Journal of Statistical Research *https://doi.org/10.3329/jsr.v59i2.88064*
2025, Vol. 59, No. 2, pp. 183-201 *ISSN 0256 - 422 X*

ON KERNEL DENSITY DERIVATIVE ESTIMATION NEAR THE BOUNDARY

HIND BOUREDJI
Laboratory of Applied Mathematics, University of mohamed khider, Biskra, Algeria
Email: hind.bouredji@univ-biskra.dz

ABDALLAH SAYAH*
Laboratory of Applied Mathematics, University of mohamed khider, Biskra, Algeria
Email: sayahabdel@yahoo.fr

SUMMARY

Estimating density derivatives is a powerful technique in statistical data analysis. It has diverse applications in machine learning, signal processing, and statistical analysis. The kernel method is one of the most popular methods in nonparametric density derivative estimation, but this estimator is biased and is not consistent when the data are near the endpoints of the support. This paper investigates the challenge of estimating the first-order derivative of an unknown probability density function defined on the interval $[0, 1]$. We focus our study near the right boundary. The asymptotic properties are derived. A Monte Carlo study and real data example are provided to illustrate the finite sample performance of the proposed estimator.

Keywords and phrases: Boundary problem; Correction methods; Density derivative estimation; Kernel method.

AMS Classification: 62G07.

1 Introduction

Density estimation is a fundamental task in statistical analysis and machine learning that aims to model the probability distribution of a dataset based on observed samples. Unlike parametric methods that assume a specific functional form for the distribution, nonparametric density estimation, such as kernel density estimation, allows for greater flexibility in capturing complex, unknown distributions. Density derivatives play a significant role in various applied fields by providing insights into the structure and behavior of complex data. For instance, the first derivative of a probability density function is used to identify peaks and valleys in data distributions, which is crucial in signal processing, medical imaging, and clustering algorithms such as mean-shift. Second and higher-order derivatives help detect curvature and inflection points, aiding in edge detection in images and in understanding changes in financial risk or environmental patterns. In machine learning, density

* Corresponding author

derivatives are applied in nonparametric clustering and anomaly detection, where rapid changes in the density function may indicate outliers or hidden subgroups. Overall, density derivatives serve as valuable tools in analyzing real-world data, revealing subtle patterns and structural features that are often hidden in raw observations. Density-derivative estimation involves calculating the derivatives of a probability density function directly from data, without explicit density estimation. These derivatives provide valuable insights into the structure of data, such as identifying modes, detecting boundaries, and analyzing gradients. Various methods, including kernel-based approaches and neural network models, are commonly employed for density-derivative estimation.

Let $(X_1, \ldots, X_n)$ be a sample, independent and identically distributed from a continuous random variable X, with unknown probability density function f with support $[0, 1]$. The classical kernel density estimator proposed by Rosenblatt (1956) and further developed by Parzen (1962) is given by

$$f_C(x) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right), \tag{1.1}$$

where h is a smoothing bandwidth such that $h \to 0$ and $nh \to \infty$ as $n \to \infty$ and k is a symmetric and non-negative function with support $[-1, 1]$ satisfying the following conditions,

$$\int_{-1}^{1} k(t)\, dt = 1, \int_{-1}^{1} tk(t)\, dt = 0, \int_{-1}^{1} t^2 k(t)\, dt = \mu_2 < \infty. \tag{1.2}$$

More precisely, $\mu_1 = 0$ since k is symmetric. With a suitable choice of h, we can divide the support of the density into distinct regions. The intervals $[0, h]$ and $[1 - h, 1]$ are called the left and the right boundary region, respectively, and the interior region is formed by the interval $]h, 1 - h[$. The performance of the kernel density estimator differs at least on one side of the support, specifically for $(x \in [0, h] \cup [1 - h, 1])$, and it differs from the interior points due to so-called boundary problems. To remove these boundary effects at the left region $(x \in [0, h])$, a diversity of methods has been developed during the past two decades. Among them,

- The reflection method (Schuster, 1985) reduces boundary bias in kernel density estimation by mirroring data points across the boundary, ensuring that the kernel mass lost outside the support is compensated by the reflected points.

- The transformation method (Marron and Ruppert, 1994) consists of a three-step process. First, a transformation is selected from a parametric family so that the density of the transformed data has a first derivative that is approximately equal to 0 at the boundaries of its support. Next, a kernel estimator with reflection is applied to the new data. Finally, this estimator is converted by the change-of-variables formula to obtain an estimate of f.

- The boundary kernel method (Jones, 1993) is more general than the reflection method in that it can adapt to any shape of densities. However a drawback of this method is that the estimates may be negative.

- The pseudo-data method (Cowling and Hall, 1996) involves constructing artificial (pseudo) observations near the boundary to correct bias problem.

Estimating the derivative of a density function is crucial in various statistical and applied fields. It helps identify modes or peaks in data distributions, which are essential for clustering algorithms like mean shift and mode estimation. In nonparametric statistics, density derivatives aid in selecting optimal bandwidths for kernel density estimation (KDE) and improving smoothing techniques. Additionally, they play a key role in anomaly detection by identifying abrupt changes in probability distributions. In information theory and Bayesian analysis, density derivatives contribute to entropy estimation and divergence measures. Moreover, they are used in statistical inference, where second derivatives relate to Fisher information and likelihood-based estimation. Overall, estimating density derivatives provides deeper insights into data structures, enabling better decision-making in various disciplines.

The first-order density derivative is defined as $f^{(1)}(x) = \dfrac{df(x)}{dx}$. A natural estimator for the first-order derivative, based on the kernel density estimator, is given by:

$$f_C^{(1)}(x) = \frac{1}{nh^2} \sum_{i=1}^{n} k^{(1)} \left(\frac{x - X_i}{h} \right), \tag{1.3}$$

where $k^{(1)}$ is the first derivative of the kernel function k. When $f^{(3)}$, the third derivative of f, exists and is continuous in a neighbourhood of x and when $nh^3 \to \infty$, the bias and variance of $f_C^{(1)}$ can be derived straightforwardly to obtain (Hansen, 2009),

$$Bias\left(f_C^{(1)}(x)\right) = \frac{1}{2} h^2 \mu_2 f^{(3)}(x) + o\left(h^2\right), \tag{1.4}$$

and

$$Var\left(f_C^{(1)}(x)\right) = \frac{f(x)}{nh^3} \int_0^1 \left(k^{(1)}(t)\right)^2 dt + o\left(\frac{1}{nh^3}\right). \tag{1.5}$$

Many researchers have studied the estimation of density derivatives. We mention among them, Jones (1994), which provides the estimators of density derivatives based on polynomial multiples of kernels and compares them with those based on differentiated kernels. Funke and Hirukawa (2023) studies the problem of estimating the first-order density derivative using asymmetric kernels. Uzuazor and Ojobor (2023) examines the performance of higher-order kernel estimation and kernel density derivatives estimation techniques with reference to the Gaussian kernel estimator owing to its wide applicability in real-life settings. Sasaki et al. (2015) give a direct method to approximate the density derivative without estimating the density itself.

