International Journal of Statistics and Applications 2019, 9(2): 45-52
DOI: 10.5923/j.statistics.20190902.01

A Study of the Ability of the Kernel Estimator of the Density Function for Triangular and Epanechnikov Kernel or Parabolic Kernel

Didier Alain Njamen Njomen*, Ludovic Kakmeni Siewe

Department of Mathematics and Computer's Science, Faculty of Science, University of Maroua, Maroua, Cameroon

Abstract In this paper, we are interested in the nonparametric estimation of probability density. From the « Rule of thumb » method, we were able to determine the smoothing parameter h_n of the Parsen-Rosenblatt kernel estimator for the density function. Our study is illustrated by numerical simulations to show the performance of the triangular core and Epanechnikov or parabolic density estimator studied.

Keywords Density function, Smoothing parameter, Method of Rule of thumb

1. Introduction

The theory of the estimator is one of the major concerns of statisticians. It is a fundamental element of statistics. It allows to generalize observed results. There are two approaches:

- the parametric approach, which considers that the models are known, with unknown parameters. The law of the studied variable is supposed to belong to a family of laws which can be characterized by a known functional form (distribution function, density f, ...) which depends on one or several unknown parameters to estimate;
- the non-parametric approach, which makes no assumptions about the law or its parameters.

Knowledge about the model (non-parametric model) is not generally accurate, i.e., we do not have enough information on this model unlike the parametric model, which is often the case in practice. In this situation, it is natural to want to estimate one of the functions describing the model, either generally the distribution function or the density function (for the continuous case): this is the objective of the functional estimation.

Since the works of Rosenblatt (1956) and Parzen (1962) on non-parametric estimators of density functions, the kernel method has been widely used in such works, as Prakasa Rao (1983), Devroye and Györfi (1985), Silverman (1986), Scott (1992), Bosq and Lecoutre (1987), Wand and Jones (1995), Benchoulak (2012), Roussas (2012) and the references cited in these publications. Based on the study of the local empirical process indexed by certain classes of functions, Deheuvels and Mason (2004) have established probability convergence speed for deviations of these estimators from their expectations.

The central purpose of this article is to show the performance of the kernel density estimator for triangular and Epanechnikov or parabolic kernels.

For us to attain this aim, the present article is divided into three (3) parts: firstly, as revision, we are going to introduce some different modes of convergences and give some Bernstein exponential inequalities which have permitted us to regulate the limit of deviations of the estimators compared to their hopes. This is why we will mention three (3) non-parametric methods of density estimation: the histogram method, the simple estimation method and the kernel method (Parzen-Rosenblatt estimator) on which will be our focus and which can be considered as an extension of the estimator by the histogram. We will also present the statistical properties of every estimation method. In the second part, using the Rule of thumb method (studied in Deheuvels (1977), and Sheather, Jones, and Marron (1996)), we will determine the h_n smoothing parameter. Finally, using numerical simulations, we will explain the performance of the studied estimator.

* Corresponding author:
didiernjamen1@gmail.com (Didier Alain Njamen Njomen)
Published online at http://journal.sapub.org/statistics

2. Density Function Estimator

2.1. The Parzen-Rosenblatt Kernel Estimator

For the fact

$$f(x) \simeq \frac{F(x+h)-F(x-h)}{2h} \; for\; h_n \; small, \tag{1}$$

Rosenblatt in 1956, has given an estimator of f by replacing F by it's estimator F_n, so:

$$f_n(x,h_n) = \frac{F_n(x+h)-F_n(x-h)}{2h}, \tag{2}$$

where F_n is the empirical function of distribution.

This estimator can also be written as:

$$\begin{aligned} f_n(x,h_n) &= \sum_{i=1}^{n} \frac{\mathbb{1}_{\{x-h<X_i\leq x+h\}}}{2nh} \\ &= \frac{1}{2nh}\sum_{i=1}^{n} \mathbb{1}_{\left\{-1<\frac{x-X_i}{h}\leq 1\right\}} \\ &= \frac{1}{2nh}\sum_{i=1}^{n} K_0\left(\frac{x-X_i}{h}\right), \end{aligned} \tag{3}$$

with $K_0(u) = \frac{1}{2}\mathbb{1}_{\{-1<u\leq 1\}}$.

In this same article, Rosenblatt (1956) measured the quality of this estimator, by calculating its bias and its variance, given respectively by

$$\begin{aligned} Bias f_n(x,h_n) &= \mathbb{E}[f_n(x,h_n) - f(x)] \\ &= \frac{1}{2h}\mathbb{E}(F_n(x+h) - F_n(x-h)) - f(x) \\ &= \frac{1}{2h}\mathbb{E}(F(x+h) - F(x-h)) - f(x) \end{aligned} \tag{4}$$

and

$$\begin{aligned} \mathbb{V}ar[\, f_n(x,h_n)] &= \frac{1}{4nh_n^2}\begin{bmatrix} F(x+h_n)(1-F(x+h_n) \\ -F(x-h_n)\big(1-F(x-h_n)\big)\end{bmatrix} \\ &\quad -\frac{1}{4nh_n^2}\begin{bmatrix} 2F(\inf((x+h_n),(x+h_n))) \\ +2\,F(x+h_n)\,F(x-h_n)\end{bmatrix} \end{aligned} \tag{5}$$

We notice that if $h_n \to 0$ and $nh_n \to \infty$ when $n \to \infty$, we have:

$$\lim_{n\to\infty} \mathbb{E}[\, f_n(x,h_n)] = f(x)$$

and

$$\lim_{n\to\infty} \mathbb{V}ar[\, f_n(x,h_n)] = 0,$$

therefore, $f_n(x,h_n)$ is a consistent estimator.

