Statistics and Probability Letters 145 (2019) 158–165

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Another bias correction for asymmetric kernel density estimation with a parametric start

Masayuki Hirukawa [a,*], Mari Sakudo [b,c,d,1,2,3]

[a] *Ryukoku University, Japan*
[b] *Development Bank of Japan, Japan*
[c] *Waseda University, Japan*
[d] *Japan Economic Research Institute, Japan*

ARTICLE INFO

Article history:
Received 13 April 2018
Received in revised form 23 July 2018
Accepted 2 September 2018
Available online 24 September 2018

MSC:
62G7
62G20

Keywords:
Asymmetric kernel
Bias reduction
Boundary effect
Semiparametric density estimation
Smoothing

ABSTRACT

This paper studies yet another semiparametric bias-corrected density estimation using asymmetric kernels. The estimator can be obtained by making a multiplicative bias correction for the initial parametric model twice, and it is shown to establish rate improvement when best implemented.

1. Introduction

This paper aims at establishing rate improvement via yet another multiplicative bias correction ("MBC") method for univariate probability density estimation using asymmetric kernels. For a random variable $X \in \mathbb{R}$ drawn from a distribution with an unknown density f, the MBC method considered throughout is based on the identity

$$f(x) \equiv g(x) \left\{ \frac{f(x)}{g(x)} \right\} := g(x)\, r(x), \tag{1}$$

where $g(x)$ is an initial density estimator and $r(x)$ serves as a correction factor.

In what follows, the support of X is assumed to have a boundary; to be more specific, supp (X) is assumed to be either $[0, 1]$ or $\mathbb{R}_+$. This type of data can be frequently observed in economics and finance. For example, recovery rates take values between 0 and 1, whereas variables such as wages, incomes and insurance claims (or financial losses) are by construction nonnegative. A convenient way of avoiding boundary bias due to kernel smoothing for such data is to employ asymmetric

* Correspondence to: Faculty of Economics, Ryukoku University, 67 Tsukamoto-cho, Fukakusa, Fushimi-ku, Kyoto 612-8577, Japan.
E-mail addresses: hirukawa@econ.ryukoku.ac.jp (M. Hirukawa), marisakudo@gmail.com (M. Sakudo).
1 Research Institute of Capital Formation, Development Bank of Japan, 9-7, Otemachi 1-chome, Chiyoda-ku, Tokyo 100-8178, Japan.
2 Waseda Institute of Political Economy, Waseda University, 6-1 Nishiwaseda 1-chome, Shinjuku-ku, Tokyo 169-8050, Japan.
3 Japan Economic Research Institute, 9-2, Otemachi 1-chome, Chiyoda-ku, Tokyo 100-0004, Japan.

https://doi.org/10.1016/j.spl.2018.09.002
0167-7152/

Table 1
Functional forms of asymmetric kernels.