In kernel density derivative estimation, accurately estimating the derivative near the boundaries of the support is challenging due to boundary bias. With appropriate choice of h, we can divide the support of the density into regions, the intervals $[0, h]$ and $[1 - h, 1]$ are referred to as the left and the right boundary region respectively and the interior region formed by the interval $]h, 1 - h[$. The performance of the kernel density derivative estimator near the boundaries of the support ($x \in [0, h] \cup [1-h, 1]$), differs from the interior points due to so-called boundary problems and the region

formed by the points with boundary problems is called the boundary region. While both left and right boundaries suffer from bias, the right boundary often poses greater difficulties, especially in truncated or skewed distributions, where the density drops sharply or is bounded from above. In this paper, we focus on the boundary bias problem in the right side of the support $[1-h,1]$, the kernel derivative density estimator has the well-known boundary problem. The bias and variance, for $x = 1 - ch$, with $c \in [0,1]$, are

$$\begin{aligned} Bias\big(f_C^{(1)}(x)\big) &= -f^{(1)}(x)\int_{-1}^{-c} k(t)\,dt - hf^{(2)}(x)\int_{-c}^{1} tk(t)\,dt \\ &+ \frac{h^2}{2}f^{(3)}(x)\int_{-c}^{1} t^2k(t)\,dt + o\left(h^2\right), \end{aligned} \tag{1.6}$$

and

$$Var\big(f_C^{(1)}(x)\big) = \frac{f(x)}{nh^3}\int_{-c}^{1}\left(k^{(1)}(t)\right)^2 dt + o\left(\frac{1}{nh^3}\right). \tag{1.7}$$

According to the expression of the bias (1.6), we can observe that $f_C^{(1)}$ is not a consistent estimator of $f^{(1)}$ and there exists an extra first-order term of h.

To resolve this boundary problem, we propose a new approach based on a generalized reflection method. The core technique involves reflecting a transformation of the observed data to construct the proposed estimator. The transformation and reflection method is a technique used to address boundary bias issues in kernel density derivative estimation. When estimating density derivatives near the boundaries of the support of a probability distribution, standard kernel methods often suffer from bias because the kernel function extends beyond the support, leading to inaccurate estimates. Our proposed method modifies the kernel function to accommodate the boundary conditions. To evaluate its effectiveness, we compare the boundary performance of our estimator with other kernel density estimators. Recognizing that meaningful comparisons between different methods require evaluating their respective optimal performances, we have adopted this strategy to ensure a fair and insightful assessment. This approach highlights the relative strengths and weaknesses of the proposed method in handling boundary issues.

The remainder of the paper is organized as follows. Notations and theoretical properties of the proposed estimator are introduced in Section 2. In Section 3, we support the theoretical results by simulation studies. Concluding remarks are given in Section 4.

2 Double Transformation-reflection Estimator

In 1985, Schuster proposed the reflection method, which is based on an intuitive idea: using the data as a mirror to correct the bias at the boundaries. When a kernel density is estimated near a boundary, such as zero, the symmetric kernel used overflows outside the data support, assigning weight to empty regions, resulting in an underestimate. To correct this, the method consists of mirroring the data in relation to the boundary, as if a mirror were placed at the end of the support. Marron and Ruppert in 1994 suggested another method called the transformation method. The principle of this method is to transform the original data into an unbounded space, where kernel estimation can be

applied without the problem of bias at the extremities. After estimating the derivative of the density in the transformed space, we then apply the change-of-variable formula to return to the original space. Karunamuni and Alberts (2005) combines those two methods to construct a new approach which corrects the boundary problem on the left side of the support for the density estimation.

We combine the reflection method with the transformation method to construct a new approach that addresses the boundary problem on the right side of the support. Assume that the transformation used in proposed estimator is a nonnegative, continuous and monotonically increasing function on $[0,1]$. In the context of bias reduction at the right boundary in kernel density derivative estimation, our proposed estimator called Double Transformation-Reflection (DTR), defined as follows,

$$f_{DTR}^{(1)}(x) = \frac{1}{nh^2}\sum_{i=1}^{n}\left(k^{(1)}\left(\frac{x-\psi(X_i)}{h}\right) + k^{(1)}\left(\frac{x-2+\psi(X_i)}{h}\right)\right), \tag{2.1}$$

where ψ is transformations that need to be determined.

Lemma 2.1. *Assume that ψ^{-1} exists and $\psi(1)=1$, $\psi^{(1)}(1)=1$ and the functions f and ψ has a third continuous derivative on $[0,1]$, with $\psi^{(0)}=\psi$. Then the bias and variance of $f_{DTR}^{(1)}$ for $x=1-ch$, $0\le c<1$ are given by,*

$$\begin{aligned} \mathit{Bias}\big(f_{DTR}^{(1)}(x)\big) = h\int_{-c}^{1} k(t)\Bigg(&\Big(-2f^{(2)}(1)\int_{-c}^{1}(t+c)k(t)\,dt\Big) \\ &+f^{(1)}(1)\psi^{(2)}(1)\left(2\int_{-c}^{1}(t+c)k(t)\,dt + c\right)\Bigg) \\ &+\frac{h^2}{2}\Bigg(-c^2f^{(3)}(1)+f^{(3)}(1)-f^{(1)}(1)\psi^{(3)}(1) \\ &-3\psi^{(2)}(1)\Big(f^{(2)}(1)-f^{(1)}(1)\psi^{(2)}(1)\Big)\int_{-1}^{1}(t+c)^2k(t)\,dt\Bigg)+o(h^2) \end{aligned} \tag{2.2}$$

and,

$$Var(f_{DTR}^{(1)}(x)) = \frac{2f(1)}{nh^3}\int_{-c}^{1}k^{(1)}(t)\,k^{(1)}(-2c-t)\,dt + o\left(\frac{1}{nh^3}\right). \tag{2.3}$$

Proof. Our first step is to proof the bias

$$\begin{aligned} E\left(f_{DTR}^{(1)}(x)\right) &= E\left(\frac{1}{nh^2}\sum_{i=1}^{n}\left(k^{(1)}\left(\frac{x-\psi(X_i)}{h}\right)\right)\right) + E\left(\frac{1}{nh^2}\sum_{i=1}^{n}k^{(1)}\left(\frac{x-2+\psi(X_i)}{h}\right)\right) \\ &= W_1 + W_2, \end{aligned}$$

where

$$\begin{aligned}
W_1 &= E\left(\frac{1}{nh^2}\sum_{i=1}^{n}\left(k^{(1)}\left(\frac{x-\psi(X_i)}{h}\right)\right)\right)\\
&= \frac{1}{h}\int_0^1 k\left(\frac{x-\psi(y)}{h}\right)f^{(1)}(y)\,dy\\
&= \int_{-c}^{1} k(t)\,\frac{f^{(1)}\left(\psi^{-1}(1-h(t+c))\right)}{\psi^{(1)}\left(\psi^{-1}(1-h(t+c))\right)}dt\\
&= \int_{-c}^{1} k(t)\left\{\frac{f^{(1)}(1)}{\psi^{(1)}\left(\psi^{-1}(1)\right)} - h(t+c)\right.\\
&\times \frac{f^{(2)}(1)\,\psi^{(1)}\left(\psi^{-1}(1)\right)-f^{(1)}(1)\,\psi^{(2)}\left(\psi^{-1}(1)\right)}{\left(\varphi^{(1)}\left(\varphi^{-1}(1)\right)\right)^3}+\frac{\left(h(t+c)\right)^2}{2}\\
&\times \left[\frac{f^{(3)}(1)\,\psi^{(1)}\left(\psi^{-1}(1)\right)-f^{(1)}(1)\,\psi^{(3)}\left(\psi^{-1}(1)\right)}{\left(\psi^{(1)}\left(\psi^{-1}(1)\right)\right)^4}\right.\\
&+ \left.\left.\frac{-3\psi^{(2)}\left(\psi^{-1}(1)\right)\left(f^{(2)}(1)\,\psi^{(1)}\left(\psi^{-1}(1)\right)-f^{(1)}(1)\,\psi^{(2)}\left(\psi^{-1}(1)\right)\right)}{\left(\psi^{(1)}\left(\psi^{-1}(1)\right)\right)^5}\right]\right\}dt+o\left(h^2\right),
\end{aligned}$$

