EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 12, No. 4, 2019, 1612-1642
ISSN 1307-5543 – www.ejpam.com
Published by New York Business Global

Density and Risk Function of Circular Kernel Study

Didier Alain Njamen Njomen[1,*], Hubert Clovis Yayebga[1]

[1] *Department of Mathematics and Computer's Science, Faculty of Science, University of Maroua, Maroua, Cameroon*

Abstract. In this article, based on the works of [18], [22] and [26] on the estimation of the survival function and the function of risk in independent cases and identically distributed with and without censorship, We manage to established the bias and variance of the density of the circular kernel. In addition, we determined the optimal window b_n^* of this estimator after having first established the mean square error (MSE) and mean integrated squared error (MISE) which are necessary conditions for obtaining the optimal window. Finally, we have established the asymptotic expression of the bias of the risk function of the circular kernel estimator.

2010 Mathematics Subject Classifications: 97K80, 62G07, 11R45

Key Words and Phrases: Nonparametric estimator, Density function, Risk function, Circular kernel, Bias, Mean squared error, Mean integrated squared error, Optimal window

---

1. Introduction

Survival analysis refers to a set of statistical techniques for processing censored data which is unknown for some of them, that a lower or higher bound and not a precise value. It represents the study of the delay of the occurrence of an event. It is called survival analysis or lifetime data analysis which is present in several areas of life among which we can mention the medicine to evaluate the effectiveness of a treatment. In demography to build life tables where these are used by actuaries to determine the amount of life insurance and life annuities, among others; actuarial tables are used when data is grouped into intervals. It is also in engineering to estimate the reliability of machines and electronic components. Survival analysis is also useful in astrophysics. Imagine that sources have been detected at a wavelength λ and that we observe at the same positions another wavelength λ'. Some sources are not detected at λ' because the ratio signal/noise (l'/b) is too low. However, how to calculate the color distribution $m - m' = 2,5\log_{10}(l'/l) + constant$? hence the interest in physics survival analysis. This analysis is linked to an event that is the subject of a study and that is how we will be able to study its lifespan or survival.

---

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v12i4.3519

Email addresses: didiernjamen1@gmail.com (D. A. N. Njamen), hyayebga@yahoo.com (H. C. Yayabga)

The term survival time refers to the time elapsed until the occurrence of a specific event. The event studied (commonly called death) is the irreversible transition between two states (commonly called living and death). The terminal event is not necessarily death: it can be the appearance of a disease (for example, the time before a relapse or rejection of a transplant), a cure (time between the diagnosis and the cure), the breakdown of a machine (duration of operation of a machine, in reliability) or the occurrence of a disaster (time between two disasters, in actuarial).

In survival analysis, the estimation of the hazard function is an interesting problem that appears in the statistical analysis of useful life to study several types of events.

Many methods have been proposed to study the estimation survival functions and risk functions in independent and identically distributed (i.i.d) cases with and without censorship. We can mention among others: [19], [17], [9], [18], [22], [20], [3], [21], [26], [8], [23], and [16]. However, the study of the determination of the mean squared error (MSE) and the mean integrated squared error (MISE) to obtain the optimum window of the circular kernel remains unclear. This article brings a relief to this problematic.

This article is subdivided into six parts: The first part is devoted to the development of the different notions studied in the density and risk functions. The second part deals with the estimation of kernel density function. The third part deals with the risk function estimation. The fourth part reminds us of the notion of mean squared error (MSE). In the fifth part, it's about the probability and risk function of the circular kernel. Finally in the last part, through simulated and real data, we show that our estimator is efficient.

2. General results on density and risk functions

In this section, we acquire various basic concepts that allow a good understanding of our work. For this purpose, we have been inspired by the founding works, articles and even courses read online. This among others include the works by: [1], [22], [3], [5] and [23].

The survival time T refers to the time elapsed since an initial moment (beginning of treatment, diagnosis, unemployment, ...) until the occurrence of an event of final interest (death of the patient, relapse, remission, healing, work, ...). It is said that the individual survives the time t if at this moment the event of final interest has not yet taken place. The variable studied is called survival time and will be marked T.

2.1. Expression of the density and risk functions in survival analysis

In what follows, it is assumed that an individual suffers once and only once a certain type of events. Observations of the occurrence (censored or uncensored) of this event in individuals constitute the samples to identify survival patterns.

Let T be the survival time that is assumed to be a positive variable. Its probability law can be defined by one of the following functions:

2.1.1. Probability density function

Definition 1. ***The probability density function of T***, *denoted $f(t)$ is defined by:*

$$f(t) = \lim_{\Delta t \to 0^+} \mathbb{P}\left(\frac{t \leq T \leq t + \Delta t}{\Delta t}\right). \tag{1}$$

For fixed t, the probability density is interpreted as the probability that the event occurs in the short time interval $]t, t + \Delta[$.

2.1.2. Risk function

Definition 2. *The risk function of T denoted $h(t)$ is defined by:*

$$h(t) = \begin{cases} \lim_{\Delta_t \to 0} \frac{\mathbb{P}(t < T \leq t + \Delta_t | T > t)}{\Delta_t}, & \textit{if } t > 0 \textit{ and such that } \mathbb{P}(T > t) > 0; \\ +\infty & \textit{if not.} \end{cases} \tag{2}$$

For fixed t, the risk function h is the instantaneous risk that the event occurs in the interval $]t, t + \Delta[$ knowing that it did not take place before time t (being in "operation" or be "alive" before time t).

Remark 1. *The risk function is also called hazard function, chance rate, failure rate or survival rate.*

The risk function may have different forms but is necessarily positive on $\mathbb{R}$.
Suppose now that T is a continuous variable, we observe that the risk function is defined by:

$$\begin{aligned} h(t) &= \lim_{\Delta_t \to 0} \frac{\mathbb{P}(t < T \leq t + \Delta_t | T > t)}{\Delta_t} \\ &= \lim_{\Delta_t \to 0} \frac{1}{\Delta_t} \frac{\mathbb{P}(t < T \leq t + \Delta_t)}{\mathbb{P}(T > t)} = \frac{f(t)}{S(t)}, \end{aligned} \tag{3}$$

where $S = \mathbb{P}(T > t)$ is the survival function.
The Hazard function characterizes the law of T because of the relation:

$$S(t) = \exp(-\int_0^t h(u)du).$$

2.1.3. Cumulative risk function

Definition 3. *The cumulative risk function $\Lambda(t)$ is defined by:*

$$\Lambda(t) = \int_0^t h(u)du. \tag{4}$$

2.2. Different forms of the risk function

In this subsection we give the main forms of law used in survival data. The usual forms of the function h are constant, monotonous (increasing or decreasing) and in the form of $\cap$.

2.2.1. Constant risk

The only distribution that has a constant h risk function is the exponential law.
The probability density and the risk function h are respectively defined by the exponential law denoted $\varepsilon(\lambda)$:

$$\begin{aligned} f(t,\lambda) &= \lambda e^{-(\lambda t)},\ t \geq 0; \\ h(t) &= \lambda. \end{aligned}$$

2.2.2. Monotonous risk

There are several life-time distributions with a monotonic risk function.

- Weibull Law $W(\lambda,\alpha)$
  There are several life-time distributions with a monotonic risk function.

$$\begin{aligned} f(t,\alpha,\lambda) &= \alpha\lambda^{\alpha}t^{\alpha-1}e^{-(\lambda t)^{\alpha}}; \\ h(t) &= \alpha\lambda^{\alpha}t^{\alpha-1}. \end{aligned}$$

  The parameter α is a form parameter of h and λ is a scale parameter.

  Remark 2. *(i) If $\alpha = 1$, we obtain the exponential law $W(\lambda,1) = \varepsilon(\lambda)$*

  (ii) If $0 < \alpha < 1$, the risk function h is decreasing from ∞ to 0

  (iii) If $\alpha > 1$, the risk function h is increasing from $(0$ to $\infty)$.

- Gamma Law $\Gamma(\lambda,\beta)$
  The probability density and the risk function are respectively defined by the Gamma law:

$$f(t,\lambda,\beta) = \lambda^{\beta}\frac{1}{\Gamma(\beta)}t^{\beta-1}e^{-\lambda t}, \quad \lambda,\beta > 0,\ t \geq 0;$$

$$h(t,\lambda,\beta) = \frac{f(t,\lambda,\beta)}{1-F(t,\lambda,\beta)}.$$

  The parameter β is the form parameter of h and λ is scale parameter.

  Remark 3. *(i) If $\beta = 1$, we obtain the exponential law with parameter λ: $\Gamma(\lambda,1) = \varepsilon(\lambda)$*

  (ii) If $\beta > 1$, the risk function h is increasing from 0 to λ.