Kernel (j)	$K_{j(x,b)}(u)$
B (Chen, 1999)	$K_{B(x,b)}(u) = \frac{u^{b/x}(1-u)^{(1-x)/b}}{B\{b/x+1,(1-x)/b+1\}} \mathbf{1}\{u \in [0,1]\}.$
MB (Chen, 1999)	$K_{MB(x,b)}(u) = \frac{u^{\varrho_{b,0}(x)-1}(1-u)^{\varrho_{b,1}(x)-1}}{B\{\varrho_{b,0}(x),\varrho_{b,1}(x)\}} \mathbf{1}\{u \in [0,1]\},$ where $\varrho_{b,0}(x) = \begin{cases} \varrho_b(x) & \text{for } x \in [0,2b) \\ x/b & \text{for } x \in [2b,1] \end{cases},$ $\varrho_{b,1}(x) = \begin{cases} (1-x)/b & \text{for } x \in [0,1-2b] \\ \varrho_b(1-x) & \text{for } x \in (1-2b,1] \end{cases}$, and $\varrho_b(x) = 2b^2 + 5/2 - \sqrt{4b^4 + 6b^2 + 9/4 - x^2 - x/b}.$
G (Chen, 2000)	$K_{G(x,b)}(u) = \frac{u^{x/b}\exp(-u/b)}{b^{x/b+1}\Gamma(x/b+1)} \mathbf{1}\{u \geq 0\}.$
MG (Chen, 2000)	$K_{MG(x,b)}(u) = \frac{u^{\rho_b(x)-1}\exp(-u/b)}{b^{\rho_b(x)}\Gamma\{\rho_b(x)\}} \mathbf{1}\{u \geq 0\},$ where $\rho_b(x) = \begin{cases} x/b & \text{for } x \geq 2b \\ (1/4)(x/b)^2 + 1 & \text{for } x \in [0,2b) \end{cases}.$
NM (Hirukawa and Sakudo, 2015)	$K_{NM(x,b)}(u) = \frac{2u^{\alpha-1}\exp\left[-\{u/(\beta\Gamma(\alpha/2)/\Gamma((\alpha+1)/2))\}^2\right]}{\{\beta\Gamma(\alpha/2)/\Gamma((\alpha+1)/2)\}^{\alpha}\Gamma(\alpha/2)} \mathbf{1}\{u \geq 0\},$ where $(\alpha,\beta) = \begin{cases} (x/b, x) & \text{for } x \geq 2b \\ \left((1/4)(x/b)^2 + 1, x^2/(4b) + b\right) & \text{for } x \in [0,2b) \end{cases}.$

kernels. Let $K_{j(x,b)}(\cdot)$ be the asymmetric kernel indexed by j that depends on a design point x and a smoothing parameter b. We exclusively consider the beta ("B") and modified beta ("MB") kernels for supp $(X) = [0,1]$ and the gamma ("G"), modified gamma ("MG") and Nakagami-m ("NM") kernels for supp $(X) = \mathbb{R}_+$. Functional forms of these kernels are presented in Table 1. Although our main focus is on the kernels, the results in this paper can be straightforwardly extended to a wide variety of asymmetric kernels that have been proposed so far (e.g., Jin and Kawczak, 2003; Scaillet, 2004). Using a random sample $\{X_i\}_{i=1}^n$ and the kernel $j \in \{B, MB, G, MG, NM\}$, we investigate the estimator of f of the form implied by the identity (1)

$$\tilde{f}_j(x) := g(x)\hat{r}_j(x) := g(x)\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{K_{j(x,b)}(X_i)}{g(X_i)}\right\}. \tag{2}$$

A few special cases of the estimator of this class have been already considered in the literature. First, if we take $g(x) \equiv 1$ (i.e., (improper) uniform density), then $\tilde{f}_j(x)$ collapses to an ordinary asymmetric kernel density estimator ("AKDE")

$$\hat{f}_j(x) = \frac{1}{n}\sum_{i=1}^{n} K_{j(x,b)}(X_i). \tag{3}$$

Second, when $g(x)$ belongs to a parametric family, $\tilde{f}_j(x)$ reduces to the Hjort and Glad (1995, "HG")-type semiparametric MBC estimator $\tilde{f}_{HG,j}(x)$ studied by Hagmann and Scaillet (2007). Third, if $g(x)$ is set equal to the AKDE (3), then $\tilde{f}_j(x)$ becomes the Jones et al. (1995, "JLN")-type fully nonparametric MBC estimator $\tilde{f}_{JLN,j}(x)$ examined by Hirukawa (2010) and Hirukawa and Sakudo (2014).

The HG-MBC estimator has a $O(b)$ bias in general, as the initial parametric model is typically misspecified. It becomes unbiased up to the order considered only under correct specification of the model. In contrast, the JLN-MBC estimator improves the bias convergence from $O(b)$ to $O(b^2)$ even in the worst-case scenario (i.e., under misspecification), provided that the true density has sufficient smoothness. Then, within the framework of density estimation using standard symmetric kernels, Jones et al. (1999, "JSH") attempt to make the most of attractive features in HG- and JLN-MBC by applying yet another MBC step to the HG-MBC estimator. The resulting estimator has a $O(b^2)$ bias in general, and it becomes unbiased in the best-case scenario (i.e., under correct specification).

