Statistics
A Journal of Theoretical and Applied Statistics

ISSN: 0233-1888 (Print) 1029-4910 (Online) Journal homepage: https://www.tandfonline.com/loi/gsta20

Moment density estimation for positive random variables

R. M. Mnatsakanov & F. H. Ruymgaart

To cite this article: R. M. Mnatsakanov & F. H. Ruymgaart (2012) Moment density estimation for positive random variables, Statistics, 46:2, 215-230, DOI: 10.1080/02331888.2010.506277

To link to this article: https://doi.org/10.1080/02331888.2010.506277

Published online: 31 Aug 2010.

Statistics, Vol. 46, No. 2, April 2012, 215–230

Moment density estimation for positive random variables

R.M. Mnatsakanov[a,b]* and F.H. Ruymgaart[c]

[a] *Department of Statistics, West Virginia University, P.O. Box 6330, Morgantown, WV 26506, USA;* [b] *Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, WV 26505, USA;* [c] *Department of Mathematics and Statistics, Texas Tech University, P.O. Box 41042, Lubbock, TX 79409, USA*

(*Received 4 December 2007; final version received 22 June 2010*)

An unknown moment-determinate cumulative distribution function or its density function can be recovered from corresponding moments and estimated from the empirical moments. This method of estimating an unknown density is natural in certain inverse estimation models like multiplicative censoring or biased sampling when the moments of unobserved distribution can be estimated via the transformed moments of the observed distribution. In this paper, we introduce a new nonparametric estimator of a probability density function defined on the positive real line, motivated by the above. Some fundamental properties of proposed estimator are studied. The comparison with traditional kernel density estimator is discussed.

Keywords: moment density estimator; mean-squared error; δ-sequence; L_1-consistency

AMS 2000 Subject Classifications: Primary: 62G05; Secondary: 62G20

1. Introduction

Goldenshluger and Spokoiny [1] consider the reconstruction of planar convex sets from noisy observations of its moments. In several indirect estimation models, like biased sampling and multiplicative censoring, the moments of unobserved distribution of actual interest can be easily estimated from the transformed moments of the observed distributions [2–4].

Another example where the moments are easily calculated but the target distribution has no closed form represents the distribution of the finite weighted sum of independent chi-squared random variables (r.v.'s) (with degree of freedom 1) [5].

It is well known that the moment-determinate probability measure or its cumulative distribution function (cdf) is uniquely defined by its moments [6–9].

It is not overly hard to construct a density estimator entirely based on estimators of the moments. One could, for instance, replace the $\mu(j)$ in Equation (1) by their estimators and subsequently differentiate the result. Some applications of a construction similar to Equation (1) have been demonstrated in the problem of estimating the mixing distribution in Poisson mixture models and in the problem of estimating the so-called structural distribution function in the multinomial

*Corresponding author. Email: rmnatsak@stat.wvu.edu

ISSN 0233-1888 print/ISSN 1029-4910 online

http://dx.doi.org/10.1080/02331888.2010.506277
http://www.tandfonline.com

scheme with a large number of a rare events (see [10,11], respectively). Although it is our ultimate purpose to construct and study density estimators when only the estimated moments are available, it seems expedient to the present authors to first investigate a hybrid form of such a moment density estimator (MDE). Such a hybrid form is obtained by replacing the underlying cdf in Equation (2) by its empirical analogue. This means, in particular, that we still assume that a sample from the unknown density is given.

In this paper, this hybrid form (for brevity still referred to as MDE) will be investigated. The derived properties are expected to be useful as intermediate results in the study of genuine MDEs, which is beyond the scope of this paper.

We will focus on some asymptotic properties of MDEs, when the underlying unknown density has support in the positive half-line $\mathbb{R}_+$. The asymptotic normality will inevitably remain restricted to the local situation because density estimators do not converge weakly in a global sense as is well known. As it turns out, the convergence rate of the mean integrated square error (MISE) is the same as that for ordinary kernel type estimator of the density defined on the entire real line $\mathbb{R}$, but there are some differences regarding the constants appearing in the first order terms.

It is well known that in a standard kernel density estimation approach for positive r.v.'s, the behaviour of the bias term near the left boundary is larger than in the interior of a support. The estimator proposed in this paper is free from this type of edge effect. See, also, Chen [12,13] where the kernel (with varying shape) density estimators (KDE) have been proposed for positive r.v.'s with the supports [0, 1] and $(0, \infty)$, respectively. These estimators have similar properties as ours. See, also, Bouezmarni and Rolin [14] where the exact asymptotic constants of uniform and L_1-errors have been derived for the KDE with Beta kernel. Let us mention also that within the class of non-negative KDEs, our estimator achieves the optimal mean square error (MSE), MISE.

This paper is organized as follows. The MDE $\hat{f}_\alpha$ is constructed in Section 2. In Section 3, the precise constants in the first-order terms of the expansions for MSE and the limiting distribution of the proposed estimator are given. The exact and asymptotic upper bounds for biases and the rates of convergence in L_1- and L_2-norms are established in Section 4. Finally, in Section 5, we applied the least-squares cross-validation (CV) technique to address the question of choice of optimal parameter α in $\hat{f}_\alpha$; the comparison of $\hat{f}_\alpha$ with the traditional KDE $\hat{f}_h$ is briefly discussed as well.

2. Construction of MDE and assumptions

Suppose that we are given a random sample $X_1, \ldots, X_n$ of independent copies of an r.v. X from cdf F and density f (with respect to Lebesgue measure μ defined on $\mathbb{R}_+ = (0, \infty)$). It will be assumed that all moments of F exist and F is moment-determinate (see [9] for conditions of moment determination of cdf). Let us define the operator $\mathcal{K}$ by

$$(\mathcal{K}F)(j) = \int_0^\infty t^j \, \mathrm{d}F(t) = \mu(j), \quad j = 0, 1, \ldots$$

To recover F from the moments $\mu(j)$, we will use the operators

$$(\mathcal{K}_\alpha^{-1}\mu)(x) = \sum_{k \le \alpha - 2} \frac{1}{k!} \left(\frac{\alpha}{x}\right)^k \sum_{j=k}^{\infty} \frac{1}{(j-k)!} \left(-\frac{\alpha}{x}\right)^{j-k} \mu(j), \tag{1}$$

where $\alpha \ge 2$ and $\alpha = \alpha(n) \to \infty$, as $n \to \infty$, at a rate to be specified below. Without explicit reference, it will be assumed henceforth that this parameter will be integer-valued:

$$\alpha = \alpha(n) \in \mathbb{N}, \quad \text{for each } n = 1, 2, \ldots$$

A minor modification of an argument in Mnatsakanov and Ruymgaart [2] yields

$$F_\alpha(x) = (\mathcal{K}_\alpha^{-1}\mathcal{K}F)(x) \longrightarrow_w F(x), \quad \text{as } n \longrightarrow \infty,$$

where the convergence is at each continuity point of x, i.e. for each x under the present assumptions. In passing it can be seen that

$$F_\alpha(x) = \int_0^\infty \int_{\alpha\lambda/x}^\infty \frac{1}{(\alpha-2)!} t^{\alpha-2} \exp(-t)\,\mathrm{d}t\,\mathrm{d}F(\lambda)$$

and has density

$$f_\alpha(x) = F'_\alpha(x) = \int_0^\infty \frac{\alpha-1}{x}\frac{1}{(\alpha-1)!}\left(\frac{\alpha}{x}\lambda\right)^{\alpha-1} \exp\left(-\frac{\alpha}{x}\lambda\right)\mathrm{d}F(\lambda), \quad x > 0. \tag{2}$$

Expression (2) suggests an estimator of f by replacing F by the empirical cdf

$$\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^n 1_{[0,t]}(X_i), \quad t \in \mathbb{R}_+.$$