  (iii) If $0 < \beta < 1$, the risk function h is decreasing from ∞ to $\frac{1}{\lambda}$.

2.2.3. Risk in $\cap$

- Log-normal law $LN(\lambda, v)$
  The probability density and the risk function are respectively defined by the Log-Normal law $LN(\mu, \sigma)$:
  $$f(t, \mu, \sigma) = \frac{1}{\sigma t}\varphi\left(\frac{lnt - \mu}{\sigma}\right);$$
  $$h(t, \mu, \sigma) = \frac{f(t, \mu, \sigma)}{S(t, \mu, \sigma)}.$$

  Remark 4. *The risk function h is in the form $\cap$: it increases from 0 to its maximum value and then decreases to 0.*

- Log-logistic Law $LL(\lambda, v)$
  The probability density and the risk function are defined respectively by the log-logistic law:
  $$f(t, \lambda, v) = (\lambda v)\lambda t^{v-1}\left(1 + (\lambda t)^v\right)^{-2};$$
  $$h(t, \lambda, v) = (\lambda v)\lambda t^{v-1}\left(1 + (\lambda t)^v\right)^{-1}.$$

  Remark 5. *The v parameter is a form parameter of h. For $v > 1$, the risk function h increases from 0 to its maximum value and then decreases to 0.*

3. Density function estimation

3.1. Kernel estimator: case of uncensored data

The concepts used in this section come from the following documents: [19], [17], [11], [24], [9], [20], [3], [20] and [23].

Let $T_1, T_2, ..., T_n$ distribution function survival durations F and let $V_{x_0} =]t_0 - \frac{b_n}{2}, t_0 + \frac{b_n}{2}[$ a neighborhood of t_0, where $(b_n)_{n \geq 0}$ is a sequence of positive parameters called window. An estimator of the density f at the point t_0 is given by:

$$\begin{aligned} f_n(t_0) &= \frac{F_n(t_0 + \frac{b_n}{2}) - F_n(t_0 - \frac{b_n}{2})}{b_n} \\ &= \frac{number\ \ of\ \ events\ \ on\ \ ]t_0 - \frac{b_n}{2}, t_0 + \frac{b_n}{2}[}{b_n}, \end{aligned}$$

where F_n is the empirical distribution function.
We can also write it in the following form:

$$f_n(t_0) = \frac{1}{nb_n}\sum_{i=1}^{n} \mathbb{1}_{]t_0 - \frac{b_n}{2}, t_0 + \frac{b_n}{2}[}(T_i)$$

$$= \frac{1}{nb_n} \sum_{i=1}^{n} \mathbb{1}_{]-\frac{1}{2},\frac{1}{2}]} \left(\frac{t_0 - T_i}{b_n} \right).$$

This writing shows a function K called kernel, which is in the form:

$$K(.) = \mathbb{1}_{]-\frac{1}{2},\frac{1}{2}]}(.).$$

Hence the definition of a kernel estimator of f, for $t \in \left]-\frac{1}{2}, \frac{1}{2}\right]$:

$$\widehat{f}_n(t) = \frac{1}{nb_n} \sum_{i=1}^{n} K\left(\frac{t - T_i}{b_n} \right), \tag{5}$$

where the kernel K is an application of $\mathbb{R}$ on $\mathbb{R}^+$, bounded, of integral equal to 1, symmetrical and of more $b_n \to 0$ when $n \to +\infty$.

3.1.1. Examples of kernels

The various kernels most used in estimating the density are:

- **Gaussian kernel**
$$\forall u \in \mathbb{R}, K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \tag{6}$$
- **Uniform kernel**
$$\forall u \in \mathbb{R}, K(u) = \frac{1}{2} \mathbb{1}_{|u| \leq 1} \tag{7}$$
- **Epanechnikov kernel**
$$\forall u \in \mathbb{R}, K(u) = \frac{3}{4}(1 - u^2) \mathbb{1}_{|u| \leq 1} \tag{8}$$
- **Triangular kernel**
$$\forall u \in \mathbb{R}, K(u) = (1 - |u|) \mathbb{1}_{|u| \leq 1} \tag{9}$$
- **Quadratic kernel**
$$\forall u \in \mathbb{R}, K(u) = \frac{15}{16}(1 - u^2)^2 \mathbb{1}_{|u| \leq 1} \tag{10}$$

Definition 4. *A kernel K is called Parzen-Rosemblatt if K is symmetric and if* $\lim_{|u| \to \infty} |u| K(u) = 0$.

If we consider $K_{b_n}(y) = \frac{1}{b_n} K(\frac{y}{b_n})$, kernel estimator f_n is written:

$$f_n(t) = \int_{-\infty}^{+\infty} K_{b_n}(t - s) d\mu_n(s) = (K_{b_n} * \mu_n)(t),$$

where $\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{T_i}$.

3.1.2. Optimal choices of kernel K and window b_n

The convergence results of the kernel estimator f_n require conditions on K and b_n, from which the problem of the optimal choice of (K, b_n).
We pose:

$$\begin{aligned}\Delta_n(t) &= \mathbb{E}\left[f_n(t) - f(t)\right]^2 \\ &= \mathbb{E}\left[f_n(t) - \mathbb{E}f_n(t) + \mathbb{E}f_n(t) - f(t)\right]^2 \\ &= \left[\mathbb{E}f_n(t) - f(t)\right]^2 + \mathbb{E}\left[f_n(t) - \mathbb{E}f_n(t)\right]^2.\end{aligned}$$

It is assumed that the kernel support K is $[-1, 1]$.

The following theorem gives the asymptotic behavior of the bias and the variance of the estimator f_n.

Theorem 1. *([3])*
If f is of class C^2 and K is a kernel of Parzen-Rosemblatt. So if $b_n \to 0$ and $nb_n \to \infty$, we have:

- $$Bias(f_n(x)) = \frac{b_n^2}{2} f''(x) \int_{\mathbb{R}} t^2 K(t)dt + O(b_n^2). \tag{11}$$

- $$\mathbb{V}ar(f_n(x)) = \frac{1}{nb_n} f(x) \int_{\mathbb{R}} K^2(t)dt + O\left(\frac{1}{nb_n}\right). \tag{12}$$

The choice of the optimal window b_n^* is done if the kernel K is given. The smoothing parameter b_n is a positive real whose choice is preponderant over that of the symmetrical continuous kernel K. Choosing a value of b_n too big leads to a curve that is too smooth. On the other hand, by choosing a very small smoothing parameter than the one previously adopted, the shape of the distribution changes.

Theorem 2. *([3])*
Under the conditions of the theorem 1, the optimal window is:

$$b_n^* = \left(\frac{f(t)}{f''^2(t)\left[\int u^2 K(u)du\right]^2} \int_{-1}^{1} K^2(u)du\right)^{\frac{1}{5}} n^{-\frac{1}{5}}. \tag{13}$$

3.2. Kernel estimator: case of censored data

The purpose of this subsection is to estimate the density of survival times T in the context of censored data.

Let $T_1, T_2, ..., T_n$ be a sequence of positive random variables i.i.d of distribution function F representing survival times and $C_1, C_2, ..., C_n$ is the sequence of random variables i.i.d representing censoring, of distribution function G. We assume that the sequence (T_i) is independent of the sequence (C_i) and we observe: $X_i = T_i \wedge C_i$, $D_i = \mathbb{1}_{\{T_i \leq C_i\}}$, $i = 1, ..., n$.

Kaplan-Meier estimator [6] of the survival function S is given by:

$$\widehat{S}_{KM}(t) = \prod_{t_{(i)} \leq t} \left(1 - \frac{1}{n-i+1}\right)^{D_{(i)}}. \tag{14}$$

It can also be written in the form:

$$\widehat{S}_{KM}(t) = \begin{cases} \prod_{X_{(i)} \leq t} \left(1 - \frac{D_{(i)}}{n-i+1}\right) & \text{si } t \leq X_{(n)} \\ 1 - \widehat{F}_n(X_{(n)}) & t > X_{(n)}, \end{cases}$$

where $X_{(1)<X_{(2)<\ldots<X_{(n)}}}$ is the order statistic of the sample $X_1, X_2, ..., X_n$ and $D_{(i)}$ is the value of D_i (the right censorship flag) associated with $X_{(i)}$.

For $t > X_{(n)}$ we considered $1-\widehat{F}_n(t) = 1-\widehat{F}_n(X_{(n)})$ since there is no more observations after t. This estimator $(\widehat{S}_{KM}(t))$ of survival function was studied by [13], [14], [15] in the context of competing risks.