This paper extends JSH's proposal to asymmetric kernel density estimation. Because the JSH-MBC technique does not affect the order of magnitude in variance, the mean integrated squared error ("MISE") of our MBC estimator is in the form of $O(b^4 + n^{-1}b^{-1/2})$. Therefore, when best implemented, the estimator can achieve a $O(n^{-8/9})$ MISE-convergence, which is faster than $O(n^{-4/5})$, the optimal convergence rate in MISE within the class of nonnegative kernel estimators. Moreover, the estimator is a semiparametrically bias-corrected one with rate improvement. In this sense, this paper can be positioned as a complement to the existing literature on semiparametric (e.g., Hagmann and Scaillet, 2007; Gustafsson et al., 2009) and nonparametric (e.g., Hirukawa, 2010; Hirukawa and Sakudo, 2014) MBC density estimations using asymmetric kernels.

The remainder of this paper is organized as follows. Section 2 proposes the JSH-MBC estimator using asymmetric kernels and develops its asymptotic properties. Section 3 conducts Monte Carlo simulations to examine finite-sample properties of the estimator. Section 4 concludes.

This paper adopts the following notational conventions: $\Gamma(a) = \int_0^\infty y^{a-1}\exp(-y)\,dy$ for $a > 0$ is the gamma function; $B(p,q) = \int_0^1 y^{p-1}(1-y)^{q-1}dy$ for $p, q > 0$ signifies the beta function; and $\mathbf{1}\{\cdot\}$ denotes an indicator function. The expression '$X_n \sim Y_n$' is used whenever $X_n/Y_n \to 1$ as $n \to \infty$. Lastly, in order to describe different asymptotic properties of an asymmetric kernel estimator across positions of the design point $x \in \mathbb{R}_+$ ($x \in [0,1]$), we denote by "interior x" and "boundary x" a design point x that satisfies $x/b \to \infty$ ($x/b, (1-x)/b \to \infty$) and $x/b \to \kappa$ ($x/b \to \kappa$ or $(1-x)/b \to \kappa$) for some $0 < \kappa < \infty$ as $n \to \infty$, respectively.

2. The JSH-MBC estimator

2.1. Definition

The JSH-MBC estimation for the unknown density f takes three steps. The first step fits a parametric model $f(x;\theta)$ to $f(x)$ and estimates the parameter $\theta \in \Theta \subseteq \mathbb{R}^p$ by maximum likelihood ("ML"). Let $\hat{\theta}$ be the ML estimator of θ. However, the initial parametric model is subject to misspecification, which tends to cause non-vanishing bias. Then, in the second step, we put $g(x) = f\left(x;\hat{\theta}\right)$ in (2) to obtain the HG-MBC estimator as

$$\tilde{f}_{HG,j}(x) := f\left(x;\hat{\theta}\right)\tilde{r}_j(x) := f\left(x;\hat{\theta}\right)\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{K_{j(x,b)}(X_i)}{f\left(X_i;\hat{\theta}\right)}\right\}.$$

The third and final step applies the MBC again to the HG-MBC estimator to accelerate the bias convergence. We define the JSH-MBC estimator formally as

$$\tilde{f}_{JSH,j}(x) := \tilde{f}_{HG,j}(x)\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{K_{j(x,b)}(X_i)}{\tilde{f}_{HG,j}(X_i)}\right\}.$$

2.2. Convergence properties

To describe convergence properties of the JSH-MBC estimator, we introduce additional notations. Let θ_0 be the pseudo true value that minimizes the Kullback–Leibler distance of $f(x,\theta)$ from the true $f(x)$, where the word "pseudo" reflects possible misspecification of the initial parametric model. We also denote $f_0(\cdot) := f(\cdot,\theta_0)$ and $r_0(\cdot) := f(\cdot)/f(\cdot,\theta_0)$. Exploring the asymptotic properties of $\tilde{f}_{JSH,j}(x)$ requires the following regularity conditions.

