Pure and Applied Mathematics Journal
2026, Vol. 15, No. 5, pp. 79–98
https://doi.org/10.11648/j.pamj.20261505.11

SciencePG
Science Publishing Group

Research Article

An Automatic Smoothing Parameter Selection Criterion for Fuzzy Set Density Estimation

Jesús A. Fajardo-González *

Department of Mathematics, University of Oriente, Cumana, Venezuela

Abstract

The fuzzy set density estimator of a probability density function is an innovative non-kernel-based estimator that exhibits superior performance compared to the classical kernel density estimator. The numerical values assigned to the smoothing parameters directly govern the performance and precision of both estimators; however, although explicit theoretical formulas exist to compute their optimal global smoothing parameters, these expressions inherently depend on knowledge of the underlying density function. This fundamental constraint renders bandwidth selection the primary operational challenge when estimating an unknown probability density function using the fuzzy set density estimator, in stark contrast to the kernel density estimator, which benefits from a wide variety of well-established parameter selection criteria in statistical literature. To address this structural limitation, this paper introduces the risk function estimator corresponding to the integrated mean squared error within the general framework of fuzzy set density estimation. Consequently, the proposed estimator is fully equipped with an automatic and autonomous smoothing parameter selection criterion, which remains fully applicable regardless of whether the density function is known or completely unknown. The theoretical findings significantly expand the practical applicability of the fuzzy set density estimator and represent a milestone contribution to fuzzy density estimation methodologies by providing its first automatic data-driven bandwidth selection rule. Finally, the theoretical properties are thoroughly illustrated through a detailed numerical example and two real-data applications.

Keywords

Fuzzy Set Estimator, Non-parametric Density Estimation, Smoothing Parameter Selection, Risk Function Estimation, Cross-validation

1. Introduction

This work is framed within the general context of non-parametric estimation theory of the density function f from independent samples. Specifically, the research focuses on analyzing the performance of estimating f through the class of fuzzy set estimators —which are not kernel-based—, representing a particular aspect of the aforementioned theory. It is worth highlighting that the fuzzy set estimator of f was introduced by Falk and Liese [6], and a particular case was presented by Fajardo et al. [4], who defined it as follows:

$$\hat{\vartheta}_n(x) = \frac{1}{n\, b_n \int \varphi(s)\, ds} \sum_{i=1}^{n} \mathbb{1}_{[0,\varphi(\frac{X_i - x}{b_n})]}(V_i),$$

where $x \in \mathbb{R}$; $X_1, \ldots, X_n$ is a random sample from a random variable X with density f; $V_1, \ldots, V_n$ are independent random variables, uniformly distributed on $(0, 1)$ and independent of $X_1, \ldots, X_n$; $b_n \in \mathbb{R}_+$ is the smoothing parameter of $\hat{\vartheta}_n$, satisfying $b_n \to 0$ as $n \to \infty$; and $\varphi : \mathbb{R} \to [0, 1]$ is a thinning

*Correspondence: Jesús A. Fajardo-González (jafajardogonzalez@udo.edu.ve.com), Jesús A. Fajardo-González (jfajardogonzalez@gmail.com)

Received: 5 August 2026; **Accepted:** 19 August 2026; **Published:** 2 November 2026

Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

function such that $0 < \int\varphi(s)\,ds < \infty$ (Reiss [Section 2.4]12). Here and hereinafter, any integral whose region is not specified is assumed to be over $\mathbb{R}$. For simplicity of notation, hereinafter $\hat{\vartheta}_n$ will be rewritten as:

$$\hat{\vartheta}_n(x) = \frac{1}{n\,b_n\int\varphi(s)\,ds}\sum_{i=1}^{n}\mathbb{1}_{\varphi}\left(\frac{X_i - x}{b_n}, V_i\right), \tag{1}$$

where

$$\mathbb{1}_{\varphi}(y, v) = \mathbb{1}_{[0,\varphi(y)]}(v), \tag{2}$$

for each $y \in \mathbb{R}$ and $v \in (0, 1)$.

Below, some of the main theoretical properties of the estimator (1) are highlighted:

1) It constitutes a particular case of the estimator introduced by Falk and Liese [6].
2) It is not kernel-based, given that the function $\mathbb{1}_{\varphi}$ in (2) is not a kernel function.
3) The term "fuzzy set" was justified in Remark 2 of Fajardo [2], being a direct consequence of Remark 1 of Fajardo et al. [4].
4) Fajardo [2] proved that the optimal global smoothing parameter (b_{asy}) and the optimal integrated mean squared error ($IMSE_{asy}$) of (1) are:

$$b_{asy} = \left[\frac{\int\varphi(u)\,du}{n\left[\int u^2\varphi(u)\,du\right]^2\int\left[f''(u)\right]^2 du}\right]^{1/5}, \tag{3}$$

and

$$IMSE_{asy}[\hat{\vartheta}_n] = \frac{5}{4}\left[\left[\int u^2\varphi(u)\,du\right]^2 * \frac{\int\left[f''(u)\right]^2 du}{\left[\int\varphi(u)\,du\right]^6}\right]^{1/5} n^{-4/5} \tag{4}$$

where f'' represents the second derivative of f and b_n^* is the value that minimizes the integrated mean squared error ($IMSE$). Furthermore, he proved that the thinning function

$$\varphi(x) = \left[1 - \left(\frac{16}{25}x\right)^2\right]\mathbb{1}_{\left[-\frac{25}{16},\frac{25}{16}\right]}(x) \tag{5}$$

minimizes (4).

It should be noted that the role of (5) was decisive in the results introduced in Fajardo [2], as its proper construction was key to deducing the following important property of (1)

$$IMSE_{asy}[\hat{\vartheta}_n] \leq IMSE_{asy}[\hat{f}_n], \tag{6}$$

where $\hat{f}_n$ is the classical kernel estimator of f with kernel K, Rosenblatt [13] and Parzen [11], with optimal global smoothing parameter h_{asy} (Silverman [pp. 39-40]15). Recall that,

$$h_{asy} = \left[\frac{\int K^2(u)du}{n\left[\int u^2K(u)du\right]^2\int\left[f''(u)\right]^2 du}\right]^{1/5} \tag{7}$$

and

$$IMSE_{asy}[\hat{f}_n] = \frac{5}{4}\left[\left[\int K^2(u)du\right]^4\left[\int u^2K(u)du\right]^2 * \int\left[f''(u)\right]^2 du\right]^{1/5} n^{-4/5}. \tag{8}$$

From (6), it is inferred that the estimator (1) provides better estimates than $\hat{f}_n$ for any kernel K. Furthermore, (5) stands out as a function that allows selecting sample points with different probabilities, unlike classical kernel estimators, which assign identical weights to all sample points (Fajardo [Remark 1]2). These aspects, together with (6), enhance the performance of the estimator (1) when using the function φ given in (5). However, the quantitative evaluation of this performance is only possible when f is known, since calculating (4) and (8) requires the value of $\int[f''(u)]^2du$. Likewise, this limitation prevents calculating (3) and makes the analytical analysis of the performance of (1) unfeasible, which reveals a potential problem in the fuzzy set density estimation method. Finally, it is worth noting that a variety of criteria have been developed in the literature to estimate f using $\hat{f}_n$ under the aforementioned restriction.

The subjective choice of the smoothing parameter is an alternative that allows obtaining an appropriate value to estimate the unknown density through $\hat{f}_n$; for a summary on the application of this method, see Ch. 3 (Sect. 3.1) Wand and Jones [16]. The reader will be able to verify that such an option applies when (1) is considered. Another proposal is framed within automatic selectors, methods that, from an observed random sample, generate smoothing parameters with the purpose of achieving the same goal. For more details, see Ch. 4 (Sect. 4.1) Gramacki [7], where the work of Heidenreich et al. [8] is highlighted, which presents —solely for the univariate kernel estimation case— a comprehensive review and comparison of fully automatic smoothing parameter selection methods, while also outlining the main features of the three predominant approaches. Regarding the choice of the smoothing parameter to estimate f with (1), the author highlights that there are currently no results associated with the subjective criterion, as well as the absence of automatic methods that make it viable. However, an option dependent on other estimation methods was considered by Fajardo [5] and Fajardo and Harmath [3] to analyze the analytical performance of the estimator (1) and the fuzzy set boundary estimator of f, respectively. This option was based on taking as the smoothing parameter the value resulting from combining

the optimal global smoothing parameters b_{asy} and h_{asy}, where h_{asy} was replaced in each analysis by the smoothing parameter computed by other authors subjectively or automatically. In this way, the idea introduced in Fajardo [5] and implemented in Fajardo and Harmath [3] provides a dependent criterion that allows solving the deficiency present in the fuzzy set density estimation method.

This work considers one of the ideas introduced by Rudemo [14] in order to equip the method developed by Fajardo et al. [4] with an automatic and autonomous criterion that enables the choice of the smoothing parameter whether f is known or not. This approach consists in implementing the risk function estimator corresponding to the $IMSE$ of (1), which allows highlighting important properties of the proposed estimator, as well as evaluating its performance using simulated and real data. It is important to point out that the achieved results can be naturally extended to the fuzzy set boundary estimation method of f by Fajardo and Harmath [3]; likewise, these findings suggest extending the presented approach by considering the second idea introduced in Rudemo [14], through the examination of the risk function estimator corresponding to the Kullback-Leibler information measure of (1).

The article is organized as follows. Section 2 describes an automatic and autonomous criterion for choosing the smoothing parameter of (1) whether f is known or not. This criterion is developed within the theoretical context introduced by Rudemo [14] for the case of the risk function estimator corresponding to the $IMSE$. Section 3 presents the calculations of all the integrals involved in each term of $\hat{Q}(\hat{\vartheta}_n)$, the risk function estimator corresponding to the $IMSE$ of (1). On the other hand, Sections 4 and 5 present the analyses of simulated and real data, respectively. In both cases, all numerical implementations and simulations were developed and executed entirely in MATLAB. Finally, the conclusions of the article are presented in Section 6.

2. Automatic Selection of the Smoothing Parameter Under Quadratic Risk Estimation

This section presents an automatic and autonomous criterion for selecting the smoothing parameter of (1) whether f is known or not. It is framed within the context of one of the general results presented by Rudemo [14] and was derived through the implementation of the risk function estimator corresponding to the $IMSE$ of (1).

Let $X_1, \ldots, X_n$ be a random sample of size n with common density f, and $V_1, \ldots, V_n$ independent random variables uniformly distributed on $(0, 1)$ and independent of $X_1, \ldots, X_n$. In what follows, $\mathbf{U} = ((X_1, V_1), \ldots, (X_n, V_n))$ will represent an observed random sample associated with the density f, or simply an observed random sample.

For a thinning function φ such that $0 < \int \varphi(s)\, ds < \infty$ and $\mathbb{I}_\varphi$ given in (2), the scalar function $g : \mathbb{R}^3 \to \mathbb{R}$ is defined as:

$$g(z, y, v) = \frac{1}{b} \frac{\mathbb{I}_\varphi\left(\frac{z-y}{b}, v\right)}{\int \varphi(s)\, ds} \tag{9}$$

for each $z, y \in \mathbb{R}$, $v \in (0, 1)$, and $b \in \mathbb{R}_+$. Now, for an observed random sample $\mathbf{U}$, (1) can be rewritten as follows:

$$\hat{\vartheta}_n(x, \mathbf{U}) = \frac{1}{n} \sum_{i=1}^{n} g(x, X_i, V_i) \tag{10}$$

for each $x \in \mathbb{R}$ and g given in (9).

Let $p \geq 1$ and ϱ be a non-negative weight function. Then the risk function of the estimator (10) is given by:

$$R^p(\hat{\vartheta}_n, f) = \mathbb{E}[\int |\hat{\vartheta}_n(x, \mathbf{U}) - f(x)|^p \varrho(x)\, dx].$$

For the particular case $p = 2$ and $\varrho(x) \equiv 1$, we have $R^2(\hat{\vartheta}_n, f) = IMSE[\hat{\vartheta}_n]$. That is, $IMSE[\hat{\vartheta}_n]$ is a risk function, commonly called the quadratic risk function of (10), and represents the most common global measure in studies of density estimators for f. For the aforementioned case, Fubini's Theorem guarantees the following expression:

$$R^2(\hat{\vartheta}_n, f) = \int \mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})]^2 dx - 2\int \mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})] f(x)\, dx + \int (f(x))^2\, dx.$$

Clearly, $R^2(\hat{\vartheta}_n, f)$ is a function of the smoothing parameter, $b \in \mathbb{R}_+$, provided that f, $\mathbf{U}$, and φ remain fixed. Furthermore, it is worth highlighting that the only terms in the previous equality that depend on the parameter b are the first and second ones. This observation, together with the equality $R^2(\hat{\vartheta}_n, f) = IMSE[\hat{\vartheta}_n]$ and the fact that a minimum $IMSE$ yields the optimal estimation of f, implies that an appropriate smoothing parameter to estimate f through (10) is given by the one that minimizes $R^2(\hat{\vartheta}_n, f)$ or, equivalently, by the parameter that minimizes $\int \mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})]^2 dx - 2\int \mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})] f(x)\, dx$. Using the idea introduced in Rudemo [14] and under the observed sample $\mathbf{U}$, the following risk function of the estimator (10) is proposed:

$$\begin{aligned} Q(\hat{\vartheta}_n, f) &= \int \mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})]^2 dx - 2\int \mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})] f(x)\, dx \\ &= R^2(\hat{\vartheta}_n, f) - \int (f(x))^2\, dx \end{aligned} \tag{11}$$

and, as an estimator of (11), we define:

$$\begin{aligned} \hat{Q}(\hat{\vartheta}_n) &= \int (\hat{\vartheta}_n(x, \mathbf{U}))^2 dx \\ &\quad - \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} g(X_i, X_j, V_j), \end{aligned} \tag{12}$$

with g given in (9). The previous analysis suggests estimating f through (10) by selecting the smoothing parameter that minimizes (12); such suggestion will be formalized later. It should be noted that (12) is a least squares cross-validation estimator, where the following expression:

$$\frac{1}{(n-1)}\sum_{\substack{j=1\\ j\neq i}}^{n} g(X_i, X_j, V_j)$$

is the fuzzy set density estimation based on the observed random sample with (X_i, V_i) removed. Although a formal notation exists for the above expression when considering the estimator $\hat{f}_n$, it will not be adapted to our case to avoid overloading the notation.

The following result establishes that (12) is an unbiased estimator of (11).

Lemma 2.1. *Let* $\mathbf{U}$ *be an observed random sample. Then:*

$$\mathbb{E}[\hat{Q}(\hat{\vartheta}_n)] = Q(\hat{\vartheta}_n, f)$$

for $\hat{Q}(\hat{\vartheta}_n)$ *and* $Q(\hat{\vartheta}_n, f)$ *given in (12) and (11), respectively.*

Proof. The independence of the random variables X_i, X_j, and V_j, with $1 \leq i \neq j \leq n$, implies that:

$$\begin{aligned}\mathbb{E}[g(X_i, X_j, V_j)] &= \int\Big(\int\int g(x, y, v)\, f(y)\, dy\, dv\Big) f(x)\, dx\\ &= \int \mathbb{E}\left[g(x, X, V)\right]\, f(x)\, dx.\end{aligned}$$

Here, X and V are independent random variables with density f and uniform distribution on $(0, 1)$, respectively. From the above equality and the identity $\mathbb{E}[g(x, X, V)] = \mathbb{E}[g(x, X_i, V_i)]$, with $1 \leq i \leq n$, it follows that:

$$\begin{aligned}&\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\mathbb{E}\left[g(X_i, X_j, V_j)\right] =\\ &n(n-1)\int \mathbb{E}\left[g(x, X, V)\right] f(x)\, dx =\\ &n(n-1)\int \mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n} g(x, X_i, V_i)\Big] f(x)\, dx =\\ &n(n-1)\int \mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})] f(x)\, dx.\end{aligned}$$

The proof concludes by combining the above equality with the mathematical expectation of (12). □

Through the following result, an explicit expression for (11) is exhibited in order to establish that it is well-defined and to guarantee that $R^2(\hat{\vartheta}_n, f)$ has a positive limit as $n \to \infty$. Furthermore, to reduce the nomenclature and facilitate the reading of the presented expressions, the real function $g_1 : \mathbb{R} \to \mathbb{R}$ defined by:

$$g_1(x) = \frac{1}{b\int \varphi(s)\, ds}\mathbb{E}\Big[\varphi\Big(\frac{x - X}{b}\Big)\Big] \tag{13}$$

is introduced for each $x \in \mathbb{R}$, $b \in \mathbb{R}_+$, $0 < \int \varphi(s)\, ds < \infty$, where X is a random variable with density f.