and

$$\begin{aligned}
W_2 &= E\left(\frac{1}{nh^2}\sum_{i=1}^{n}k^{(1)}\left(\frac{x-2+\psi(X_i)}{h}\right)\right)\\
&= \frac{1}{h}\int_0^1 k\left(\frac{x-2+\psi(y)}{h}\right)f^{(1)}(y)\,dy\\
&= \int_{-1}^{-c} k(t)\,\frac{f^{(1)}\left(\psi^{-1}(1+h(t+c))\right)}{\psi^{(1)}\left(\psi^{-1}(1+h(t+c))\right)}dt\\
&= \int_{-1}^{-c} k(t)\left\{\frac{f^{(1)}(1)}{\psi^{(1)}\left(\psi^{-1}(1)\right)} + h(t+c)\right.\\
&\times \frac{f^{(2)}(1)\,\psi^{(1)}\left(\psi^{-1}(1)\right)-f^{(1)}(1)\,\psi^{(2)}\left(\psi^{-1}(1)\right)}{\left(\psi^{(1)}\left(\psi^{-1}(1)\right)\right)^3}+\frac{\left(h(t+c)\right)^2}{2}\\
&\times \left[\frac{f^{(3)}(1)\,\psi^{(1)}\left(\psi^{-1}(1)\right)-f^{(1)}(1)\,\psi^{(3)}\left(\psi^{-1}(1)\right)}{\left(\psi^{(1)}\left(\psi^{-1}(1)\right)\right)^4}\right.\\
&+ \left.\left.\frac{-3\psi^{(2)}\left(\psi^{-1}(1)\right)\left(f^{(2)}(1)\,\psi^{(1)}\left(\psi^{-1}(1)\right)-f^{(1)}(1)\,\psi^{(2)}\left(\psi^{-1}(1)\right)\right)}{\left(\psi^{(1)}\left(\psi^{-1}(1)\right)\right)^5}\right]\right\}dt+o\left(h^2\right).
\end{aligned}$$

Using the condition $\psi^{-1}(1)=1$ and $\psi^{(1)}(1)=1$, we get

$$\begin{aligned} W_1 &= f^{(1)}(1)\int_{-c}^{1} k(t)\,dt - h\left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right)\int_{-c}^{1}(t+c)\,k(t)\,dt \\ &+ \frac{h^2}{2}\left(f^{(3)}(1)-f^{(1)}(1)\,\psi^{(3)}(1)-3\psi^{(2)}(1)\left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right)\right) \\ &\times \int_{-c}^{1}(t+c)^2\,k(t)\,dt + o\left(h^2\right), \end{aligned} \tag{2.4}$$

and

$$\begin{aligned} W_2 &= f^{(1)}(1)\int_{-1}^{-c} k(t)\,dt + h\left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right)\int_{-1}^{-c}(t+c)\,k(t)\,dt \\ &+ \frac{h^2}{2}\left(f^{(3)}(1)-f^{(1)}(1)\,\psi^{(3)}(1)-3\psi^{(2)}(1)\left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right)\right) \\ &\times \int_{-1}^{-c}(t+c)^2\,k(t)\,dt + o\left(h^2\right). \end{aligned} \tag{2.5}$$

By the existence and continuity of $f^{(3)}(.)$ near 1, for $x=1-ch$, we have

$$f^{(1)}(1) = f^{(1)}(x) + chf^{(2)}(x) + \frac{(ch)^2}{2}f^{(3)}(x) + o\left(h^2\right),$$

$$f^{(2)}(x) = f^{(2)}(1) - chf^{(3)}(1) + o(h),$$

and

$$f^{(3)}(x) = f^{(3)}(1) + o(1). \tag{2.6}$$

Substituting (2.6) into (2.5) and (2.4), we obtain

$$\begin{aligned} W_1 &= \int_{-c}^{1} k(t)\left\{f^{(1)}(x) + chf^{(2)}(1) - \frac{(ch)^2}{2}f^{(3)}(1) - h(t+c)\right. \\ &\times \left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right) + \frac{h^2}{2}(t+c)^2 \\ &\times \left.\left(f^{(3)}(1)-f^{(1)}(1)\,\psi^{(3)}(1)-3\psi^{(2)}(1)\left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right)\right)\right\}dt + o\left(h^2\right), \end{aligned}$$

and

$$\begin{aligned} W_2 &= \int_{-1}^{-c} k(t)\left\{f^{(1)}(x) + chf^{(2)}(1) - \frac{(ch)^2}{2}f^{(3)}(1) + h(t+c)\right. \\ &\times \left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right) + \frac{h^2}{2}(t+c)^2 \\ &\times \left.\left(f^{(3)}(1)-f^{(1)}(1)\,\psi^{(3)}(1)-3\psi^{(2)}(1)\left(f^{(2)}(1)-f^{(1)}(1)\,\psi^{(2)}(1)\right)\right)\right\}dt + o\left(h^2\right). \end{aligned}$$

Hence

$$\begin{aligned}\text{Bias}\left(f_{DTR}^{(1)}(x)\right) = h\int_{-c}^{1}k(t)\Bigg(&-2f^{(2)}(1)\int_{-c}^{1}(t+c)k(t)dt\\&+f^{(1)}(1)\psi^{(2)}(1)\left(2\int_{-c}^{1}(t+c)k(t)dt+c\right)dt\\&+\frac{h^2}{2}\Big(-c^2f^{(3)}(1)+\Big(f^{(3)}(1)-f^{(1)}(1)\psi^{(3)}(1)\\&-3\psi^{(2)}(1)\left(f^{(2)}(1)-f^{(1)}(1)\psi^{(2)}(1)\right)\Big)\times\int_{-1}^{1}(t+c)^2k(t)dt\Bigg)+o\left(h^2\right).\end{aligned}$$

The task now is to prove the variance. We have

$$\begin{aligned}Var\left(f_{DTR}^{(1)}(x)\right) &= \frac{1}{n^2h^4}Var\left(\sum_{i=1}^{n}\left(k^{(1)}\left(\frac{x-\psi(X_i)}{h}\right)\right)+\sum_{i=1}^{n}k^{(1)}\left(\frac{x-2+\psi(X_i)}{h}\right)\right)\\&= U_1+U_2.\end{aligned}$$