By putting $h_n = a_{k+1} - a_k$, we notice that the estimator of f on $[a_k, a_{k+1}]$ does not present the problem of the choice of origin a_0 as is the case of the histogram but it has the disadvantage of being discontinuous at points $X_i \pm h$.

The generalization of this estimator had been introduced by Parzen since 1963 by performing

$$f_n(x,h_n) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{X_i-x}{h}\right), \tag{6}$$

where h_n, called window, is a strict sequence of positive real tending to zero when $n \to \infty$ (called window) and K is a measurable function defined from $\mathbb{R} \to \mathbb{R}$, called kernel.

2.2. Properties of the Estimator

The pillar of the first results of the convergence of this estimator is the theorem of Bochner (1955). The estimator kernel of the density depends on two (2) parameters: the window h_n and the kernel K. The kernel K establishe he aspect of neighborhood of x and h_n, controls the wideness of this neighborhood, so h_n is the first parameter to have good asymptotic properties. Nevertheless, the kernel K must not be neglected. As the works of Parzen (1962) on the consistence of this estimator shows, this properties is obtained after having studied the asymptotic bias of the variance and the following decomposition:

$$\begin{aligned} \mathbb{E}[f_n(x,h_n) - f(x)]^2 &= \mathbb{V}ar[f_n(x,h_n)] \\ &+[Biais\{f_n(x,h_n)\}]^2. \end{aligned} \tag{7}$$

After that, we suppose that K is a kernel verifying the following conditions:

(K.1) K is limitted, which means $Sup_{x\in\mathbb{R}}|K(x)| < \infty$;

(K.2) $\lim |x|K(x) = 0$, when $|x| \to \infty$;

(K.3) $K \in L_1(\mathbb{R})$, meaning that $\int_{\mathbb{R}}|K(x)|dx < \infty$;

(K.4) $\int_{\mathbb{R}} K(x)dx = 1$;

(K.5) K is bounded, integrable and a compactly support.

2.2.1. Study of Bias

The bias of $f_n(x, h_n)$ is given by the following result:

Proposition 1 (Parzen, 1962)

Under the hypothesis (K.1), (K.2), (K.3) and (K.4) above, and if f is continuous, then:

$$\forall\, x \in \mathbb{R}, \lim_{n\to\infty} \mathbb{E}[f_n(x, h_n)] = f(x). \tag{8}$$

We notice that the bias of the estimator converges to zero when the window turns to zero. Futhermore, in view of his expression we notice that it does not depend on the number of variables, but mostly on the kernel K.

2.2.2. Study of the Variance of $f_n(x, h_n)$

The variance of $f_n(x, h_n)$ is given by the following result:

Proposition 2 (**Parzen, 1962**)

Under the conditions K.1), (K.2), (K.3) and (K.4), and if f is continuous in all the points x of $\mathbb{R}$, then we have:

$$\lim_{n\to\infty} \mathbb{V}ar[f_n(x, h_n)] = 0. \tag{9}$$

These two (2) propositions imply the convergence in quadratic average, and thus principally the consistence of the estimator.

3. Choice of Smoothing Parameter

In this section, we study the choice of the smoothing parameter h_n by the « Rule of thumb » method and give it's result for the triangular kernel K and the Epanechnikov or parabolic kernel. In order to obtain these results, we determine at priori the mean and integral quadratic errors of $f_n(x, h_n)$.

3.1. Method of the Mean Quadratic Error Criteria of $f_n(x, h_n)$

The mean square error (MSE) is a measure permitting the evaluation of the similarities of f_n relative to the unknown density function f at a given point x of $\mathbb{R}$. Our aim being to minimize the following quantities:

$$MSE\big(f_n(x,h_n)\big) = \mathbb{E}(f\_n\ (x,h\_n\ ) - f(x))^2. \quad (10)$$

The development of this expression gives us:

$$MSE\big(f_n(x,h_n)\big) = \mathbb{V}ar\big(f_n(x,h_n)\big) + (Bias(f_n(x,h_n)))^2. \quad (11)$$

For us to attain our aim, we calculate the mean and the variance of $f_n(x,h_n)$:

a) Calculating the Mean of $\mathbf{f_n(x,h_n)}$

The calculation of the means gives us:

$$\begin{aligned}\mathbb{E}\big(f_n(x,h_n)\big) &= \mathbb{E}\left[\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)\right]\\ &= \frac{1}{h_n}\mathbb{E}\left[\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)\right]\\ &= \frac{1}{h_n}\int_{\mathbb{R}} K\left(\frac{x-X_i}{h_n}\right) f(u)du.\end{aligned}$$

In putting: $y = \frac{x-u}{h_n}$ i.e. $dy = -\frac{du}{h_n}$, we have:

$$\mathbb{E}\big(f_n(x,h_n)\big) = \int_{\mathbb{R}} K(y) f(x-h_n y)dy. \quad (12)$$

By making Taylor's limit development at order 2 at the point y = 0 of $f(x-h_n y)$, we obtain:

$$f(x-h_n y) = f(x) - \frac{h_n y}{1!}f'(x) + \frac{h_n^2 y^2}{2!}f''(x) + O(h_n^2). \quad (13)$$

So

$$\begin{aligned}&\mathbb{E}\big(f_n(x,h_n)\big) =\\ &\int_{\mathbb{R}} K(y)\left[f(x) - \frac{h_n y}{1!}f'(x) + \frac{h_n^2 y^2}{2!}f''(x) + O(h_n^2)\right]dy\\ &= f(x)\int_{\mathbb{R}} K(y)dy - h_n f'(x)\int_{\mathbb{R}} yK(y)\\ &\quad + \frac{h_n^2}{2}f''(x)\int_{\mathbb{R}} y^2K(y) + O(h_n^2)\\ &= f(x) + \frac{h_n^2}{2}f''(x)\int_{\mathbb{R}} y^2K(y)\,dy + O(h_n^2). \quad (14)\end{aligned}$$