Assumption 1. $\{X_i\}_{i=1}^n$ are *i.i.d.* random variables drawn from a univariate distribution having a density f with $\mathrm{supp}(X)$ either on $[0,1]$ or $\mathbb{R}_+$.

Assumption 2. For a given design point $x \in \mathrm{supp}(X)$, $f(x), f_0(x) > 0$, and $r_0(x)$ has four continuous and bounded derivatives in the neighborhood of x.

Assumption 3. The smoothing parameter $b = b_n\,(> 0)$ satisfies $b + \left(nb^3\right)^{-1} \to 0$ as $n \to \infty$.

Assumption 4. **(i)** The parameter space Θ is a compact subset of $\mathbb{R}^p$. **(ii)** $f(x;\theta)$ is continuous in θ for every $x \in \mathrm{supp}(X)$. **(iii)** θ_0 is interior in Θ and uniquely minimizes the Kullback–Leibler distance. **(iv)** The negative log-likelihood of the sample $\{X_i\}_{i=1}^n$ uniformly converges in probability to the Kullback–Leibler distance. **(v)** $\sqrt{n}\left(\hat{\theta}-\theta_0\right) = O_p(1)$.

Assumptions 1–3 are standard in the literature on MBC methods for density estimation using asymmetric kernels. Positivity of both $f(x)$ and $f_0(x)$ in Assumption 2 ensures finiteness of the asymptotic variance of the JSH-MBC estimator, whereas the fourth-order smoothness of $r_0(x)$ is required for its bias approximation. Assumption 3 implies that the shrinkage rate of b must be slower than $O\left(n^{-1/3}\right)$. We require this condition to control the order of magnitude in remainder terms when approximating the bias in the third step. It will be shown shortly that the MSE-optimal smoothing parameter for the JSH-MBC estimator becomes $b^* = O\left(n^{-2/9}\right)$ for interior x and $b^\dagger = O\left(n^{-1/5}\right)$ for boundary x; these convergence rates are indeed within the required range. First four conditions of Assumption 4 establish (weak) consistency of $\hat{\theta}$ to θ_0. The final condition on $\sqrt{n}$-convergence reflects that what is required is the convergence rate of $\hat{\theta}$, not its asymptotic distribution; interested readers may consult, for example, White (1982) for sufficient conditions for asymptotic normality of $\hat{\theta}$.

Assumptions 1–3 also suffice for the bias and variance approximations to the AKDE (3) on which our main result is built. Specifically, the bias of $\hat{f}_j(x)$ admits the expansion

$$Bias\left\{\hat{f}_j(x)\right\} = a_{1,j}(x,f)\,b + a_{2,j}(x,f)\,b^2 + o\left(b^2\right),$$

Table 2
Explicit forms of $a_{1,j}(x,f)$ and $v_j(x)$.

Kernel (j)	$a_{1,j}(x,f)$	$v_j(x)$
B	$(1-2x)f^{(1)}(x)+(x/2)(1-x)f^{(2)}(x)$	$1/\{2\sqrt{\pi}\sqrt{x(1-x)}\}$
MB	$\begin{cases}\varsigma_b(x)f^{(1)}(x) & \text{for } x\in[0,2b)\\ (x/2)(1-x)f^{(2)}(x) & \text{for } x\in[2b,1-2b]\\ -\varsigma_b(1-x)f^{(1)}(x) & \text{for } x\in(1-2b,1]\end{cases}$	$1/\{2\sqrt{\pi}\sqrt{x(1-x)}\}$
G	$f^{(1)}(x)+(x/2)f^{(2)}(x)$	$1/(2\sqrt{\pi}\sqrt{x})$
MG	$\begin{cases}(x/2)f^{(2)}(x) & \text{for } x\geq 2b\\ \xi_b(x)f^{(1)}(x) & \text{for } x\in[0,2b)\end{cases}$	$1/(2\sqrt{\pi}\sqrt{x})$
NM	$\begin{cases}(x/4)f^{(2)}(x) & \text{for } x\geq 2b\\ \xi_b(x)f^{(1)}(x) & \text{for } x\in[0,2b)\end{cases}$	$1/\left(\sqrt{2\pi}\sqrt{x}\right)$