Then

$$\begin{aligned}U_1 &= \frac{1}{nh^4}E\left(k^{(1)}\left(\frac{x-\psi(X_i)}{h}\right)+k^{(1)}\left(\frac{x-2+\psi(X_i)}{h}\right)\right)^2\\&= \frac{1}{nh^4}\int_0^1\left(\left(k^{(1)}\left(\frac{x-\psi(y)}{h}\right)\right)^2+\left(k^{(1)}\left(\frac{x-2+\psi(y)}{h}\right)\right)^2\right.\\&+ \left.2k^{(1)}\left(\frac{x-\psi(y)}{h}\right)k^{(1)}\left(\frac{x-2+\psi(y)}{h}\right)\right)f(y)\,dy\\&= U_{11}+U_{12}+U_{13},\end{aligned}$$

where

$$\begin{aligned}U_{11} &= \frac{1}{nh^4}\int_0^1\left(k^{(1)}\left(\frac{x-\psi(y)}{h}\right)\right)^2f(y)\,dy\\&= \frac{1}{nh^3}\int_{-c}^{1}\left(k^{(1)}(t)\right)^2\frac{f^{(1)}\left(\psi^{-1}(x-ht)\right)}{\psi^{(1)}\left(\psi^{-1}(1-h(t+c))\right)}dt\\&= \frac{f(1)}{nh^3}\int_{-c}^{1}\left(k^{(1)}(t)\right)^2dt+o\left(\frac{1}{nh^3}\right),\\U_{12} &= \frac{1}{nh^4}\int_0^1\left(k^{(1)}\left(\frac{x-2+\psi(y)}{h}\right)\right)^2f(y)\,dy\\&= \frac{1}{nh^3}\int_{-1}^{-c}\left(k^{(1)}(t)\right)^2\frac{f\left(\psi^{-1}(1+h(t+c))\right)}{\psi^{(1)}\left(\psi^{-1}(1+h(t+c))\right)}dt\\&= \frac{f(1)}{nh^3}\int_{-1}^{-c}\left(k^{(1)}(t)\right)^2dt+o\left(\frac{1}{nh^3}\right),\end{aligned}$$

and

$$\begin{aligned}U_{13} &= \frac{2}{nh^4}\int_0^1k^{(1)}\left(\frac{x-\psi(y)}{h}\right)k^{(1)}\left(\frac{x-2+\psi(y)}{h}\right)f(y)\,dy\\&= \frac{2f(1)}{nh^3}\int_{-c}^{1}k^{(1)}(t)\,k^{(1)}(-2c-t)\,dt+o\left(\frac{1}{nh^3}\right).\end{aligned}$$

Also, it can be easily shown that

$$U_2 = \frac{1}{nh^4}\left[E\left(k^{(1)}\left(\frac{x-\psi\left(X_i\right)}{h}\right)+k^{(1)}\left(\frac{x-2+\psi\left(X_i\right)}{h}\right)\right)\right]^2 = o\left(\frac{1}{nh^3}\right).$$

Summing U_1 and U_2 gives us the desired result for the variance.

The main objective of our transformation ψ is to eliminate the first-order term in the bias expression (2.2). From the bias expression (2.2), it is evident that the term multiplied by h depends on c, in other words, the transformation function is influenced by the location of the estimation point within the right boundary region. The transformation function ψ is locally adaptive and must satisfy

$$\psi^{(2)}(1) = \frac{2f^{(2)}(1)\int_{-c}^{1}(t+c)\,k(t)\,dt}{f^{(1)}(1)\left(2\int_{-c}^{1}(t+c)\,k(t)\,dt+c\right)}. \tag{2.7}$$

We can easily construct the function ψ that verifies certain conditions. We employ the following transformation in our investigation

$$\psi(x) = dq_c - \beta(dq_c)^2 + \left(1-2\left(dq_c\right)+3\beta(dq_c)^2\right)x + \left(\left(dq_c\right)-3\beta(dq_c)^2\right)x^2 + \beta(dq_c)^2x^3, \tag{2.8}$$

where

$$d = \frac{f^{(2)}(1)}{f^{(1)}(1)}, \qquad q_c = \frac{2\int_{-1}^{-c}(t+c)\,k^{(1)}(t)\,dt}{2\int_{-1}^{-c}(t-c)\,k^{(1)}(t)\,dt - c}, \qquad \beta > \frac{1}{3}.$$

This condition on β is necessary for $\psi(x)$ to be an increasing function of x. We construct the transformation function as a cubic polynomial because it is easy to decide whether its derivative is always positive. For ψ be defined by (2.8), we have

$$Bias\left(f_{DTR}^{(1)}(x)\right) = \frac{h^2}{2}\left(-c^2f^{(3)}(1)+L(1)\int_{-1}^{1}(t+c)^2\,k(t)\,dt\right)+o\left(h^2\right),$$

where

$$L(1) = f^{(3)}(1) - 6dq_c\left(2dq_cf^{(1)}(1) - f^{(2)}(1) + \beta dq_cf^{(1)}(1)\right).$$

The mean integrated squared error $(MISE)$ of $f_{DTR}^{(1)}$ can be expressed as the sum of the integrated squared bias and the integrated variance, given by

$$\begin{aligned} MISE(f_{DTR}^{(1)}) &= \int_0^1 \mathbb{E}\left(f_{DTR}^{(1)}(x)-f(x)\right)^2 dx \\ &= \int_0^1 \left(Bias(f_{DTR}^{(1)}(x))\right)^2 dx + \int_0^1 \mathrm{Var}(f_{DTR}^{(1)}(x))\,dx \\ &= ISB(f_{DTR}^{(1)}) + IV(f_{DTR}^{(1)}). \end{aligned} \tag{2.9}$$

Estimation of d In practice, ψ is not available because it is defined by an unknown term d. So d can be written as the derivative of $logf^{(1)}(x)$ evaluated at $x = 1$, and hence it can be estimated by

$$d_n = \frac{\log f_C^{(1)}(1) - \log f_C^{(1)}(1-h)}{h}.$$

The proposed estimator is not very delicate to the precise details of the pilot estimate of d. Therefore, any appropriate estimate can be used. So, we define

$$\psi_n(x) = d_n q_c - \beta (d_n q_c)^2 + \Big(1 - 2(d_n q_c) + 3\beta (d_n q_c)^2\Big) x + \Big((d_n q_c) - 3\beta (d_n q_c)^2\Big) x^2 + \beta (d_n q_c)^2 x^3.$$

3 Numerical Analysis

In this section, we perform a Monte Carlo study to implement our proposed estimator and illustrate its finite sample performance against the other well-known estimators. In addition, we provide a real example that shows the effectiveness of our estimator. All computations were done by utilizing R Software.

3.1 Simulated data

To evaluate the estimator performance at the boundaries, we focused on sample sizes of 200 and 500, simulating $R = 1000$ iterations for each distribution. Throughout our study, we use the Epanechnikov kernel $k(t) = (3/4)(1 - t^2)I(|t| \leq 1)$, where I denotes the indicator function. We evaluate the mean integrated squared error $(MISE)$ for all estimators. In addition to the $MISE$, we also evaluate the Integrated Squared Bias (ISB). To be more specific, let $\hat{\theta}(x)$ be the estimate of $\theta(x)$, where x is a specific point on the support (a, b). Let $\Delta = (b - a)/m$ be the length of m intervals $I_j, j = 1, ..., m$, and $\hat{\theta}_r(x)$ be the estimate of $\theta(x)$ based on the r^{th} generated random numbers of size n. Also, let R be the number of replications. The Monte Carlo estimator of the $MISE$ is

$$MISE(\hat{\theta}) = \frac{1}{R} \sum_{r=1}^{R} \sum_{x \in I_1}^{I_m} \Big(\hat{\theta}_r(x) - \theta(x)\Big)^2 \Delta.$$

Then the Monte Carlo estimate of ISB is

$$ISB(\hat{\theta}) = \sum_{x \in I_1}^{I_m} \big(\bar{\theta}(x) - \theta(x)\big)^2 \Delta,$$

where $\bar{\theta}(x) = \frac{1}{R} \sum_{r=1}^{R} \hat{\theta}_r(x)$.