This yields the MDE on $\mathbb{R}_+$:

$$\hat{f}_\alpha(x) = \frac{1}{n}\sum_{i=1}^n \frac{1}{x}\cdot\frac{1}{\Gamma(\alpha-1)}\left(\frac{\alpha}{x}X_i\right)^{\alpha-1}\exp\left(-\frac{\alpha}{x}X_i\right) = \int_0^\infty \frac{1}{x}L_\alpha\left(\frac{\tau}{x}\right) f(\tau)\,\mathrm{d}\hat{F}_n(\tau),$$

with $L_\alpha(u) = (\alpha u)^{\alpha-1}\exp(-\alpha u)/\Gamma(\alpha-1)$, $u \in \mathbb{R}_+$. Of course, $\hat{f}_\alpha(x) \geq 0$ for each $x > 0$ and since it is easily seen that $\int_0^\infty \hat{f}_\alpha(x)\,\mathrm{d}x = 1$, the estimator $\hat{f}_\alpha$ is itself a probability density. But it is mathematically convenient to replace this estimator with

$$\begin{aligned}\hat{f}_\alpha(x) &= \frac{1}{n}\sum_{i=1}^n \frac{\alpha}{x}\cdot\frac{1}{(\alpha-1)!}\left(\frac{\alpha}{x}X_i\right)^{\alpha-1}\exp\left(-\frac{\alpha}{x}X_i\right)\\ &= \frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}\cdot\frac{1}{(\alpha-1)!}\left(\frac{\alpha}{x}X_i\right)^{\alpha}\exp\left(-\frac{\alpha}{x}X_i\right) = \frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}K_\alpha\left(\frac{x}{X_i}\right) = \frac{1}{n}\sum_{i=1}^n M_i,\end{aligned} \tag{3}$$

where $K_\alpha(u) = (\alpha/u)^\alpha \exp(-\alpha/u)/\Gamma(\alpha)$, $u \in \mathbb{R}_+$. It is worth to note that $\hat{f}_\alpha = (\alpha-1)\hat{f}_\alpha/\alpha$ for any $\alpha > 1$.

Remark 2.1 The problem of estimation of a cdf F when only the empirical moments

$$\hat{\mu}(j) = \frac{1}{n}\sum_{i=1}^n X_i^j, \quad j \in \mathbb{N},$$

of F are available can be solved by using the construction (1) with $\hat{\mu}$ instead of μ. Some asymptotic properties of the so-called moment-empirical cdf, defined as $\hat{P}_n = \mathcal{K}_\alpha^{-1}\hat{\mu}$, with $\alpha = n$, have been studied in Mnatsakanov and Ruymgaart [2]. In the framework of Hausdorff moment problem, Mnatsakanov [15] investigated the asymptotic properties of moment-recovered density function derived directly from its assigned moments. The asymptotic properties of corresponding MDE given the estimated moments of f will be studied in the forthcoming paper. Recently, applying the expansion of f by means of the Legendre polynomials in $[-1, 1]$, Ngoc [16] obtained the MISE rate of estimate of f given the noisy observations of moments.

Remark 2.2 Note that the construction (3) is different from the traditional KDE (see, e.g. [17] or [18]). Namely, in the kernel density estimation, the convolution is considered with respect to addition as the group operation on the entire real line $\mathbb{R}$ and with a fixed kernel. Our estimator $\hat{f}_\alpha$ from Equation (3) turns out to be of kernel type with asymmetric and varying kernel that is the probability density function (pdf) of Gamma($\cdot$, shape $= \alpha$, scale $= x/\alpha$), provided that convolution is considered on the space of positive half-line $(\mathbb{R}_+, \mathrm{d}H)$ equipped with the multiplication as a group operation and with the Haar measure $\mathrm{d}H(t) = \mathrm{d}t/t$. This does not seem to be unnatural when densities on the positive half-line $\mathbb{R}_+$ are to be estimated. Note that the mean of $\hat{f}_\alpha(x)$ can be written as the convolution operator

$$f_\alpha(x) = \mathbf{E}\hat{f}_\alpha(x) = \int_0^\infty K_\alpha\left(\frac{x}{t}\right) f(t)\,\mathrm{d}H(t), \quad x \in \mathbb{R}_+.$$

In Theorem 4.2 (see Equations (20)–(22)) and in Lemma 4.1 (see Section 4), it is proved that the sequence of functions $\{K_\alpha(\cdot/t)/t, t \in \mathbb{R}_+, \alpha \in \mathbb{N}\}$ with $K_\alpha(\cdot)$ introduced in Equation (3) forms the δ-sequences in L_2- and L_1-norms defined on $L_2(\mathbb{R}_+, \mathrm{d}\mu)$ and $L_1(\mathbb{R}_+, \mathrm{d}\mu)$, respectively, where μ is the Lebesgue measure on $\mathbb{R}_+$.

3. Local properties: MSE and asymptotic normality

Let us study at first the local properties of $\hat{f}_\alpha(x)$. Without explicit reference, it will be assumed in this section that the following conditions are satisfied. Namely, we will assume that the underlying density satisfies

$$f \in C^{(2)}(\mathbb{R}_+), \quad \text{with} \sup_{t>0} |f''(t)| = M < \infty. \tag{4}$$

Throughout this section, the MDE will be considered at a fixed point $x > 0$, where

$$f(x) > 0.$$

Writing

$$M_i = M_{n,x,i} = \frac{1}{X_i} \cdot \frac{1}{(\alpha-1)!} \left(\frac{\alpha}{x} X_i\right)^\alpha \exp\left(-\frac{\alpha}{x} X_i\right) = \frac{1}{X_i} K_\alpha\left(\frac{x}{X_i}\right),$$

with $K_\alpha(u) = (\alpha/u)^\alpha \exp(-\alpha/u)/\Gamma(\alpha)$, $u \in \mathbb{R}_+$, we have $\mathbf{E}\, M_i = f_\alpha(x)$ for each n and

$$\hat{f}_\alpha(x) - f_\alpha(x) = \sum_{i=1}^n \frac{1}{n}\{M_i - f_\alpha(x)\}$$

are the averages of i.i.d. r.v.'s centred at 0.

THEOREM 3.1 *Under the assumptions* (4), *the bias of* $\hat{f}_\alpha$ *satisfies*

$$f_\alpha(x) - f(x) = \frac{x^2 f''(x)}{2\cdot\alpha} + o\left(\frac{1}{\alpha}\right), \quad \text{as } n \longrightarrow \infty. \tag{5}$$

For the MSE, we have

$$\mathrm{MSE}\{\hat{f}_\alpha(x)\} = n^{-4/5}\left[\frac{f(x)}{2\sqrt{\pi}x} + \frac{x^4\{f''(x)\}^2}{4}\right] + o(n^{-4/5}), \tag{6}$$

provided that we choose $\alpha = \alpha(n) \sim n^{2/5}$.

Remark 3.1 Since the magnitude of the bias$\{\hat{f}_\alpha(x)\} = f_\alpha(x) - f(x)$ is of the same order $O(\alpha^{-1})$ near the origin and in the interior of the support of F, the MDE $\hat{f}_\alpha$ is free from the bias effect. Note also that under the condition $\int_0^\infty \{x^2 f''(x)\}^2 \,\mathrm{d}x < \infty$ (see Equation (17) in Section 4), the expression $x^2 f''(x)$ tends towards zero as $x \to \infty$, i.e. the bias becomes smaller when x increases. Our simulations proved this phenomenon as well (Figure 1).

Proof of Theorem 3.1 Let us note that for each $k \in \mathbb{N}$ and $x \in \mathbb{R}_+$, the function

$$h_{\alpha,x,k}(u) = \frac{1}{\{k(\alpha-1)\}!}\left(\frac{k\alpha}{x}\right)^{k(\alpha-1)+1} u^{k(\alpha-1)} \exp\left(-\frac{k\alpha}{x}u\right), \quad u \geq 0, \tag{7}$$

is a gamma density with mean $\{k(\alpha-1)+1\}x/(k\alpha)$ and variance $\{k(\alpha-1)+1\}x^2/(k\alpha)^2$. For each $k \in \mathbb{N}$, moreover, these densities form as well a delta sequence as $n \to \infty$ (and consequently $\alpha \to \infty$). In addition, for $k = 1$ we have $M_i = h_{\alpha,x,1}(X_i)$ and

$$\int_0^\infty u\, h_{\alpha,x,1}(u)\,\mathrm{d}u = x,$$
$$\int_0^\infty (u-x)^2 h_{\alpha,x,1}(u)\,\mathrm{d}u = \frac{x^2}{\alpha}.$$

