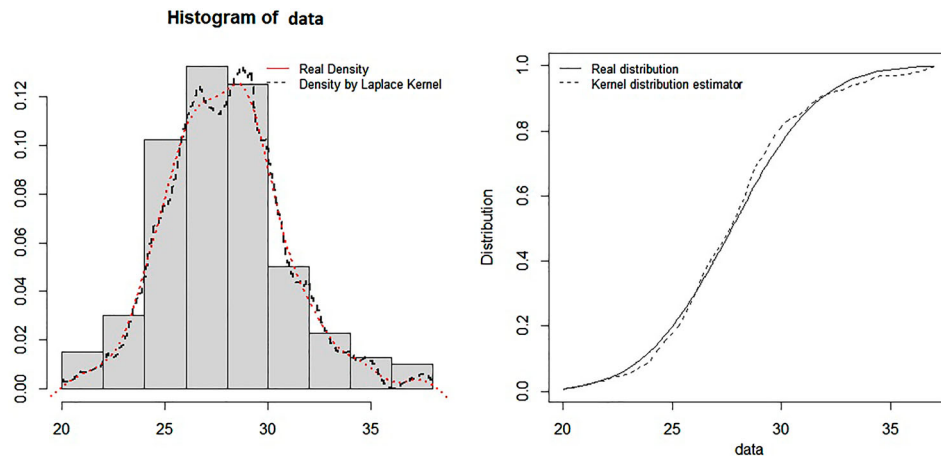


Figure 1. Density and distribution real data with density and distribution kernel estimate.

the MSE quantiles given by (25%, 0.0260281), (50%, 0.04186125), (75%, 0.058767206). Note that the results of the real data come supported those of the simulation.

6. Conclusion

The present work proposed a new method to obtain the convergence rate of the MSE that is much more efficient in the CSMC case. Indeed, the CSM convergence rate gives better results than that of AC convergence. The previous results indicate that the choice of this type of convergence using kernel method estimate is a good alternative to the almost complete one. We can apply this type of convergence in any estimation that requires the study of the MSE, such as for example neural network, least squares method, ..., and following the suggestion, it can be applied in neutrosophic statistics developed in Smarandache [27] and Afzal et al [28]. Moreover, maybe we can apply this type of convergence to extend the theorem of the law of large numbers.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- [1] Chow D. On the rate of moment convergence of sample function of a random variables with bounded support. *Bull Inst Math Acad Sin.* 1988;16:177–201.
- [2] Liang H, Li D, Rosalsky A. Complete moment and integral convergence for sums of negatively associated random variables. *Acta Math Sin (Engl Ser).* 2010;26(3):419–432.
- [3] Qui D, Chen P. Complete moment convergence for i.i.d random variables. *Statist Probab Lett.* 2014;91:76–82.
- [4] Hsu P, Robbins H. Complete convergence and the law of large numbers. *Proc Natl Acad Sci USA.* 1947;33(2):25–31.
- [5] Gut A, Stadtmüller U. An intermediate Baum–Katz theorem. *Statist Probab Lett.* 2011;81(10):1486–1492.
- [6] Gut A. Marcinkiewicz laws and convergence rates in the law of large numbers for random variables with multidimensional indices. *Ann Probab.* 1978;6:469–482.
- [7] Gut A. Convergence rates for probabilities of moderate deviations for sums of random variables with multidimensional indices. *Ann Probab.* 1980;8(2):298–313.
- [8] Li D, Rao MB, Jiang T, et al. Complete convergence and almost sure convergence of weighted sums of random variables. *J Theoret Probab.* 1995;8(1):49–76.
- [9] Sung SH. Complete convergence for weighted sums of random variables. *Statist Probab Lett.* 2007;77(3):303–311.
- [10] Sung SH, Volodin A. On the rate of complete convergence for weighted sums of arrays of random elements. *J Korean Math Soc.* 2006;43(4):815–828.
- [11] Rosenblatt M. Remarks on some non parametric estimates of a density function. *Ann Math Statist.* 1956;27(3):832–837.
- [12] Parzen E. On estimation of a probability density function and mode. *Ann Math Stat.* 1962;33(3):1065–1076.
- [13] Habbema JDF, Hermans J, Vanden Broeke K. A stepwise discriminant analysis program using density estimation. In: Bruckmann G, editor. *Comp stat 1974, Proceedings in Computational Statistics.* Vienna: Physica Verlag; 1974. p. 101–110.
- [14] Hall P, Kang K-H. Bandwidth choice for non parametric classification. *Ann Stat.* 2005;33:284–306.
- [15] Hall P, Wand MP. On non parametric discrimination using density differences. *Biometrika.* 1988;75(3):541–547.
- [16] Ghosh AK, Chaudhuri P. Optimal smoothing in kernel analysis discriminant. *Stat Sin.* 2004;14:457–483.
- [17] Ghosh AK, Hall P. On error-rate estimation in nonparametric classification. *Stat Sin.* 2008;18:1081–1100.
- [18] Nadaraya EA. Some new estimates for distribution function. *Theory of Probab. Appl.* 1964;9:497–500.
- [19] Parzen E. Non parametric statistical data modelling. *J Amer Stat Assoc.* 1979;74(365):105–131.
- [20] Baszczynska A. Kernel estimation of cumulative distribution function of a random variable with bounded support. *Statist Trans.* 2016;17:541–556.
- [21] Azzalini A. A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika.* 1981;68(1):326–328.
- [22] Sheather SJ, Marron JS. Kernel quantile estimators. *J Amer Statist Assoc.* 1990;85(410):410–416.
- [23] Falk M. Relative deficiency of Kernel type estimators of quantiles. *Ann Statist.* 1984;12:261–268.
- [24] David HA. *Order statistics.* 2nd ed. New York: John Wiley; 1981.
- [25] Khan JA, Akbar A. Density estimation by Laplace kernel. Working paper. Department of Statistics, Bahauddin Zakariya, Multan, Pakistan. 2021.

- [26] Chen SX. Probability density function estimation using gamma kernels. *Ann Inst Statist Math.* [2000](#);52(3): 471–480.
- [27] Smarandache F. Neutrosophic statistics vs. classical statistics, section in *Nidus Idearum/superluminal physics*. Vol. 7, 3rd ed. 2019. p. 117.
- [28] Afzal U, Alrweili H, Ahamd N, et al. Neutrosophic statistical analysis of resistance depending on the temperature variance of conducting material. *Sci Rep.* [2021](#);11(1):Article ID 23939.
- [29] Yu Y, et al. On the complete convergence for uncertain random variables. *Soft Comput.* [2022](#);26(3):1025–1031.