Lemma 2.2. *Let* $\mathbf{U}$ *be an observed random sample. Then:*

$$\begin{aligned}Q(\hat{\vartheta}_n, f) = \frac{1}{n\, b\, \int\varphi(s)\, ds}\int g_1(x)\, dx + \frac{(n-1)}{n}\int g_1^2(x)\, dx\\ -2\int g_1(x) f(x)\, dx\end{aligned} \tag{14}$$

for each $b \in \mathbb{R}_+$, $0 < \int\varphi(s)\, ds < \infty$, *and* g_1 *given in (13) with* $\mathbb{E}[\varphi\left(\frac{x-X}{b}\right)] \in (0, 1)$.

Proof. Let $x \in \mathbb{R}$ and g be given in (9). The following expansion:

$$\Big(\sum_{i=1}^{n} g(x, X_i, V_i)\Big)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} g(x, X_i, V_i) g(x, X_j, V_j)$$

implies that:

$$\begin{aligned}\mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})]^2 = \frac{1}{n^2}\Big\{&\sum_{i=1}^{n}\mathbb{E}[g^2(x, X_i, V_i)]\\ &+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\mathbb{E}[g(x, X_i, V_i)\, g(x, X_j, V_j)]\Big\}\end{aligned}$$

for each $x \in \mathbb{R}$. By (2),

$$\mathbb{I}_{\varphi}^{2}(y, v) = \mathbb{I}_{\varphi}(y, v)$$

for each $y \in \mathbb{R}$, $v \in (0, 1)$. The above equality and the independence of $(X_1, V_1), \ldots, (X_n, V_n)$ imply that:

$$\begin{aligned}\mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})]^2 = \frac{1}{n^2}\Big\{&\frac{1}{b\int\varphi(s)\, ds}\sum_{i=1}^{n}\mathbb{E}\left[g(x, X_i, V_i)\right]\\ &+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\mathbb{E}\left[g(x, X_i, V_i)\right]\mathbb{E}\left[g(x, X_j, V_j)\right]\Big\}\end{aligned}$$

for each $x \in \mathbb{R}$, $b \in \mathbb{R}_+$ and $0 < \int\varphi(s)\, ds < \infty$. Using the fact that X and V are independent random variables with density f and uniform distribution on $(0, 1)$, we have:

$$\begin{aligned}\mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})]^2 = \frac{1}{n^2}\Big\{&\frac{n}{b\int\varphi(s)\, ds}\mathbb{E}[g(x, X, V)]\\ &+n(n-1)(\mathbb{E}[g(x, X, V)])^2\Big\}\end{aligned} \tag{15}$$

for each $x \in \mathbb{R}$, $b \in \mathbb{R}_+$ and $0 < \int\varphi(s)\, ds < \infty$. Furthermore, (2) in Fajardo et al. [4] guaranties that:

$$\begin{aligned}\mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})] &= \frac{1}{n}\mathbb{E}\Big[\sum_{i=1}^{n} g(x, X_i, V_i)\Big]\\ &= \frac{1}{b\int\varphi(s)\, ds}\mathbb{E}\Big[\varphi\Big(\frac{x - X}{b}\Big)\Big]\end{aligned}$$

for each $x \in \mathbb{R}$, $b \in \mathbb{R}_+$, $0 < \int\varphi(s)\, ds < \infty$ and $\mathbb{E}[\varphi(\frac{x-X}{b})] \in (0, 1)$. By (13) and using the fact that X and V are independent

random variables with density f and uniform distribution on $(0, 1)$, we have:

$$\mathbb{E}\left[\hat{\vartheta}_n(x, \mathbf{U})\right] = \mathbb{E}\left[g(x, X, V)\right] = g_1(x)$$

for each $x \in \mathbb{R}$. The proof concludes by combining (11) and (15) along with the above equality.

□

The fact that φ and f are integrable guarantees that (14) is well-defined. Furthermore, since the integrals in (14) do not depend on n, it follows that:

$$\lim_{n\to\infty} Q(\hat{\vartheta}_n, f) = \int g_1^2(x)\, dx - 2\int g_1(x)\, f(x)\, dx.$$

From the above equality, it is deduced that:

$$\begin{aligned}\lim_{n\to\infty} R^2(\hat{\vartheta}_n, f) &= \lim_{n\to\infty}\left[Q(\hat{\vartheta}_n, f) + \int f^2(x)\, dx\right] \\ &= \int \left(g_1(x) - f(x)\right)^2 dx\end{aligned}$$

with g_1 given in (13). Furthermore, for the observed random sample $\mathbf{U}$, the bias of (10) is given by $(\mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})] - f(x))$ for each $x \in \mathbb{R}$. However, in the proof of Lemma 2.2 we saw that $\mathbb{E}[\hat{\vartheta}_n(x, \mathbf{U})] = g_1(x)$ for each $x \in \mathbb{R}$. From this, $([\hat{\vartheta}_n(x, \mathbf{U})] - f(x)) = (g_1(x) - f(x))$ for each $x \in \mathbb{R}$. By $(A.4)$ in Fajardo [2], we have that $\left(g_1(x) - f(x)\right)^2 > 0$ for each $x \in \mathbb{R}$. Therefore, the risk function $R^2(\hat{\vartheta}_n, f)$ has the positive limit $\int (g_1(x) - f(x))^2 dx$ as $n \to \infty$.

The following result establishes an explicit expression for $\int(\hat{\vartheta}_n(x, \mathbf{U}))^2 dx$, which will enable the practical implementation of (12). Once again, to reduce nomenclature and facilitate the reading of the presented expressions, the scalar function $\mathbb{1}_\varphi * \mathbb{1}_\varphi : \mathbb{R} \times (0, 1) \times (0, 1) \to \mathbb{R}$ defined by:

$$\mathbb{1}_\varphi * \mathbb{1}_\varphi(x, u, v) = \int \mathbb{1}_\varphi(s, u) \cdot \mathbb{1}_\varphi(s - x, v)\, ds \qquad (16)$$

is introduced for each $x \in \mathbb{R}$, $u, v \in (0, 1)$ and $\mathbb{1}_\varphi$ given in (2). It is worth highlighting that the explicit expressions associated with (16) will be calculated in the next section.

Theorem 2.3. *Let* $\mathbf{U}$ *be an observed random sample. Then:*

$$\begin{aligned}\int (\hat{\vartheta}_n(x, \mathbf{U}))^2 dx = &\frac{1}{b(n\int\varphi(u)\, du)^2} \\ &*\sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_\varphi * \mathbb{1}_\varphi\left(\frac{X_j - X_i}{b}, V_i, V_j\right)\end{aligned} \qquad (17)$$

for each $b \in \mathbb{R}_+$, $0 < \int\varphi(s)\, ds < \infty$, *and* $\mathbb{1}_\varphi * \mathbb{1}_\varphi$ *given in (16).*

Proof. Notice that:

$$\begin{aligned}\left[\sum_{i=1}^n \mathbb{1}_\varphi\left(\frac{x - X_i}{b}, V_i\right)\right]^2 = &\sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_\varphi\left(\frac{x - X_i}{b}, V_i\right) \\ &*\mathbb{1}_\varphi\left(\frac{x - X_j}{b}, V_j\right)\end{aligned}$$

for each $x \in \mathbb{R}$, $b \in \mathbb{R}_+$ and $\mathbb{1}_\varphi$ given in (2). Combining the above equality with (10), it is obtained that:

$$\begin{aligned}\int (\hat{\vartheta}_n(x, \mathbf{U}))^2 dx = &\frac{1}{\left(n\, b \int\varphi(u)\, du\right)^2} \\ &*\sum_{i=1}^n \sum_{j=1}^n \int \mathbb{1}_\varphi\left(\frac{x - X_i}{b}, V_i\right) \mathbb{1}_\varphi\left(\frac{x - X_j}{b}, V_j\right) dx\end{aligned}$$

for each $b \in \mathbb{R}_+$, $0 < \int\varphi(s)\, ds < \infty$ and $\mathbb{1}_\varphi$ given in (2). The change of variable $s = \frac{x - X_i}{b}$ implies that:

$$\begin{aligned}\int (\hat{\vartheta}_n(x, \mathbf{U}))^2 dx = &\frac{1}{b\, \left(n\, \int\varphi(u)\, du\right)^2} \\ &*\sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_\varphi * \mathbb{1}_\varphi\left(\frac{X_j - X_i}{b}, V_i, V_j\right)\end{aligned}$$

for each $b \in \mathbb{R}_+$, $0 < \int\varphi(s)\, ds < \infty$ and $\mathbb{1}_\varphi * \mathbb{1}_\varphi$ given in (16), which concludes the proof.

□

Corollary 2.4. *Let* $\mathbf{U}$ *be an observed random sample and* φ *a thinning function symmetric about zero. Then:*

$$\begin{aligned}\hat{Q}(\hat{\vartheta}_n) = &\\ &\frac{1}{b(n\int\varphi(s)\, ds)^2} \sum_{i=1}^n \sum_{j=1}^n \mathbb{1}_\varphi * \mathbb{1}_\varphi\left(\frac{X_j - X_i}{b}, V_i, V_j\right) \\ &-\frac{1}{b\int\varphi(s)\, ds} \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ j\neq i}}^n \mathbb{1}_\varphi\left(\frac{X_j - X_i}{b}, V_j\right),\end{aligned} \qquad (18)$$

for each $b \in \mathbb{R}_+$, $0 < \int\varphi(s)\, ds < \infty$, $\mathbb{1}_\varphi * \mathbb{1}_\varphi$ *given in (16), and* $\mathbb{1}_\varphi$ *given in (2).*

Proof. Equation (12) is combined with (17), and the symmetry of φ about zero is used.

□

The following result will justify one of the necessary conditions to establish the finite variance of (18). For this purpose, and in what follows, the thinning function φ given in (5) will be considered in order to preserve the property that originates (6), namely, that the fuzzy set density estimator yields better estimates than classical kernel density estimators.

Lemma 2.5. *Let* $\mathbb{1}_\varphi$ *and* φ *be given in (2) and (5), respectively. Then:*

$$\int \mathbb{1}_\varphi\left(\frac{x - y}{b}, v\right) dx = \frac{25}{16}\, b\, \sqrt{1 - v}$$

for each $y \in \mathbb{R}$, $b \in \mathbb{R}_+$, *and* $v \in (0, 1)$.

Proof. Notice that the value of the considered integral is zero when $\frac{x-y}{b} \notin [-\frac{25}{16}, \frac{25}{16}]$. Suppose that $\frac{x-y}{b} \in [-\frac{25}{16}, \frac{25}{16}]$. Now, consider that $\frac{x-y}{b} \in [-\frac{25}{16}, 0]$. Then the left branch of φ given in (5), denoted by φ_{-} with $\varphi_{-} : [-\frac{25}{16}, 0] \longrightarrow [0, 1]$, is increasing. Consequently, there exists the increasing inverse function φ_{-}^{-1} given by:

$$\varphi_{-}^{-1}(s) = -\frac{25}{16}\sqrt{1-s}\,,$$

for each $s \in [0, 1]$. From (2), it follows that:

$$\mathbb{I}_{\varphi}\Big(\frac{x-y}{b}, v\Big) = \begin{cases} 1 & \text{if} \quad v \leq \varphi\big(\frac{x-y}{b}\big) \\ 0 & \text{otherwise,} \end{cases} \tag{19}$$

for each $x, y \in \mathbb{R}$, $b \in \mathbb{R}_{+}$, and $v \in (0, 1)$. Using the inequality $\frac{x-y}{b} \leq 0$, we have:

$$\mathbb{I}_{\varphi}\Big(\frac{x-y}{b}, v\Big) = \begin{cases} 1 & \text{if} \quad \varphi_{-}^{-1}(v)\, b + y \leq x \leq y \\ 0 & \text{otherwise,} \end{cases}$$

for each $x, y \in \mathbb{R}$, $b \in \mathbb{R}_{+}$, and $v \in (0, 1)$. From which:

$$\int \mathbb{I}_{\varphi}\Big(\frac{x-y}{b}, v\Big)\, dx = \frac{25}{16}\, b\, \sqrt{1-v}$$

for each $y \in \mathbb{R}$, $b \in \mathbb{R}_{+}$, and $v \in (0, 1)$. Now, consider that $\frac{x-y}{b} \in [0, \frac{25}{16}]$. Then, the right branch of φ in (5), denoted by φ_{+} with $\varphi_{+} : [0, \frac{25}{16}] \longrightarrow [0, 1]$, is decreasing. Therefore, there exists the decreasing inverse function φ_{+}^{-1} given by:

$$\varphi_{+}^{-1}(s) = \frac{25}{16}\sqrt{1-s}\,,$$

for each $s \in [0, 1]$. Combining (19) with the inequality $\frac{x-y}{b} \geq 0$, we obtain:

$$\mathbb{I}_{\varphi}\Big(\frac{x-y}{b}, v\Big) = \begin{cases} 1 & \text{if} \quad y \leq x \leq \varphi_{+}^{-1}(v)\, b + y \\ 0 & \text{otherwise,} \end{cases}$$

for each $x, y \in \mathbb{R}$, $b \in \mathbb{R}_{+}$, and $v \in (0, 1)$. The proof concludes by integrating the above expression with respect to x.

□

On the other hand, in the general results introduced by Rudemo [14], the following two conditions are imposed:

$$\begin{cases} \mathbb{E}[\varrho(X_1)\, \mathsf{g}(X_1, X_2)]^2 < \infty \\ \mathbb{E}[\int \varrho(x)\, \mathsf{g}^2(x, X_1)\, dx]^2 < \infty \end{cases} \tag{20}$$

to guarantee that $\mathrm{Var}[\hat{Q}(\hat{f})] < \infty$, where $\hat{f}(x, \mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} \mathsf{g}(x, X_i)$ with $\mathbf{X} = (X_1, \ldots, X_n)$, $\mathsf{g}(x, y) = (1/h)K((x-y)/h)$ for each $x, y \in \mathbb{R}$, $h \in \mathbb{R}_{+}$, and K is a kernel. Furthermore, it is highlighted that through Theorem 5.2 in Hoeffding [9], it can be proved that $\mathrm{Var}[\hat{Q}(\hat{f})] = O(n^{-1})$ as $n \longrightarrow \infty$ with g fixed.