For each $k \in \mathbb{N}$ we have

$$\begin{aligned}
\mathbf{E}M_i^k &= \int_0^\infty \frac{1}{\{(\alpha-1)!\}^k}\left(\frac{\alpha}{x}\right)^k \left(\frac{\alpha}{x}u\right)^{k(\alpha-1)} \exp\left(-\frac{k\alpha}{x}u\right) f(u)\,\mathrm{d}u,\\
&= \int_0^\infty \frac{\{k(\alpha-1)\}!}{\{(\alpha-1)!\}^k}\left(\frac{\alpha}{x}\right)^{k\alpha}\left(\frac{x}{k\alpha}\right)^{k(\alpha-1)+1} h_{\alpha,x,k}(u) f(u)\,\mathrm{d}u,\\
&= \left(\frac{\alpha}{x}\right)^{k-1}\frac{\{k(\alpha-1)\}!}{\{(\alpha-1)!\}^k}\frac{1}{k^{k(\alpha-1)+1}}\int_0^\infty h_{\alpha,x,k}(u) f(u)\,\mathrm{d}u.
\end{aligned} \tag{8}$$

In particular, for $k = 1$

$$f_\alpha(x) = \mathbf{E}\hat{f}_\alpha(x) = \mathbf{E}M_i = \int_0^\infty h_{\alpha,x,1}(u) f(u)\,\mathrm{d}u. \tag{9}$$

This yields for the bias ($\mu = x, \sigma^2 = x^2/\alpha$)

$$\begin{aligned}
f_\alpha(x) - f(x) &= \int_0^\infty h_{\alpha,x,1}(u)\{f(u) - f(x)\}\,\mathrm{d}u,\\
&= \int_0^\infty h_{\alpha,x,1}(u)\{f(x) + (u-x)f'(x) + \frac{1}{2}(u-x)^2 f''(\tilde{u}) - f(x)\}\,\mathrm{d}u,\\
&= \frac{1}{2}\int_0^\infty (u-x)^2 h_{\alpha,x,1}(u) f''(x)\,\mathrm{d}u\\
&\quad + \frac{1}{2}\int_0^\infty (u-x)^2 h_{\alpha,x,1}(u)\{f''(\tilde{u}) - f''(x)\}\,\mathrm{d}u,\\
&= \frac{1}{2}\frac{x^2}{\alpha} f''(x) + o\left(\frac{1}{\alpha}\right), \qquad \text{as } \alpha \longrightarrow \infty.
\end{aligned}$$

For the variance we have

$$\mathrm{Var}\hat{f}_\alpha(x) = \frac{1}{n}\mathrm{Var}M_i = \frac{1}{n}\{\mathbf{E}M_i^2 - f_\alpha^2(x)\}. \tag{10}$$

Applying Equation (8) with $k = 2$ and $B_\alpha = 2^{-(2\alpha-1)}\alpha\Gamma(2\alpha-1)/[\Gamma(\alpha)]^2$ yields

$$\begin{aligned}
\mathbf{E}M_i^2 &= \frac{1}{x}\frac{\alpha\Gamma(2\alpha-1)}{[\Gamma(\alpha)]^2}\frac{1}{2^{2\alpha-1}}\int_0^\infty h_{\alpha,x,2}(u)f(u)\,\mathrm{d}u,\\
&= \frac{B_\alpha}{x}\int_0^\infty h_{\alpha,x,2}(u)f(u)\,\mathrm{d}u,\\
&\sim \frac{\alpha}{x}\frac{1}{\sqrt{2\pi}}\frac{\mathrm{e}^{-2(\alpha-1)}\{2(\alpha-1)\}^{2(\alpha-1)+1/2}}{\mathrm{e}^{-2(\alpha-1)}\{(\alpha-1)\}^{2(\alpha-1)+1}}\frac{1}{2^{2(\alpha-1)+1}}\int_0^\infty h_{\alpha,x,2}(u)f(u)\,\mathrm{d}u,\\
&= \frac{1}{\sqrt{2\pi}}\frac{\alpha}{x}\frac{1}{\sqrt{2}}\frac{1}{\sqrt{\alpha-1}}\int_0^\infty h_{\alpha,x,2}(u)f(u)\,\mathrm{d}u = \frac{\alpha}{\sqrt{\alpha-1}}\frac{1}{2\sqrt{\pi}}\frac{1}{x}\{f(x)+o(1)\},\\
&= \left\{\sqrt{\alpha}+O\left(\frac{1}{\sqrt{\alpha}}\right)\right\}\frac{1}{2\sqrt{\pi}}\frac{1}{x}\{f(x)+o(1)\},\\
&= \frac{1}{2\sqrt{\pi}}\frac{\sqrt{\alpha}}{x}f(x)+o(\sqrt{\alpha}), \qquad \text{as } \alpha \longrightarrow \infty.
\end{aligned} \tag{11}$$

Inserting Equation (11) in Equation (10), we obtain

$$\begin{aligned}
\mathrm{Var}\hat{f}_\alpha(x) &= \frac{1}{n}\left[\frac{1}{2\sqrt{\pi}}\frac{\sqrt{\alpha}}{x}f(x)+o(\sqrt{\alpha})-\left\{f(x)+O\left(\frac{1}{\alpha}\right)\right\}^2\right],\\
&= \frac{\sqrt{\alpha}}{2n\sqrt{\pi}}\frac{f(x)}{x}+o\left(\frac{\sqrt{\alpha}}{n}\right), \qquad \text{as } \alpha, n \longrightarrow \infty.
\end{aligned} \tag{12}$$

Finally, this leads to the MSE of $\hat{f}_\alpha(x)$:

$$\begin{aligned}
\mathrm{MSE}\{\hat{f}_\alpha(x)\} = \mathrm{Var}\hat{f}_\alpha(x) + \text{ bias }^2\{\hat{f}_\alpha(x)\}, &= \frac{\sqrt{\alpha}}{2n\sqrt{\pi}}\frac{f(x)}{x}+\frac{1}{4}\frac{x^4}{\alpha^2}\{f''(x)\}^2\\
&+o\left(\frac{\sqrt{\alpha}}{n}\right)+o\left(\frac{1}{\alpha^2}\right), \qquad \text{as } \alpha, n \longrightarrow \infty.
\end{aligned} \tag{13}$$

For the optimal rate, we may take

$$\alpha = \alpha(n) \sim n^{2/5}, \tag{14}$$

assuming that n is such that $\alpha(n)$ is an integer. By substituting $\alpha = n^{2/5}$ in Equation (13), we find Equation (6). ■

To derive the limiting distribution of $\hat{f}_\alpha$, let us first prove the following.

THEOREM 3.2 *Under the assumptions* (4) *and choosing* $\alpha = \alpha(n) \sim n^\delta$ *for any* $0 < \delta < 2$, *we have*

$$\frac{\hat{f}_\alpha(x) - f_\alpha(x)}{\sqrt{Var\hat{f}_\alpha(x)}} \longrightarrow_d \text{ Normal } (0, 1), \tag{15}$$

as $n \to \infty$.

Proof of Theorem 3.2 In this proof, let $0 < C < \infty$ denote a generic constant whose value may differ from line to line, but which does not depend on n. For arbitrary $k \in \mathbb{N}$, the 'c_r-inequality' entails that $\mathbf{E}|M_i - f_\alpha(x)|^k \leq C\mathbf{E}M_i^k$, in view of Equations (8) and (9). For an arbitrary $k > 2$, we thus obtain from Equations (8) and (12)

$$\frac{\sum_{i=1}^{n} \mathbf{E}|(1/n)\{M_i - f_\alpha(x)\}|^k}{\{\operatorname{Var}\hat{f}_\alpha(x)\}^{k/2}} \leq C\frac{n^{1-k}k^{-1/2}\alpha^{k/2-1/2}}{(n^{-1}\alpha^{1/2})^{k/2}}$$
$$= C\frac{1}{\sqrt{k}}\frac{\alpha^{k/4-1/2}}{n^{k/2-1}} \longrightarrow 0, \quad \text{as } n \longrightarrow \infty,$$

for $\alpha \sim n^\delta$, $0 < \delta < 2$. Apparently Lyapunov's condition for the central limit theorem is fulfilled and Equation (15) follows. ∎

THEOREM 3.3 *Under the assumptions* (4), *we have*

$$\frac{n^{1/2}}{\alpha^{1/4}}\{\hat{f}_\alpha(x) - f(x)\} \longrightarrow_d \text{ Normal } \left(0, \frac{f(x)}{2x\sqrt{\pi}}\right),$$

as $n \to \infty$, provided that we take $\alpha = \alpha(n) \sim n^\delta$ for some $\frac{2}{5} < \delta < 2$.