We assume that F admits a density f in relation to Lebesgue's measure proposes to estimate using the observations (X_i, D_i), $I = 1, ..., n$. Based on the Kaplan-Meier estimator, [2] proposed an estimator of the density f by the kernel method given by:

$$f_n(t) = \frac{1}{b_n} \int_0^{+\infty} K\left(\frac{t-s}{b_n}\right) d\widehat{F}_n(s), \tag{15}$$

where $\widehat{F}_n$ is an empirical distribution function, $(b_n)_{n\geq 1}$ is the window with $b_n \to 0$ when $n \to \infty$ and K a kernel of support $[-1, 1]$.

This estimator f_n has been studied in particular by [2], [4] and [10].

4. Risk function estimation

4.1. Kernel estimator: case of uncensored data

In this part, we will study an estimator of the risk function in the case of uncensored data, by the method of kernel by developing the article written by [18].

In 1983, Henrik Ramlau-Hansen had proposed a methodology, cited by [7], to construct an estimator $\widehat{h}_n$ of the risk function by smoothing. This smoothing occurs through convolution by a kernel and by use of a window b_n.

Let $T_1, T_2, ..., T_n$ the random variables i.i.d of density f and F distribution function. The kernel estimator of the density proposed by [19] is defined by:

$$\widehat{f}_n(t) = \frac{1}{b_n} \int_{-\infty}^{+\infty} K\left(\frac{t-s}{b_n}\right) dF_n(s), \tag{16}$$

where F_n is the empirical distribution function of T_i and K a Parzen-Rosenblatt kernel of support $[-1, 1]$ (K is bounded, integrable, symmetric, $\int_{-1}^{1} K d\mu = 1$ where μ is a measure on $\mathbb{R}$ and $\lim_{|x|\to+\infty} |x||K(x)| = 0$).

Recall that the risk function

$$h(t) = \frac{f(t)}{1-F(t)}, \tag{17}$$

with $t \geq 0$ and $F(t) < 1$.
[25] studied a kernel estimator of risk function h given by:

$$\widehat{h}_n(t) = \frac{1}{n-i+1}\sum_{i=1}^{n}\frac{1}{b_n}K\left(\frac{t-T_i}{b_n}\right), \tag{18}$$

where $T_{(1)}, T_{(2)}, ..., T_{(n)}$ is the order statistic of T_i.
If we pose

$$\overline{N}_n(t) = nF_n(t) = \sum_{i=1}^{n} \mathbb{1}_{\{T_i \leq t\}}$$

and

$$\overline{Y}_n(t) = n - \overline{N}_n(t)(t^-) = \sum_{i=1}^{n} \mathbb{1}_{\{T_i \geq t\}},$$

then, the expression (18) can be written in the form:

$$\widehat{h}_n(t) = \frac{1}{b_n}\int_0^{+\infty} K\left(\frac{t-s}{b_n}\right) d\widehat{H}_{NA}(s), \tag{19}$$

where

$$\widehat{H}_{NA}(t) = \int_0^t \frac{d\overline{N}_n(s)}{\overline{Y}_n(s)} \tag{20}$$

is the Nelson-Aalen estimator of the cumulative hazard rate (see [12], [1] and [13]).

Definition 5. *Let K be a bounded function of integral 1 and b_n a positive window. The kernel estimator corresponding to the h hazard rate is given by:*

$$\widehat{h}_n(t) = \frac{1}{b_n}\int_0^{1} K\left(\frac{t-s}{b_n}\right) d\widehat{H}_{NA}(s). \tag{21}$$

The instants of process N jumps are $T_{(1)}, T_{(2)}, ..., T_{(n)}$.
So,

$$\widehat{h}_n(t) = \frac{\frac{1}{b_n}\sum_i K(\frac{t-T_{(i)}}{b_n})}{\overline{Y}_n(T_{(i)})}. \tag{22}$$

4.2. Mean and variance of the kernel estimator $\widehat{h}_n$

We assumed that the support of the kernel K is $[-1,1]$ and $0 < b_n < \frac{1}{2}$. Let consider $t \in [b_n, 1-b_n]$. Then we have:

$$\widetilde{h}_n(t) = \frac{1}{b_n}\int_0^1 K\left(\frac{t-s}{b_n}\right) d\widetilde{H}_n(s) = \int_{-1}^1 K(u)h(t-b_nu)J(t-b_nu)du,$$

with $J(t) = \mathbb{1}_{\{Y(t)>0\}}$.

Proposition 1. *([8])*
By taking,

$$K_{b_n}(s) = \frac{1}{b_n}K\left(\frac{s}{b_n}\right)$$

and

$$j_n(s) = \mathbb{E}(J_n(s))$$

then for all $t \geq 0$, we have:

$$\mathbb{E}(\widehat{h}_n(t)) = \mathbb{E}(\widetilde{h}_n(t)) = (K_{bn} * hj_n)(t), \tag{23}$$

and,

$$\sigma_n^2(t) = \mathbb{E}\left(\widehat{h}_n(t) - \widetilde{h}_n(t)\right)^2 = \frac{1}{b_n}\int_{-1}^1 K^2(u)h(t-b_nu)\mathbb{E}\left(\frac{J_n(t-b_nu)}{\overline{Y}_n(t-b_nu)}\right) du, \tag{24}$$

with $j_n(t)) = \mathbb{1}_{\{\overline{Y}_n(t)>0\}}$.

Proposition 2. *([8])*
For all $t \in [b_n, 1-b_n]$, an unbiased estimator of σ^2 is

$$\widehat{\sigma}^2(t) = \frac{1}{b_n^2}\int_{-1}^1 K^2(\frac{t-s}{b_n})\left(\frac{J_n(s)}{\overline{Y}_n^2(s)}\right) d\overline{N}_n(s), \tag{25}$$

with $j_n(t)) = \mathbb{1}_{\{\overline{Y}_n(t)>0\}}$.

4.3. Kernel estimation of the risk function: case of censored data

In this subsection, we drew on the work of [22].

Let the survival durations $T_1, T_2, ..., T_n$ of distribution function F and density f and $C_1, C_2, ..., C_n$ the censorship random variables and F_C their distribution function. It is assumed that T_i are independent of C_i for all $i = 1, ..., n$ and we observe $X_i = T_i \wedge C_i$ and $D_i = \mathbb{1}_{\{T_i \leq C_i\}}$. We denote f_X the density of X_i and F_X their distribution function.
Recall that the kernel estimator of the risk function of T_n is given by:

$$\widehat{h}_n(t) = \sum_{j=1}^n \frac{D_{(j)}}{n-j+1}K_{b_n}(t-X_{(j)}) = \sum_{i=1}^n \frac{D_{(i)}}{n-R_i+1}K_{b_n}(t-X_i), \tag{26}$$

where $X_{(1)}, X_{(2)}, ..., X_{(n)}$ is the statistic of order and R_i the rank of X_i.

This estimator appeared for the first time in the work of [25] and developed [18].

For the rest, we assume that:

(i) The risk function h is continuous and $0 < F(y) < 1$.

(ii) The kernel K is symmetric, positive and $K(t) = O(t^{-1})$ when $t \to +\infty$ and $\int K(t)dt = 1$.

4.4. Mean and variance of the estimator $\widehat{h}_n$

Let

$$m(y) = \frac{f(y)\,[1 - F_C(y)]}{f_X(y)}$$

with $f_X(y) > 0$.

To determine the mean and the variance of $\widehat{h}_n$, we need the following lemma:

Lemma 1. *([22])*
For all j, we have:

$$\mathbb{E}(D_{(j)}/X_{(j)} = y) = m(y), \tag{27}$$

and for all $r < s, y < z$, we have:

$$\mathbb{E}(D_{(r)}D_{(s)}/X_{(r)} = y, X_{(s)} = z) = m(y)m(z). \tag{28}$$

The following theorem gives the mean and the variance of $\widehat{h}_n$.

Theorem 3. *([22])*
We have:

$$\mathbb{E}\left(\widehat{h}_n(t)\right) = \int (1 - F_X^n(y))\, h(y) K_{b_n}(t-y)dy, \tag{29}$$

and,

$$\begin{aligned} \mathbb{V}\left(\widehat{h}_n(t)\right) &= \int I_n\left(F_X(y)\right) h(y) K_{b_n}^2(t-y)dy + 2\int\int_{y\le z} \{F_X^n(z) - F_X^n(y)F_X^n(z) & (30) \\ &- \frac{1-F_X(y)}{F_X(z)-F_X(y)}\left[F_X^n(z) - F_X^n(y)\right]\} h(y)h(z)K_{b_n}(t-y)K_{b_n}(t-z)dydz, & (31) \end{aligned}$$

where

$$I_n(F_X) = \sum_{k=0}^{n} (n-k)^{-1} C_n^k F_X^k (1-F)^{n-k}.$$

5. Mean Squared Error

In statistics, the mean squared error (MSE) of an estimator $\widehat{\theta}$ with a parameter θ of dimension 1 is a measure characterizing the precision of this estimator. It is more often called quadratic risk.