Now, for the particular case $\varrho(x) \equiv 1$ and $\hat{\vartheta}_n$ given in (10), it will suffice to justify the conditions given in (20) to establish the asymptotic representation $\mathrm{Var}[\hat{Q}(\hat{\vartheta})] = O(n^{-1})$ as $n \longrightarrow \infty$ with g given in (9) fixed. For this purpose, consider $\mathbb{I}_{\varphi}$ and φ given in (2) and (5), respectively. Clearly, the non-negative function:

$$\frac{\mathbb{I}_{\varphi}}{\int \varphi(s)\, ds} : \mathbb{R} \times (0, 1) \longrightarrow \Big\{0, \frac{1}{\int \varphi(s)\, ds}\Big\}$$

is not a kernel; however, it satisfies the following equality:

$$\iint \frac{\mathbb{I}_{\varphi}(x, v)}{\int \varphi(s)\, ds}\, dx\, dv = 1.$$

It is worth noting that the value of $\int \varphi(s)\, ds$ exists and is non-zero if φ is the thinning function given in (5). Furthermore, observe that the bivariate function $\frac{\mathbb{I}_{\varphi}}{\int \varphi(s)\, ds}$ is fixed; indeed, φ is fixed and is given in (5). Consequently, g given in (10) is fixed. Since $\mathbb{I}_{\varphi}^{2}(y, v) = \mathbb{I}_{\varphi}(y, v)$ for each $y \in \mathbb{R}$ and $v \in (0, 1)$, we have:

$$\begin{aligned} \mathbb{E}\left[g(X_1, X_2, V_2)\right]^2 &= \frac{1}{b\, \int \varphi(s)\, ds}\, \mathbb{E}\left[g(X_1, X_2, V_2)\right] \\ &\leq \Big(\frac{1}{b\, \int \varphi(s)\, ds}\Big)^2 \\ &< \infty, \end{aligned}$$

for each $b \in \mathbb{R}_{+}$. Again, combining the equality $\mathbb{I}_{\varphi}^{2}(y, v) = \mathbb{I}_{\varphi}(y, v)$ for $y \in \mathbb{R}$ and $v \in (0, 1)$ with Lemma 2.5, and recalling that V_1 is a uniform random variable on $(0, 1)$, we obtain:

$$\begin{aligned} \mathbb{E}[\int g^2(x, X_1, V_1)\, dx]^2 &= \Big(\frac{1}{b\, \int \varphi(s) ds}\Big)^2 \mathbb{E}[\int g(x, X_1, V_1) dx]^2 \\ &= \Big(\frac{25}{16\, \int \varphi(s)\, ds}\Big)^2 \mathbb{E}[\sqrt{1-V_1}\,]^2 \\ &= \Big(\frac{25}{16\sqrt{2}\, \int \varphi(s)\, ds}\Big)^2 \\ &< \infty, \end{aligned}$$

verifying (20). Thus:

$$\mathrm{Var}[\hat{Q}(\hat{\vartheta})] = O(n^{-1}) \tag{21}$$

as $n \longrightarrow \infty$ with g fixed.

It was highlighted that $\frac{\mathbb{I}_{\varphi}}{\int \varphi(s)\, ds}$ is fixed. Therefore, the expression given in (18) is a function of the smoothing parameter b. To estimate f, the notation and criterion introduced in Rudemo [14] will be adopted. To this end, we define the class $(\hat{\vartheta}_{\beta})_{\beta \in \mathcal{B}}$ such that $\hat{\vartheta}_{\beta}(x, \mathbf{U}) = \hat{\vartheta}_{n}(x, \mathbf{U})$ with $\beta = (\varphi, b)$ and φ given in (5).

Notice that the estimators defining the class $(\hat{\vartheta}_{\beta})_{\beta \in \mathcal{B}}$ depend solely on the parameter b. Now, we will select the estimator

$\hat{\vartheta}_\beta(x,\mathbf{U})$ whose smoothing parameter minimizes $\hat{Q}(\hat{\vartheta}_\beta)$, which we will denote by $\tilde{b}_n$. It is worth pointing out that the task focuses on calculating $\tilde{b}_n$ and using the estimator $\hat{\vartheta}_\beta(x,\mathbf{U})$ with $\beta=(\varphi,\tilde{b}_n)$ to estimate f with φ given in (5). This is equivalent to computing $\tilde{b}_n$ via (18) to estimate f using (10) with the observed random sample $\mathbf{U}$ and the thinning function φ given in (5). It is opportune to highlight that the resulting smoothing parameter $\tilde{b}_n$ turns out to be automatic and autonomous.

Recall that Rudemo [14], through (21) –within the context of kernel density estimation of f–, established that for a finite and fixed number of candidate estimators, the probability of choosing one that minimizes the risk function $R^2(\hat{f}_\alpha,f)$ approaches one as $n\to\infty$. Here, $\alpha=(K,h)\in\mathcal{A}$ and $\hat{f}_\alpha(x,\mathbf{X})=\hat{f}_n(x,\mathbf{X})$ for $x\in\mathbb{R}$ and $\mathbf{X}=(X_1,\ldots,X_n)$, where $\hat{f}_n$ is the classical kernel estimator of f with kernel K. However, it was stressed that the above conclusion does not constitute a fully satisfactory convergence theorem, since the result should allow for an infinite number of candidate estimators or for the number of elements in $\mathcal{A}$ to increase with the sample size, n, at an adequate rate. Clearly, what was established and highlighted by Rudemo [14] also applies to the theory developed in the present article.

Finally, it should be noted that, from now on, the estimation of f for both simulated and real data cases using the estimator $\hat{\vartheta}_n$ relies on the automatic and autonomous smoothing parameter $\tilde{b}_n$ to enable such estimation. Note that, for the first case, the smoothing parameters b_{asy} and $\tilde{b}_n$ strengthen the criterion that allows displaying the results achieved when implementing the estimator $\hat{\vartheta}_n$. For the second case, and with the purpose of consolidating the criterion for presenting the obtained results, it is suggested to consider subjective smoothing parameters, b_s, obtained from $\tilde{b}_n$ by slight increases or decreases. Of course, each modification of $\tilde{b}_n$ must be monitored by evaluating the smoothness of the estimated density. Here, b_{asy} is given in (3).

3. Explicit Expressions for $\mathbb{I}_\varphi * \mathbb{I}_\varphi$

In order to enable the practical implementation of the estimator (18), this section provides explicit expressions for (16). This will be based on the calculation of the set $\mathrm{I}_{u,0}^{v,x}$, defined below, such that for each $x\in\mathbb{R}$, $u,v\in(0,1)$, and $s\in\mathrm{I}_{u,0}^{v,x}$, the following two inequalities are satisfied: $v\le\varphi_x(s)$ and $u\le\varphi(s)$, where $\varphi_x(s)=\varphi(s-x)$.

The above begins with the following results.

Lemma 3.1. *Let $v\in(0,1)$, $x\in\mathbb{R}$, and φ given in (5). Then there exist $s'_{v,x},s_{v,x}\in(x-\frac{25}{16},x+\frac{25}{16})$ such that $s'_{v,x}<s_{v,x}$, and $v\le\varphi_x(s)$ for each $s\in[s'_{v,x},s_{v,x}]$, where $\varphi_x(s)=\varphi(s-x)$,*

$$s'_{v,x}=x-\frac{25}{16}\sqrt{1-v}\quad\text{and}\quad s_{v,x}=x+\frac{25}{16}\sqrt{1-v}\,. \tag{22}$$

Proof. For each $x\in\mathbb{R}$, φ_x represents a rigid transformation of the graph of φ, since φ_x is the graph of φ horizontally shifted to the right by x units if $x>0$, or to the left by $|x|$ units if $x<0$. Clearly, the symmetry of φ with respect to zero implies that φ_x is symmetric with respect to x. Furthermore, for each $x\in\mathbb{R}$ and $v\in(0,1)$, the expressions $s_1=x-\frac{25}{16}\sqrt{1-v}$ and $s_2=x+\frac{25}{16}\sqrt{1-v}$ are the solutions to the equation $v=\varphi_x(s)$ for each $s\in\mathbb{R}$ and φ given in (5). Notice that $s_1<s_2$. Now, since $\sqrt{1-v}<1$, we have $\frac{25}{16}\sqrt{1-v}<\frac{25}{16}$ for each $v\in(0,1)$. Therefore, $x-\frac{25}{16}<x-\frac{25}{16}\sqrt{1-v}<x+\frac{25}{16}\sqrt{1-v}<x+\frac{25}{16}$ for each $x\in\mathbb{R}$. That is, $s_1,s_2\in(x-\frac{25}{16},x+\frac{25}{16})$ for each $x\in\mathbb{R}$. Thus, for each $x\in\mathbb{R}$ and $v\in(0,1)$, there exist $s'_{v,x}=x-\frac{25}{16}\sqrt{1-v}$ and $s_{v,x}=x+\frac{25}{16}\sqrt{1-v}$ such that $s'_{v,x},s_{v,x}\in(x-\frac{25}{16},x+\frac{25}{16})$ and $s'_{v,x}<s_{v,x}$. However, for each $x\in\mathbb{R}$ and $s\in[s'_{v,x},s_{v,x}]$, the symmetry of $\varphi_x(s)$ with respect to x guarantees that $s\in[s'_{v,x},x]$ or $s\in[x,s_{v,x}]$. Then, the increasing property of $\varphi_x(s)$ on $[s'_{v,x},x]$ and the decreasing property of $\varphi_x(s)$ on $[x,s_{v,x}]$ imply that $v=\varphi_x(s'_{v,x})\le\varphi_x(s)$ with $s\in[s'_{v,x},x]$ and $v=\varphi_x(s_{v,x})\le\varphi_x(s)$ with $s\in[x,s_{v,x}]$, for each $x\in\mathbb{R}$ and $v\in(0,1)$, completing the proof. □

Corollary 3.2. *Let $u\in(0,1)$ and φ given in (5). Then there exist $s'_{u,0},s_{u,0}\in(-\frac{25}{16},\frac{25}{16})$ such that $s'_{u,0}<s_{u,0}$ and $u\le\varphi_0(s)$ for each $s\in[s'_{u,0},s_{u,0}]$, where $\varphi_0(s)=\varphi(s)$,*

$$s'_{u,0}=-\frac{25}{16}\sqrt{1-u}\quad\text{and}\quad s_{u,0}=\frac{25}{16}\sqrt{1-u}\,. \tag{23}$$

Proof. It suffices to set $x=0$ in Lemma 3.1, substitute v by u, and use the equality $\varphi_0(s)=\varphi(s)$ for each $s\in[s'_{u,0},s_{u,0}]$, which concludes the proof. □

Lemma 3.3. *Let $\mathbb{I}_\varphi*\mathbb{I}_\varphi$ be given in (16) with $\mathbb{I}_\varphi$ and φ given in (2) and (5), respectively. Let $x\in\mathbb{R}$ and $u,v\in(0,1)$. Then:*

$$\mathbb{I}_\varphi*\mathbb{I}_\varphi(x,u,v)=\begin{cases}\int_{\mathrm{I}_{u,0}^{v,x}}ds & \text{if}\quad \mathrm{I}_{u,0}^{v,x}\neq\emptyset,\\ 0 & \text{if}\quad \mathrm{I}_{u,0}^{v,x}=\emptyset,\end{cases}$$

where $\mathrm{I}_{u,0}^{v,x}=[s'_{v,x},s_{v,x}]\cap[s'_{u,0},s_{u,0}]$, with $s'_{v,x},s_{v,x}$ and $s'_{u,0},s_{u,0}$ given in (22) and (23), respectively.

Proof. Combining Lemma 3.1, Corollary 3.2, and (2) implies that:

$$\mathbb{I}_\varphi(s-x,v)=\begin{cases}1 & \text{if}\quad s\in[s'_{v,x},s_{v,x}],\\ 0 & \text{otherwise},\end{cases}$$

and

$$\mathbb{I}_\varphi(s,u)=\begin{cases}1 & \text{if}\quad s\in[s'_{u,0},s_{u,0}],\\ 0 & \text{otherwise},\end{cases}$$

for each $x\in\mathbb{R}$ and $u,v\in(0,1)$, where $s'_{v,x},s_{v,x}$ and $s'_{u,0},s_{u,0}$ are given in (22) and (23), respectively. Let $\mathrm{I}_{u,0}^{v,x}=[s'_{v,x},s_{v,x}]\cap[s'_{u,0},s_{u,0}]$ for each $x\in\mathbb{R}$ and $u,v\in(0,1)$. Clearly:

$$\mathbb{I}_\varphi(s,u)\cdot\mathbb{I}_\varphi(s-x,v)=0$$

for each $s \in (I_{u,0}^{v,x})^{\complement}$ with $x \in \mathbb{R}$ and $u, v \in (0,1)$. Combining the above equality with (16), it follows that:

$$\mathbb{I}_{\varphi} * \mathbb{I}_{\varphi}(x,u,v) = \int_{I_{u,0}^{v,x}} \mathbb{I}_{\varphi}(s,u) \cdot \mathbb{I}_{\varphi}(s-x,v)\, ds \qquad (24)$$

for each $x \in \mathbb{R}$ and $u, v \in (0,1)$. If $I_{u,0}^{v,x} = \emptyset$, by (24) it directly follows that:

$$\mathbb{I}_{\varphi} * \mathbb{I}_{\varphi}(x,u,v) = 0.$$

On the other hand, if $I_{u,0}^{v,x} \neq \emptyset$, then:

$$\mathbb{I}_{\varphi}(s,u) \cdot \mathbb{I}_{\varphi}(s-x,v) = 1$$

for each $s \in I_{u,0}^{v,x}$ with $x \in \mathbb{R}$ and $u, v \in (0,1)$. This relationship together with (24) implies that:

$$\mathbb{I}_{\varphi} * \mathbb{I}_{\varphi}(x,u,v) = \int_{I_{u,0}^{v,x}} ds,$$

which completes the proof. □

Lemma 3.3 guarantees that, through the calculation of $I_{u,0}^{v,x}$, the explicit formula for (16) can be obtained. Therefore, the next step will focus on deriving an expression for $I_{u,0}^{v,x}$ by considering the cases presented below. These are governed by the simultaneous comparison of the graphical representation of the functions φ_x and φ through the progressive shift, from left to right, of the function φ_x, recalling that both functions are symmetric with respect to x and zero, respectively, and that their support intervals have the same length.