Hence, we have:

$$\mathbb{E}\big(f_n(x,h_n)\big) = f(x) + \frac{h_n^2}{2}f''(x)\int_{\mathbb{R}} y^2K(y)\,dy + O(h_n^2). \quad (15)$$

According to the expression of the mean above we have:

$$\mathbb{E}\big(f_n(x,h_n)\big) - f(x) = \frac{h_n^2}{2}f''(x)\int_{\mathbb{R}} y^2K(y)\,dy + O(h_n^2). \quad (16)$$

But the bias is given by:

$$B\text{ias}\big(f_n(x,h_n)\big) = \mathbb{E}\big(f_n(x,h_n)\big) - f(x), \quad (17)$$

So

$$Bias\big(f_n(x,h_n)\big) = \frac{h_n^2}{2}f''(x)\int_{\mathbb{R}} y^2K(y)\,dy + O(h_n^2). \quad (18)$$

b) Calculating of the Variance of $\mathbf{f_n(x,h_n)}$

The variance of $f_n(x,h_n)$ is given by:

$$\begin{aligned}&\mathbb{V}ar\big(f_n(x,h_n)\big) = \mathbb{V}ar\left[\frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)\right]\\ &= \frac{1}{n^2h_n^2}\sum_{i=1}^{n}\mathbb{V}ar\left[K\left(\frac{x-X_i}{h_n}\right)\right]\\ &= \frac{1}{n^2h_n^2}\sum_{i=1}^{n}\left[\mathbb{E}\left(K\left(\frac{x-u}{h_n}\right)\right)^2\right] - \frac{1}{n^2h_n^2}\sum_{i=1}^{n}\left[\mathbb{E}\left(K\left(\frac{x-u}{h_n}\right)\right)\right]^2\\ &= \frac{1}{nh_n^2}\int_{\mathbb{R}}\left[\left(K\left(\frac{x-u}{h_n}\right)\right)^2\right]f(u)du - \frac{1}{nh_n^2}\left[\int_{\mathbb{R}} K\left(\frac{x-u}{h_n}\right)f(u)du\right]^2 \quad (19)\end{aligned}$$

In putting $y = \frac{x-u}{h_n}$ i.e. $dy = -\frac{du}{h_n}$, we have:

$$\mathbb{V}ar\big(f_n(x,h_n)\big) = \frac{1}{nh_n}\int_{\mathbb{R}}[K(y)]^2 f(x-h_n y)dy - \frac{1}{n}\left[\int_{\mathbb{R}} K(y)\,f(x-h_n y)dy\right]^2.$$

Using an analogue working with the calculation of the mean above and using the Proposition 2, we obtain a new expression of the variance:

$$\mathbb{V}ar\big(f_n(x,h_n)\big) = \frac{f(x)}{nh_n}\int_{\mathbb{R}} K^2(y)dy + O(h_n^{-1}). \quad (20)$$

So, the Mean Square Error (MSE) is:

$$\begin{aligned}MSE\big(f_n(x,h_n)\big) &= \frac{f(x)}{nh_n}\int_{\mathbb{R}} K^2(y)dy + O(h_n^{-1})\\ &\quad + \frac{h_n^4}{4}\big(f''(x)\big)^2\left(\int_{\mathbb{R}} y^2K(y)\,dy\right)^2 + O(h_n^4). \quad (21)\end{aligned}$$

To find out a compromise between the bias and the variance, we minimise relatively to h_n the expression of the Asymptotic Mean Squared Error (**AMSE**) given by:

$$\begin{aligned}MSE\big(f_n(x,h_n)\big) &= \frac{f(x)}{nh_n}\int_{\mathbb{R}} K^2(y)dy\\ &\quad + \frac{h_n^4}{4}\big(f''(x)\big)^2\left(\int_{\mathbb{R}} y^2K(y)\,dy\right)^2. \quad (22)\end{aligned}$$

Since **AMSE** is a convex function, so the window $h_{opt\,(f_n(x,h_n))}^{MSE}$ is a solution to the equation: $\frac{\partial}{\partial h_n}\left[\frac{f(x)}{nh_n}\int_{\mathbb{R}} K^2(y)dy + \frac{h_n^4}{4}\big(f''(x)\big)^2\left(\int_{\mathbb{R}} y^2K(y)\,dy\right)^2\right] = 0.$

Thus, the smoothing parameter in the case of the estimator of the density function of the Parzen-Rosenblatt kernel is given by:

$$h_{opt\,(f_n(x,h_n))}^{MSE} = n^{-\frac{1}{5}}\left\{\frac{f(x)\int_{\mathbb{R}} K^2(y)dy}{\big(f''(x)\big)^2\left(\int_{\mathbb{R}} y^2K(y)dy\right)^2}\right\}^{\frac{1}{5}}, \quad (23)$$

with $f''(x) \neq 0$.

c) Global Approach

We will now focus on the global approach to select h_n parameter. For this, we introduce the mean integrated square error (MISE) of $f_n(x,h_n)$. We obtain:

$$\begin{aligned}MISE\big(f_n(x,h_n)\big) &= \int MSE\big(f_n(x,h_n)\big)\,dx\\ &= \int\left[\frac{f(x)}{nh_n}\int_{\mathbb{R}} K^2(y)dy + \frac{h_n^4}{4}\big(f''(x)\big)^2\left(\int_{\mathbb{R}} y^2K(y)\,dy\right)^2\right]dx\\ &= \frac{1}{nh_n}\int_{\mathbb{R}} K^2(y)dy + \frac{h_n^4}{4}\int_{\mathbb{R}}(f''(x))^2\left(\int_{\mathbb{R}} y^2K(y)\,dy\right)^2 dx, \quad (24)\end{aligned}$$

because $\int f(x)dx = 1$.