Definition 6. *The mean squared error is defined by:*

$$MSE(\widehat{\theta}) = \mathbb{E}\left[(\widehat{\theta} - \theta)^2\right]. \tag{32}$$

Mean squared error can be expressed as a function of bias and the variance of the estimator hence the following theorem:

Theorem 4. *([5])*

$$MSE(\widehat{\theta}) = Biais(\widehat{\theta})^2 + Var(\widehat{\theta}). \tag{33}$$

Remark 6. *As the name suggests, the mean integrated squared error (MISE) is in fact the integration made on the mean squared error.*

6. Probability Density and Risk Function of the Circular Kernel

In this section, we consider a sequence $(x_i)_{i=1}^n$ of random variables according to an unknown probability density f and such that f has a continuous second derivative. Moreover, we work within the strict case of non censored data.

6.1. Circular Kernel

The circular kernel still called the cosine kernel is defined by:

$$K(t) = \frac{\pi}{4}\cos\left(\frac{\pi}{2}t\right).\mathbb{1}_{\{|u|\leq 1\}}. \tag{34}$$

The expectation and variance of the circular kernel, were defined by [3].

Proposition 3. *Let X be a random variable following a circular law then its expectation and its variance are given by:*

- $\mathbb{E}(X) = 0$
- $\mathbb{V}(X) = 1 - \frac{8}{\pi^2}$.

6.2. Circular kernel estimator

The density estimator probability f associated to a circular kernel is given by:

$$\begin{aligned} \widehat{f}(x) &= \frac{1}{nb_n}\sum_{i=1}^n K\left(\frac{x_i - x}{b_n}\right) \\ &= \frac{1}{nb_n}\sum_{i=1}^n \frac{\pi}{4}\cos\left[\frac{\pi}{2}(\frac{x - x_i}{b_n})\right]. \end{aligned} \tag{35}$$

6.3. Risk function estimation of the circular kernel estimator

In this section, we are talking about estimating the risk function of the circular kernel.

Definition 7. *The kernel estimator of the hazard function is defined by:*

$$\widehat{h}(x) = \frac{\widehat{f}(x)}{1 - \widehat{F}(x)} = \frac{\widehat{f}(x)}{\widehat{S}(x)},$$

where $F(.)$ and $S(.)$ respectively designate the distribution and survival functions; $\widehat{F}(.)$ and $\widehat{S}(.)$ their respective estimators.

7. Main results

7.1. Case of the circular kernel estimator

7.1.1. Bias and Variance of the circular kernel estimator

The bias and variance of the circular kernel estimator is given by the following proposition and is the first result of this article.

Proposition 4. *The bias and variance of the circular kernel estimator $\widehat{f}$ of the relation (35) is given respectively by:*

$$\mathbb{E}\left(\widehat{f}(x) - f(x)\right) = \frac{b_n^2}{2} f''(x)\left[1 - \frac{8}{\pi^2}\right] + O(b_n^2). \quad (36)$$

$$(37)$$

and

$$\mathbb{V}(\widehat{f}(x)) = \frac{1}{nb_n} f(t)\frac{\pi^2}{16} + O\left(\frac{1}{nb_n}\right). \quad (38)$$

Proof. Let $X_1, X_2, ..., X_n$ be random variables that are i.i.d. We then have:

•

$$\begin{aligned}
\mathbb{E}(\widehat{f}(x)) &= \mathbb{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b_n}K\left(\frac{x - X_i}{b_n}\right)\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left\{\frac{1}{b_n}K\left(\frac{x - X_i}{b_n}\right)\right\} \\
&= \mathbb{E}\left\{\frac{1}{b_n}K\left(\frac{x - X_1}{b_n}\right)\right\} \\
&= \int_{\mathbb{R}}\frac{1}{b_n}K\left(\frac{x - x_1}{b_n}\right) f(x_1)dx_1
\end{aligned}$$

We perform the following variable change: $-t = \frac{x - x_1}{b_n}$.

Hence, $x_1 = b_n t + x$.
From the expression of x_1 above, we obtain:

$$\mathbb{E}(\widehat{f}(x)) = \int_{\mathbb{R}} K(-t) f(x + b_n t) dt.$$

Thus,

$$\begin{aligned} Bias(\widehat{f}(x)) &= \mathbb{E}(\widehat{f}(x) - f(x)) \\ &= \int_{\mathbb{R}} K(-t) f(x + b_n t) dt - f(x) \\ &= \int_{\mathbb{R}} K(t) f(x + b_n t) dt - f(x) \quad \text{because } K \text{ is symmetrical.} \end{aligned}$$

In an aim to obtain a simpler form, which depends only on the parameter b_n, we approximate the bias formula using the Taylor-Lagrange formula:

$$f(x + b_n t) = f(x) + b_n t f'(x) + \frac{b_n^2 t^2}{2} f''(x) + O(b_n^2 t^2).$$

Thus, replacing the expression of $f(x + b_n t)$ in the expression of $Bias(\widehat{f}(x))$ above, we get:

$$\begin{aligned} Bias(\widehat{f}(x)) &= \int_{\mathbb{R}} K(t) \left(f(x) + b_n t f'(x) + \frac{b_n^2 t^2}{2} f''(x) + O(b_n^2 t^2) \right) dt - f(x) \\ &= f(x) \int_{\mathbb{R}} K(t) dt + b_n f'(x) \int_{\mathbb{R}} t K(t) dt \\ &+ \frac{1}{2} b_n^2 f''(x) \int_{\mathbb{R}} t^2 K(t) dt - f(x) + O(b_n^2). \end{aligned}$$

And according to our assumptions on the K kernel, we finally get:

$$Bias(\widehat{f}(x)) = \frac{b_n^2}{2} f''(x) \int_{\mathbb{R}} t^2 K(t) dt + O(b_n^2),$$

so,

$$\mathbb{E}\widehat{f}(x) - f(x) = \frac{b_n^2}{2} f''(x) \int_{\mathbb{R}} t^2 K(t) dt + O(b_n^2).$$

Yet,

$$\int_{\mathbb{R}} t^2 K(t) dt = \int_{-1}^{1} t^2 K(t) dt = 1 - \frac{8}{\pi^2},$$

so, by replacing it in our expression above, we obtain:

$$\mathbb{E}\left(\widehat{f}(x) - f(x)\right) = \frac{b_n^2}{2} f''(x) \left[1 - \frac{8}{\pi^2}\right] + O(b_n^2).$$

Hence,

$$Bias(\widehat{f}(x)) = \mathbb{E}\left(\widehat{f}(x) - f(x)\right) = \frac{b_n^2}{2} f''(x)\left[1 - \frac{8}{\pi^2}\right] + O(b_n^2).$$

Hence the bias of $\widehat{f}$.

- Starting from the assumption of independence between the X_i, we have:

$$\begin{aligned}
\mathbb{V}ar(\widehat{f}(x)) &= \mathbb{V}ar\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{b_n}K\left(\frac{x-X_i}{b_n}\right)\right\} \\
&= \frac{1}{n}Var\left\{\frac{1}{b_n}K\left(\frac{x-X_1}{b_n}\right)\right\} \\
&= \frac{1}{n}\mathbb{E}\left[\left\{\frac{1}{b_n}K\left(\frac{x-X_1}{b_n}\right)\right\}^2\right] - \frac{1}{n}\left[\mathbb{E}\left\{\frac{1}{b_n}K\left(\frac{x-X_1}{b_n}\right)\right\}\right]^2 \\
&= \frac{1}{n}\int_{\mathbb{R}}\frac{1}{h^2}K^2\left(\frac{x-x_1}{b_n}\right)f(x_1)dx_1 - \frac{1}{n}\left\{\int_{\mathbb{R}}\frac{1}{b_n}K\left(\frac{x-x_1}{b_n}\right)f(x_1)\right\}^2.
\end{aligned}$$

By performing the $-t = \frac{x-x_1}{b_n}$ variable change, we get:

$$\begin{aligned}
\mathbb{V}(\widehat{f}(x)) &= \frac{1}{nb_n^2}\int_{\mathbb{R}}K(-t)^2 f(x+b_n t)b_n dt - \frac{1}{n}\left\{\int_{\mathbb{R}}K(-t)f(x+b_n t)b_n dt\right\}^2 \\
&= \frac{1}{nb_n}\int_{\mathbb{R}}K(t)^2 f(b_n t + x)dt - \frac{1}{n}\left[Biais(\widehat{f}(x)) + f(x)\right]^2 \\
&= \frac{1}{nb_n}\int_{\mathbb{R}}K(t)^2 f(b_n t + x)dt - \frac{1}{n}\left\{O(b_n^2) + f(x)\right\}^2 \\
&= \frac{1}{nb_n}\int_{\mathbb{R}}K^2(t)\left(f(x) + b_n t f'(x) + \frac{b_n}{2}t^2 f''(x) + O(b_n^2 t^2)\right) \\
&- \frac{1}{n}\left\{O(b_n^2) + f(x)\right\}^2 \\
&= \frac{1}{nb_n}f(x)\int_{\mathbb{R}}K^2(t)dt + \frac{1}{n}f'(x)\int_{\mathbb{R}}tK^2(t)dt + \frac{1}{n}b_n f''(x)\int_{\mathbb{R}}t^2K^2(t)dt \\
&- \frac{1}{n}\left\{O(b_n^2) + f(x)\right\}^2.
\end{aligned}$$