3.1.1 Compared Estimators

For a complete picture and to test the performance of our estimator on the right boundary support, we compare it with the other density derivative estimators. Among these methods, we compare

the performance of the kernel density derivative estimator, the reflection kernel density derivative estimator, defined by

$$f_{RE}^{(1)}(x) = \frac{1}{nh^2}\sum_{i=1}^{n}\left(k^{(1)}\left(\frac{x-X_i}{h}\right) + \sum_{i=1}^{n}k^{(1)}\left(\frac{x-2+X_i}{h}\right)\right),$$

and the general boundary kernel density derivative estimator, given by

$$f_{B}^{(1)}(x) = \frac{1}{nh^2}\sum_{i=1}^{n}k_B^{(1)}\left(\frac{x-X_i}{h}\right),$$

where the modified kernel function is given at the right boundary region based on the Epanechnikov kernel (see georg Müller (1991)) by

$$k_B(t) = \frac{6}{(1+c)^3}(1-t)(c+t)\left(1+5\left(\frac{1-c}{1+c}\right)^2 - 10\frac{1-c}{(1+c)^2}t\right)I(-c \le t \le 1).$$

This kernel satisfies the following conditions

$$\int_{-1}^{-c} k_B(t)\,dt = 0,\ \int_{-c}^{1} k_B(t)\,dt = 1,\ \int_{-c}^{1} tk_B(t)\,dt = 0,\ \int_{-c}^{1} t^2k_B(t)\,dt < \infty. \tag{3.1}$$

Note that k_B is a natural continuation of the Epanechnikov's kernel. When putting $c = 1$ in $k_B(t)$, one obtains $k(t)$. Tenreiro (2018) and Zhang et al. (2020) argue that k_B must satisfy

$$\int_{-c}^{1}\frac{c+t}{c}k_B(t)dt = 1.$$

3.1.2 Results and discussions

We consider eight distributions with support $[0, 1]$, which contains various shapes of curves (Simple (Beta, Symmetric), Skewed (Beta, Asymmetric), Mixtures, Heavy-tailed (Lomax, Gamma)), listed in Table 1. For each simulated sample and each estimator, the smoothing bandwidth was chosen by an unbiased cross-validation method. The results of the Monte Carlo study, which compare the performance of the four estimators, are given in Tables 2 and 3. To confirm the results and affirm the effectiveness of our estimator, we add the graphs and the boxplots of the absolute bias included at the Figures 1-2.

It can be found that from Table 2, the classical estimator $f_C^{(1)}$ tends to perform worse. Across most distributions, the classical estimator has the highest $MISE$, indicating lower estimation accuracy, especially in skewed distributions like Truncated Lomax $(2, 2)$ or Truncated Gamma $(2, 6)$. Then, the reflection estimator $f_{RE}^{(1)}$ performs well in some cases, particularly efficient when the distribution is skewed (e.g., Truncated Lomax $(2, 2)$ or Truncated Normal $(2, 5)$), but not always superior. Next, we can see that the double transformation-reflection estimator $f_{DTR}^{(1)}$ is often the best.

It exhibits the smallest $MISE$ in 4 out of 6 distributions, especially when the density is smooth and the domain is bounded and it shows very low ISB, suggesting effective bias reduction. Finally, the boundary estimator $f_B^{(1)}$ is competitive; it performs closely to $f_{DTR}^{(1)}$ and $f_{RE}^{(1)}$, especially when the density is near boundaries (like in Beta distributions).

From Table 3, we conclude that the double transformation estimator $f_{DTR}^{(1)}$ consistently achieves the lowest $MISE$ values in most cases, particularly for smooth, bounded, or mixture distributions such as Beta$(4, 4)$, 0.5Beta$(3, 7)$ + 0.5Beta$(2, 5)$, and 0.25Beta$(4, 5)$ + 0.75Beta$(2, 8)$. This highlights its robustness and adaptability to different data shapes. The boundary estimator $f_B^{(1)}$ also shows strong performance, especially in distributions with support near the boundaries, such as Beta$(4, 7)$ and Truncated Normal$(2, 5)$, where boundary effects could otherwise introduce significant bias. The reflection estimator $f_{RE}^{(1)}$ performs notably well in skewed or heavy-tailed cases like Truncated Lomax$(2, 2)$ and Truncated Gamma$(2, 6)$, effectively minimizing the integrated squared bias. In contrast, the classical estimator $f_C^{(1)}$ exhibits the highest $MISE$ and ISB in nearly all scenarios, confirming its limitations in practical applications where boundary effects or skewness are present. Overall, the results indicate that while the classical approach lacks flexibility, advanced estimators like $f_{DTR}^{(1)}$, $f_{RE}^{(1)}$, and $f_B^{(1)}$ provide more accurate and reliable estimates depending on the nature of the underlying distribution.

Figure 1 presents boxplots of estimation errors for the first derivative of density functions, using four estimators across a variety of distributions. Each subplot corresponds to a specific distribution, including symmetric and asymmetric Beta distributions, mixtures, and truncated forms of Lomax, Gamma, and Normal. A clear pattern emerges across nearly all distributions: the $f_{DTR}^{(1)}$ estimator consistently shows the lowest error with tight variability, especially in the Beta$(4, 7)$, Truncated Normal$(2, 5)$, and Mixture distributions, indicating both accuracy and robustness. In contrast, the classical $f_C^{(1)}$ and boundary $f_B^{(1)}$ estimators tend to show larger error spreads, particularly under complex or asymmetric shapes like Truncated Lomax and Truncated Gamma, where the variability is most pronounced. The $f_{RE}^{(1)}$ estimator performs moderately well, generally improving upon the classical method but not reaching the low error levels of $f_{DTR}^{(1)}$. These results underscore the effectiveness of the $f_{DTR}^{(1)}$ approach, particularly for challenging boundary conditions and skewed or truncated distributions.

Figure 2 presents a comparison of four methods for estimating the first derivative of density functions across various distributions, including symmetric, asymmetric, truncated, and mixture cases. The true derivative is shown in black, while the estimators $f_C^{(1)}$, $f_{RE}^{(1)}$, $f_{DTR}^{(1)}$, and $f_B^{(1)}$ are plotted with different line styles and colors. In symmetric cases like Beta$(4, 4)$ and Beta$(3, 4)$, most estimators perform reasonably well in the interior but tend to deviate near the right boundary. The classical estimator $f_C^{(1)}$ and the reflection estimator $f_{RE}^{(1)}$ often show reduced bias compared to $f_{DTR}^{(1)}$ and $f_B^{(1)}$, which can overshoot or oscillate, especially near boundaries. In more complex distributions such as mixtures (e.g., 0.5Beta(3,7) + 0.5Beta(2,5)) and truncated forms, the performance differences become more pronounced. For instance, in the truncated normal and truncated gamma cases, all estimators exhibit high variability, while $f_{DTR}^{(1)}$ tends to follow the true derivative more closely. Overall, the figure illustrates that while all methods have comparable behavior in the central region of the distribution, the $f_{DTR}^{(1)}$ estimator offers more stable and accurate boundary correction.