Thus, the Asymptotic Mean Integrated Error (**AMISE**) is

$$\begin{aligned}\text{AMISE}\big(f_n(x,h_n)\big) &= \frac{1}{nh_n}\int_{\mathbb{R}} K^2(y)dy\\ &\quad + \frac{h_n^4}{4}\int_{\mathbb{R}}(f''(x))^2\left(\int_{\mathbb{R}} y^2K(y)\,dy\right)^2 dx, \quad (25)\end{aligned}$$

and the window minimising the **AMISE** of the global criteria is:

$$h_{opt\,(f_n(x,h_n))}^{AMISE} = n^{-\frac{1}{5}}\left\{\frac{\int_{\mathbb{R}}K^2(y)dy}{\left(\int_{\mathbb{R}}y^2K(y)dy\right)^2\int\left(f^{''}(x)\right)^2dx}\right\}^{\frac{1}{5}}, \quad (26)$$

with $f^{''}(x) \neq 0$.

3.2. Method of Optimisation of h_n

3.2.1. Introduction

There are several methods of optimizing h_n in the literature. The most used are: The Plug-in method (Shealter, Jones, and Marron, 1996), the thumb method still called Rule of thumb (Deheuvels, 1977) and the method of cross validation (Rudemo, 1982; Bowmann, 1985 and Scott-Terrel, 1987). The method used in this article is that of Rule of thumb because it is best suited for calculating densities.

3.2.2. Rule of Thumb Method

The optimal smoothing parameter with respect to the integrated root mean square contains the unknown term $f''(x)$. This method proposed by Deheuvels (1977) consists in supposing that $f(x)$ is the Gauss density of mean 0 and variance σ_n^2, if we use the Gauss kernel, we obtain the window:

$$h_{opt} = 1.06\hat{\sigma}_n n^{-\frac{1}{5}},$$

with $\hat{\sigma}_n$ the emperical estimator of $\hat{\sigma}$.

If the true density is no't Gauss, this estimation of the windows does not gives good results.

3.3. Fundamental Results

To have our results, we will need the following hypotheses:

$(\boldsymbol{H.1})$ $f(.)$ is a function of class $\mathbb{C}^2(\mathbb{R})$;

$(\boldsymbol{H.2})$ $lim_{n\to\infty}h_n = 0$ when n→ ∞;

$(\boldsymbol{H.3})$ $\int_{\mathbb{R}}uK(u)du = 0$;

$(\boldsymbol{H.4})$ $\int_{\mathbb{R}}u^2K(u)du < \infty$;

$(\boldsymbol{H.5})$ K fits the property **(K.1) − (K.5);**

$(\boldsymbol{H.6})$ K fits the property (**K.4**).

3.3.1. Triangular Kernel

Let the triangular kernel be defined by:

$$K(x) = (1-|x|)\mathbb{1}_{[-1,1]}(x). \quad (27)$$

The following technical lemma will help us as we proceed.

Lemma 1 Under the hypothesis $(\boldsymbol{H.4})$ **and** $(\boldsymbol{H.6})$**,** we have:

$$\int_{\mathbb{R}}K^2(y)dy = \frac{2}{3} \quad (28)$$

and

$$\int_{\mathbb{R}}y^2K(y)dy = \frac{1}{6}. \quad (29)$$

Proof

Calculations are done at the -1 and 1 terminals.

We have:

$$\int_{-1}^{1}K^2(y)dy = \int_{-1}^{1}(1-|y|)^2dy$$
$$= \int_{-1}^{0}(1+y)^2dy + \int_{0}^{1}(1-y)^2dy$$
$$= \int_{-1}^{0}(1+2y+y^2)dy + \int_{0}^{1}(1-2y+y^2)dy$$
$$= \left[y+y^2+\frac{1}{3}y^3\right]_{-1}^{0} + \left[y-y^2+\frac{1}{3}y^3\right]_{0}^{1}$$
$$= 1-1+\frac{1}{3}+1-1+\frac{1}{3}$$
$$= \frac{2}{3}.$$

Similarly, we have:

$$\int_{-1}^{1}y^2K(y)dy = \int_{-1}^{1}y^2(1-|y|)\,dy$$
$$= \int_{-1}^{0}y^2(1+y)\,dy + \int_{0}^{1}y^2(1-y)\,dy$$
$$= \int_{-1}^{0}(y^2+y^3)dy + \int_{0}^{1}(y^2-y^3)dy$$
$$= \left[\frac{1}{3}y^3+\frac{1}{4}y^4\right]_{-1}^{0} + \left[\frac{1}{3}y^3-\frac{1}{4}y^4\right]_{0}^{1}$$
$$= \frac{1}{3}-\frac{1}{4}+\frac{1}{3}-\frac{1}{4}$$
$$= \frac{2}{3}-\frac{1}{2}$$
$$= \frac{1}{6}.$$

This fundamental result specifies the choice of the window h_n of the triangular kernel by Rule of thumb method.

Theorem: Under the hypotheses $(\boldsymbol{H.1}) - (\boldsymbol{H.5})$ and if we choose f as the unknown normal distribution of mean 0 and variance σ^2, the value of h_{RT}^{AMISE} is given by:

$$h_{opt} = 2.58\hat{\sigma}\,n^{-\frac{1}{5}}, \quad (30)$$

where $\hat{\sigma} = min\left\{\hat{S}, \frac{IQ}{1.349}\right\}$ and where $\hat{S}$ is the estimator of the standard deviation and IQ is the estimator of the interquartile deviation.