Proof of Theorem 3.3 This is immediate from Equations (12) and (15), and because Equation (5) entails that $n^{1/2}\alpha^{-1/4}\{f_\alpha(x) - f(x)\} = O(n^{1/2}\alpha^{-5/4}) = o(1)$, as $n \to \infty$, for the present choice of α. ∎

4. Global properties

4.1. *MISE rate of convergence*

Throughout this section again F concentrates mass 1 on $(0, \infty)$ but it is also supposed to have a sufficiently smooth density. Let us consider the following conditions:

$$\int_0^\infty \frac{f(x)}{x}\,\mathrm{d}x = C_0 < \infty, \tag{16}$$

and

$$\int_0^\infty \{x^2 f''(x)\}^2\,\mathrm{d}x = C_1 < \infty. \tag{17}$$

One can very easily obtain the optimal rate $n^{-4/5}$ for MISE $\{\hat{f}_\alpha\}$ as $\alpha, n \to \infty$ by integrating the terms on the right-hand side of Equation (13). Namely, the following statement is true.

THEOREM 4.1 *Under the assumptions Equations* (4), (16), *and* (17), *we have*

$$\text{MISE}\{\hat{f}_\alpha\} = \int_0^\infty \operatorname{Var}\hat{f}_\alpha(x)\,dx + \int_0^\infty \text{bias}^2\{\hat{f}_\alpha(x)\}\,\mathrm{d}x$$
$$\sim \frac{C_0\sqrt{\alpha}}{2n\sqrt{\pi}} + \frac{C_1}{4\alpha^2},$$

as $\alpha, n \to \infty$. While for optimal $MISE$, we have

$$\text{MISE}\{\hat{f}_\alpha\} \sim n^{-4/5}\left(\frac{C_0}{2\sqrt{\pi}}\right)^{4/5} \cdot \frac{5C_1^{1/5}}{4}, \quad \text{as } \alpha, n \longrightarrow \infty,$$

provided that we choose $\alpha = \alpha(n) = n^{2/5}(2C_1\sqrt{\pi}/C_0)^{2/5}$.

Actually we can weaken the conditions on f and obtain the exact upper bound for MISE$\{\hat{f}_\alpha\}$ and show that the corresponding asymptotic rate is $n^{-2/3}$. Indeed, let us denote again by $B_\alpha = 2^{-(2\alpha-1)}\alpha\Gamma(2\alpha-1)/[\Gamma(\alpha)]^2$ and consider the following condition (instead of Equation (17)):

$$\int_0^\infty \{xf'(x)\}^2\,\mathrm{d}x = C_2 < \infty. \tag{18}$$

THEOREM 4.2 *Under the assumptions* (16) *and* (18)*, we have*

$$\begin{aligned} \mathrm{MISE}\{\hat{f}_\alpha\} &\le \frac{B_\alpha C_0}{n} + \frac{C_2\alpha}{(\alpha-1)(\alpha+1)}, \quad \alpha > 1, \\ \mathrm{MISE}\{\hat{f}_\alpha\} &\le \frac{C_0\sqrt{\alpha}}{2n\sqrt{\pi}} + \frac{C_2}{\alpha} + o\left(\frac{\sqrt{\alpha}}{n}\right) + o\left(\frac{1}{\alpha}\right), \end{aligned} \tag{19}$$

as $\alpha, n \to \infty$. *While for optimal MISE, we have*

$$\mathrm{MISE}\{\hat{f}_\alpha\} \le n^{-2/3}\left(\frac{C_0}{2\sqrt{\pi}}\right)^{2/3} \cdot \frac{3C_2^{1/3}}{2^{2/3}} + o(n^{-2/3}), \quad \text{as } n \longrightarrow \infty,$$

provided that we choose $\alpha = \alpha(n) = n^{2/3}(4\,C_2\sqrt{\pi}/C_0)^{2/3}$.

Proof of Theorem 4.2 Let us denote by $\bar{\xi}_\alpha$ the sample mean of i.i.d. r.v.'s $\xi_1, \ldots, \xi_\alpha$ with distribution Exp(1). Then after a simple algebra combined with application of the Cauchy–Schwarz's inequality, we obtain

$$\begin{aligned} \int_0^\infty \mathrm{bias}^2\{\hat{f}_\alpha(x)\}\,\mathrm{d}x &= \int_0^\infty \left[\int_0^\infty \{f(s)-f(x)\}h_{\alpha,x,1}(s)\,\mathrm{d}s\right]^2 \mathrm{d}x \\ &= \int_0^\infty \left[\mathbf{E}(f(x\bar{\xi}_\alpha) - f(x))\right]^2 \mathrm{d}x = \int_0^\infty \left[\mathbf{E}\int_x^{x\bar{\xi}_\alpha} f'(s)\,\mathrm{d}s\right]^2 \mathrm{d}x \\ &\le \mathbf{E}\int_0^\infty \left[\int_x^{x\bar{\xi}_\alpha} (f'(s))^2\,\mathrm{d}s x(\bar{\xi}_\alpha - 1)\right] \mathrm{d}x \\ &= \mathbf{E}\left\{ I_{[\bar{\xi}_\alpha \le 1]} \int_0^\infty \left[(f'(s))^2 \int_s^{s/\bar{\xi}_\alpha} x(1-\bar{\xi}_\alpha)\,\mathrm{d}x\right] \mathrm{d}s \right. \\ &\quad \left. + I_{[\bar{\xi}_\alpha \ge 1]} \int_0^\infty \left[(f'(s))^2 \int_{s/\bar{\xi}_\alpha}^{s} x(\bar{\xi}_\alpha - 1)\,\mathrm{d}x\right] \mathrm{d}s \right\} \\ &= \mathbf{E}\left\{ I_{[\bar{\xi}_\alpha \le 1]} \int_0^\infty (f'(s))^2 \frac{1}{2}s^2 \left(\frac{1}{\bar{\xi}_\alpha^2} - 1\right)(1-\bar{\xi}_\alpha)\,\mathrm{d}s \right. \\ &\quad \left. + I_{[\bar{\xi}_\alpha \ge 1]} \int_0^\infty (f'(s))^2 \frac{1}{2}s^2 \left(1 - \frac{1}{\bar{\xi}_\alpha^2}\right)(\bar{\xi}_\alpha - 1)\,\mathrm{d}s \right\} \\ &= \frac{1}{2}\mathbf{E}\left(\frac{(\bar{\xi}_\alpha - 1)^2(\bar{\xi}_\alpha + 1)}{\bar{\xi}_\alpha^2}\right) \int_0^\infty \{sf'(s)\}^2 \mathrm{d}s. \end{aligned} \tag{20}$$

But

$$\mathbf{E}\left(\frac{(\bar{\xi}_\alpha - 1)^2(\bar{\xi}_\alpha + 1)}{\bar{\xi}_\alpha^2}\right)$$
$$= \int_0^\infty \frac{(u/\alpha - 1)^2(u/\alpha + 1)}{(u/\alpha)^2} \cdot \frac{u^{\alpha-1}}{\Gamma(\alpha)} \mathrm{e}^{-u}\,\mathrm{d}u = \frac{1}{\alpha\Gamma(\alpha)}\int_0^\infty (u-\alpha)^2(u+\alpha)u^{\alpha-3}\mathrm{e}^{-u}\,\mathrm{d}u$$
$$= \frac{2\alpha}{(\alpha-1)(\alpha+1)} \sim \frac{2}{\alpha}, \quad \text{as } \alpha \longrightarrow \infty. \tag{21}$$

Combination of Equations (20) and (21) gives

$$\int_0^\infty \mathrm{bias}^2\{\hat{f}_\alpha(x)\}\,\mathrm{d}x \le \frac{\alpha}{(\alpha-1)(\alpha+1)}\int_0^\infty s^2(f'(s))^2\,\mathrm{d}s$$
$$= \frac{1}{\alpha}\int_0^\infty s^2\{f'(s)\}^2\mathrm{d}s(1+o(1)), \quad \text{as } \alpha \longrightarrow \infty. \tag{22}$$