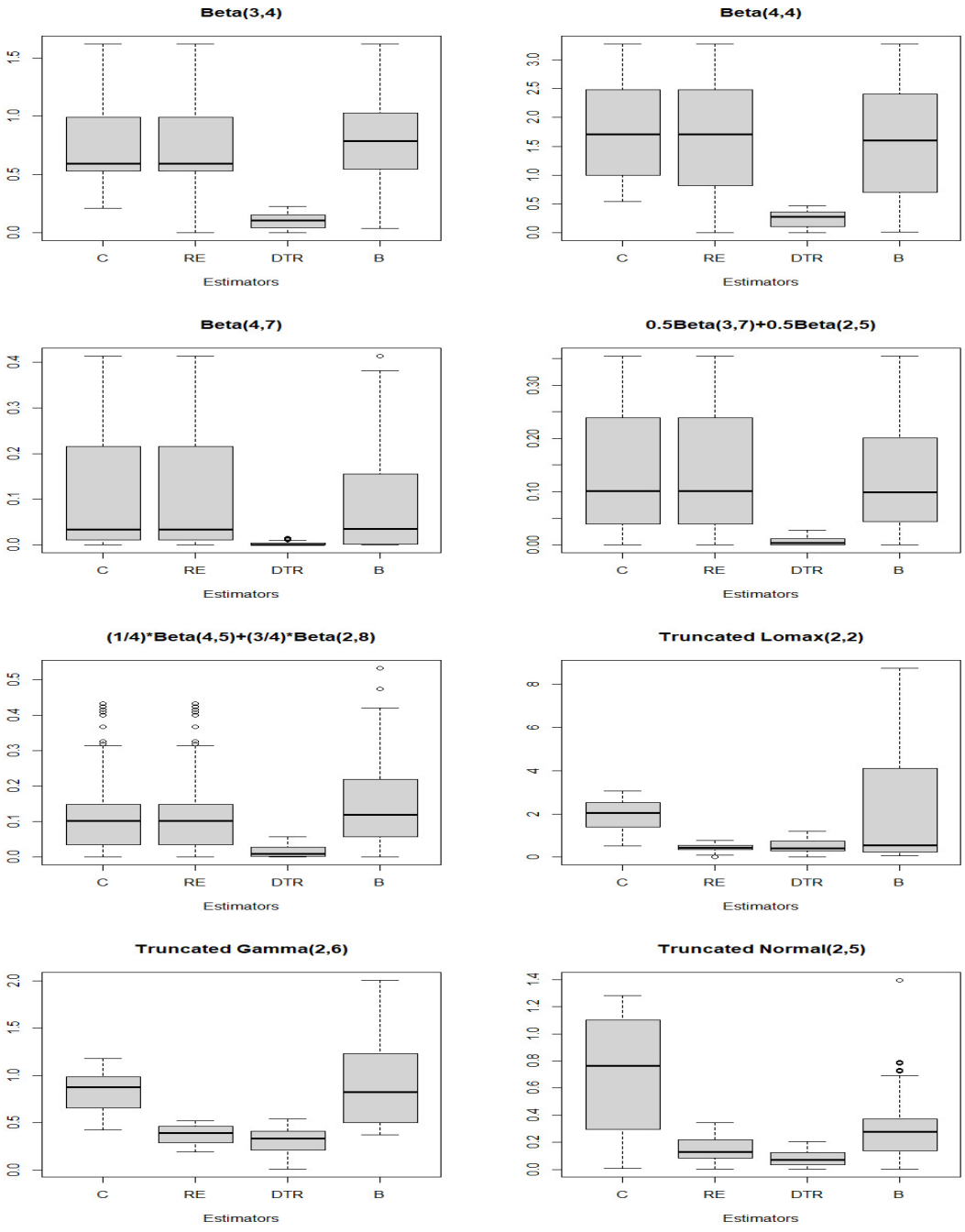

Figure 1: Boxplot of the absolute bias of estimators.

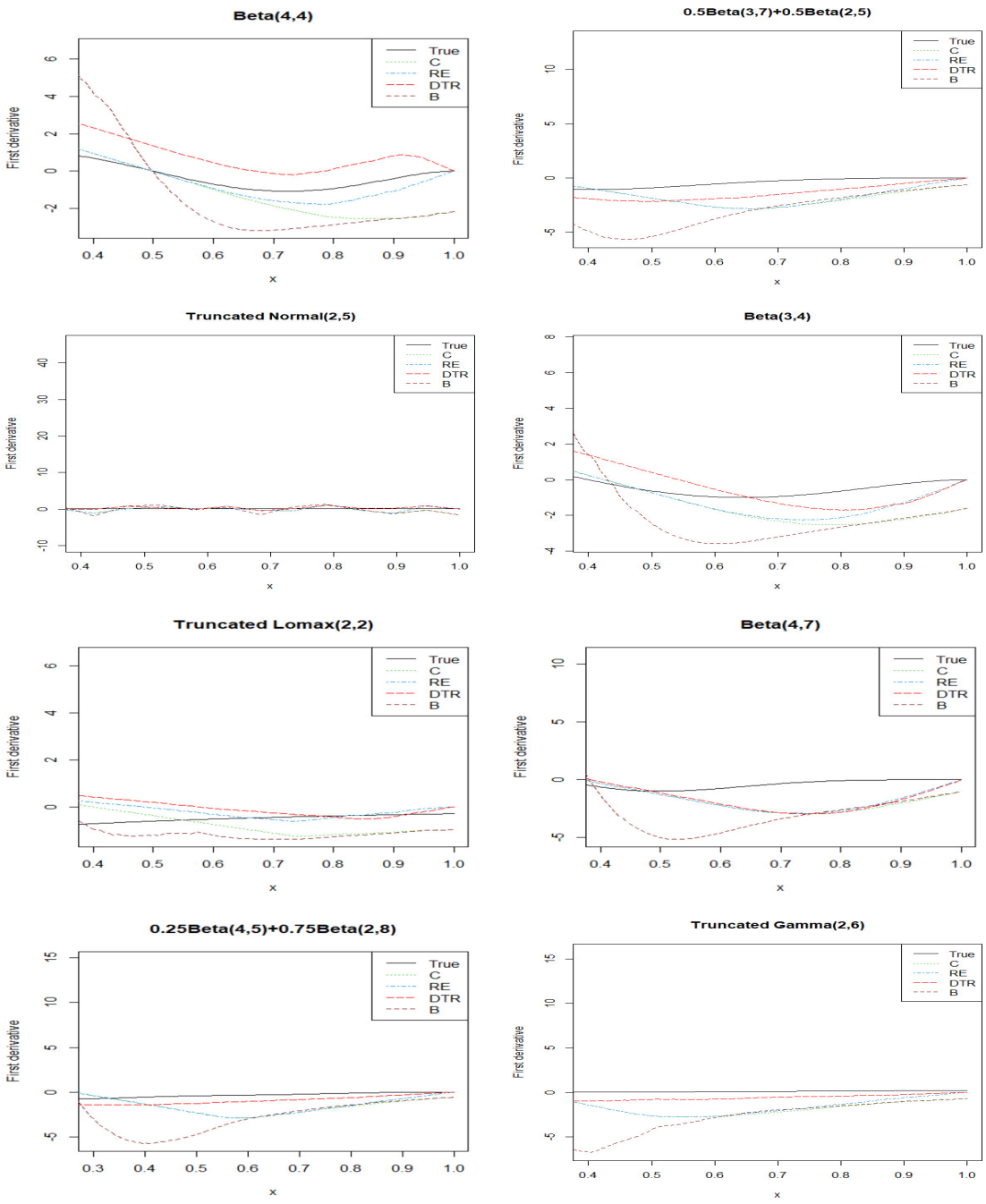

Figure 2: Comparison of first derivative estimates of the density function under different estimation methods (C, RE, DTR, B) across various distributions.