Proof The value of AMISE is given by:

$$h_{opt\,(f_n(x,h_n))}^{AMISE} = n^{-\frac{1}{5}}\left\{\frac{\int_{\mathbb{R}}K^2(y)dy}{(\int_{\mathbb{R}}y^2K(y)dy)^2\int_{\mathbb{R}}(f^{''}(x))^2dx}\right\}^{\frac{1}{5}}.$$

On the other hand, f being an unknown normal distribution of mean 0 and variance σ^2,

we have:

$$\int_{\mathbb{R}}\left(f^{''}(x)\right)dx = \frac{3}{8\sqrt{\pi}}\sigma^{-5}.$$

Thus,

$$h_{opt} = n^{-\frac{1}{5}}\left\{\frac{8\sqrt{\pi}\int_{\mathbb{R}}K^2(y)dy}{3\sigma^{-5}(\int_{\mathbb{R}}y^2K(y)dy)^2}\right\}^{\frac{1}{5}} \text{considering lemma1}$$
$$= n^{-\frac{1}{5}}\left\{\frac{\frac{16\sqrt{\pi}}{3}}{\frac{3}{36}\sigma^{-5}}\right\}^{\frac{1}{5}}$$
$$= n^{-\frac{1}{5}}\left\{\frac{16\sqrt{\pi}}{3}\right\}^{\frac{1}{5}}\cdot\left\{\frac{3}{36}\sigma^5\right\}^{\frac{1}{5}}$$

$$= n^{-\frac{1}{5}}\left\{\frac{16}{3}\right\}^{\frac{1}{5}} \cdot \{\sqrt{\pi}\}^{\frac{1}{5}} \cdot \left\{\frac{3}{36}\right\}^{\frac{1}{5}} \cdot \{\sigma^5\}^{\frac{1}{5}}$$

$$= \left\{\frac{16}{3}\right\}^{\frac{1}{5}} \cdot \left\{\frac{3}{36}\right\}^{\frac{1}{5}} \cdot \{\sqrt{\pi}\}^{\frac{1}{5}} \cdot \{\sigma^5\}^{\frac{1}{5}} \cdot n^{-\frac{1}{5}}$$

$$= 2.58\sigma n^{-\frac{1}{5}}$$

$$= 2.58\hat{\sigma} n^{-\frac{1}{5}} \text{ by the method of the Rule of thumb,}$$

where $\hat{\sigma} = min\left\{\hat{S}, \frac{IQ}{1.349}\right\}$ and where $\hat{S}$ is the estimator of the standard deviation and IQ is the estimator of the interquartile deviation.

3.3.2. Epanechnikov Kernel or Parabolic Kernel

Let the Epanechnikov kernel or parabolic kernel be defined by:

$$K(x) = \frac{3}{4}(1 - x^2)\mathbb{1}_{[-1,1]}(x). \quad (31)$$

The following technical lemma will be necessary for us:

Lemma: Under the hypotheses **(H.4)** and **(H.6),** we have:

$$\int_{\mathbb{R}} K^2(y)dy = \frac{3}{5}, \quad (32)$$

and

$$\int_{\mathbb{R}} y^2K(y)dy = \frac{1}{5}. \quad (33)$$

Proof The calculation is done at the limits -1 and 1. We have:

$$\int_{\mathbb{R}} K^2(y)dy = \int_{-1}^{1} (\frac{3}{4}(1 - x^2))^2 dy$$

$$= \frac{9}{16}\int_{-1}^{1} (1 - y^2)^2 dy$$

$$= \frac{9}{16}\int_{-1}^{1} (1 - 2y^2 + y^4)\, dy\,.$$

Finally, we have:

$$\int_{\mathbb{R}} K^2(y)dy = \frac{9}{16}\left[y - \frac{2}{3}y^3 + \frac{1}{5}y^5\right]_{-1}^{1}$$

$$= \frac{9}{16}\left[1 - \frac{2}{3} + \frac{1}{5} - (-1 + \frac{2}{3} - \frac{1}{5})\right]$$

$$= \frac{9}{16}\left[2 - \frac{4}{3} + \frac{2}{5}\right]$$

$$= \frac{9}{16} * \left(\frac{16}{15}\right)$$

$$= \frac{3}{5}.$$

Similarly, we have:

$$\int_{\mathbb{R}} y^2K(y)dy = \int_{-1}^{1} y^2\frac{3}{4}(1 - y^2)dy$$

$$= \frac{3}{4}\int_{-1}^{1} (y^2 - y^4)dy$$

$$= \frac{3}{4}\left[\frac{1}{3}y^3 - \frac{1}{5}y^5\right]_{-1}^{1}$$

$$= \frac{3}{4}\left[\frac{1}{3} - \frac{1}{5} + \frac{1}{3} - \frac{1}{5})\right]$$

$$= \frac{3}{4} * \left(\frac{4}{15}\right)$$

$$= \frac{1}{5}.$$

This fundamental result following precise the choice of the window h_n and of the Epanechnikov kernel by the Rule of thumb method.

Theorem

Under the hypothesis $(\boldsymbol{K}.\mathbf{1}) - (\boldsymbol{K}.\mathbf{5})$, and if we choose f like the normal unknown distribution of mean 0 and variance σ^2 the value of h_{RT}^{AMISE} is given by:

$$h_{opt} = 4.90\hat{\sigma} n^{-\frac{1}{5}}, \quad (34)$$

where $\hat{\sigma} = min\left\{\hat{S}, \frac{IQ}{1.349}\right\}$ and where $\hat{S}$ is the estimator of the standard deviation and IQ is the estimator of the interquartile deviation.