Yet,

$$\int_{\mathbb{R}} tK^2(t)dt = \int_{-1}^{1} t\left(\frac{\pi}{4}\cos(\frac{\pi}{2}t)\right)^2(t)dt = 0,$$

We have:

$$\mathbb{V}(\widehat{f}(x)) = \frac{1}{nb_n}f(x)\int_{\mathbb{R}}K^2(t)dt + \frac{1}{n}b_n f''(x)\int_{\mathbb{R}}t^2K^2(t)dt - \frac{1}{n}\left\{O(b_n^2) + f(x)\right\}^2.$$

Finally, under the condition of having $\int K(t)^2 dt < +\infty$ and for n large enough, we have:

$$\mathbb{V}\left(\widehat{f}(x)\right) = \frac{1}{nb_n} f(x) \int_{\mathbb{R}} K(t)^2 dt.$$

By elsewhere,

$$\int_{\mathbb{R}} K^2(t)dt = \int_{-1}^{1} K^2(t)dt = \int_{-1}^{1} \frac{\pi^2}{16} \cos^2(\frac{\pi}{2}t)dt = \frac{\pi^2}{16},$$

Hence the variance of $\widehat{f}$ is:

$$\mathbb{V}(\widehat{f}(x)) = \frac{1}{nb_n} f(x) \frac{\pi^2}{16} + O\left(\frac{1}{nb_n}\right).$$

To determine the optimal window, we first need to find the MSE and the MISE of the estimator $\widehat{f}$. In statistics, the mean squared error is a way of estimating the difference between an estimator and the actual value of the quantity to be calculated. It provides a way to choose the best estimator. The proposition gives the mean squared error (MSE) and the integrated mean squared error (MISE) of the circular kernel estimator, and is therefore the second result of this paper.

7.1.2. Mean squared error (MSE) and mean integrated squared error (MISE)

The mean squared error (MSE) and the mean integrated squared error (MISE) of the circular kernel estimator are given by the following result:

Proposition 5. *The Mean Square Error (MSE) and the Mean Integrated Squared Error (MISE) of the circular kernel estimator $\widehat{f}$ of the relation (35) are given respectively by:*

(i) $MSE(\widehat{f}) \approx \frac{b_n^4}{4}(1 - \frac{8}{\pi^2})^2 f''^2(x) + \frac{\pi^2}{16nb_n} f(x)$;

(ii) $MISE(\widehat{f}) \approx \frac{b_n^2}{4}(1 - \frac{8}{\pi^2})^2 I_1 + \frac{\pi^2}{16nb_n} I_2$,

where $I_1 = \int f''^2(x)dx$ *and* $I_2 = \int f(x)dx$.

Proof.

(i) According to the definition of the mean squared error, we have:

$$\begin{aligned}
MSE(\widehat{f}) &= \mathbb{V}(f(x)) + Bias(\widehat{f}(x))^2 \\
&= \frac{1}{nb_n} f(x) \frac{\pi^2}{16} + O(\frac{1}{nb_n}) + \left[\frac{b_n^2}{2} f''(x) \left[1 - \frac{8}{\pi^2}\right] + O(b_n^2)\right]^2 \\
&\approx \frac{1}{nb_n} f(x) \frac{\pi^2}{16} + \left[\frac{b_n^2}{2} f''(x) \left[1 - \frac{8}{\pi^2}\right]\right]^2
\end{aligned}$$

$$\approx \frac{b_n^4}{4}(1-\frac{8}{\pi^2})^2 f''^2(x) + \frac{\pi^2}{16nb_n} f(x).$$

So, the $MSE(\widehat{f})$.

(ii) We know that $MISE(\widehat{f}) = \int MSE(\widehat{f})dx$, then we have:

$$\begin{aligned} MISE(\widehat{f}) &= \int MSE(\widehat{f})dx \\ &\approx \int \left[\frac{b_n^4}{4}(1-\frac{8}{\pi^2})^2 f''^2(x) + \frac{\pi^2}{16nb_n} f(x)\right] dx \\ &\approx \frac{b_n^4}{4}(1-\frac{8}{\pi^2})^2 \int f''^2(x)dx + \frac{\pi^2}{16nb_n} \int f(x)dx. \end{aligned}$$

By posing $I_1 = \int f''^2(x)dx$ and $I_2 = \int f(x)dx$, we obtain:

$$MISE(\widehat{f}) \approx \frac{b_n^4}{4}(1-\frac{8}{\pi^2})^2 I_1 + \frac{\pi^2}{16nb_n} I_2.$$

7.1.3. Optimal window of the circular kernel estimator

The following result gives the optimal window of the estimator (35) and is the third result of this article:

Theorem 5. *The optimal window of the kernel estimator $\widehat{f}$ of the relation (35) is given by:*

$$b_n^* = \pi \sqrt[5]{\frac{\pi I_2}{16nI_1(\pi^2-8)^2}}. \tag{39}$$

Proof. To determine the optimal window, we look for the value of the window which minimizes the integrated mean squared error.

For that, let's look for the value of b_n that minimizes the integrated mean squared error.

According to the above proposal, we have:

$$MISE(\widehat{f}) \approx \frac{b_n^4}{4}(1-\frac{8}{\pi^2})^2 I_1 + \frac{\pi^2}{16nb_n} I_2.$$

By deriving $MISE(\widehat{f})$, we have:

$$\frac{\partial}{\partial b_n} MISE(\widehat{f}) = \frac{\partial}{\partial b_n}\left(\frac{b_n^4}{4}(1-\frac{8}{\pi^2})^2 I_1 + \frac{\pi^2}{16nb_n} I_2\right)$$

$$= \quad b_n^3\left(1-\frac{8}{\pi^2}\right)^2 I_1 - \frac{\pi^2}{16nb_n^2} I_2.$$

By canceling the derivative above, we have:

$$\begin{aligned}
\frac{\partial}{\partial b_n} MISE(\widehat{f}) = 0 \quad &\Leftrightarrow \quad b_n^3\left(1-\frac{8}{\pi^2}\right)^2 I_1 - \frac{\pi^2}{16nb_n^2} I_2 = 0\\
&\Leftrightarrow \quad b_n^5\left(1-\frac{8}{\pi^2}\right)^2 I_1 = \frac{\pi^2}{16n} I_2\\
&\Leftrightarrow \quad b_n^5 = \frac{\frac{\pi^2}{16n} I_2}{\left(1-\frac{8}{\pi^2}\right)^2 I_1}\\
&\Leftrightarrow \quad b_n = \sqrt[5]{\frac{\frac{\pi^2}{16n} I_2}{\left(1-\frac{8}{\pi^2}\right)^2 I_1}}.
\end{aligned}$$

Moreover,

$$\frac{\partial^2}{\partial b_n^2} MISE(\widehat{f}) = 3b_n^2\left(1-\frac{8}{\pi^2}\right)^2 I_1 + \frac{\pi^2}{8nb_n^3} I_2 > 0.$$

As the quantities $3b_n^2\left(1-\frac{8}{\pi^2}\right)^2$ and $8nb_n^3$ are positive, and since I_1 and I_2 are also positive according to condition (C.4), then we conclude that the value of b_n that minimizes the integrated mean squared error is:

$$b_n^* = \pi \sqrt[5]{\frac{\pi I_2}{16nI_1(\pi^2-8)^2}}.$$

Hence the result.

7.1.4. Optimal mean integrated squared error ($MISE_{opt}$)

The following result gives the optimal integrated mean squared error ($MISE_{opt}$) of the estimator (35) and represents the 4th result of this article.

Theorem 6. *The optimal mean integrated squared error denoted ($MISE_{opt}$) of the estimator (35) is given by:*

$$MISE_{opt}(\widehat{f}) = 5\left(\frac{\pi^4 I_1 I_2^4(\pi^2-8)^2}{4^{13}\times n^4}\right)^{\frac{1}{5}}. \tag{40}$$

Proof.