Case 1. Let $x \in (-\infty, -\frac{25}{8}]$; the functions φ_x and φ intersect at the point $(-\frac{25}{16}, 0)$ when $x = -\frac{25}{8}$. Indeed, the symmetry of φ with respect to zero and the fact that $s = -\frac{25}{16}$ is a root of φ guarantee the following equalities: $\varphi_{-\frac{25}{8}}(-\frac{25}{16}) = \varphi(-\frac{25}{16} + \frac{25}{8}) = \varphi(\frac{25}{16} - \frac{25}{8}) = \varphi(-\frac{25}{16}) = 0$ and $\varphi_{-\frac{25}{8}}(-\frac{75}{16}) = \varphi(-\frac{75}{16} + \frac{25}{8}) = \varphi(-\frac{25}{16}) = 0$. Clearly, $\varphi_{-\frac{25}{8}}$ intersects the X-axis at $s = -\frac{75}{16}$ and $s = -\frac{25}{16}$. Since $-\frac{75}{16} \notin [-\frac{25}{16}, \frac{25}{16}]$, the above statement holds. Consequently, a general graphical representation of both functions, in this case, is as follows:

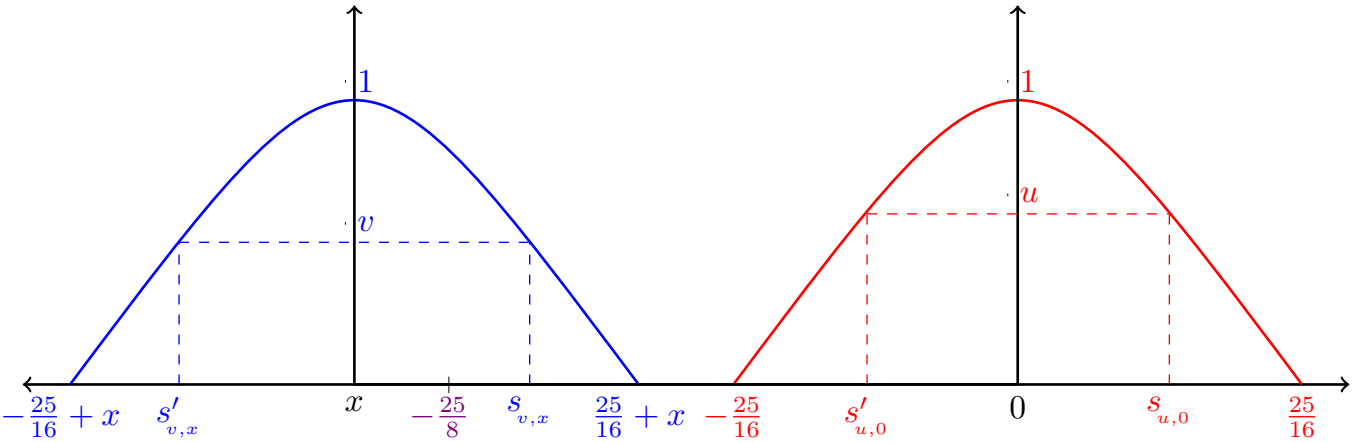


Figure 1. *General graphical representation of the functions φ_x and φ for each $x \in (-\infty, -\frac{25}{8}]$, where $(-\frac{25}{16}, 0)$ is their point of intersection when $x = -\frac{25}{8}$.*

Observe that $-\frac{25}{16} < s'_{u,0}$ for each $u > 0$. Now, with the support of Figure 1, we can guarantee that $I_{u,0}^{v,x} \in \{\emptyset, \{-\frac{25}{8}\}\}$ for each $x \in (-\infty, -\frac{25}{8}]$ and $u, v \in (0,1)$. By Lemma 3.3, we have:

$$\mathbb{I}_{\varphi} * \mathbb{I}_{\varphi}(x,u,v) = 0$$

for each $x \in (-\infty, -\frac{25}{8}]$ and $u, v \in (0,1)$.

Case 2. Let $x \in (-\frac{25}{8}, 0)$; the functions φ_x and φ intersect at the point $(\frac{x}{2}, \varphi(\frac{x}{2}))$, where $s = \frac{x}{2}$ is the unique solution to $\varphi_x(s) = \varphi(s)$ for each $x \in (-\frac{25}{8}, 0)$. Indeed, the latter equality can be written as follows: $-x \cdot (2s - x) = 0$. Since $x \neq 0$, it follows that $s = \frac{x}{2}$ is the unique solution to $\varphi_x(s) = \varphi(s)$, verifying the previous statement. Furthermore, the fact that $x \in (-\frac{25}{8}, 0)$ allows us to establish the following inequalities: $-\frac{25}{16} < \frac{x}{2} < 0$ and $-\frac{25}{16} < \frac{25}{16} + x < \frac{25}{16}$. Now, and without loss of generality, suppose that $x \in [-\frac{25}{16}, 0)$. Then a general graphical representation of both functions, in this case, is as follows:

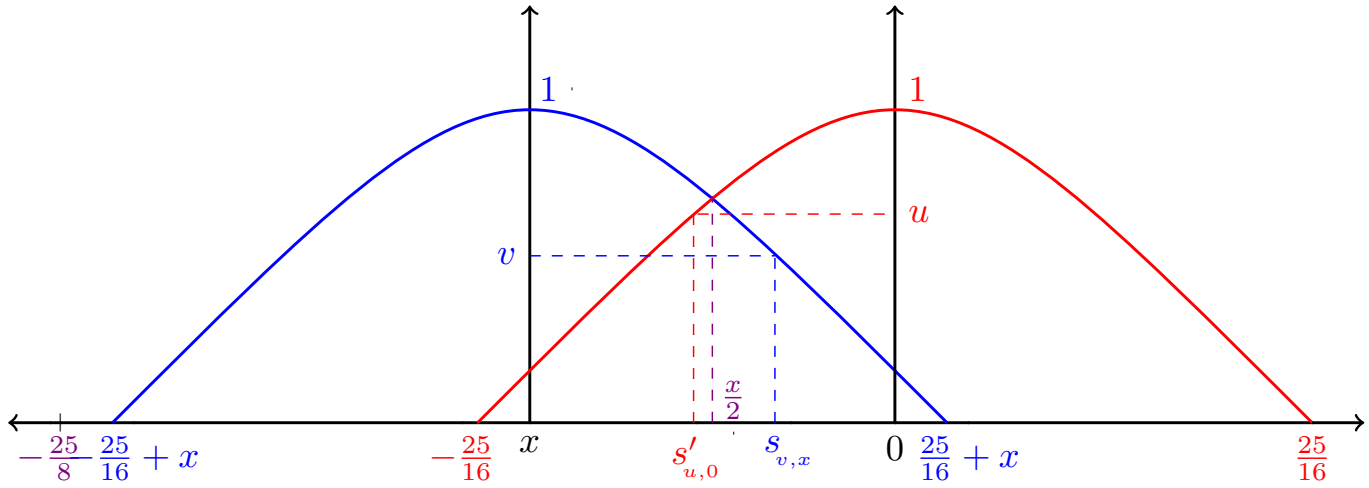


Figure 2. *General graphical representation of the functions φ_x and φ for each $x \in (-\frac{25}{8}, 0)$, where $(\frac{x}{2}, \varphi(\frac{x}{2}))$ is their point of intersection when $s = \frac{x}{2}$.*

With the support of Figure 2, we can guarantee that:

1) $\mathrm{I}^{v,x}_{u,0} = [s'_{u,0}, s_{v,x}]$ if $s'_{u,0} < s_{v,x}$, $s'_{u,0} \neq \frac{x}{2}$, and $s_{v,x} \neq \frac{x}{2}$: $\mathbf{C}_{2_i}$;
2) $\mathrm{I}^{v,x}_{u,0} = [s'_{u,0}, \frac{x}{2}]$ if $s'_{u,0} < s_{v,x}$, $s'_{u,0} \neq \frac{x}{2}$, and $s_{v,x} = \frac{x}{2}$: $\mathbf{C}_{2_{ii}}$;
3) $\mathrm{I}^{v,x}_{u,0} = [\frac{x}{2}, s_{v,x}]$ if $s'_{u,0} < s_{v,x}$, $s'_{u,0} = \frac{x}{2}$, and $s_{v,x} \neq \frac{x}{2}$: $\mathbf{C}_{2_{iii}}$;
4) $\mathrm{I}^{v,x}_{u,0} \in \{\emptyset, \{s'_{u,0}\}, \{s_{v,x}\}, \{\frac{x}{2}\}\}$ if $s_{v,x} \leq s'_{u,0}$: $\mathbf{C}_{2_{iv}}$;

for each $x \in (-\frac{25}{8}, 0)$ and $u, v \in (0, 1)$. Combining Lemma 3.3, Corollary 3.2, and the equalities in items (i)–(iv) implies that:

$$\mathbb{I}_\varphi * \mathbb{I}_\varphi(x, u, v) = \begin{cases} x + \frac{25}{16}(\sqrt{1-u} + \sqrt{1-v}) & \text{if } \mathbf{C}_{2_i}, \\ \frac{x}{2} + \frac{25}{16}\sqrt{1-u} & \text{if } \mathbf{C}_{2_{ii}}, \\ \frac{x}{2} + \frac{25}{16}\sqrt{1-v} & \text{if } \mathbf{C}_{2_{iii}}, \\ 0 & \text{if } \mathbf{C}_{2_{iv}}, \end{cases}$$

for each $x \in (-\frac{25}{8}, 0)$ and $u, v \in (0, 1)$.

Case 3. Let $x = 0$, then $\varphi_0(s) = \varphi(s)$ for each $s \in \mathbb{R}$, so that the general graphical representation of both functions, in this case, is as follows:

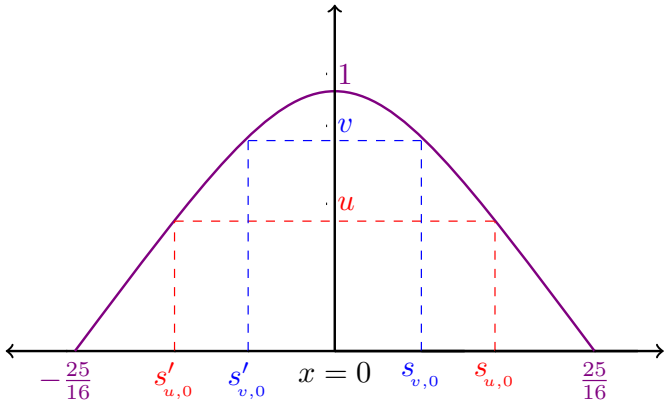


Figure 3. *General graphical representation of the functions φ_x and φ for $x = 0$.*

With the support of Figure 3, we have:

i. $\mathrm{I}^{v,0}_{u,0} = [s'_{v,0}, s_{v,0}]$ if $s'_{u,0} \leq s'_{v,0}$: $\mathbf{C}_{3_i}$;

ii. $\mathrm{I}^{v,0}_{u,0} = [s'_{u,0}, s_{u,0}]$ if $s'_{v,0} < s'_{u,0}$: $\mathbf{C}_{3_{ii}}$;

for each $u, v \in (0, 1)$. Combining Lemma 3.3, Corollary 3.2, and the equalities in items (i)–(ii) implies that:

$$\mathbb{I}_\varphi * \mathbb{I}_\varphi(0, u, v) = \begin{cases} \frac{25}{8}\sqrt{1-v} & \text{if } \mathbf{C}_{3_i}, \\ \frac{25}{8}\sqrt{1-u} & \text{if } \mathbf{C}_{3_{ii}}, \end{cases}$$

for each $u, v \in (0, 1)$.

Case 4. Let $x \in (0, \frac{25}{8})$. An analysis analogous to Case 2 guarantees that the functions φ_x and φ have as a common point $(\frac{x}{2}, \varphi(\frac{x}{2}))$, where $s = \frac{x}{2}$ is the unique solution to the equality $\varphi_x(s) = \varphi(s)$ for each $x \in (0, \frac{25}{8})$. Furthermore, the fact that $x \in (0, \frac{25}{8})$ allows us to establish the following inequality: $0 < \frac{x}{2} < \frac{25}{16}$. Now, without loss of generality, suppose that $x \in (0, \frac{25}{16}]$. Then, $\frac{25}{16} < x + \frac{25}{16} \leq \frac{25}{8}$. Consequently, to obtain the particular graphical representation of both functions in this case, it suffices to swap x and 0 in the general graphical representation of Case 2 (Figure 2), considering the adjustments involved in such a swap. Thus, said graphical representation is:

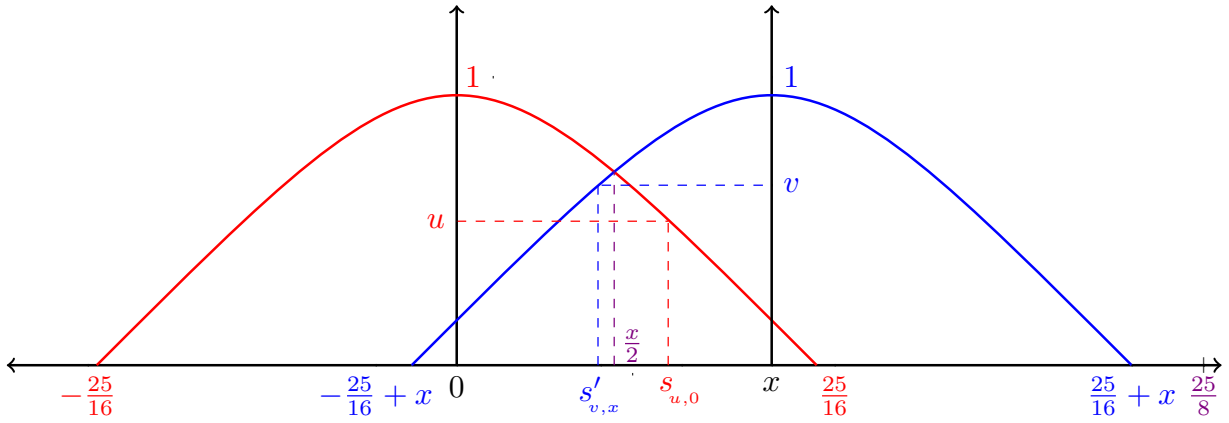


Figure 4. *General graphical representation of the functions φ_x and φ for each $x \in \left(0, \frac{25}{8}\right)$, where $(\frac{x}{2}, \varphi(\frac{x}{2}))$ is their point of intersection when $s = \frac{x}{2}$.*

With the support of Figure 4, we have:

1) $\mathrm{I}^{v,x}_{u,0} = [s'_{v,x}, s_{u,0}]$ if $s'_{v,x} < s_{u,0}$, $s'_{v,x} \neq \frac{x}{2}$, and $s_{u,0} \neq \frac{x}{2}$: $\mathbf{C}_{4_i}$;
2) $\mathrm{I}^{v,x}_{u,0} = [\frac{x}{2}, s_{u,0}]$ if $s'_{v,x} < s_{u,0}$, $s'_{v,x} = \frac{x}{2}$, and $s_{u,0} \neq \frac{x}{2}$: $\mathbf{C}_{4_{ii}}$;
3) $\mathrm{I}^{v,x}_{u,0} = [s'_{v,x}, \frac{x}{2}]$ if $s'_{v,x} < s_{u,0}$, $s'_{v,x} \neq \frac{x}{2}$, and $s_{u,0} = \frac{x}{2}$: $\mathbf{C}_{4_{iii}}$;
4) $\mathrm{I}^{v,x}_{u,0} \in \{\emptyset, \{s'_{v,x}\}, \{s_{u,0}\}, \{\frac{x}{2}\}\}$ if $s_{u,0} \leq s'_{v,x}$: $\mathbf{C}_{4_{iv}}$;

for each $u, v \in (0,1)$ and $x \in (0, \frac{25}{8})$. Combining Lemma 3.3, Corollary 3.2, and the equalities in items (i)–(iv) implies that:

$$\mathbb{I}_\varphi * \mathbb{I}_\varphi(x,u,v) = \begin{cases} -x + \frac{25}{16}\left(\sqrt{1-u} + \sqrt{1-v}\right) & \text{if } \mathbf{C}_{4_i}, \\ -\frac{x}{2} + \frac{25}{16}\sqrt{1-u} & \text{if } \mathbf{C}_{4_{ii}}, \\ -\frac{x}{2} + \frac{25}{16}\sqrt{1-v} & \text{if } \mathbf{C}_{4_{iii}}, \\ 0 & \text{if } \mathbf{C}_{4_{iv}}, \end{cases}$$

for each $x \in \left(0, \frac{25}{8}\right)$ and $u, v \in (0,1)$.