Note: $\varsigma_b(x)=(1-x)\{\varrho_b(x)-x/b\}/[1+b\{\varrho_b(x)-x/b\}]$ for $\varrho_b(x)$ given in Table 1, and $\xi_b(x)=\{(1/2)(x/b)-1\}^2$.

where $a_{1,j}(x,f)$ and $a_{2,j}(x,f)$ are kernel-specific functions that depend on the design point x and derivatives of f. The variance of $\hat{f}_j(x)$ can be approximated by

$$Var\left\{\hat{f}_j(x)\right\}=\begin{cases}\left(n^{-1}b^{-1/2}\right)v_j(x)f(x)+o\left(n^{-1}b^{-1/2}\right) & \text{for interior } x\\ O\left(n^{-1}b^{-1}\right) & \text{for boundary } x\end{cases}.$$

Table 2 presents explicit forms of $a_{1,j}(x,f)$ and $v_j(x)$. Notice that the table does not provide $a_{2,j}(x,f)$, because the dominant bias term of $\tilde{f}_{JSH,j}(x)$ is free of this coefficient, as will be seen shortly. We also omit the analytical expression of the dominant term in $Var\left\{\hat{f}_j(x)\right\}$ for boundary x and report only its order of magnitude in the sense that the inferior rate does not affect the global property of $\tilde{f}_{JSH,j}(x)$.

The theorem below documents the main result of this paper.

Theorem 1. *If Assumptions 1–4 hold, then the bias of $\tilde{f}_{JSH,j}(x)$ for $j\in\{B,MB,G,MG,NM\}$ can be approximated as*

$$Bias\left\{\tilde{f}_{JSH,j}(x)\right\}=-f(x)\,a_{1,j}\left\{x,h_j(x,r_0)\right\}b^2+o\left(b^2\right),$$

where $a_{1,j}\left\{x,h_j(x,r_0)\right\}$ can be obtained by replacing $f=f(x)$ in $a_{1,j}(x,f)$ given in Table 2 with $h_j(x,r_0):=a_{1,j}(x,r_0)/r_0(x)$. In addition, $Var\left\{\tilde{f}_{JSH,j}(x)\right\}\sim Var\left\{\hat{f}_j(x)\right\}$ regardless of the position of x.

Proof. To save space, we only consider the case for $j=G$. The theorem can be immediately established by combining the proof of Theorem 2 in Hirukawa and Sakudo (2014) with Eqs. (4.15) and (4.16) of Hagmann and Scaillet (2007). ■

A few remarks are in order. First, Theorem 1 explains how the JSH-MBC estimator succeeds appealing bias properties of the HG- and JLN-MBC estimators. It is demonstrated that the bias convergence in $\tilde{f}_{JSH,j}(x)$ is accelerated from $O(b)$ to $O\left(b^2\right)$ under sufficient smoothness of f, even when the initial parametric model is misspecified. This is also the property of JLN-MBC. Indeed, the JLN-MBC estimator may be viewed as a special case of the JSH-MBC estimator; the former corresponds with the latter when the (improper) uniform density is chosen as the initial parametric model, to be more precise. On the other hand, if the parametric proposal is close enough to (coincides with) the truth, then $r_0(x)\approx 1$ ($r_0(x)\equiv 1$) so that $\tilde{f}_{JSH,j}(x)$ has a very small leading bias coefficient (is unbiased up to the order considered). This property is what HG- and JSH-MBC have in common.