Let the following

$$g(x, a_1, b_1) = \frac{1}{x^2} \cdot \frac{b_1^{a_1}(1/x)^{a_1-1}\mathrm{e}^{-b_1/x}}{\Gamma(a_1)}, \quad t > 0,$$

denote the inverse gamma density with the shape a_1 and the rate b_1. According to the definitions of the inverse gamma $g(\cdot, a_1, b_1)$ and gamma $h_{\alpha,x,2}$ density (Equation (7)), we have

$$\int_0^\infty \frac{1}{x}h_{\alpha,x,2}(u)\,\mathrm{d}x = \frac{1}{u}\int_0^\infty g(x, a_1 = 2\alpha - 1, b_1 = 2\alpha u)\,\mathrm{d}x = \frac{1}{u}.$$

So that integration of the both sides of the first equation in Equation (11) combined with $B_\alpha \sim \alpha^{1/2}/2\sqrt{\pi}$, as $\alpha \to \infty$, yields

$$\frac{1}{n}\int_0^\infty \mathbf{E}M_{n,x,i}^2\,\mathrm{d}x = \frac{B_\alpha}{n}\int_0^\infty \frac{1}{x}\left[\int_0^\infty h_{\alpha,x,2}(u)\,f(u)\,\mathrm{d}u\right]\mathrm{d}x$$
$$= \frac{B_\alpha}{n}\int_0^\infty f(u)\left[\int_0^\infty \frac{1}{x}h_{\alpha,x,2}(u)\,\mathrm{d}x\right]\mathrm{d}u = \frac{B_\alpha}{n}\int_0^\infty \frac{f(u)}{u}\,\mathrm{d}u$$
$$\sim \frac{\sqrt{\alpha}}{2n\sqrt{\pi}}\int_0^\infty \frac{f(u)}{u}\,\mathrm{d}u, \quad \text{as } \alpha \to \infty. \tag{23}$$

Hence, it is proved

$$\int_0^\infty \mathrm{Var}\hat{f}_\alpha(x)\,\mathrm{d}x \le \int_0^\infty \frac{1}{n}\{\mathbf{E}M_{n,x,i}^2\}\mathrm{d}x$$
$$\sim \frac{\sqrt{\alpha}}{2n\sqrt{\pi}}\int_0^\infty \frac{f(u)}{u}\,\mathrm{d}u + o\left(\frac{\sqrt{\alpha}}{n}\right), \quad \text{as } n, \alpha \longrightarrow \infty. \tag{24}$$

Finally, from Equations (22) and (24), we obtain the statement (19) of the Theorem 4.2. ■

4.2. L_1-consistency

Theorem 4.3 establishes the L_1-consistency and corresponding optimal rate of MDE $\hat{f}_\alpha$ under the mild conditions on f. In Lemma 4.1, we provide the exact L_1-rate for $f_\alpha = \mathbf{E}\hat{f}_\alpha$ under the smoothed conditions on f:

$$\text{the second derivative } f'' \text{ exists and } \int_0^\infty x^2|f''(x)|\,\mathrm{d}x = C_3 < \infty. \tag{25}$$

Denote the L_1-distance on $L_1(\mathbb{R}_+, \mathrm{d}\mu)$ (μ is the Lebesgue measure on $\mathbb{R}_+$) as follows:

$$\| f_\alpha - f \|_L = \int_0^\infty |f_\alpha(x) - f(x)|\,\mathrm{d}\mu(x).$$

Here

$$f_\alpha(x) = \mathbf{E}\hat{f}_\alpha(x) = \int_0^\infty K_\alpha\left(\frac{x}{t}\right)\frac{1}{t} f(t)\,\mathrm{d}t = \mathbf{E}[f(x\bar{\xi}_\alpha)], \quad x \in \mathbb{R}_+ \tag{26}$$

with $K_\alpha(u) = (\alpha/u)^\alpha \exp(-\alpha/u)/\Gamma(\alpha)$, $u \in \mathbb{R}_+$. Recall that the r.v. $\bar{\xi}_\alpha$ introduced in the proof of Theorem 4.2 has pdf $K_\alpha^*(u) = (1/u)\,K_\alpha(1/u)$, $u > 0$, with

$$\mathbf{E}(\bar{\xi}_\alpha) = 1 \quad \text{and} \quad \mathrm{Var}(\bar{\xi}_\alpha) = \frac{1}{\alpha}.$$

Now let us show that the functions $\{(1/t)K_\alpha(\cdot/t), t > 0\}$ form a δ-sequence in L_1- norm and establish corresponding rate as $\alpha \to \infty$. Namely, the following statement is true.

LEMMA 4.1 *Under the assumptions* (25), *we have*

$$\| f_\alpha - f \|_L \le C_3 \frac{1}{2}\mathbf{E}\left(\frac{(\bar{\xi}_\alpha - 1)^2(\bar{\xi}_\alpha + 1)}{\bar{\xi}_\alpha^2}\right) = C_3 \cdot \frac{\alpha}{(\alpha-1)(\alpha+1)}, \quad \alpha > 1. \tag{27}$$

Proof of Lemma 4.1 Combination of Equation (26) and the equations

$$\int_0^\infty K_\alpha(x/s)\frac{1}{s}\,\mathrm{d}s = 1, \quad \mathbf{E}[x(\bar{\xi}_\alpha - 1)] = 0,$$

$$f(x\bar{\xi}_\alpha) - f(x) = f'(x)(x\bar{\xi}_\alpha - x) + \int_x^{x\bar{\xi}_\alpha} \mathrm{d}s \int_x^s f''(y)\,\mathrm{d}y$$

gives

$$\begin{aligned}
\| f_\alpha - f \|_L &= \int_0^\infty \left| \int_0^\infty \{f(s) - f(x)\} K_\alpha(x/s)\frac{1}{s}\,\mathrm{d}s \right| \mathrm{d}x = \int_0^\infty \left|\mathbf{E}(f(x\bar{\xi}_\alpha) - f(x))\right| \mathrm{d}x \\
&= \int_0^\infty \left| \mathbf{E}\int_x^{x\bar{\xi}_\alpha} \mathrm{d}s \int_x^s f''(y)\,\mathrm{d}y \right| \mathrm{d}x \le \int_0^\infty \Bigg[\mathbf{E} I_{[\bar{\xi}_\alpha > 1]}(x\bar{\xi}_\alpha - x) \int_x^{x\bar{\xi}_\alpha} |f''(y)|\,\mathrm{d}y \\
&\quad + \mathbf{E} I_{[\bar{\xi}_\alpha > 1]}(x - x\bar{\xi}_\alpha) \int_{x\bar{\xi}_\alpha}^x |f''(y)|\,\mathrm{d}y \Bigg] \mathrm{d}x.
\end{aligned} \tag{28}$$

Now in a similar way as we did in Equation (20), changing the integrations in Equation (28) yields

$$\| f_\alpha - f \|_L \le \frac{1}{2}\int_0^\infty y^2|f''(y)|\,\mathrm{d}y \cdot \mathbf{E}\left(\frac{(\bar{\xi}_\alpha - 1)^2(\bar{\xi}_\alpha + 1)}{\bar{\xi}_\alpha^2}\right). \tag{29}$$

Finally, from Equations (21) and (29), we obtain Equation (27). Lemma 4.1 is proved. ■

To prove L_1-consistency of MDE $\hat{f}_\alpha$ in statement (i), one can very easily apply the result from Mnatsakanov and Khmaladze [19] where the necessary and sufficient conditions for L_1-consistency of general density estimators are obtained. Besides, under the smooth conditions on f, we obtain the optimal L_1-rate of $O(n^{-2/5})$ in (ii).

THEOREM 4.3 (i) *If f is a bounded and continuous function and the condition* (16) *is satisfied, then*

$$\mathbf{E}\|\hat{f}_\alpha - f\|_L \longrightarrow 0, \tag{30}$$

as $\sqrt{\alpha}/n \to 0, \alpha, n \to \infty$.