We know that the optimal mean integrated squared error($MISE_{opt}$) of the estimator (35) can be written in the form:

$$MISE_{opt}(\widehat{f}) = \frac{b_n^{*4}}{4}(1-\frac{8}{\pi^2})^2 I_1 + \frac{\pi^2}{16nb_n^*} I_2.$$

Then, by replacing the optimal window b_n^* by its value, we obtain:

$$
\begin{aligned}
MISE_{opt}(\widehat{f}) &= \frac{b_n^{*4}}{4}(1-\frac{8}{\pi^2})^2 I_1 + \frac{\pi^2}{16nb_n^*} I_2 \\
&= \frac{\pi^4}{4}\left(1-\frac{8}{\pi^2}\right)^2 I_1 \sqrt[5]{\left(\frac{\pi I_2}{16n(\pi^2-8)^2}\right)^4} + \frac{\pi^2 I_2}{16n\pi}\sqrt[5]{\frac{16nI_1(\pi^2-8)^2}{\pi I_2}} \\
&= \frac{1}{4}\frac{(\pi^2-8)^2\pi^{\frac{4}{5}} I_2^{\frac{4}{5}} I_1^{\frac{1}{5}}}{(16n)^{\frac{4}{5}}(\pi^2-8)^{\frac{8}{5}}} + \frac{\pi}{16n}\frac{I_2^{\frac{4}{5}} I_1^{\frac{1}{5}}(16n)^{\frac{1}{5}}(\pi^2-8)^{\frac{2}{5}}}{\pi^{\frac{1}{5}}} \\
&= \frac{1}{4}\frac{(\pi^2-8)^{\frac{2}{5}}\pi^{\frac{4}{5}} I_2^{\frac{4}{5}} I_1^{\frac{1}{5}}}{(16n)^{\frac{4}{5}}} + \frac{\pi^{\frac{4}{5}} I_2^{\frac{4}{5}} I_1^{\frac{1}{5}}(\pi^2-8)^{\frac{2}{5}}}{(16n)^{\frac{4}{5}}} \\
&= \frac{5}{4}\frac{\pi^{\frac{4}{5}} I_2^{\frac{4}{5}} I_1^{\frac{1}{5}}(\pi^2-8)^{\frac{2}{5}}}{(16n)^{\frac{4}{5}}} \\
&= 5\left(\frac{\pi^4 I_1 I_2^4(\pi^2-8)^2}{4^{13}\times n^4}\right)^{\frac{1}{5}}.
\end{aligned}
$$

Hence the result.

7.2. Case of the risk function

The following theorem gives the bias and variance of the circular kernel risk function estimator and is the 5th result of this article.

Theorem 7. *The bias and variance of the risk function estimator $\widehat{h}$ of the circular kernel are respectively given by:*

$$
Biais(\widehat{h}(x)) = \frac{b_n^2 f''(x)(\pi^2-8)}{2\pi^2 S(x)} + O(1), \tag{41}
$$

and,

$$
\mathbb{V}(\widehat{h}(x)) = \frac{\pi^2}{16}.\frac{1}{nb_n}.\frac{h(x)}{S(x)} + O\left((nb_n^2)^{-1}\right).
$$

Proof.
The expectancy of the risk function estimator of the circular kernel is:

$$
\begin{aligned}
\mathbb{E}(\widehat{h}(x)) &= \mathbb{E}\left(\frac{\widehat{f}(x)}{\widehat{S}(x)}\right) \\
&= \frac{\mathbb{E}(\widehat{f}(x))}{\mathbb{E}(\widehat{S}(x))} \\
&= \frac{f(x) + \frac{b_n^2}{2}\left(1-\frac{8}{\pi^2}\right) f''(x)}{S(x)} + O(1)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{f(x)}{S(x)} + \frac{\frac{b_n^2}{2}\left(1-\frac{8}{\pi^2}\right)f''(x)}{S(x)} \\
&= h(x) + \frac{\frac{b_n^2}{2}\left(1-\frac{8}{\pi^2}\right)f''(x)}{S(x)}
\end{aligned}
$$

So,

$$
\begin{aligned}
Bias\left(\widehat{h}(x)\right) &= \frac{\frac{b_n^2}{2}\left(1-\frac{8}{\pi^2}\right)f''(x)}{S(x)} + O(1) \\
&= \frac{b_n^2 f''(x)(\pi^2-8)}{2\pi^2 S(x)} + O(1).
\end{aligned}
$$

Hence the result.

For the calculation of the variance, under Theorem 1 and we have:

$$
\begin{aligned}
\mathbb{V}(\widehat{h}(x)) &= \mathbb{V}\left(\frac{\widehat{f}(x)}{\widehat{S}(x)}\right) \\
&= \frac{1}{nb_n}\left(\int K^2(t)dt\right)\frac{h(x)}{S(x)} + O\left((nb_n^2)^{-1}\right).
\end{aligned}
$$

However, according to the above, we have:

$$
\int_{-1}^{1} K^2(t)dt = \int_{-1}^{1}\left(\frac{\pi}{4}\cos(\frac{\pi}{2}t)\right)dt = \frac{\pi^2}{16}.
$$

So,

$$
\mathbb{V}(\widehat{h}(x)) = \frac{\pi^2}{16nb_n}.\frac{h(x)}{S(x)} + O\left((nb_n)^{-1}\right)^2.
$$

Hence the result.

8. Applications

In this section, we want to see if our estimator is robust. For this purpose, the efficiency of the circular core density function estimator will be highlighted through the density curves and the kernel estimator which will be performed on the basis of simulated data and actual data.

8.1. Choice of the programming language

In this article, we used the programming language **R** not only for the simulation of our estimator but also, for the realization of our different curves. It is a programming language for manipulating data and performing statistical and graphical analysis. This Scheme-inspired language and software, abbreviated as **S**, is distributed under the GNU (General Public License) terms. This structure is responsible for distributing and developing the **R** software with many other contributors from around the world.

8.2. Simulation: Case of Simulated Data

Let X be a continuous random variable whose law has the density f defined on $\mathbb{R}$ by:

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}.$$

The function f is called density function of the reduced normal centered law denoted $\mathcal{N}(0, 1)$. It is a positive function.

If X follows a "model" distribution, it is associated with a curve:

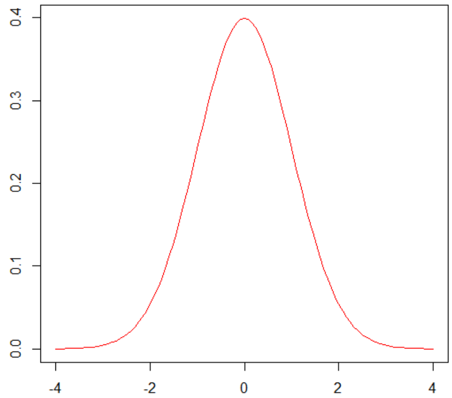

Figure 1: Density of the Normal Law Centered Reduced

The curve above is a symmetrical curve with respect to the origin which has a bell-shaped shape. It is commonly called Gaussian curve.

The normal law is used in this subsection because it is the most commonly used probabilistic model to describe many phenomena observed in practice.

In what follows, we will generate the data from the **R** software to show the influence of the h smoothing parameter in evaluating the performance of the circular kernel estimator.

8.2.1. Result of the simulation

For a sample $N = 500$, we assign the values to the smoothing parameter, which allows us to obtain the following curves:

For $h = 2$, we have the following curve:

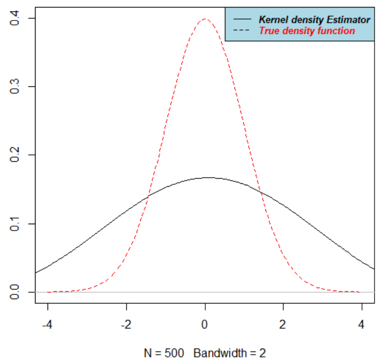

Figure 2: Comparison of the theoretical curve and the kernel density estimator at h = 2

For $h = 1$, we have the following curve:

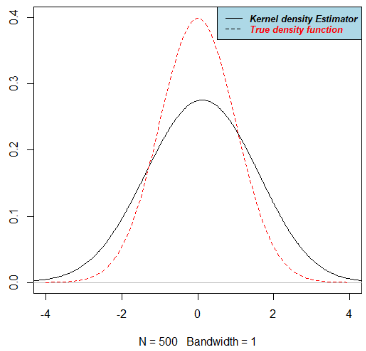

Figure 3: Comparison of the theoretical curve and the kernel density estimator at h = 1

For $h = 0.5$, we have the following curve:

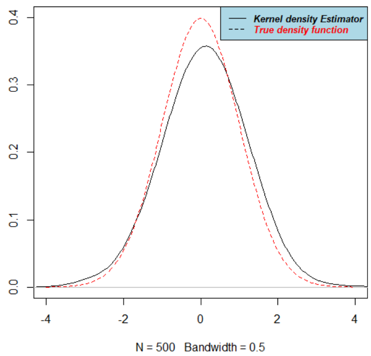

Figure 4: Comparison of the theoretical curve and the kernel density estimator at h = 0.5