Case 5. Let $x \in [\frac{25}{8}, \infty)$. An analysis analogous to Case 1 guarantees that the functions φ_x and φ intersect at the point $(\frac{25}{16}, 0)$ when $x = \frac{25}{8}$. Consequently, a general graphical representation of both functions, in this case, is as follows:

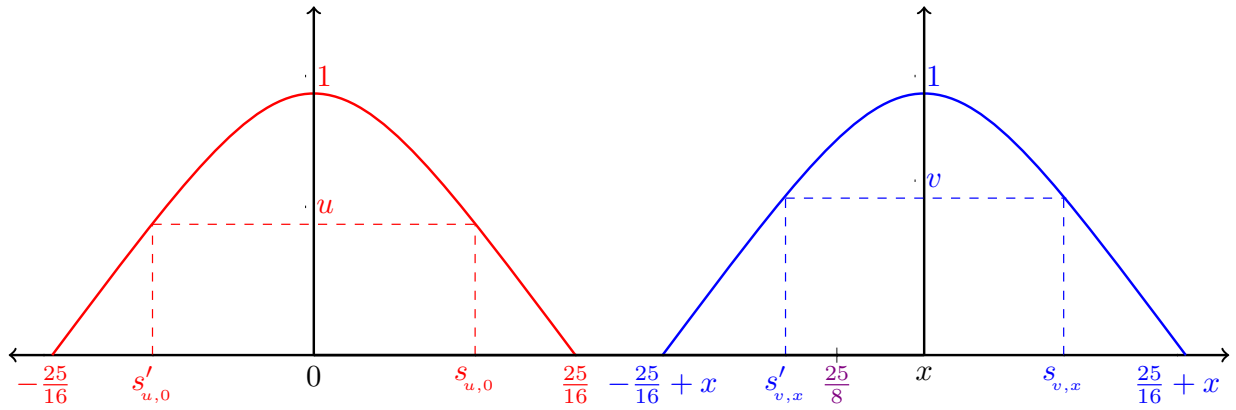


Figure 5. *General graphical representation of the functions φ_x and φ for each $x \in [\frac{25}{8}, \infty)$, where $(\frac{25}{16}, 0)$ is their point of intersection when $x = \frac{25}{8}$.*

Observe that $x - \frac{25}{16} < s'_{v,x}$ for each $v > 0$. Now, with the support of Figure 5, we can guarantee that $\mathrm{I}^{v,x}_{u,0} \in \{\emptyset, \{\frac{25}{8}\}\}$ for each $u, v \in (0,1)$ and $x \in [\frac{25}{8}, \infty)$. By Lemma 3.3, we have:

$$\mathbb{I}_\varphi * \mathbb{I}_\varphi(x,u,v) = 0$$

for each $x \in [\frac{25}{8}, \infty)$ and $u, v \in (0,1)$.

4. Simulated Data Analysis

For the purpose of studying the practical behavior of the estimator (18) with simulated data, this section presents the choice of the smoothing parameter $\tilde{b}_n$ to estimate a known density f via (10). To this end, we will consider the density of the log-normal distribution with $\mu = 50$ and $\sigma = 0.3$, given by:

$$f(x) = \frac{1}{x\,\sigma\sqrt{2\,\pi}}\,\exp\left[-\frac{1}{2\,\sigma^2}\left(\log\left[\frac{x}{\mu}\right]\right)^2\right], \quad x > 0.$$

For the choice of the smoothing parameter $\tilde{b}_n$ through (18), φ given in (5) was implemented and four observed random samples $\mathbf{U}$, associated with f, of size $n = 500$ were considered such that $\mathbf{U} = ((X_{1_k}, V_{1_{k_i}}) \ldots, (X_{n_k}, V_{n_{k_i}}))$, where X_{j_k} has density f and $V_{j_{k_i}} \sim \mathcal{U}(0,1)$ for $1 \leq k \leq 4$, $1 \leq j \leq n$, and $1 \leq i \leq 1000$. The smoothing parameters $\tilde{b}_n$ obtained through the average estimates of (18) are exhibited in Table 1, for $b \in \{0.10, 0.25, 0.40, 0.55, \ldots, 25\}$ in each estimation. Note that one thousand estimates of (18) were averaged across each of the four observed random samples. Furthermore, Figure 6 displays the graphical representation of the average estimate of (18) for the observed random sample No. 4.

On the other hand, to estimate the density f with (10), φ given in (5) was implemented, along with the smoothing parameters $b_{asy} = 5.1325$ and $\tilde{b}_n \in \{2.6500, 4.9000, 5.6500, 6.2500\}$. Additionally, an observed random sample $\mathbf{U}$, associated with f, of size $n = 1000$ was considered such that $\mathbf{U} = ((X_1, V_{1_i}) \ldots, (X_n, V_{n_i}))$, where X_j has density f and $V_{j_i} \sim \mathcal{U}(0,1)$ for $1 \leq j \leq n$ and $1 \leq i \leq 1000$. Figure 7 presents the graphs of f and the average estimates of (10) for $x \in \{0.1, 1.1, 3.1, 4.1, \ldots, 125\}$ in each case. Note that one thousand estimates of (10) were averaged for the considered sample. It is worth emphasizing that, once the intensive simulations were carried out and completed, the subfigures in Figure 7 were subjectively selected by choosing the graph of the average estimate of (10) with smoothing parameter b_{asy} that presented the best fit with respect to f.

Table 1. *Choice of the smoothing parameter $\tilde{b}_n$ via $\hat{Q}(\hat{\vartheta}_n)$.*

Observed random sample No.	Selected smoothing parameter $\tilde{b}_n$
1	6.2500 with $\min(\hat{Q}(\hat{\vartheta}_n)) = -0.0173$
2	2.6500 with $\min(\hat{Q}(\hat{\vartheta}_n)) = -0.0178$
3	4.9000 with $\min(\hat{Q}(\hat{\vartheta}_n)) = -0.0175$
4	5.6500 with $\min(\hat{Q}(\hat{\vartheta}_n)) = -0.0179$

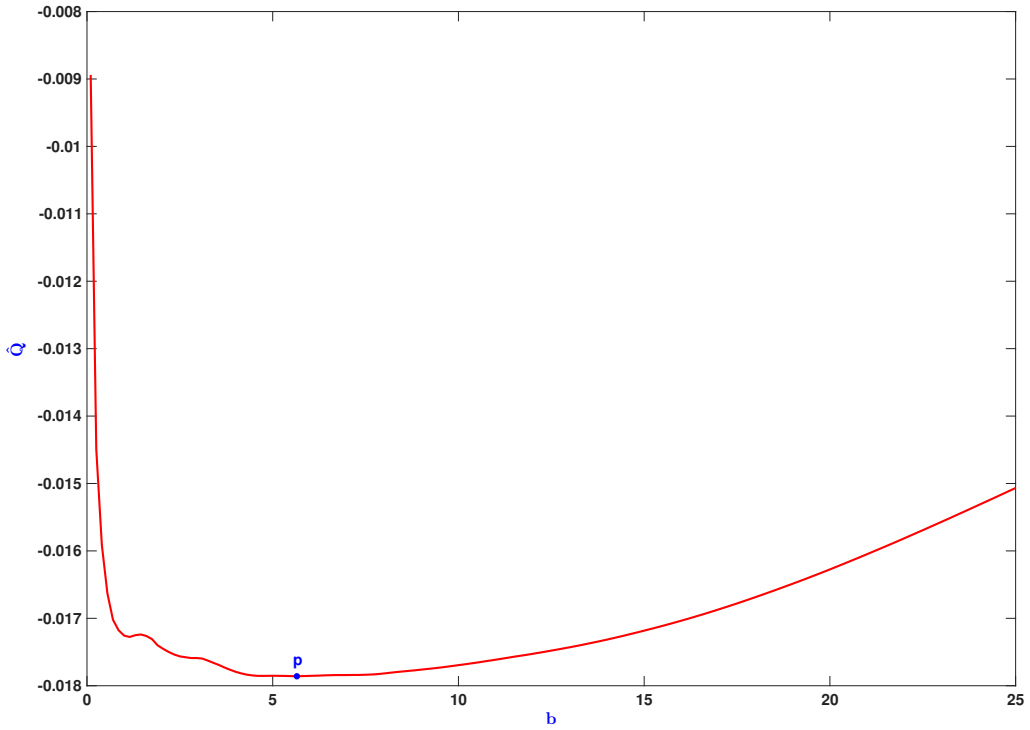
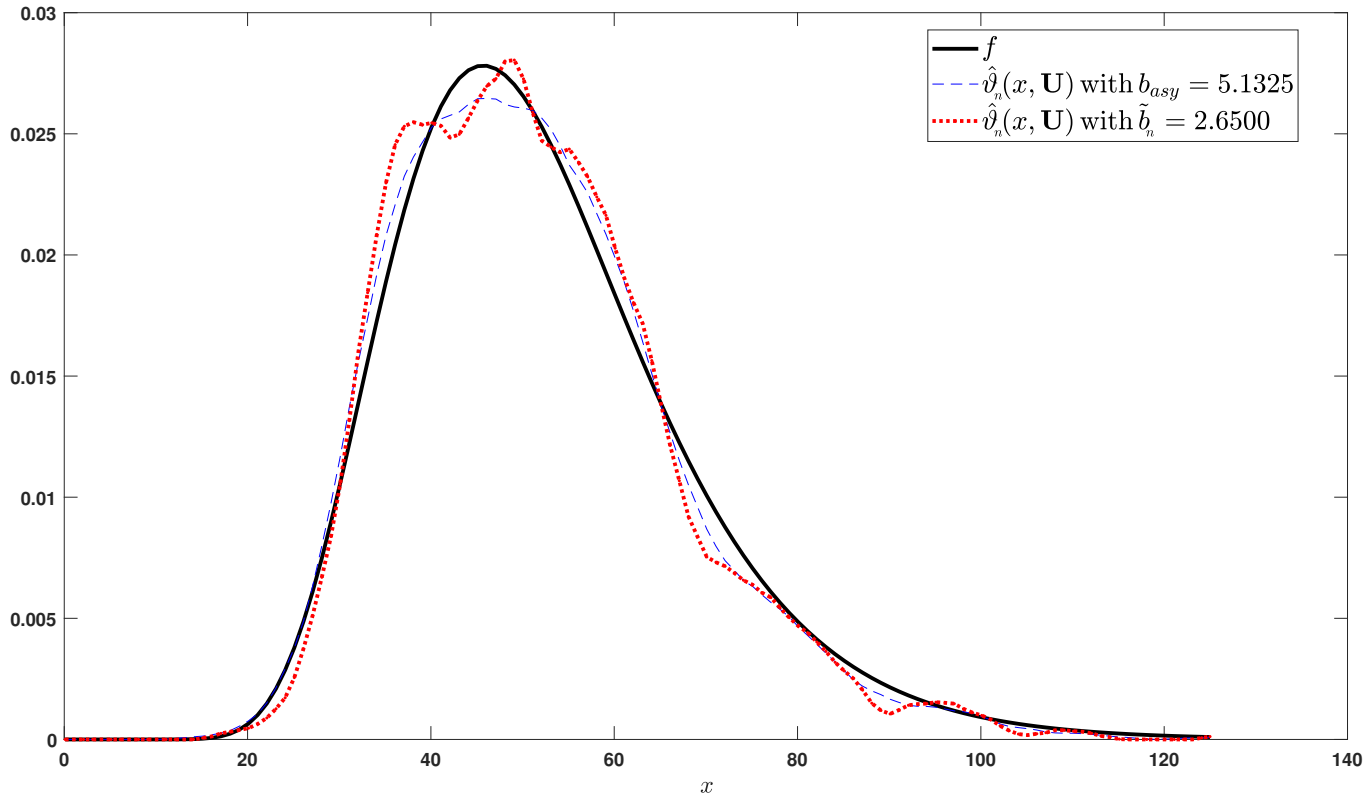
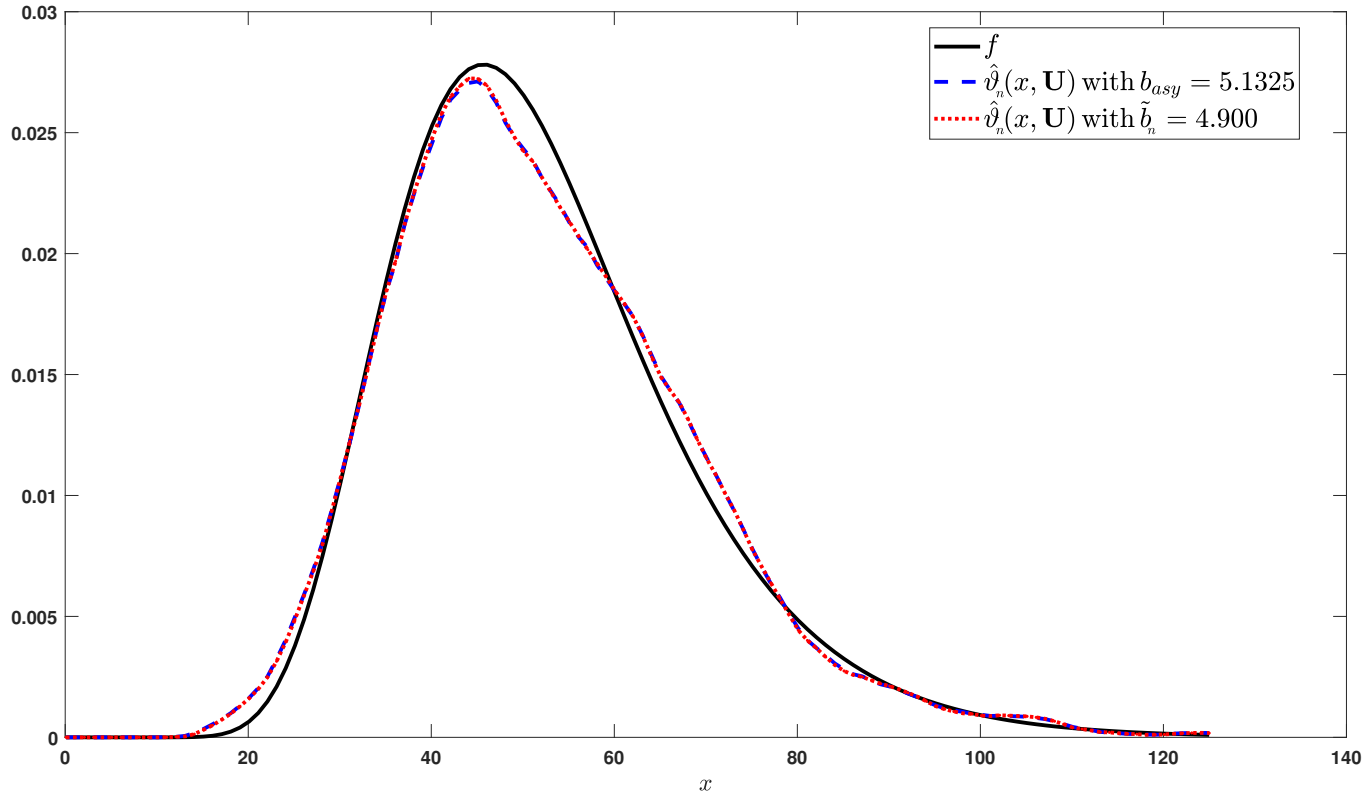