(ii) *If the condition* (25) *is satisfied and* $f(x) \le C\min(x^\delta, x^\varepsilon)$, *for some* $\delta > 0, \varepsilon > 1$, *and some* $0 < C < \infty$, *then*

$$\mathbf{E}\|\hat{f}_\alpha - f\|_L \le C\sqrt{\frac{B_\alpha}{n}} + \frac{C_3\alpha}{(\alpha-1)(\alpha+1)}. \tag{31}$$

While for optimal L_1-error, we have

$$\mathbf{E}\|\hat{f}_\alpha - f\|_L \sim n^{-2/5}\frac{5C^{4/5}}{4}\left(\frac{C_3}{\pi}\right)^{1/5}, \quad \text{as } n \longrightarrow \infty,$$

provided that we choose $\alpha = \alpha(n) = n^{2/5}4\pi^{1/5}(C_3/C)^{4/5}$.

Proof of (*i*) Recall that $\hat{f}_\alpha = (\alpha-1)\hat{f}_\alpha/\alpha$, where $\hat{f}_\alpha$ is defined in Section 2 with

$$\hat{f}_\alpha(x) = \int_0^\infty \frac{1}{x}L_\alpha(\tau/x)\,\mathrm{d}\hat{F}_n(\tau).$$

Here $L_\alpha(u) = (\alpha u)^{\alpha-1}\exp(-\alpha u)/\Gamma(\alpha-1)$, $u \in \mathbb{R}_+$. Also note that $\mathbf{E}\hat{f}_\alpha(x)$ represent a PDF on $\mathbb{R}_+$. Hence, combining Lemma 1 from Feller [8, v. II, Ch. 7, p. 219], and the Scheffe's theorem (see Theorem 7 in [20, Ch. 2]), we derive $\|\mathbf{E}\hat{f}_\alpha - f\|_L \to 0$ and

$$\|\mathbf{E}\hat{f}_\alpha - f\|_L \le \left\|\frac{\alpha-1}{\alpha}\mathbf{E}\hat{f}_\alpha - \mathbf{E}\hat{f}_\alpha\right\|_L + \|\mathbf{E}\hat{f}_\alpha - f\|_L = \frac{1}{\alpha} + \|\mathbf{E}\hat{f}_\alpha - f\|_L \longrightarrow 0.$$

Now to prove Equation (30), it is sufficient to show

$$F\{A_n(\delta)\} = F\left\{x : \int_0^\infty \frac{1}{x^2}L_\alpha^2(\tau/x)f(\tau)\,\mathrm{d}\tau \ge n\delta\right\} \longrightarrow 0,$$

for all δ and $\alpha, n \to \infty$ (see Theorem 1 in [19]). But F is an absolutely continuous distribution with respect to Lebesgue measure μ; so, let us establish

$$\mu\{A_n(\delta)\} \longrightarrow 0,$$

for all δ and $\alpha, n \to \infty$. Indeed, application of the steps similar to Equation (23) with $\tilde{B}_\alpha = \alpha^{-1}2^{-(2\alpha-1)}\Gamma(2\alpha-1)/[\Gamma(\alpha-1)]^2 \sim \alpha^{1/2}/2\sqrt{\pi}$ instead of B_α yields

$$\begin{aligned}
\mu\{A_n(\delta)\} &\le \frac{1}{n\delta}\int_{A_n(\delta)}\mathrm{d}x\int_0^\infty \frac{1}{x^2}L_\alpha^2(\tau/x)f(\tau)\,\mathrm{d}\tau \\
&\le \frac{1}{n\delta}\int_0^\infty \mathrm{d}x\int_0^\infty \frac{1}{x^2}L_\alpha^2(\tau/x)f(\tau)\,\mathrm{d}\tau = \frac{\tilde{B}_\alpha}{n\delta}\int_0^\infty \frac{f(\tau)}{\tau}\,\mathrm{d}\tau \\
&\sim \frac{\sqrt{\alpha}}{2n\delta\sqrt{\pi}}\int_0^\infty \frac{f(\tau)}{\tau}\,\mathrm{d}\tau, \quad \text{as } \alpha \longrightarrow \infty.
\end{aligned} \tag{32}$$

Hence, the proof of (i) follows from Equations (16) and (32) and $\sqrt{\alpha}/n \to 0$. ∎

Proof of (ii) Throughout, C will be used as a generic constant. Let us introduce a r.v. η_α with Gamma$(2\alpha - 1, 2\alpha)$ distribution. From Equations (10) and (11), we have

$$\operatorname{Var} \hat{f}_\alpha(x) \leq \frac{B_\alpha}{nx} \mathbf{E} f(x\eta_\alpha). \tag{33}$$

Next let us observe that

$$\begin{aligned}
\int_0^\infty \sqrt{\operatorname{Var} \hat{f}_\alpha(x)}\, \mathrm{d}x &\leq \sum_{k=0}^{\infty} \int_k^{k+1} \sqrt{\frac{B_\alpha}{n\,x} \mathbf{E} f(x\eta_\alpha)}\, \mathrm{d}x \\
&\leq \sum_{k=0}^{\infty} \sqrt{\int_k^{k+1} \frac{B_\alpha}{nx} \mathbf{E} f(x\eta_\alpha)\, \mathrm{d}x} \leq C\sqrt{\frac{B_\alpha}{n}} \sqrt{\int_0^1 \frac{1}{x} x^\delta \mathbf{E}(\eta_\alpha^\delta)\, \mathrm{d}x} \\
&+ C\sqrt{\frac{B_\alpha}{n}} \sum_{k=1}^{\infty} \sqrt{\int_k^{k+1} \frac{1}{x} x^{-\varepsilon} \mathbf{E}(\eta_\alpha^{-\varepsilon})\, \mathrm{d}x} \\
&\leq C\sqrt{\frac{B_\alpha}{n}} \left\{ \sqrt{\mathbf{E}(\eta_\alpha^\delta)} + \sqrt{\mathbf{E}(\eta_\alpha^{-\varepsilon})} \cdot \sum_{k=1}^{\infty} \sqrt{\int_k^{k+1} \frac{1}{x^{1+\varepsilon}}\, \mathrm{d}x} \right\} \leq C\sqrt{\frac{B_\alpha}{n}},
\end{aligned} \tag{34}$$

assuming that $\mathbf{E}(\eta_\alpha^r) \approx 1$ for $r \in \mathbb{R}$, if α is sufficiently large. Finally, note that

$$\begin{aligned}
\mathbf{E}\|\hat{f}_\alpha - f\|_L &\leq \mathbf{E} \int_0^\infty |\hat{f}_\alpha(x) - \mathbf{E}\hat{f}_\alpha(x)|\, \mathrm{d}x + \|\mathbf{E}\hat{f}_\alpha - f\|_L \\
&\leq \left[\mathbf{E} \left\{ \int_0^\infty |\hat{f}_\alpha(x) - f_\alpha(x)|\, \mathrm{d}x \right\}^2 \right]^{1/2} + \|f_\alpha - f\|_L \\
&= \left[\int_0^\infty \int_0^\infty \mathbf{E}|\hat{f}_\alpha(x) - f_\alpha(y)||\hat{f}_\alpha(y) - f_\alpha(x)|\, \mathrm{d}x\, \mathrm{d}y \right]^{1/2} + \|f_\alpha - f\|_L \\
&\leq \left[\int_0^\infty \sqrt{\operatorname{Var} \hat{f}_\alpha(x)}\, \mathrm{d}x \cdot \int_0^\infty \sqrt{\operatorname{Var} \hat{f}_\alpha(y)}\, \mathrm{d}y \right]^{1/2} + \|f_\alpha - f\|_L \\
&\leq C\sqrt{\frac{B_\alpha}{n}} + \frac{C_3 \alpha}{(\alpha - 1)(\alpha + 1)}.
\end{aligned} \tag{35}$$

The last inequality of Equation (35) is derived according to Equations (34) and (27). ■

Remark 4.1 Taking $\alpha = h^{-2}$, one can see that the condition $\sqrt{\alpha}/n \to 0$ from Theorem 4.3 corresponds to the condition $nh \to \infty$ in traditional kernel density estimation.