Table 1: Densities considered in Monte Carlo simulations.

$f(x), x \in [0,1]$	Description
$f_1(x) = 140x^3(1-x)^3$	$Beta(4,4)$
$f_2(x) = 3x(x-1)^4\left(42x^3 - 84x^2 + 42x + 5\right)$	$\frac{1}{2}Beta(3,7) + \frac{1}{2}Beta(2,5)$
$f_3(x) = (1/\sqrt{10\pi})\exp\left(-(1/2)\left((x-2)/(\sqrt{5})\right)^2\right)$	$Truncated\ Normal(2,5)$
$f_4(x) = 60x^2(1-x)^3$	$Beta(3,4)$
$f_5(x) = (1/0.5555556)(1+(x/2))^{-3}$	$Truncated\ \ Lomax(2,2)$
$f_6(x) = 840x^3(1-x)^6$	$Beta(4,7)$
$f_7(x) = -2x(x-1)^4\left(27x^3 - 116x^2 + 81x - 27\right)$	$\frac{1}{4}Beta(4,5) + \frac{3}{4}Beta(2,8)$
$f_8(x) = (2/117.92)\exp(-2x)(2x)^5$	$Truncated\ Gamma(2,6)$

3.2 Real data

As mentioned in the previous section, the double transformation estimator is the most effective. Therefore, we applied this estimator to real data to evaluate its performance. The data set examined is the number of confirmed cases of coronavirus in 2022 in Constantine Hospital, Algeria Constantine University Hospital, Epidemiology Department (2022). For this data set, we can map the unit interval by the transformation $Y_i = (X_i - a)/(b-a)$, where X_i a real observation in $[a,b]$. The smoothing parameter for the density and the derivative estimators is obtained by using UCV method. The proposed estimator is plotted in Figure 3. Table 4 shows the descriptive statistics of the data.

Figure 3 presents the double transformation kernel estimator applied to the COVID-19 dataset. The left panel shows the estimated density function, which appears to be unimodal, with a clear peak around $x = 0.8$. This indicates that the most frequent normalized value in the data lies near this point, possibly reflecting a period of relatively high case frequency or concentration. The right panel, showing the first derivative of the density, provides further evidence for this peak: the derivative changes sign from negative to positive at approximately $x = 0.8$. This sign change confirms the presence of a local maximum in the density estimate, thereby supporting the mode's location. Such behavior reflects the estimator's ability to not only smooth the data effectively but also capture meaningful structural features. Furthermore, the estimator remains stable and free from excessive bias near both boundaries, demonstrating the advantage of the double transformation technique in real-world, bounded data settings. This result suggests that the proposed method is well-suited for analyzing epidemiological data where accurate boundary behavior is crucial.

4 Conclusion

In this work, we have investigated a new approach for estimating the first density derivative with support $[0,1]$. We focus our study on the right boundary region. The main theoretical character-

Table 2: Values of MISE(ISB) of different kernel density derivative estimators for $n = 200$

$n = 200$		
	$Beta(4,4)$	$0.5Beta(3,7) + 0.5Beta(2,5)$
$f_C^{(1)}$	$0.58354\,(0.06629)$	$2.29717 \times 10^{-3}\,(2.06414 \times 10^{-4})$
$f_{RE}^{(1)}$	$0.55662\,(0.05768)$	$2.28081 \times 10^{-3}\,(1.97021 \times 10^{-4})$
$f_{DTR}^{(1)}$	$0.00779\,(0.00171)$	$2.30200 \times 10^{-6}\,(1.29061 \times 10^{-6})$
$f_B^{(1)}$	$0.49452\,(0.03863)$	$1.78838 \times 10^{-3}\,(8.10363 \times 10^{-5})$
	$Beta(3,4)$	$Truncated\ Lomax(2,2)$
$f_C^{(1)}$	$0.05994\,(4.91606 \times 10^{-3})$	$0.66891\,(8.07726 \times 10^{-2})$
$f_{RE}^{(1)}$	$0.05791\,(4.35572 \times 10^{-3})$	$0.00437\,(1.83781 \times 10^{-5})$
$f_{DTR}^{(1)}$	$0.00029\,(7.71595 \times 10^{-5})$	$0.04586\,(3.74787 \times 10^{-3})$
$f_B^{(1)}$	$0.04780\,(2.73724 \times 10^{-3})$	$0.00467\,(4.92360 \times 10^{-4})$
	$Truncated\ Normal(2,5)$	$0.25Beta(4,5) + 0.75Beta(2,8)$
$f_C^{(1)}$	$0.09201\,(1.10211 \times 10^{-2})$	$8.42341 \times 10^{-3}\,(8.53166 \times 10^{-4})$
$f_{RE}^{(1)}$	$0.00146\,(1.99785 \times 10^{-4})$	$8.34785 \times 10^{-3}\,(8.11997 \times 10^{-4})$
$f_{DTR}^{(1)}$	$0.00084\,(1.10826 \times 10^{-4})$	$2.00043 \times 10^{-5}\,(6.73101 \times 10^{-6})$
$f_B^{(1)}$	$0.00178\,(8.97835 \times 10^{-5})$	$6.96035 \times 10^{-3}\,(2.95171 \times 10^{-4})$
	$Truncated\,Gamma(2,6)$	$Beta(4,7)$
$f_C^{(1)}$	$0.16691\,(0.02304)$	$2.77203 \times 10^{-3}\,(2.00267 \times 10^{-4})$
$f_{RE}^{(1)}$	$0.06885\,(0.00836)$	$2.77046 \times 10^{-3}\,(1.97803 \times 10^{-4})$
$f_{DTR}^{(1)}$	$0.04872\,(0.00575)$	$2.18748 \times 10^{-6}\,(3.02430 \times 10^{-7})$
$f_B^{(1)}$	$0.10037\,(0.01372)$	$2.20675 \times 10^{-3}\,(7.47137 \times 10^{-5})$

Table 3: Values of MISE(ISB) of different kernel density derivative estimators for $n = 500$