Proof The value of the AMISE is given:

$$h_{opt\,(f_n(x,h_n))}^{AMISE} = n^{-\frac{1}{5}}\left\{\frac{\int_{\mathbb{R}} K^2(y)dy}{(\int_{\mathbb{R}} y^2K(y)dy)^2 \int_{\mathbb{R}} (f''(x))^2dx}\right\}^{\frac{1}{5}}.$$

On the other hand, f being an unknown normal distribution of mean 0 and variance σ^2, we have:

$$\int_{\mathbb{R}} (f''(x))dx = \frac{3}{8\sqrt{\pi}}\sigma^{-5}.$$

Thus,

$$h_{opt} = n^{-\frac{1}{5}}\left\{\frac{8\sqrt{\pi}\int_{\mathbb{R}} K^2(y)dy}{3\sigma^{-5}(\int_{\mathbb{R}} y^2K(y)dy)^2}\right\}^{\frac{1}{5}}$$

$$= n^{-\frac{1}{5}}\left\{\frac{\frac{24\sqrt{\pi}}{5}}{\frac{3}{25}\sigma^{-5}}\right\}^{\frac{1}{5}} \text{ according to lemma 1}$$

$$= n^{-\frac{1}{5}}\left\{\frac{24\sqrt{\pi}}{5}\right\}^{\frac{1}{5}} \cdot \left\{\frac{3}{25}\sigma^5\right\}^{\frac{1}{5}}$$

$$= n^{-\frac{1}{5}}\left\{\frac{24}{5}\right\}^{\frac{1}{5}} \cdot \{\sqrt{\pi}\}^{\frac{1}{5}} \cdot \left\{\frac{3}{25}\right\}^{\frac{1}{5}} \cdot \{\sigma^5\}^{\frac{1}{5}}$$

$$= \left\{\frac{24}{5}\right\}^{\frac{1}{5}} \cdot \left\{\frac{3}{25}\right\}^{\frac{1}{5}} \cdot \{\sqrt{\pi}\}^{\frac{1}{5}} \cdot \{\sigma^5\}^{\frac{1}{5}} \cdot n^{-\frac{1}{5}} = 4.90\sigma n^{-\frac{1}{5}}$$

$$= 4.90\hat{\sigma} n^{-\frac{1}{5}} \text{ by the method of the Rule of thumb,}$$

where $\hat{\sigma} = min\left\{\hat{S}, \frac{IQ}{1.349}\right\}$ and where $\hat{S}$ is the estimator of the standard deviation and IQ is the estimator of the interquartile deviation.

4. Simulation

We present in this section a simulation study carried out using the R software, to try to illustrate the different theoretical aspects discussed in the previous section. This numerical illustration will allow us to see the result of the density estimation by the method of the Role of thumb of the smoothing parameter.

4.1. Introduction

We consider a sample $(X_i)_{i=1,\ldots,n}$, series of random

independent and identical distributed variables (i.i.d) of probability density f which obey the law $(X_i \rightsquigarrow \mathcal{N}(\mu, \sigma^2))$. To estimate f in a given interval, we suppose that F represents the function of distribution and f their density function in the form:

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

where K is the chosen kernel and h_n is the window parameter.

If K is a triangular kernel, then the value of optimal h_n noted h_{opt} is given according to section 3.3.1 by:

$$h_{opt} = 2.58\hat{\sigma}\, n^{-\frac{1}{5}}.$$

On the other hand, if k is a parabolic or Epanechnikov kernel, then the value of optimal h_n noted h_{opt} is given according to section 3.3.2. by:

$$h_{opt} = 4.90\hat{\sigma}\, n^{-\frac{1}{5}}.$$

We will generate for every of the applications which we propose the samples of height $n = 10, n = 100, n = 1\,000, n = 10\,000, n = 100\,000\ and\ n = 1\,000\,000$ respectively.

4.2. Simulation Algorithm

In other to simulate the sample defined above and to evaluate the performances in a given interval, we go through the following steps:

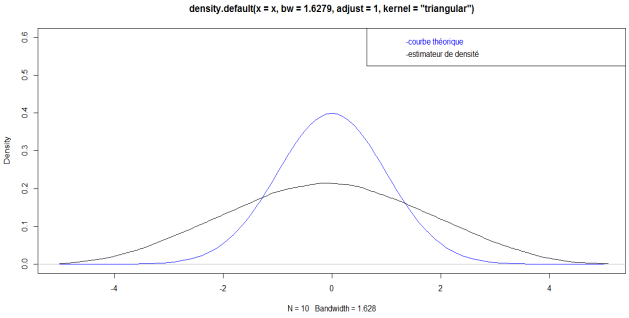
1. Generate the sample X_i according to the normal law;
2. Give the number of observation n of the simulation;
3. Give the interval of the simulated space;
4. Choose the kernel $K(.)$;
5. Choose the smoothing window h_n;
6. Estimate $f(x)$ with their estimator;
7. Draw the graph of the estimated densities.

4.3. Results of Simulation

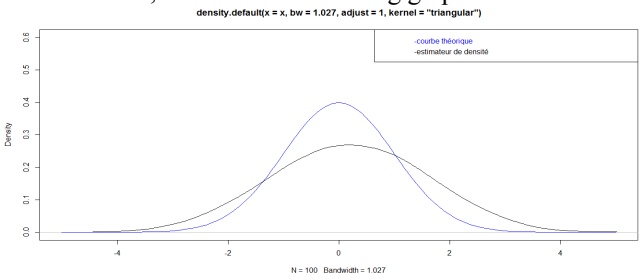
The following simulation curves obtained in this section are conceived in the software **R**. the construction codes are given in the annex.