5. Comparison with KDE

Now let us compare the expressions of the main terms of MSE for MDE $\hat{f}_\alpha(x)$ and the KDE $\hat{f}_h(x)$ studied, say, in Silverman [18]. Under the same assumptions as stated in Equation (4), we have for the traditional kernel estimator

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} r\left(\frac{x - X_i}{h}\right), \quad x \in \mathbb{R},$$

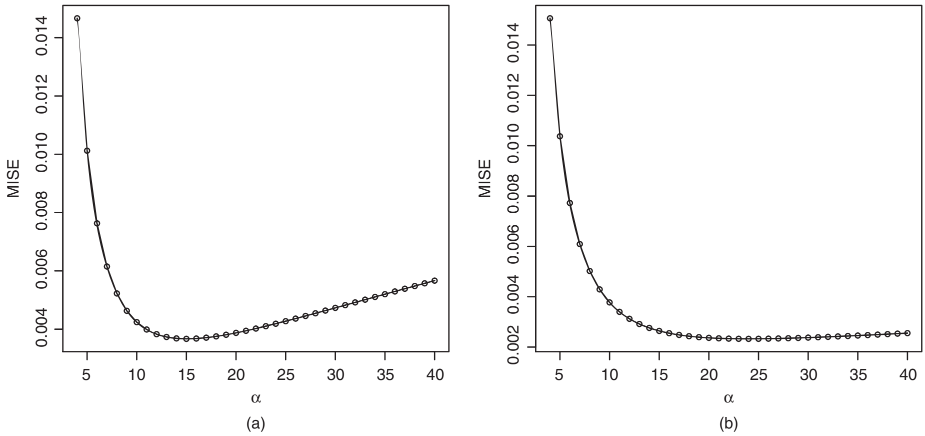

Figure 1. The simulated $\widehat{\mathrm{MISE}}$ for Gamma(2, 1) with sample size (a) $n = 200$ and $N = 50$ and (b) $n = 600$ and $N = 50$.

with kernel r and bandwidth h:

$$\mathrm{MSE}\{\hat{f}_h(x)\} = \frac{1}{nh} f(x) \int_{-\infty}^{\infty} r^2(y)\,\mathrm{d}y + o\left(\frac{1}{nh}\right) + \frac{1}{4}h^4 \left\{ f''(x) \int_{-\infty}^{\infty} y^2 r(y)\,\mathrm{d}y \right\}^2 + o(h^4). \tag{36}$$

For optimal rate, we may take

$$h = h(n) \sim n^{-1/5}, \tag{37}$$

and by substitution, we now obtain

$$\mathrm{MSE}\{\hat{f}_h(x)\} = n^{-4/5} \left[\left\{ \int_{-\infty}^{\infty} r^2(y)\,\mathrm{d}y \right\} f(x) + \frac{1}{4} \left\{ \int_{-\infty}^{\infty} y^2 r(y)\,\mathrm{d}y \right\}^2 \{f''(x)\}^2 \right] + o(n^{-4/5}).$$

The differences in the constants for asymptotics of α defined in Equation (14) and h from Equation (37) would not change the rate of MSE as a function of x.

In a similar way, one can derive the optimal bandwidth h when minimizing the MISE$\{\hat{f}_h\}$. In this case, we have

$$h_{\mathrm{opt}} = \left\{ \int_{-\infty}^{\infty} y^2 r(y)\,\mathrm{d}y \right\}^{-2/5} \left\{ \int_{-\infty}^{\infty} r^2(y)\,\mathrm{d}y \right\}^{1/5} \left\{ \int_{-\infty}^{\infty} [f''(x)]^2\,\mathrm{d}x \right\}^{-1/5} n^{-1/5}.$$

In our simulation study, we assumed r to be a standard normal density function. We also took the standard normal bandwidth $h_{\mathrm{opt}} = 1.06\hat{\sigma}n^{-1/5}$, where $\hat{\sigma}$ is the sample standard deviation of the sampled data (cf. [18, Ch. 3.4]).

Now to chose the optimal $\alpha = \alpha^*$, let us use the least-squares CV algorithm [18]. We simulated the r.v.'s X_i, $i = 1, \ldots, n$, with $n = 200k (1 \leq k \leq 4)$ sample sizes from three different distributions: Gamma(2, 1), Exp(1), and Log-Normal(0, 1). Also we repeated these simulations $N = 50$ times and studied the performance of MDE $\hat{f}_\alpha$ defined in Equation (3) via MISE for different sample sizes. Let us measure the accuracy of MDE $\hat{f}_\alpha$ in terms of the estimated MISE:

$$\widehat{MISE} := \hat{E}(\mathrm{ISE})\{\hat{f}_\alpha\} = \frac{1}{N} \sum_{j=1}^{N} \int_0^{\infty} |\hat{f}_{\alpha,j}(x) - f(x)|^2\,\mathrm{d}x. \tag{38}$$

Here, the expectation $\hat{E}$ is calculated with respect to the empirical cdf of $N = 50$ values of ISE's, while $\hat{f}_{\alpha,j}$ denotes the MDE derived on the jth replication. In other words, the optimal α^*

minimizes the expression $M_n(\hat{f}_\alpha)$, i.e.

$$\alpha^* = \text{argmin}_\alpha M_n(\hat{f}_\alpha) = \text{argmin}_\alpha \left[\int_0^\infty [\hat{f}_\alpha(x)]^2 \, dx - 2 \int_0^\infty \hat{f}_\alpha(x) \, d\hat{F}_n(x) \right], \qquad (39)$$

where $\alpha \in \{4, \dots, 40\}$ for each $n = 200k$ with $1 \le k \le 4$. In the second term of the right-hand side of Equation (39), let us apply the leave-one-out construction instead of $\hat{f}_\alpha$. This yields the following expression of

$$M_n(\hat{f}_\alpha) = \frac{\alpha \Gamma(2\alpha - 1)}{n^2 \Gamma^2(\alpha)} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{(X_i X_j)^{\alpha-1}}{(X_i + X_j)^{2\alpha-1}} - \frac{2}{n(n-1)\Gamma(\alpha)} \sum_{i=1}^{n} \sum_{j \ne i} \frac{1}{X_j} \left(\frac{\alpha X_j}{X_i} \right)^{\alpha} e^{-(\alpha X_j)/X_i}.$$

In Table 1, we recorded the values of α^* and corresponding $\widehat{\text{MISE}}$ for three different distributions mentioned above when the sample sizes are $n = 200k$, $1 \le k \le 4$, and the number of replications $N = 50$. The simulations study justifies that $\widehat{\text{MISE}}$ is a decreasing function of n when $\alpha = \alpha^*$. See also Figure 1(a)–(b) where we plotted two curves of $\widehat{\text{MISE}}$ for values of $\alpha \in \{4, \dots, 40\}$ when the sampled distribution is Gamma(2, 1) with the sample sizes $n = 200$ and $n = 600$, respectively.

Table 1. The optimal α^* and the simulated $\widehat{\text{MISE}}$.

	Optimal α^* and $\widehat{\text{MISE}}$			
Model	$n = 200$	$n = 400$	$n = 600$	$n = 800$
Gamma(2, 1)	15	20	24	26
	0.00367	0.00241	0.00233	0.00148
Exp(1)	5	7	6	9
	0.04965	0.02824	0.02071	0.01416
Log-Normal(0, 1)	12	15	18	20
	0.00331	0.00163	0.00150	0.00478

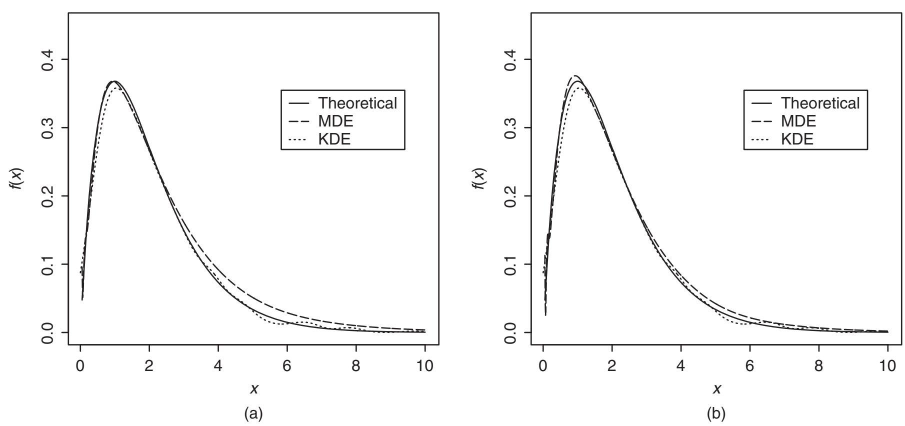

Figure 2. Estimation of Gamma(2, 1) density function f by (a) $\hat{f}_\alpha$ and $\hat{f}_h$ with $\alpha = n^{2/5}$ and bandwidth $h = h_{\text{opt}}$ and (b) $\hat{f}_{\alpha^*}$ and $\hat{f}_h$ with $\alpha^* = 24$. In both plots $n = 600$.