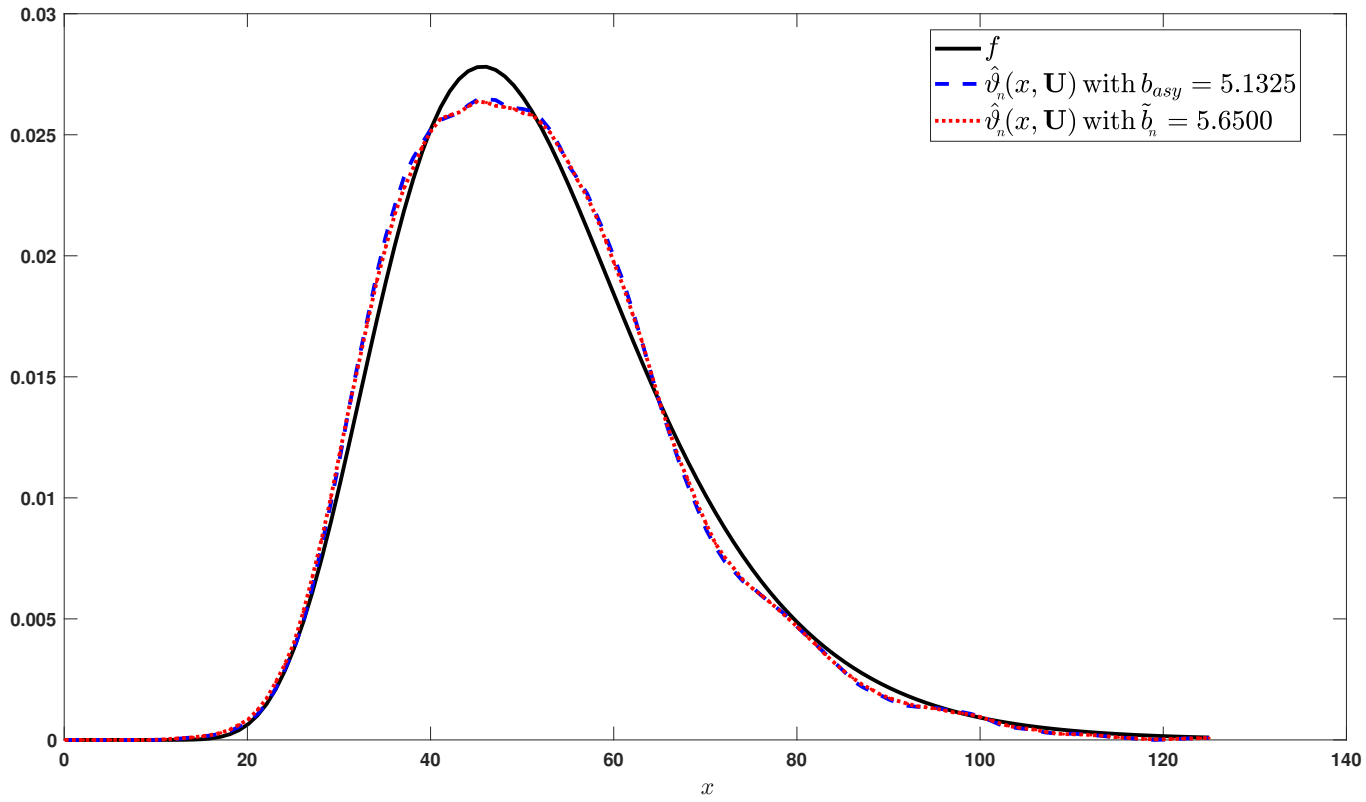
Figure 6. *Graphical representation of $\hat{Q}(\hat{\vartheta}_n)$ for the observed random sample No. 4, where p has coordinates $(\tilde{b}_n, \min(\hat{Q}(\hat{\vartheta}_n))) = (5.6500, -0.0179)$.*


0.03
0.025
0.02
0.015
0.01
0.005
0
0
20
40
60
80
100
120
140
x
f
$\hat{\vartheta}_n(x, \mathbf{U})$ with $b_{asy} = 5.1325$
$\hat{\vartheta}_n(x, \mathbf{U})$ with $\tilde{b}_n = 2.6500$

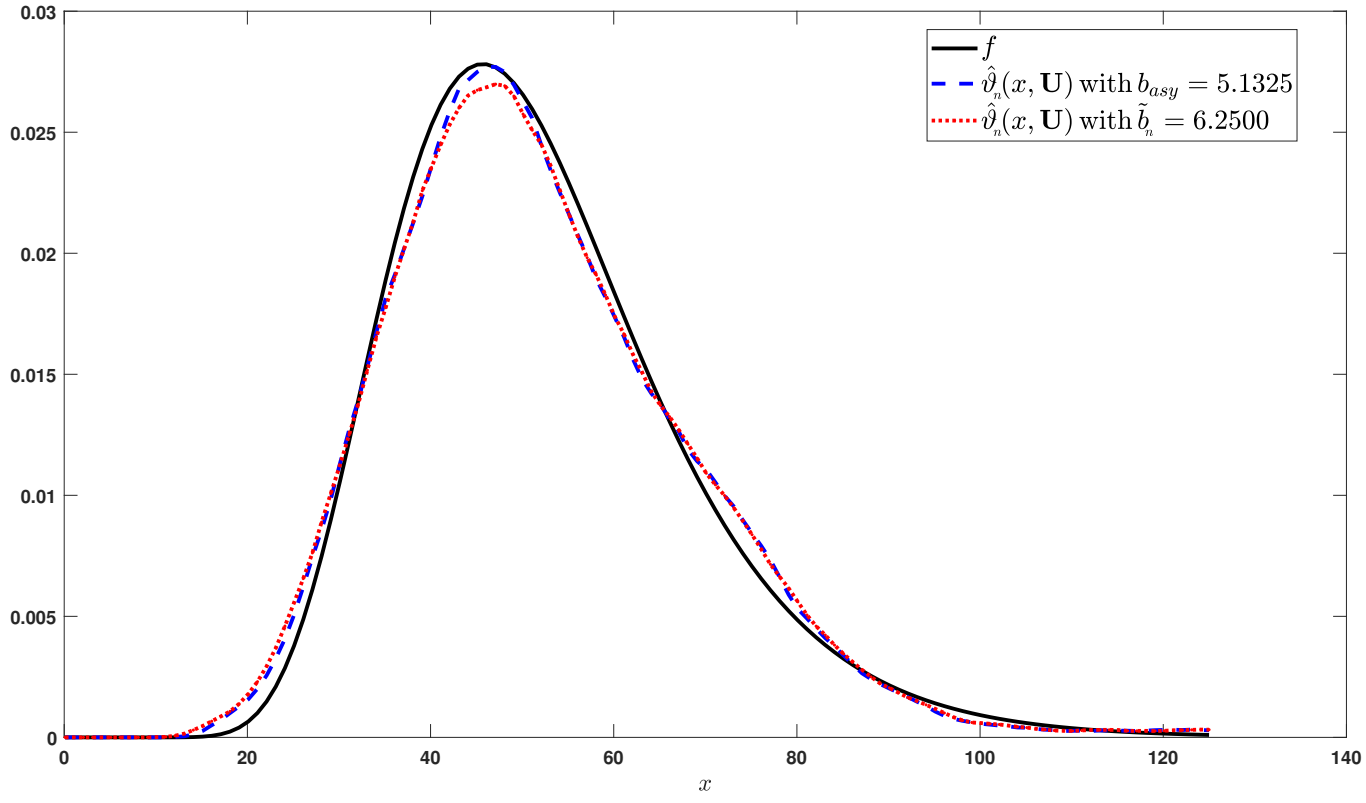
(a)


0.03
0.025
0.02
0.015
0.01
0.005
0
0
20
40
60
80
100
120
140
x
f
$\hat{\vartheta}_n(x, \mathbf{U})$ with $b_{asy} = 5.1325$
$\hat{\vartheta}_n(x, \mathbf{U})$ with $\tilde{b}_n = 4.900$

(b)



(c)



(d)

Figure 7*. Graphs of f and its estimates $\hat{\vartheta}_n$ with $b_{asy} = 5.1325$ and $\tilde{b}_n \in \{2.6500, 4.9000, 5.6500, 6.2500\}$.*

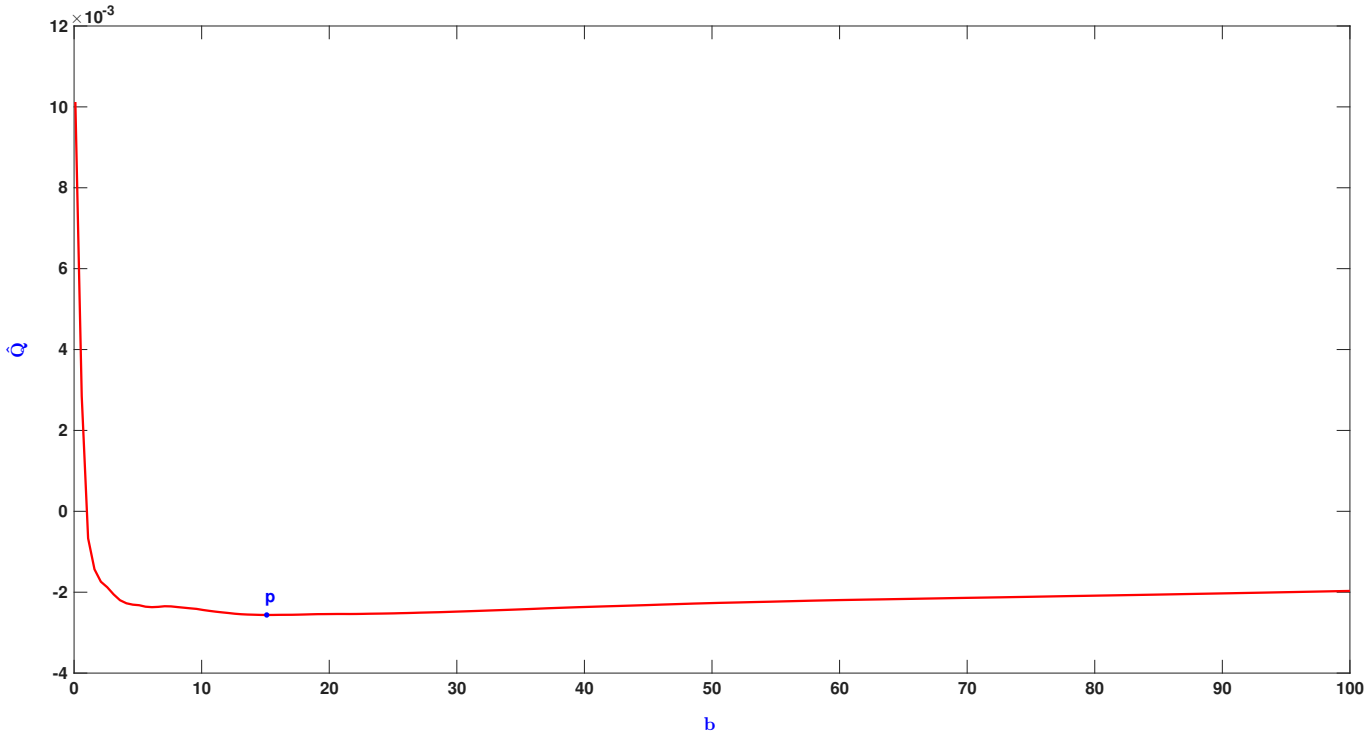
5. Real Data Analysis

In this section, the practical utility of (18) will be tested with two real datasets, where the smoothing parameter $\tilde{b}_n$ and three subjectively chosen bandwidths will be implemented to estimate the unknown density of the data via (10).

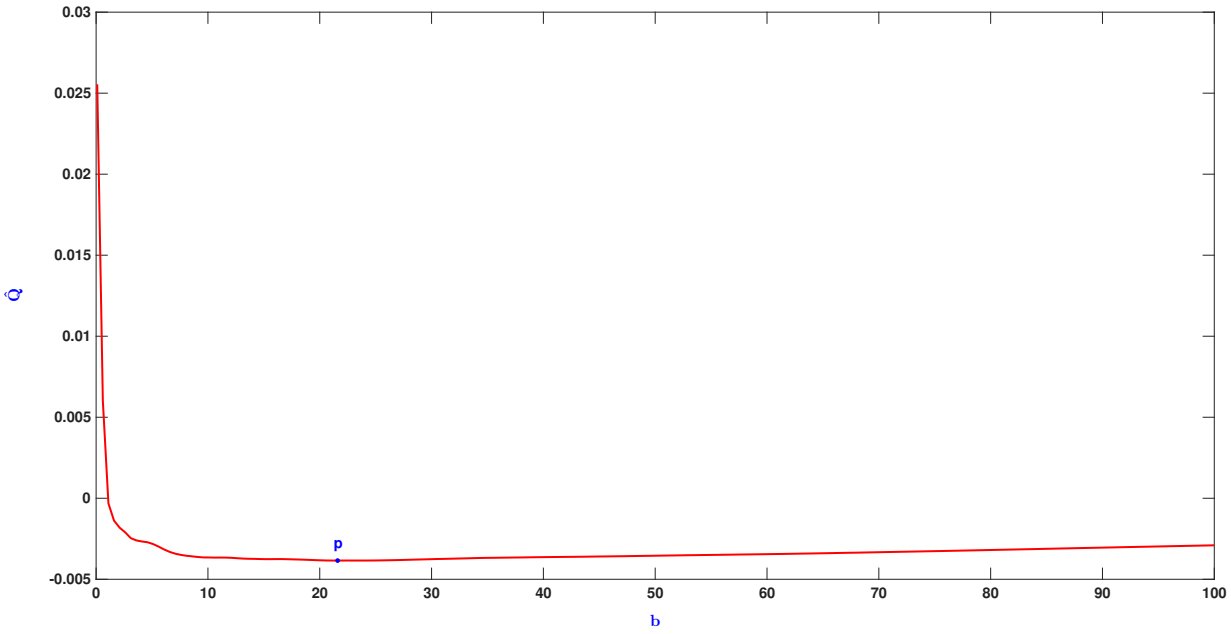
The first dataset corresponds to the one hundred and ninety expanded and corrected observations in Section 7 of Jarrett [10], which characterize the intervals in days between successive coal-mining disasters in Great Britain, from March 15, 1851 to March 22, 1962. In this case, we considered the following observed random samples $\mathbf{U} = ((X_1, V_{1_i}), (X_2, V_{2_i}), \ldots, (X_n, V_{n_i}))$, where $n = 190$, X_j is the j-th observation in [10], and $V_{j_i} \sim \mathcal{U}(0,1)$ for $1 \leq j \leq n$ and $1 \leq i \leq 1000$. The average estimate of (18) with φ given in (5) provided the smoothing parameter $\tilde{b}_n = 15.1000$ and $\min(\hat{Q}(\hat{\vartheta}_n)) = -0.0026$. The graph of the previous average is displayed in Figure 10a, for $b \in \{0.1, 0.6, 1.1, 1.6, \ldots, 100\}$. Simultaneously and under the previous practical scheme, Figure 9 displays the graphs of the average estimated densities with (10) for the real dataset considered, using $\tilde{b}_n = 15.1000$, the subjective smoothing parameters $b_s \in \{60.2000, 120.0000, 180.0000\}$, and $x \in \{0.1, 0.125, 0.150, 0.275, \ldots, 2000\}$. Note that one thousand estimates of (18) and (10) were averaged. It is worth emphasizing that once the intensive simulations were completed, each graph was selected considering its smoothness.

The second dataset considered corresponds to the observations published by Copas and Fryer [1], which characterize the duration of eighty-six psychiatric treatment periods in patients used as controls in a study on suicide risks. For $n = 86$ and under the same practical scheme applied to the first dataset, the average estimate of (18) provided the smoothing parameter $\tilde{b}_n = 21.6000$ and $\min(\hat{Q}(\hat{\vartheta}_n)) = -0.0038$. The graph of the obtained average is displayed in Figure 10b for $b \in \{0.1, 0.6, 1.1, 1.6, \ldots, 100\}$. Next, Figure 10 presents the graphs of the average estimated densities with (10) for the real dataset considered, using $\tilde{b}_n = 21.6000$, the subjective smoothing parameters $b_s \in \{43.2000, 63.2000, 90.2000\}$, and $x \in \{0.1, 0.125, 0.150, 0.275, \ldots, 900\}$. Note that one thousand estimates of (18) and (10) were averaged. It should be noted that, once the intensive simulations were completed, each graph was selected considering its smoothness.

Notice that progressive increases in $\tilde{b}_n$ in both datasets allowed us to obtain subjective smoothing parameters, b_s, that significantly improved the smoothness of the estimated density. A decrease in such parameter will distort it.