Second, the theorem also states that the variance of $\tilde{f}_{JSH,j}(x)$ is first-order asymptotically equivalent to those of $\hat{f}_j(x)$ and $\tilde{f}_{JLN,j}(x)$. In contrast, while the variance of the original JSH-MBC estimator using a symmetric kernel is still first-order asymptotically equivalent to that of the JLN-MBC estimator using the same kernel, it is larger than that of the ordinary, bias-uncorrected density estimator using the same kernel. This is because the dominant variance term of the original JSH-MBC estimator involves the roughness of the 'twiced' kernel generated by the symmetric kernel (Stuetzle and Mittal, 1979).

Third, we should briefly mention the local and global properties of $\tilde{f}_{JSH,j}(x)$. Its mean squared error ("MSE") is of the following form:

$$MSE\left\{\tilde{f}_{JSH,j}(x)\right\}=\begin{cases}O\left(b^4+n^{-1}b^{-1/2}\right) & \text{for interior } x\\ O\left(b^4+n^{-1}b^{-1}\right) & \text{for boundary } x\end{cases}.$$

Substituting the MSE-optimal smoothing parameter $b^*=O\left(n^{-2/9}\right)$ and $b^\dagger=O\left(n^{-1/5}\right)$ into the corresponding MSE leads to the best-possible MSE convergence of $O\left(n^{-8/9}\right)$ and $O\left(n^{-4/5}\right)$ for interior and boundary x, respectively. Moreover, the inferior MSE convergence for boundary x does not affect the global property of $\tilde{f}_{JSH,j}(x)$. By the trimming argument by Chen (1999, p.136) and Chen (2000, p.476), the MISE of $\tilde{f}_{JSH,j}(x)$ is

$$MISE\left\{\tilde{f}_{JSH,j}(x)\right\} = b^4 E\left\{f(X)\, a_{1,j}^2\left(X, h_j(X, r_0)\right)\right\} + \frac{E\left\{v_j(X)\right\}}{nb^{1/2}} + o\left(b^4 + \frac{1}{nb^{1/2}}\right)$$

provided that two expectations are finite, where $a_{1,j}$ for $j = MB$ and for $j \in \{MG, NM\}$ refers to the one for $x \in [2b, 1-2b]$ and for $x \geq 2b$, respectively. The MISE-optimal smoothing parameter that minimizes two dominant terms is given by

$$b^{**} = \left[\frac{E\left\{v_j(X)\right\}}{8E\left\{f(X)\, a_{1,j}^2\left(X, h_j(X, r_0)\right)\right\}}\right]^{2/9} n^{-2/9}.$$

Therefore, the optimal MISE becomes

$$MISE^{**}\left\{\tilde{f}_{JSH,j}(x)\right\} \sim \frac{9}{8^{8/9}}\left[E\left\{f(X)\, a_{1,j}^2\left(X, h_j(X, r_0)\right)\right\}\right]^{1/9}\left[E\left\{v_j(X)\right\}\right]^{8/9} n^{-8/9}.$$

It is noteworthy that the best possible convergence rate of JSH-MBC is $n^{4/9}$, which is the best-possible nonparametric one under fourth-order smoothness of the unknown density; see Stone (1980) for more details.

3. Finite-sample performance

3.1. Monte Carlo setup

We conduct a small Monte Carlo study to assess finite-sample properties of the JSH-MBC estimator. Among all asymmetric kernels, the gamma kernel is exclusively considered due to its popularity. The simulation study compares the following four estimators as variants of (2): (i) the ordinary AKDE [G]; (ii) the JLN-MBC estimator [JLN]; (iii) the HG-MBC estimator with a gamma start [HG]; and (iv) the JSH-MBC estimator with a gamma start [JSH].

Four distributions below are chosen as the truth. All these distributions are popularly chosen as models for income distributions, loss distributions, baseline hazard, and stochastic production frontier analysis.