For $h = 0.3$, we have the following curve:

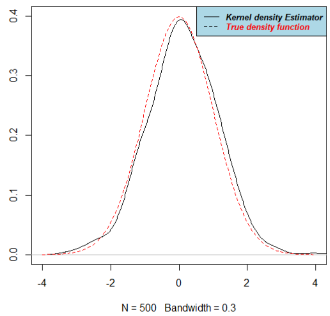

Figure 5: Comparison of the theoretical curve and the kernel density estimator at h = 0.3

For $h = h_{opt}$, we have the following curve:

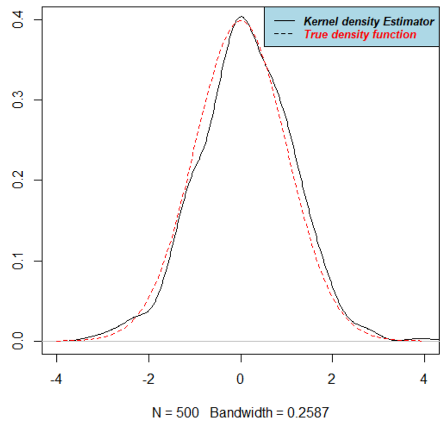

Figure 6: Comparison of the theoretical curve and the kernel density estimator at $h = h_{opt}$

For $h = 0.15$, we have the following curve:

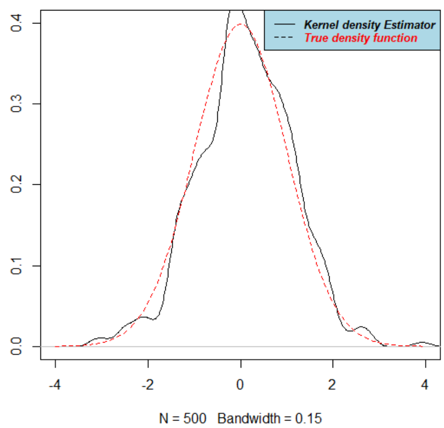

Figure 7: Comparison of the theoretical curve and the kernel density estimator at h=0.15

8.2.2. Interpretation

As a reminder, bw refers to the window or smoothing parameter that we have noted h in all our work. The density function is estimated using the circular core estimator. This estimator and the function to be estimated for the density function of the reduced normal centered law, if any, are represented according to different values of the smoothing parameter. We noticed that for every larger values of the smoothing window, our estimator has a rounded shape and approaches the line of equation $y = 0$ (see Figure 2 and Figure 3). On the other hand, for values of the window smaller and smaller, our estimator approaches the true density function of the reduced normal centered law (see Figure 4, Figure 5 and Figure 7). Thus, if the choice of the kernel is considered as having little influence on the estimator, it is not the same for the smoothing parameter. A parameter that is too weak causes the appearance of artificial details appearing on the graph of the estimator. For a value of h that is too large (see Figure 2, Figure 3), the majority of the characteristics are on the contrary erased. The choice of h is therefore a central question in estimating density. Hence the importance of determining an optimal smoothing parameter. The calculation of this optimal parameter gives us for a sample of size $N = 500$, $h_{opt} = 0.2587$, (see Figure 6). At this value, our estimator is almost similar to the true density, which shows us the efficiency of our estimator and so this estimator can therefore substitute for the density function of the reduced normal centered law.

8.3. Real Data Case

In this subsection, we use data collected during a survey to study the concentration of CO_2 in the atmosphere around a very active volcano called Mauna Loa in the Hawaiian Archipelago in the United States of America culminating at 4, 100 meters above level. This study was conducted between 1959 and 1997 from January to December of those years of study. The catches below present all these data expressed in ppm (parts per million):

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct
1959	315.42	316.31	316.50	317.56	318.13	318.00	316.39	314.65	313.68	313.18
1960	316.27	316.81	317.42	318.87	319.87	319.43	318.01	315.74	314.00	313.68
1961	316.73	317.54	318.38	319.31	320.42	319.61	318.42	316.63	314.83	315.16
1962	317.78	318.40	319.53	320.42	320.85	320.45	319.45	317.25	316.11	315.27
1963	318.58	318.92	319.70	321.22	322.08	321.31	319.58	317.61	316.05	315.83
1964	319.41	320.07	320.74	321.40	322.06	321.73	320.27	318.54	316.54	316.71
1965	319.27	320.28	320.73	321.97	322.00	321.71	321.05	318.71	317.66	317.14
1966	320.46	321.43	322.23	323.54	323.91	323.59	322.24	320.20	318.48	317.94
1967	322.17	322.34	322.88	324.25	324.83	323.93	322.38	320.76	319.10	319.24
1968	322.40	322.99	323.73	324.86	325.40	325.20	323.98	321.95	320.18	320.09
1969	323.83	324.26	325.47	326.50	327.21	326.54	325.72	323.50	322.22	321.62
1970	324.89	325.82	326.77	327.97	327.91	327.50	326.18	324.53	322.93	322.90
1971	326.01	326.51	327.01	327.62	328.76	328.40	327.20	325.27	323.20	323.40
1972	326.60	327.47	327.58	329.56	329.90	328.92	327.88	326.16	324.68	325.04
1973	328.37	329.40	330.14	331.33	332.31	331.90	330.70	329.15	327.35	327.02
1974	329.18	330.55	331.32	332.48	332.92	332.08	331.01	329.23	327.27	327.21
1975	330.23	331.25	331.87	333.14	333.80	333.43	331.73	329.90	328.40	328.17
1976	331.58	332.39	333.33	334.41	334.71	334.17	332.89	330.77	329.14	328.78
1977	332.75	333.24	334.53	335.90	336.57	336.10	334.76	332.59	331.42	330.98
1978	334.80	335.22	336.47	337.59	337.84	337.72	336.37	334.51	332.60	332.38
1979	336.05	336.59	337.79	338.71	339.30	339.12	337.56	335.92	333.75	333.70
1980	337.84	338.19	339.91	340.60	341.29	341.00	339.39	337.43	335.72	335.84
1981	339.06	340.30	341.21	342.33	342.74	342.08	340.32	338.26	336.52	336.68
1982	340.57	341.44	342.53	343.39	343.96	343.18	341.88	339.65	337.81	337.69

Figure 8: Concentration data in CO_2 from January to October

1983	341.20	342.35	342.93	344.77	345.58	345.14	343.81	342.21	339.69	339.82
1984	343.52	344.33	345.11	346.88	347.25	346.62	345.22	343.11	340.90	341.18
1985	344.79	345.82	347.25	348.17	348.74	348.07	346.38	344.51	342.92	342.62
1986	346.11	346.78	347.68	349.37	350.03	349.37	347.76	345.73	344.68	343.99
1987	347.84	348.29	349.23	350.80	351.66	351.07	349.33	347.92	346.27	346.18
1988	350.25	351.54	352.05	353.41	354.04	353.62	352.22	350.27	348.55	348.72
1989	352.60	352.92	353.53	355.26	355.52	354.97	353.75	351.52	349.64	349.83
1990	353.50	354.55	355.23	356.04	357.00	356.07	354.67	352.76	350.82	351.04
1991	354.59	355.63	357.03	358.48	359.22	358.12	356.06	353.92	352.05	352.11
1992	355.88	356.63	357.72	359.07	359.58	359.17	356.94	354.92	352.94	353.23
1993	356.63	357.10	358.32	359.41	360.23	359.55	357.53	355.48	353.67	353.95
1994	358.34	358.89	359.95	361.25	361.67	360.94	359.55	357.49	355.84	356.00
1995	359.98	361.03	361.66	363.48	363.82	363.30	361.94	359.50	358.11	357.80
1996	362.09	363.29	364.06	364.76	365.45	365.01	363.70	361.54	359.51	359.65
1997	363.23	364.06	364.61	366.40	366.84	365.68	364.52	362.57	360.24	360.83

Figure 9: Concentration data in CO_2 from November to December

In this study, 468 data was collected. The curve of the function of the reduced normal centered law for these real data is obtained through the software **R**:

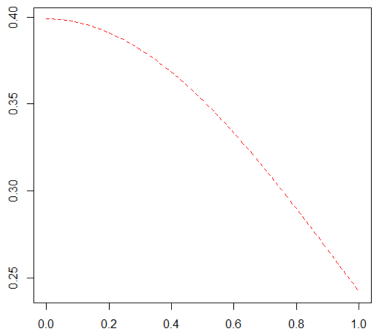

Figure 10: True density function with real data

The curve of the density of the normal law above is positive and is taken in half.