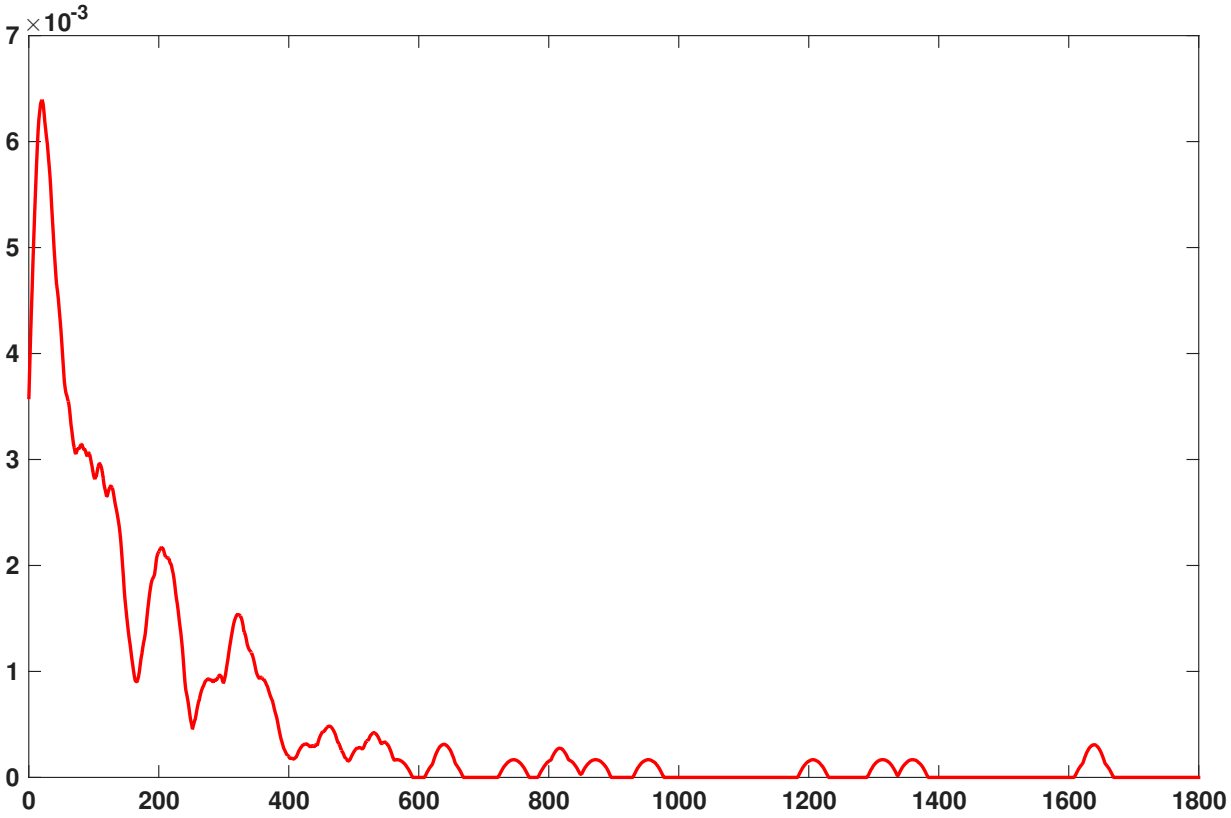


(a) $\mathrm{p} = (\tilde{b}_n, \min(\hat{Q}(\hat{\vartheta}))) = (15.1000, -0.0026)$

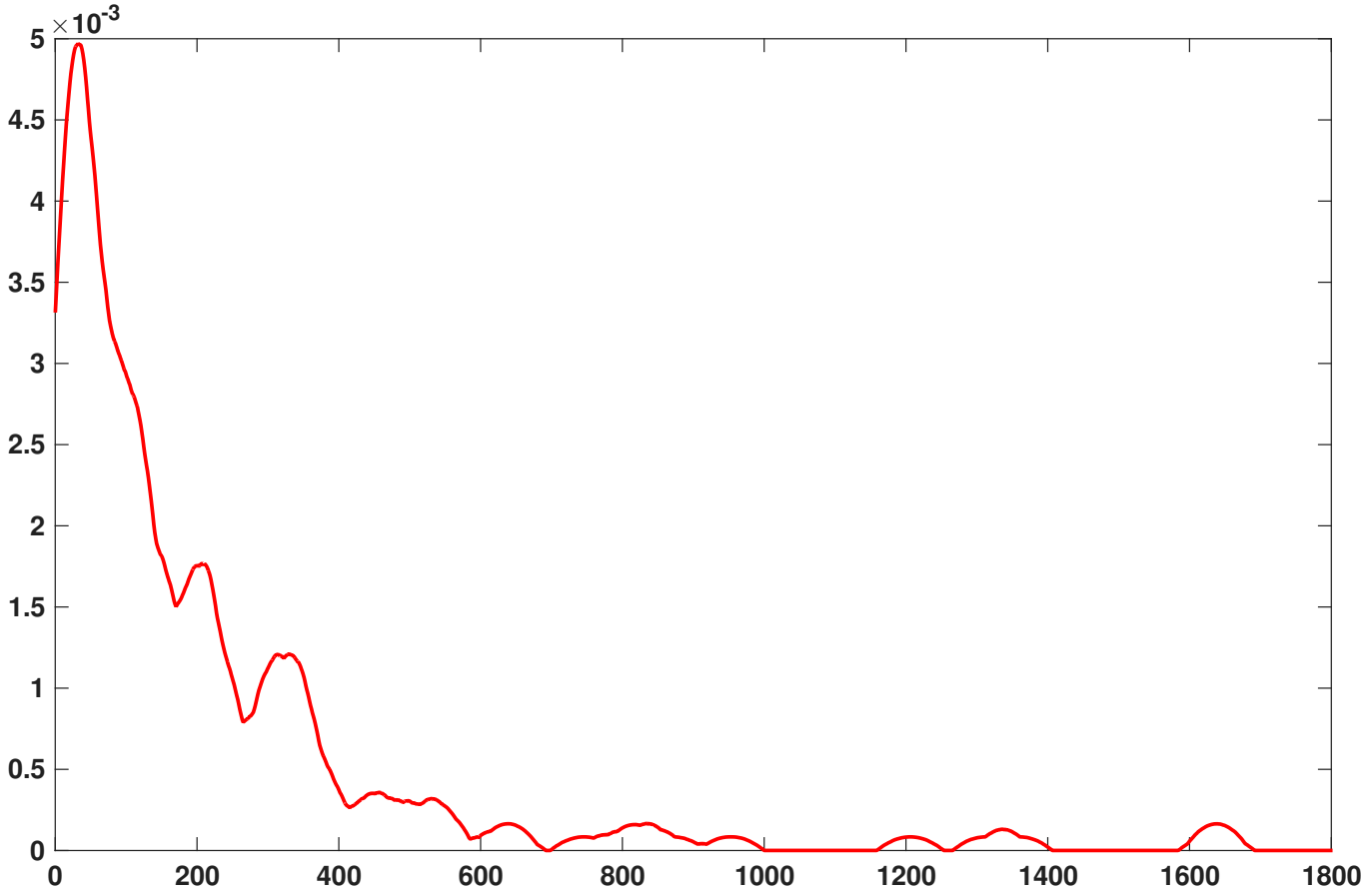


(b) p = $(\tilde{b}_n, \min(\hat{Q}(\hat{\vartheta})) = (21.6000, -0.0038)$

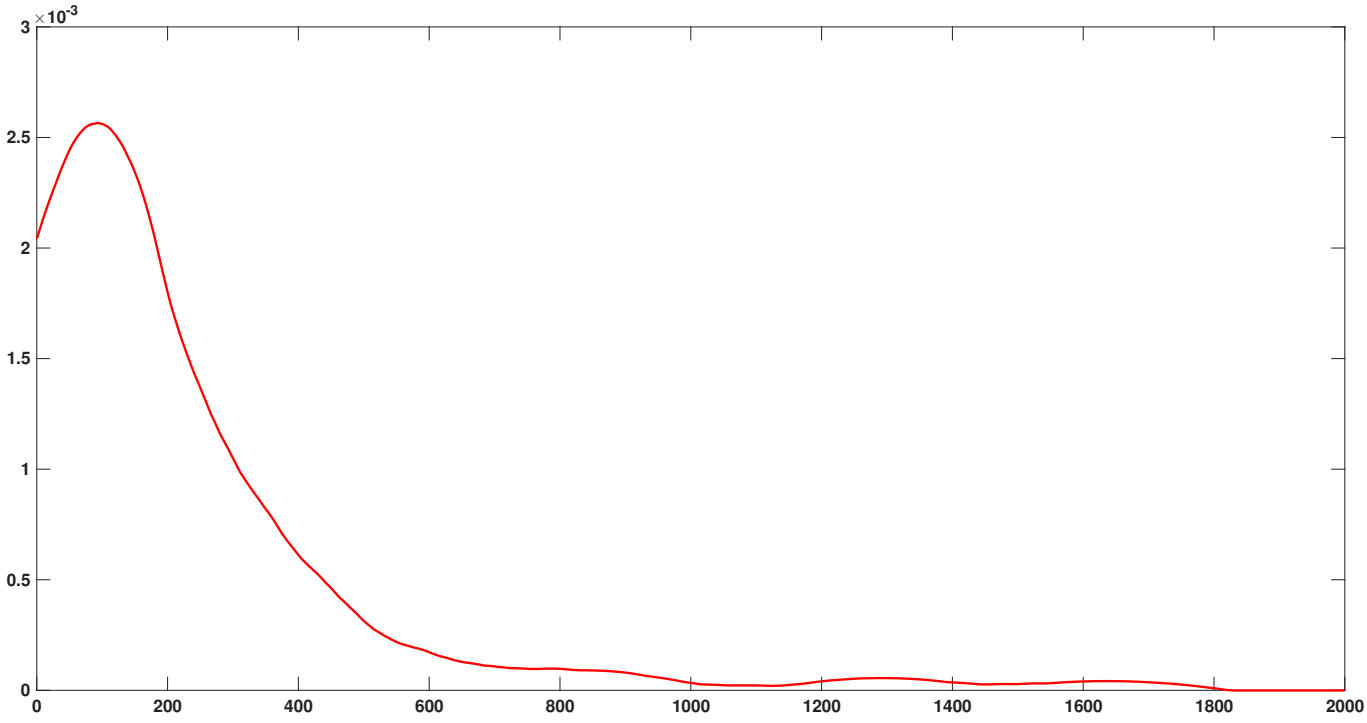
Figure 8. *Estimation of $\hat{Q}(\hat{\vartheta}_n)$ for the observations in Jarrett [10], and Copas and Fryer [1], respectively.*



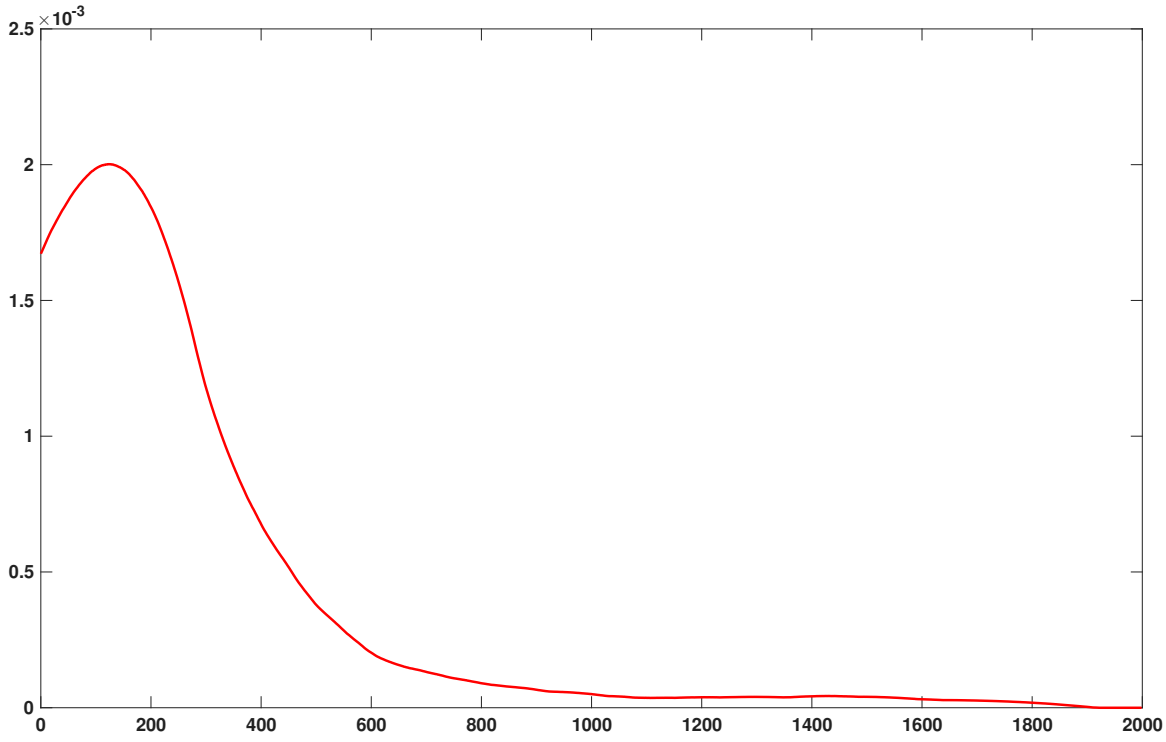
(a) *Estimation with $\tilde{b}_n$* = 15.1000