$n = 500$		
	$Beta(4,4)$	$0.5Beta(3,7) + 0.5Beta(2,5)$
$f_C^{(1)}$	$0.59421\,(0.06768)$	$2.17658 \times 10^{-3}\,\left(1.96959 \times 10^{-4}\right)$
$f_{RE}^{(1)}$	$0.56724\,(0.05902)$	$2.16710 \times 10^{-3}\,\left(1.89399 \times 10^{-4}\right)$
$f_{DTR}^{(1)}$	$0.00826\,(0.00176)$	$4.92331 \times 10^{-6}\,\left(1.41880 \times 10^{-6}\right)$
$f_B^{(1)}$	$0.49848\,(0.03851)$	$1.71888 \times 10^{-3}\,\left(7.66209 \times 10^{-5}\right)$
	$Beta(3,4)$	$Truncated\ Lomax(2,2)$
$f_C^{(1)}$	$0.05585\,\left(4.58409 \times 10^{-3}\right)$	$0.67975\,\left(8.20370 \times 10^{-2}\right)$
$f_{RE}^{(1)}$	$0.05399\,\left(4.06979 \times 10^{-3}\right)$	$0.00474\,\left(6.80026 \times 10^{-5}\right)$
$f_{DTR}^{(1)}$	$0.00032\,\left(7.42965 \times 10^{-5}\right)$	$0.05114\,\left(4.71317 \times 10^{-3}\right)$
$f_B^{(1)}$	$0.04482\,\left(2.50235 \times 10^{-3}\right)$	$0.00834\,\left(1.13186 \times 10^{-3}\right)$
	$Norm\ tronq(2,5)$	$0.25Beta(4,5) + 0.75Beta(2,8)$
$f_C^{(1)}$	$0.09233\,\left(1.10679 \times 10^{-2}\right)$	$8.53162 \times 10^{-3}\,\left(8.61793 \times 10^{-4}\right)$
$f_{RE}^{(1)}$	$0.00119\,\left(1.64764 \times 10^{-4}\right)$	$8.43545 \times 10^{-3}\,\left(8.10116 \times 10^{-4}\right)$
$f_{DTR}^{(1)}$	$0.00295\,\left(3.92102 \times 10^{-4}\right)$	$2.89978 \times 10^{-5}\,\left(7.54768 \times 10^{-6}\right)$
$f_B^{(1)}$	$0.00088\,\left(9.65712 \times 10^{-5}\right)$	$6.80084 \times 10^{-3}\,\left(3.55194 \times 10^{-4}\right)$
	$Truncated\ Gamma(2,6)$	$Beta(4,7)$
$f_C^{(1)}$	$0.17425\,(0.02404)$	$3.09647 \times 10^{-3}\,\left(2.30374 \times 10^{-4}\right)$
$f_{RE}^{(1)}$	$0.07488\,(0.00910)$	$3.09478 \times 10^{-3}\,\left(2.27837 \times 10^{-4}\right)$
$f_{DTR}^{(1)}$	$0.06116\,(0.00733)$	$5.78254 \times 10^{-4}\,\left(6.42944 \times 10^{-5}\right)$
$f_B^{(1)}$	$0.09426\,(0.01250)$	$2.47650 \times 10^{-3}\,\left(8.07615 \times 10^{-5}\right)$

Table 4: Descriptive statistics of the measures intervals of confirmed cases of coronavirus in 2022 in Constantine Hospital, Algeria

Min	1st Qu	Median	Mean	3rd Qu	Max	kurtosis	skewness
0.0	5.0	10.0	144.3	67.5	2521.0	18.6	3.9

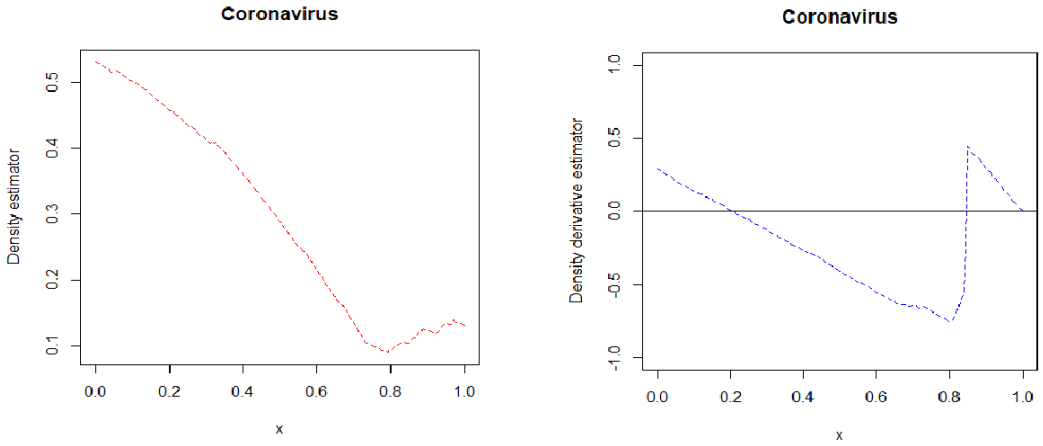

Figure 3: Double transformation-reflection estimator of density and density derivative from Coronavirus data.

istics of the proposed estimator are introduced. Moreover, the Monte Carlo simulations and real data example reveal the good results of the proposed estimator compared to the other well-known estimators. We carried out the simulation study in terms of ISB and $MISE$. Generally, our estimator gave superior performance among the other compared estimators. Note that the plots and simulations were carried out using the R software.

References

Constantine University Hospital, Epidemiology Department (2022), "Daily Data of Confirmed COVID-19 Cases (January 2022 – December 2022)," Unpublished data.

Cowling, A. and Hall, P. (1996), "On Pseudodata Methods for Removing Boundary Effects in Kernel Density Estimation," *Journal of the Royal Statistical Society: Series B (Methodological)*, 58, 551–563.

Funke, B. and Hirukawa, M. (2023), "Density Derivative Estimation Using Asymmetric Kernel," *Journal of Nonparametric Statistics*, 36, 994–1017.

georg Müller, H. (1991), "Smooth optimum kernel estimators near endpoints," *Biometrika*, 78, 521–530.

Hansen, B. (2009), "Lecture Notes on Nonparametrics," `http://www.ssc.wisc.edu/~bhansen/718/NonParametrics1.pdf`, lecture notes.

Jones, M. (1993), "Simple Boundary Correction for Kernel Density Estimation," *Statistics and Computing*, 3, 135–146.

— (1994), "On Kernel Density Derivative Estimation," *Communications in Statistics - Theory and Methods*, 23, 2133–2139.

Karunamuni, R. J. and Alberts, T. (2005), "On boundary correction in kernel density estimation," *Statistical Methodology*, 2, 191–212.

Marron, J. and Ruppert, D. (1994), "Transformations to Reduce Boundary Bias in Kernel Density Estimation," *Journal of the Royal Statistical Society: Series B (Methodological)*, 56, 653–671.

Parzen, E. (1962), "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, 33, 1065–1076.

Rosenblatt, M. (1956), "Remarks on Some Nonparametric Estimates of a Density Function," *The Annals of Mathematical Statistics*, 27, 832–837.

Sasaki, H., Noh, Y.-K., and Sugiyama, M. (2015), "Direct Density-Derivative Estimation and Its Application in KL-Divergence Approximation," in *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Proceedings of Machine Learning Research (PMLR), pp. 809–818.

Schuster, E. F. (1985), "Incorporating support constraints into nonparametric estimators of densities," *Communications in Statistics-Theory and methods*, 14, 1123–1136.

Tenreiro, C. (2018), "A new class of boundary kernels for distribution function estimation," *Communications in Statistics-Theory and Methods*, 47, 5319–5332.

Uzuazor, S. and Ojobor, S. (2023), "A Comparative Study of Higher Order Kernel Estimation and Kernel Density Derivative Estimation of the Gaussian Kernel Estimator with Data Application," *Pakistan Journal of Statistics and Operation Research*, 19, 299–311.

Zhang, S., Li, Z., and Zhang, Z. (2020), "Estimating a distribution function at the boundary," *Austrian Journal of Statistics*, 49, 1–23.

Received: March 15, 2025

Accepted: September 7, 2025