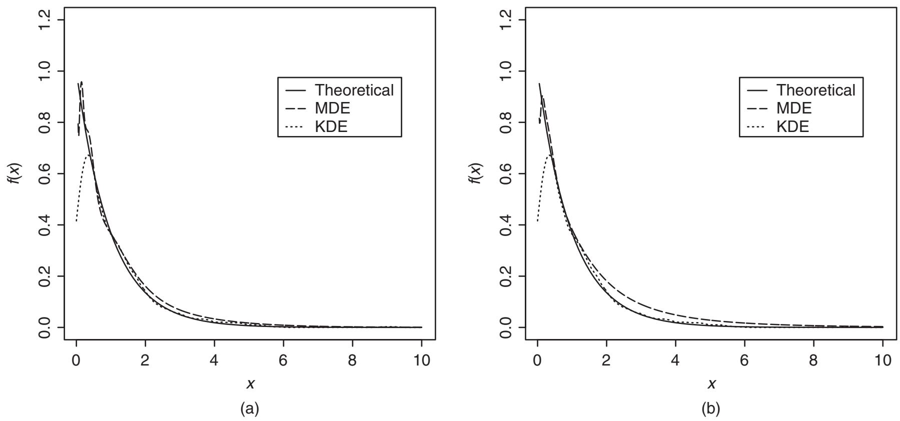

Figure 3. Estimation of Exp(1) density function f by (a) $\hat{f}_\alpha$ and $\hat{f}_h$ with $\alpha = n^{2/5}$ and bandwidth $h = h_{\text{opt}}$ and (b) $\hat{f}_{\alpha^*}$ and $\hat{f}_h$ with $\alpha^* = 6$. In both plots $n = 600$.

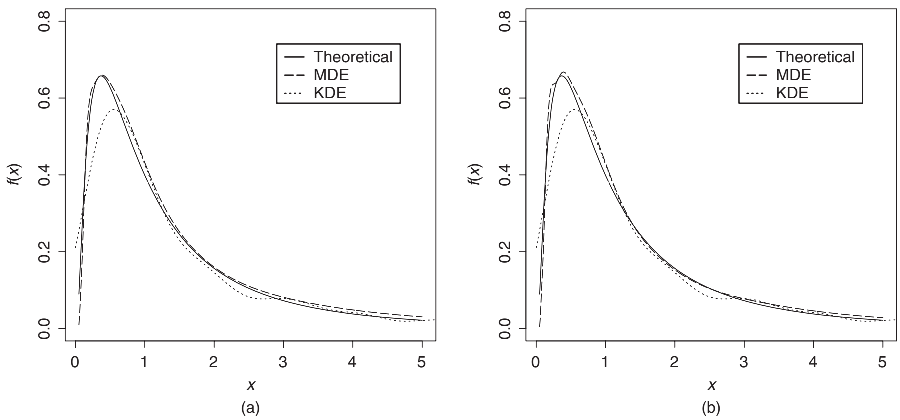

Figure 4. Estimation of Log-Normal(0, 1) density function f by (a) $\hat{f}_\alpha$ and $\hat{f}_h$ with $\alpha = n^{2/5}$ and bandwidth $h = h_{\text{opt}}$ and (b) $\hat{f}_{\alpha^*}$ and $\hat{f}_h$ with $\alpha^* = 18$. In both plots $n = 600$.

Finally, let us simulate the data $X_i, i = 1, \ldots, n$, from Gamma(2, 1), Exp(1), and Log-Normal(0, 1) distributions when $n = 600$. The graphs of estimators $\hat{f}_\alpha$ (the dashed curves) with $\alpha = n^{2/5}$ and $\hat{f}_h$ (the dotted curves) with $h = h_{\text{opt}}$, respectively, are plotted in Figures 2(a)–4(a), while in Figures 2(b)–4(b), we plotted the graphs of estimators $\hat{f}_\alpha$ and $\hat{f}_h$ for $\alpha = \alpha^*$ and $h = h_{\text{opt}}$, respectively. The corresponding sampled pdfs f (the solid curves) are plotted as well. Based on graphical illustrations of the MDE and KDE estimators for three underlying distributions, we conclude that the performance of $\hat{f}_\alpha$ is improved (except for Exp(1)) when $\alpha = \alpha^*$ is used instead of $\alpha = n^{-2/5}$.

Acknowledgements

We thank Estate V. Khmaladze and Chris A.J. Klaassen for helpful discussions which lead to improvement of the constants in the MISE rate of convergence. The authors are also thankful to S. Li for conducting the simulations presented in Section 5. The research of Robert M. Mnatsakanov associated with the revision of manuscript was supported by the NSF grant DMS-0906639. The work of Frits H. Ruymgaart was partially supported by NSF grant DMS-0203942.

References

[1] A. Goldenshluger and V. Spokoiny, *On the shape-from-moments problem and recovering edges from noisy Radon data*, Probab. Theory Relat. Fields 128 (2004), pp. 123–140.

[2] R.M. Mnatsakanov and F.H. Ruymgaart, *Some properties of moment-empirical cdf's with application to some inverse estimation problems*, Math. Methods Stat. 12 (2003), pp. 478–495.

[3] R.M. Mnatsakanov and F.H. Ruymgaart, *Some results for moment-empirical cumulative distribution functions*, J. Nonparametr. Stat. 17 (2005), pp. 733–744.

[4] R.M. Mnatsakanov and F.H. Ruymgaart, *On Moment-Density Estimation in Some Biased Models*, 2nd Lehmann Symposium – Optimality Vol. 49, IMS Lecture Notes-Monograph Series, Institute of Mathematical Statistics, Beachwood, OH, 2006, pp. 322–333.

[5] B.G. Lindsay, R.S. Pilla, and P. Basak, *Moment-based approximations of distributions using mixtures: Theory and applications*, Ann. Inst. Statist. Math. 52 (2000), pp. 215–230.

[6] N.I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Oliver and Boyd, Edinburgh, 1965.

[7] J.A. Shohat and J.D. Tamarkin, *The Problem of Moments*, American Mathematical Society, New York, 1943.

[8] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, Wiley, New York, 1971.

[9] J. Stoyanov, *Krein condition in probabilistic moment problems*, Bernoulli 6 (2000), pp. 939–949.

[10] R.M. Mnatsakanov and C.A.J. Klaassen, *Estimation of the mixing distribution in the Poisson mixture models. Uncensored and censored samples*, Proceedings of Hawaii International Conference on Statistics and Related Fields, Honolulu, Hawaii, June 4–8, 2003, pp. 1–18, CD: ISSN# 1539-7211. Available at http://www.hicstatistics.org/2003StatsProceedings/

[11] C.A.J. Klaassen and R.M. Mnatsakanov, *Consistent estimation of the structural distribution function*, Scand. J. Stat. 27 (2000), pp. 773–746.

[12] S.X. Chen, *A beta kernel estimation for density functions*, Comput. Stat. Data Anal. 31 (1999), pp. 131–145.

[13] S.X. Chen, *Probability density function estimation using gamma kernels*, Ann. Inst. Statist. Math. 52 (2000), pp. 471–480.

[14] T. Bouezmarni and J.-M. Rolin, *Consistency of the beta kernel density function estimator*, Canad. J. Stat. 31 (2003), pp. 89–98.

[15] R.M. Mnatsakanov, *Hausdorff moment problem: Reconstruction of density functions*, Stat. Probab. Lett. 78 (2008), pp. 1868–1877.

[16] T.M.P. Ngoc, *A statistical minimax approach to the Hausdorff moment problem*, Inverse Probl. 24 (2008), pp. 1–13.

[17] E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Stat. 33 (1962), pp. 1065–1076.

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

[19] R.M. Mnatsakanov and E.V. Khmaladze, *On L_1-convergence of statistical kernel estimators of distribution densities*, Soviet Math. Dokl. 23 (1981), pp. 633–636.

[20] L. Devroye and L. Györfi, *Nonparametric Density Estimation*, John Wiley and Sons, New York, 1985.