In what follows, we will generate the data from the **R** software

8.3.1. Result of Simulation

The simulation was performed for several values of the smoothing parameter.

For $h = 50$, we have the following curve:

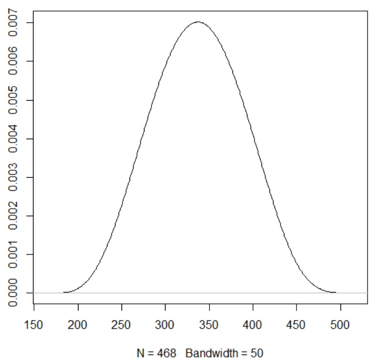

Figure 11: Curve of the kernel estimator for h = 50

For $h = 20$, we have the following curve:

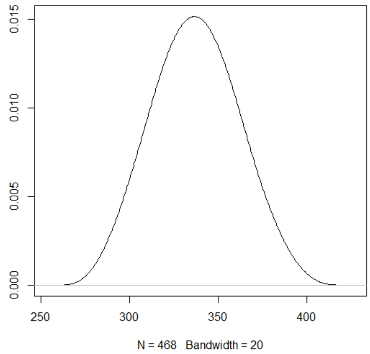

Figure 12: Curve of the kernel estimator for h = 20

For $h = 10$, we have the following curve:

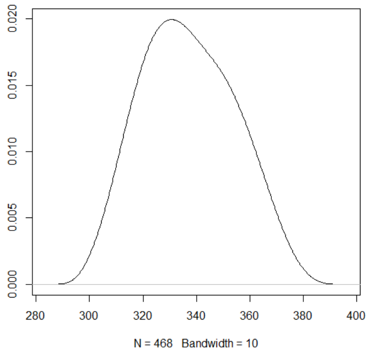

Figure 13: Curve of the kernel estimator for h = 10

For $h = 3$, we have the following curve:

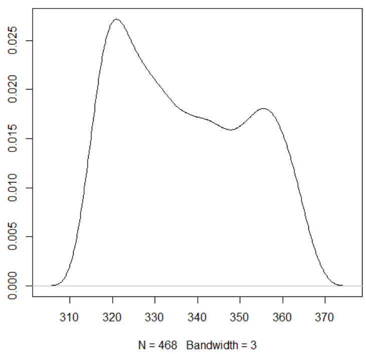

Figure 14: Curve of the kernel estimator for h = 2

For $h = 2$, we have the following curve:

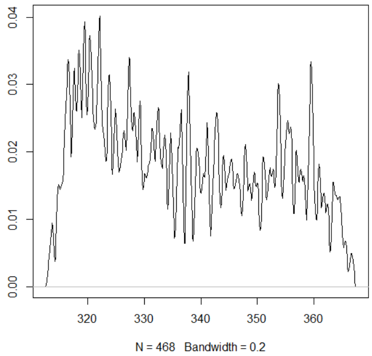

Figure 15: Curve of the kernel estimator for h = 2

For $h = 1$, we have the following curve:

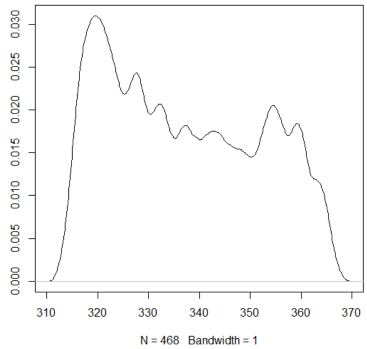

Figure 16: Curve of the kernel estimator for h = 1

8.3.2. Interpretation

As in the case of simulated data, according to different values of the smoothing parameter,we have a curve which has a different look and tends to move away or closer to the true graph of the density function of the reduced normal centered law (see Figure 12 and Figure 13) represented under the basis of our 468 observations.

On the other hand, we found that at first, the density curve of the normal law is taken almost half (see Figure 10). And secondly, our different curves have several fluctuations and we observe rounded shapes of their representations (see Figure 14, Figure 15 and Figure 16), this is due to the fact that all the different data we are working on have very small values and so we have used a smaller scale.

However, whether in the context of simulated data or in the context of real data, the estimation of probability densities by the kernel method requires a reliable estimation of the smoothing parameter. The choice of the window remains so paramount. The quality of non-parametric kernel-based estimators is closely related. Hence the importance is given to the determination of an appropriate smoothing window to remain in perfect adequacy in the handling of our raw data.

9. Conclusion and Perspectives

In this paper, we have established the bias and variance of circular kernel density. In addition, we have determined the optimal b_n^* window of this estimator after first establishing the mean squared error (MSE) and the mean integrated squared error (MISE) of this estimator. Finally, we have established the asymptotic expression of the bias of the circular kernel risk function estimator. This work is part of the continuity of the work of [18], [22] and [26] on the estimation of the survival function and the risk function in independent cases and identically distributed with and without censorship.

Through simulated data and real data, we have shown that the curve of the non-parametric estimator of the density function of the circular kernel is confounded with that for the optimal window this which proves the efficiency of our estimator.

However, some improvements remain to be done in this article. Indeed, it is a question of establishing the confidence intervals of the kernel estimator of the density function and the risk function of the circular kernel; to redo this work in the context of asymmetric nuclei and if possible to extend the study of the estimation of the conditional risk function in the case of the discrete nucleus.

Acknowledgements

The authors thank the reviewers of the European Journal of Pure and Applied Mathematics for the contributions and suggestions they will make on this article.

References

[1] Odd Aalen. Nonparametric inference for a family of counting processes. *The Annals of Statistics*, pages 701–726, 1978.

[2] JR Blum and V Susarla. Maximal deviation theory of density and failure rate function estimates based on censored data. *Multivariate analysis*, 5:213–222, 1980.

[3] D Bosq and JF Lecoutre. ((th eorie de l'estimation fonctionnelle)) economica, 1987.

[4] Antonia Foldes and Lidia Rejto. Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *The Annals of Statistics*, pages 122–129, 1981.

[5] William H Greene. Censored data and truncated distributions. *Available at SSRN 825845*, 2005.

[6] Edward L Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

[7] John P Klein and Melvin L Moeschberger. Semiparametric proportional hazards regression with fixed covariates. *Survival analysis: techniques for censored and truncated data*, pages 243–293, 2003.

[8] John P Klein and Melvin L Moeschberger. *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media, 2006.

[9] Jean-Pierre Lecoutre. *Contribution à l'estimation non paramétrique de la régression*. PhD thesis, 1982.

[10] Jan Mielniczuk et al. Some asymptotic properties of kernel estimators of a density function in case of censored data. *The Annals of Statistics*, 14(2):766–773, 1986.

[11] Elizbar A Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964.

[12] Wayne Nelson. Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14(4):945–966, 1972.

[13] Didier Alain Njamen-Njomen and Joseph Ngatchou-Wandji. Nelson-aalen and kaplan-meier estimators in competing risks. *Applied Mathematics*, 5(04):765, 2014.

[14] DA Njamen Njomen. Convergence of the nelson-aalen estimator in competing risks. *International Journal of Statistics and Probability*, 6(3):9–23, 2017.

[15] Didier Alain Njamen Njomen and Joseph Wandji Ngatchou. Consistency of the kaplan-meier estimator of the survival function in competiting risks. *The Open Statistics & Probability Journal*, 9(1), 2018.

[16] Didier Alain Njamen Njomen and Ludovic Kakmeni Siewe. A study of the ability of the kernel estimator of the density function for triangular and epanechnikov kernel or parabolic kernel. *International Journal of Statistics and Applications*, 9(2):45–52, 2019.

[17] Emanuel Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076, 1962.

[18] Henrik Ramlau-Hansen. Smoothing counting process intensities by means of kernel functions. *The Annals of Statistics*, pages 453–466, 1983.

[19] Murray Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, pages 832–837, 1956.

[20] Bernard W Silverman. *Density Estimation for Statistics and Data Analysis*, volume 26. CRC Press, 1986.

[21] Jeffrey S Simonoff. *Smoothing Methods in Statistics.* Springer Science & Business Media, 1996.

[22] Martin A Tanner and Wing Hung Wong. The estimation of the hazard function from randomly censored data by the kernel method. *The Annals of Statistics*, pages 989–993, 1983.

[23] Alexandre B Tsybakov. Introduction to nonparametric estimation, 2009. *URL https://doi. org/10.1007/b13794. Revised and extended from the*, 2004.

[24] Geoffrey S Watson. Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372, 1964.

[25] GS Watson and MR Leadbetter. Hazard analysis. i. *Biometrika*, 51(1/2):175–184, 1964.

[26] Biao Zhang. Some asymptotic results for kernel density estimation under random censorship. *Bernoulli*, pages 183–198, 1996.