(b) Estimation with b_s = 60.2000

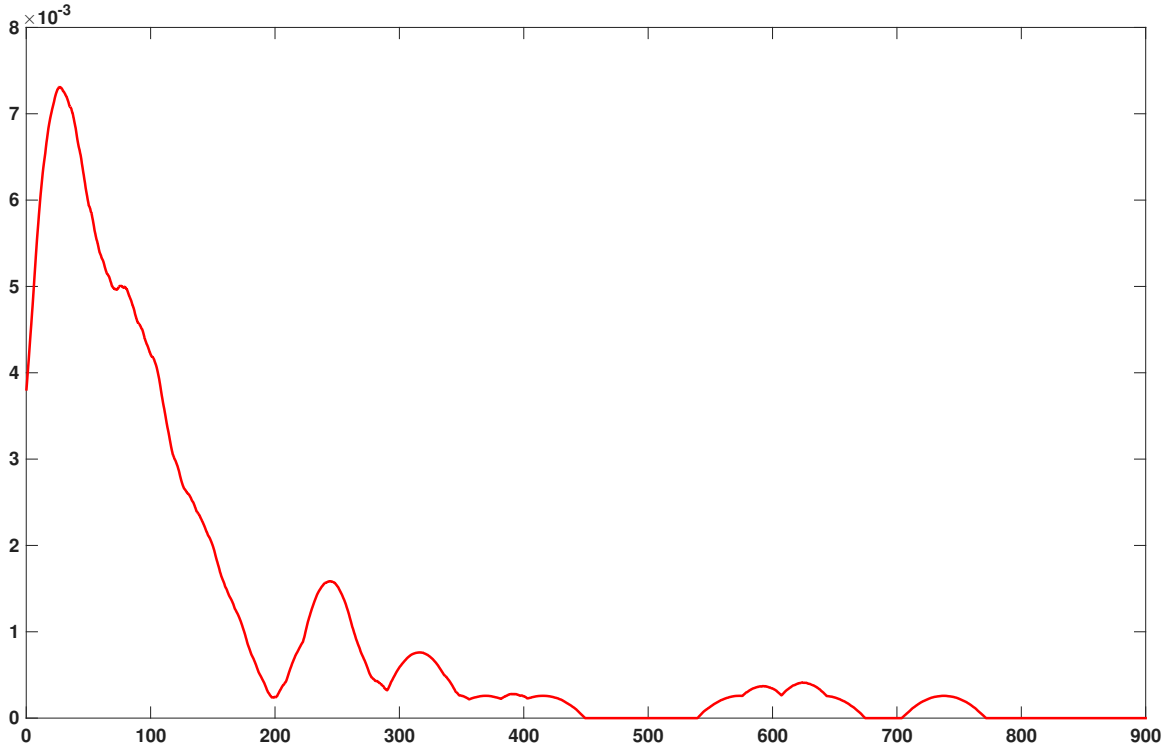


(c) Estimation with b_s = 120.0000

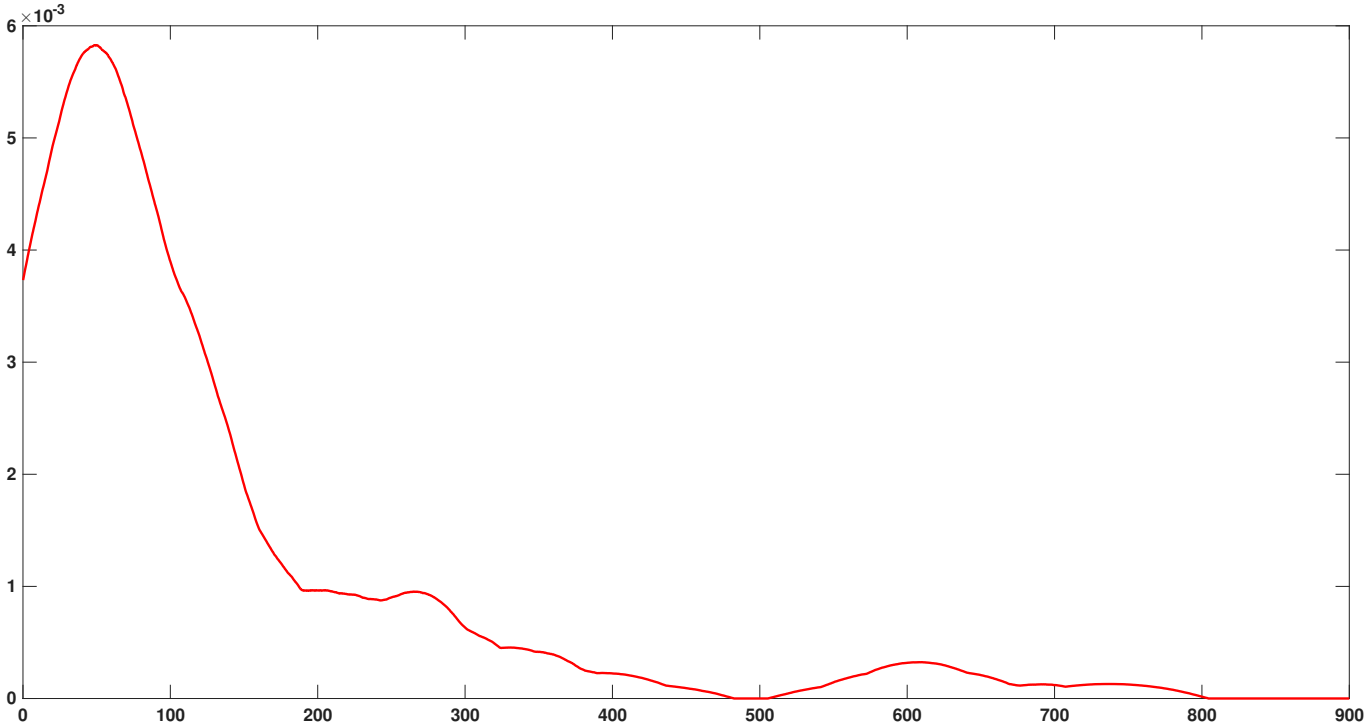


(d) Estimation with b_s = 180.0000

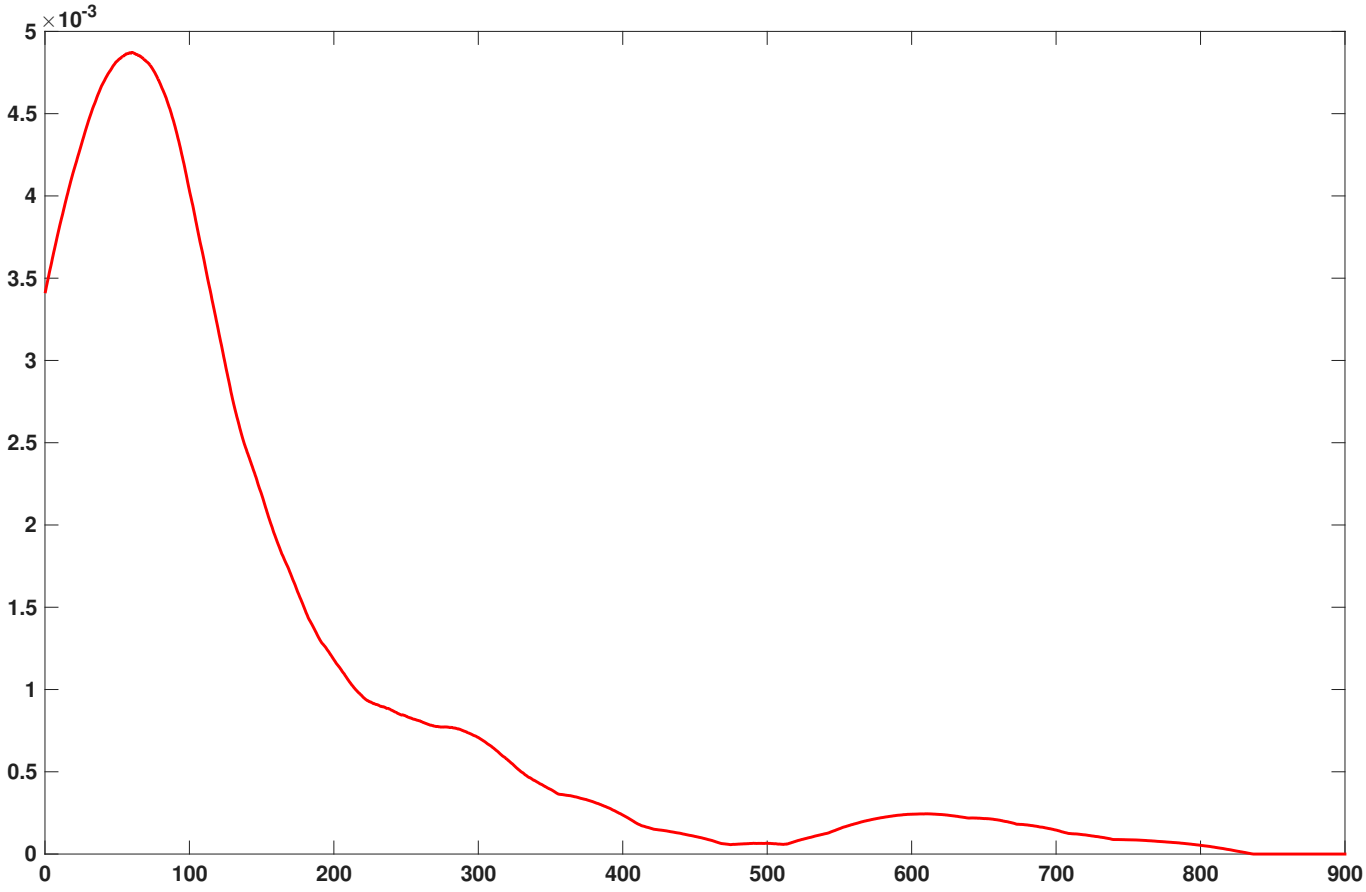
Figure 9. *Estimated densities for the dataset in Jarrett [10].*



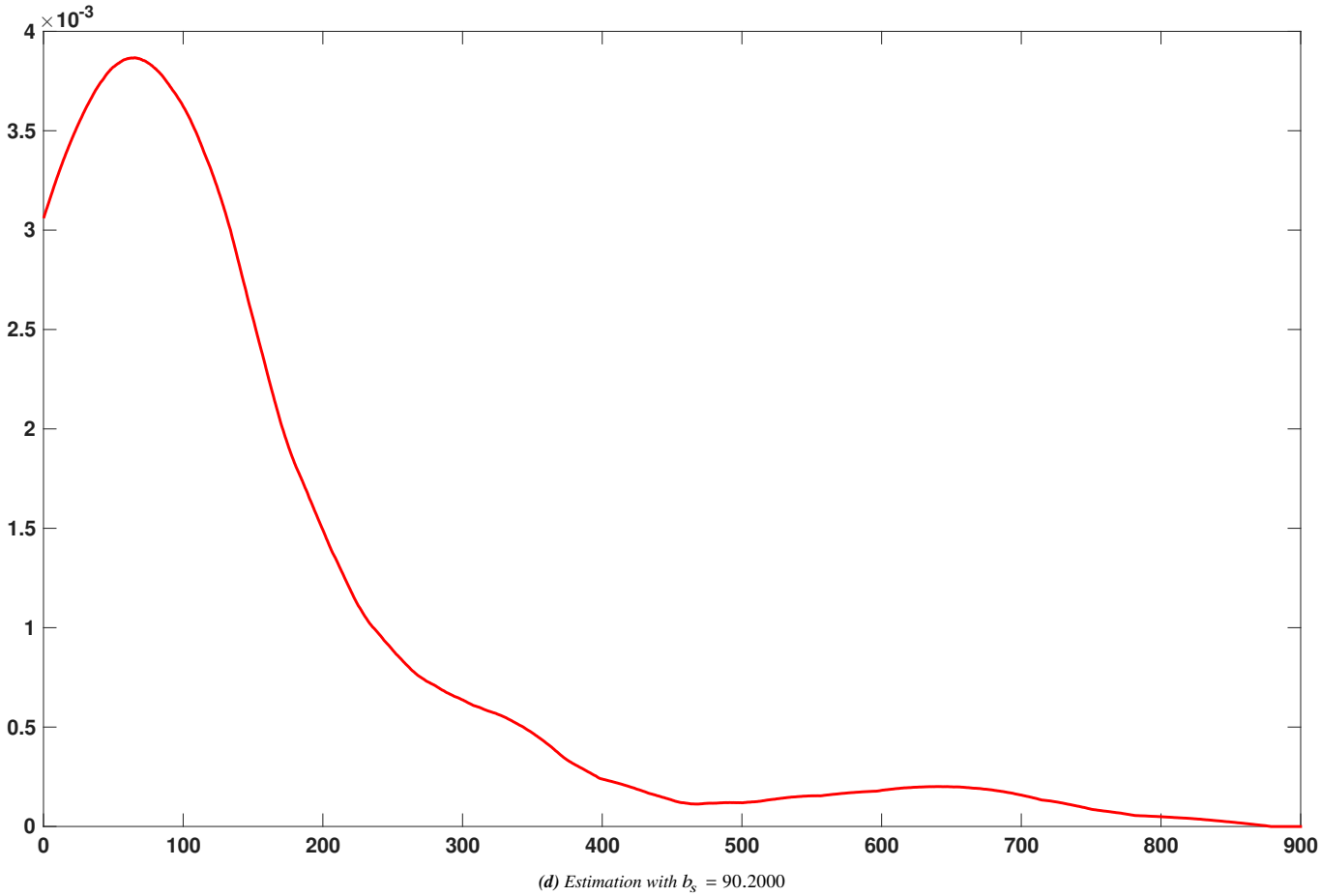
(a) Estimation with $\tilde{b}_n$ = 21.6000



(b) Estimation with $b_s = 43.2000$



(c) Estimation with $b_s = 63.2000$



(d) Estimation with b_s = 90.2000

Figure 10. *Estimated densities for the dataset in Copa and Fryer [1].*

6. Conclusions

The results obtained in this paper allow us to extend the theory introduced by Rudemo [14] to the case of the risk function estimator corresponding to the integrated mean squared error. Furthermore, they provide an important contribution to the fuzzy set density estimation method, as they equip it with an automatic and autonomous smoothing parameter to estimate the density both when it is known and when it is unknown. This contribution will undoubtedly find a natural application in fuzzy set boundary estimation.

Finally, it is worth emphasizing that a review of the implementation of the risk function estimator corresponding to the Kullback-Leibler information measure Rudemo [14] will optimize the criterion for selecting the smoothing parameter of the fuzzy set density estimator. However, further research is needed to broaden the criteria for selecting the smoothing parameter in this context, to the point of establishing the asymptotic optimality of the chosen smoothing parameter. Currently, the author is actively working in this direction.

Abbreviations

IMSE	Integrated Mean Squared Error
CV	Cross-Validation

Acknowledgments

The author would like to express his sincere gratitude to Editorial Assistant Alina Carlson and the editorial team for their valuable support, guidance, and assistance throughout the submission and editorial process. The author is also deeply grateful for the insightful comments and suggestions provided by the anonymous reviewers, which significantly contributed to improving the clarity and quality of this paper.

Author Contributions

Jesús A. Fajardo-González: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

References

[1] J. B. Copas and M. J. Fryer, "Density estimation and suicide risks in psychiatric treatment," *J. Roy. Statist. Soc. A*, 143(2): 167–176, 1980. https://api.semanticscholar.org/CorpusID:125890360

[2] J. Fajardo, "A criterion for the fuzzy set estimation of the density function," *Braz. J. Probab. Stat.*, 28(3): 301–312, 2014. https://doi.org/10.1214/12-BJPS208

[3] J. Fajardo and P. Harmath, "Boundary estimation with the fuzzy set density estimator," *METRON*, 79(3): 285–302, 2021. https://doi.org/10.1007/s40300-021-00210-z

[4] J. Fajardo, R. Ríos, and L. Rodríguez, "Properties of convergence of a fuzzy set estimator of the density function," *Braz. J. Probab. Stat.*, 26(2): 208–217, 2012. http://www.jstor.org/stable/43601209

[5] J. A. Fajardo, "Notas sobre el desempeño de los estimadores fronteras de la densidad con núcleo localmente adaptable y con conjunto difuso," *Divulg. Mat.*, 22(2): 48–65, 2022. https://doi.org/10.5281/zenodo.7487484

[6] M. Falk and F. Liese, "LAN of thinned empirical processes with an application to fuzzy set density estimation," *Extremes*, 1(3): 323–349, 1998. https://doi.org/10.1023/A:1009981817526

[7] A. Gramacki, *Nonparametric Kernel Density Estimation and Its Computational Aspects*, volume 37. Springer, 2018. https://doi.org/10.1007/978-3-319-71688-6

[8] N.-B. Heidenreich, A. Schindler, and S. Sperlich, "Bandwidth selection for kernel density estimation: a review of fully automatic selectors," *AStA Adv. Stat. Anal.*, 97(4): 403–433, 2013. https://doi.org/10.1007/s10182-013-0216-y

[9] W. Hoeffding, "A class of statistics with asymptotically normal distribution," *Ann. Math. Statist.*, 19(3): 293–325, 1948. http://www.jstor.org/stable/2235637

[10] R. G. Jarrett, "A note on the intervals between coalmining disasters," *Biometrika*, 66(1): 191–193, 1979. https://doi.org/10.1093/biomet/66.1.191

[11] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, 33(3): 1065–1076, 1962. http://www.jstor.org/stable/2237880

[12] R. D. Reiss, *A Course on Point Processes*. Springer Series in Statistics, Springer, New York, 1993. https://doi.org/10.1007/978-1-4613-9308-5

[13] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Ann. Math. Statist.*, 27(3): 832–837, 1956. https://doi.org/10.1214/aoms/1177728190

[14] M. Rudemo, "Empirical choice of histograms and kernel density estimators," *Scand. J. Statist.*, 9(2): 65–78, 1982. https://www.jstor.org/stable/4615859

[15] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, New York, 1986. https://doi.org/10.1007/978-1-4899-3324-9

[16] M. P. Wand and M. C. Jones, *Kernel Smoothing* (1st ed.). Chapman and Hall/CRC, 1994. https://doi.org/10.1201/b14876