4.3.1. Triangular Kernel

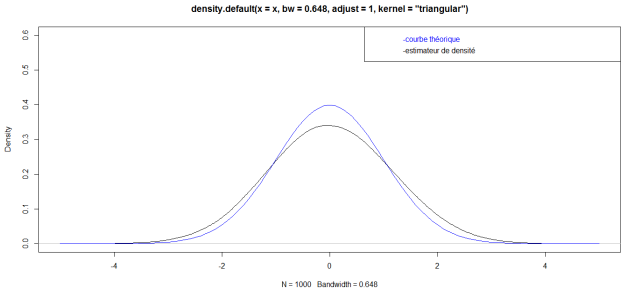
For n = 10, we have the following graph:



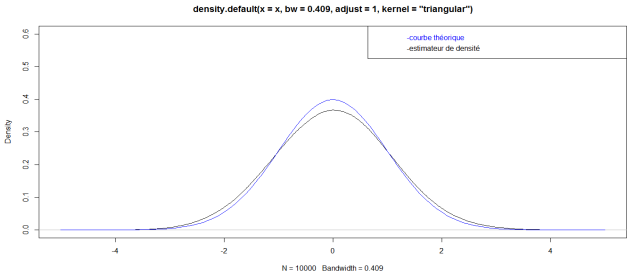
For n = 100, we have the following graph:



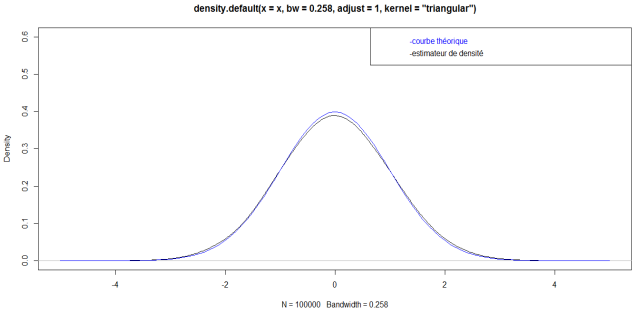
For n = 1 000, we have the following graph:



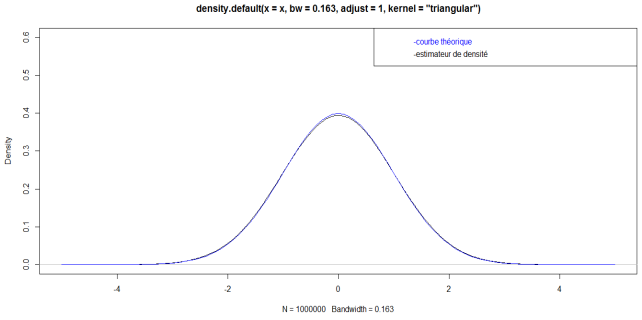
For n = 10 000, we have the following graph:



For $n = 100\,000$, we have the following graph:



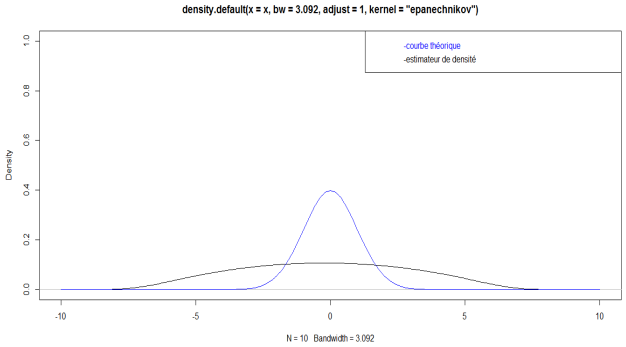
For n = 1 000 000, we have the following graph:



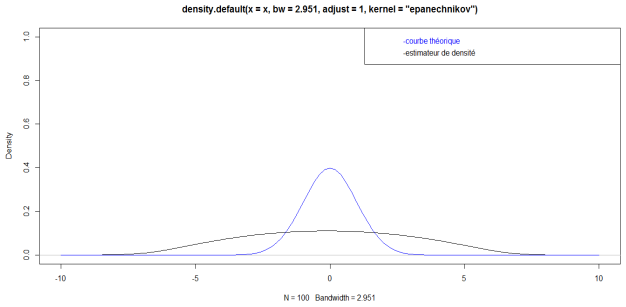
We notice that the theoretical curves greatly differ from those of the density estimators for small values ($n = 10, n = 100, n = 1\,000$) while for great values ($n = 10\,000, n = 100\,000$), they are almost identical. Finally, for a very big value (n=1 000 000), the curves are identical, which confirms the robustness of our density estimator in the case of triangular kernel.

4.3.2. Epanechnikov Kernel or Parabolic Kernel

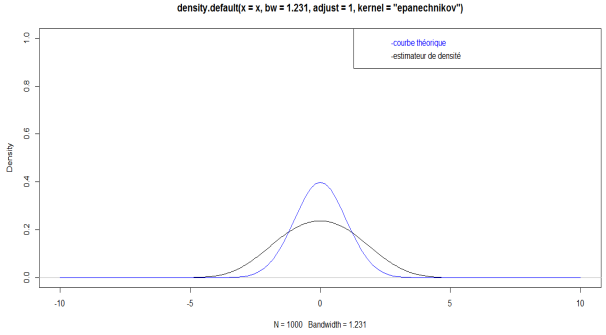
For $n = 10$, we have the following graph:



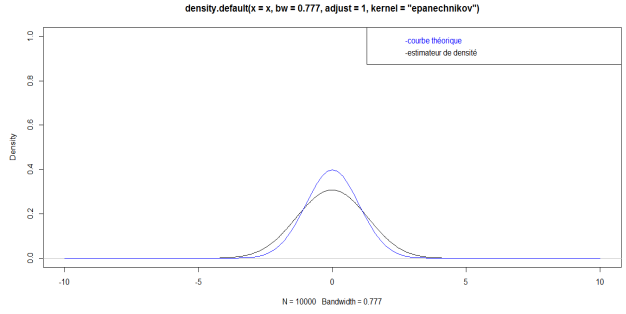
For $n = 100$, we have the following graph:



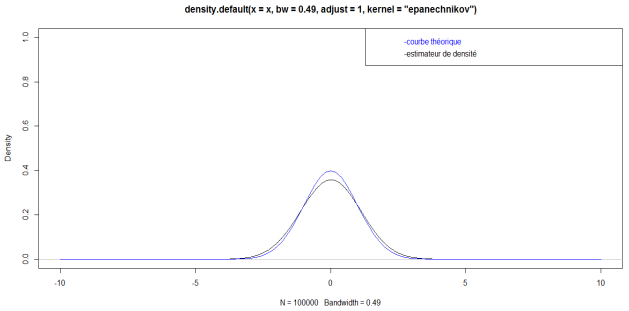
For $n = 1\,000$, we have the following graph:



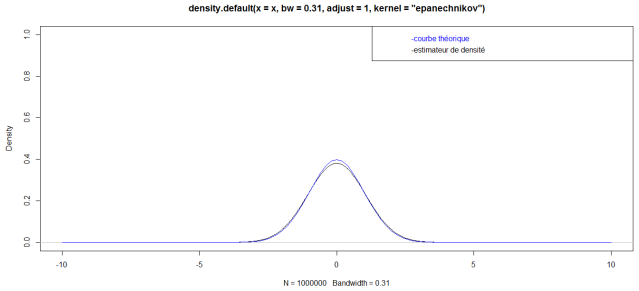
For $n = 10\,000$, we have the following graph:



For $n = 100\,000$, we have the following graph:



For $n = 1\,000\,000$, we have the following graph:



We notice that the theoretical curves differ greatly from those of the density estimator for small values (n=10, n=100, n=1000) while being identical to high value ones (n=10 000, n=100 000). Finally, for a very high value (n=1 000 000), the curves are almost identical, which confirms the performance of our density estimator in the case of the Epanechnikov kernel.

5. Conclusions

In this paper, by studying the nonparametric estimate of the probability density of the triangular core and the Epanechnikov kernel by the "Rule of thumb" method, we have succeeded in determining the smoothing parameter h_n of the kernel estimator of Parsen-Rosenblatt. We notice that when we increase the number of observations N, the error decreases and the information of the estimator is almost the same as the theoretical information. The results obtained from the software **R** perfectly illustrate this reduction of the error. By comparing them, we clearly see that the shape of the Parzen-Rosenblatt estimator approaches the shape of the theoretical probability density when the number of observations N increases and the window h decreases. In general, the performance characteristics obtained in the different observations of this sample with the Parzen-Rosenblatt estimator are very close to the theoretical ones. The higher $\boldsymbol{N}$ is, the better the estimate of densities.

We plan to study the kernel estimator of Nadaraya-Watson using the regression function and evaluate the quality of the estimation, to treat the asymptotic properties of these estimators, namely the convergence in quadratic average. This will allow us to study the convergence almost complete punctual as well as uniform.

The study of this kernel estimator of the Nadaraya-Watson density function will be studied in the context of competiting risks such as defined by Njamen and Ngatchou (2014) in order to compare the robustness of the two methods.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the reviewers who reviewed this article.

REFERENCES

[1] H. Benchoulak, "Bande de confiance pour les fonctions de densités et de régression", Mémoire de Magistère, Université Mentouri-Contantine, 2012.

[2] S. Bochner, "Harmonic Analysis and the Theory of probability", University of Chicago Press, Chicogo, Illinois, 1955.

[3] D. Bosq, and J.P. Lecoutre, "Théorie de l'estimation Fonctionnelle'', Economica, Paris, 1987.

[4] A.W. Bowman, "A comparative study of some kernel-based non-parametric density estimators", Journal of Statistical Computation and Simulation, 21, 313–327, 1985.

[5] P. Deheuvels, "Estimation non paramétrique de la densité par histogrammes généralisés'', Revue de Statistique Appliquée, XXV, 5-42, 1977.

[6] P. Deheuvels, and D. M. Mason, "General asymptotic confidence bands based on kernel-type function estimators'', Stat. Infer. Stoc. Processes, 225-277, 2004.

[7] L. Devroye, and L. Györfi, "Nonparametric Density Estimation", The view. Wiley, New York, 1985

[8] D. A. Njamen-Njomen and J. Ngatchou-Wandji, "Nelson-Aalen and Kaplan-Meier estimators in competing risks", Applied Mathematics, Vol. 5, No 4, 765-776, 2014.

[9] E. Parzen, "On estimation of a probability density function and mode", Ann. Maths. Statist., 33, 1065-1076, 1962.

[10] E. Parzen, "A new approach to the synthesis of optimal smoothing and prediction systems", Mathematical Optimization Techniques, 75-108, 1963.

[11] Praskassa, and Rao, "Nonparametric Functional Estimation", Academic Press, New York, 1983.

[12] G. G. Roussas, "Nonparametric functional estimation and related topics", Springer Science & Business Media. Vol. 335, 2012.

[13] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function", Ann. Math. Statist., 27, 832-837, 1956.

[14] M. Rudemo, "Empirical choice of histograms and kernel density estimators", Scandinavian Journal of Statistics, Vol.9, 65-78, 1982.

[15] D. W. Scott, "Multivariate Density Estimation-Theory, Practice and Visualization", Wiley, New York, 1992.

[16] D. W. Scott and G. R. Terrell, "Biased and unbiased cross validation in density estimation", Journal of the American Statistical Association, Vol.82. No.400, 1131-1146, 1987.

[17] J. Sheather, M. C. Jones and J. S. Marron, "A Brief survey of bandwidth selection for density estimation", Journal of the American Statistical Association. Vol. 91, No. 433, 401-407, 1996.

[18] B. W. Silverman, "Density Estimation for Statistics and Data Analysis", Chapman and Hall, London, 1986.

[19] M. P. Wand and M. C. Jones, "Kernel Smoothing", Chapman and Hall, London, 1995.