1. Gamma $\quad x^{\alpha-1}\exp(-x/\beta)/\{\beta^\alpha\Gamma(\alpha)\},\ (\alpha,\beta) = (1.5, 1)$.
2. Weibull $\quad (\alpha/\beta)(x/\beta)^{\alpha-1}\exp\{-(x/\beta)^\alpha\},\ (\alpha,\beta) = (1.5, 1.5)$.
3. Half-Normal $\quad \left\{2/\left(\sqrt{2\pi}\sigma\right)\right\}\exp\left\{-(x-\mu)^2/\left(2\sigma^2\right)\right\},\ (\mu,\sigma) = (0, 1.5)$.
4. Generalized Gamma $\quad \gamma x^{\alpha-1}\exp\{-(x/\beta)^\gamma\}/\{\beta^\alpha\Gamma(\alpha)\},\ (\alpha,\beta,\gamma) = (2, 2, 2.5)$.

For each distribution, 1000 data sets of sample size $n = 100$, 200 or 500 are simulated. All density estimates are evaluated on an equally spaced grid of 500 points over the interval [0, 5]. Performance of an estimator $\bar{f}(x)$ is evaluated by three measures, namely, the integrated absolute deviation ("IAD") $IAD\left\{\bar{f}(x)\right\} = \int_0^\infty \left|\bar{f}(x) - f(x)\right| dx$, the root integrated squared error ("RISE") $RISE\left\{\bar{f}(x)\right\} = \sqrt{\int_0^\infty \left\{\bar{f}(x) - f(x)\right\}^2 dx}$, and the integrated absolute bias ("IAB") $IAB\left\{\bar{f}(x)\right\} = \int_0^\infty \left|E\left\{\bar{f}(x)\right\} - f(x)\right| dx$. In our reports, the integrals are approximated over the 500 points. As regards choices of smoothing parameters, our pilot simulations indicate that plug-in methods such as the one in Hirukawa and Sakudo (2014) do not work well. Then, we turn to the simple rule of thumb as in Gustafsson et al. (2009). Specifically, smoothing parameter values are set equal to (i) $\hat{b}_{ROT_A} = \hat{\sigma}_x n^{-2/5}$ for G and HG and (ii) $\hat{b}_{ROT_B} = \hat{\sigma}_x n^{-2/9}$ for JLN and JSH, where $\hat{\sigma}_x$ is the standard deviation of each Monte Carlo sample.

3.2. Results and guidance for practitioners

Table 3 presents simulation averages and standard deviations of IAD and RISE as well as IAB. The results are mixed. It is no surprise that HG outperforms others in the Gamma case, because the gamma start correctly specifies the truth. In the remaining three cases, the gamma start is misspecified. When the deviation of the truth from the gamma start is small (i.e., the Weibull case), HG still performs best and JSH follows. On the other hand, when the deviation is substantial (i.e., the Half-Normal and Generalized Gamma cases), G and JLN appear to have advantage.

JSH does not outperform HG in terms of IAD or RISE except for the Generalized Gamma case. Most possibly, the extra MBC step from HG to JSH generates additional higher-order terms that can be omitted in its first-order asymptotics but may be non-negligible in finite samples. However, as a positive side of the MBC step, we can find that JSH often yields the smallest IAB.

Because of the mixed results, practitioners may wonder which estimator(s) to employ. Then, following Hagmann and Scaillet (2007, Section 4.6), we attempt to provide guidance on choosing among the fully parametric ML, fully nonparametric

Table 3
Averages and standard deviations of performance measures.

n	Distribution	Estimator	IAD		RISE		IAB
			Ave	SD	Ave	SD	
100	Gamma	G	0.02583	(0.00875)	0.03893	(0.01203)	0.01745
		JLN	0.02337	(0.00834)	0.03879	(0.01264)	0.01561
		HG	0.02187	(0.00852)	0.03092	(0.01229)	0.00087
		JSH	0.02391	(0.00900)	0.03340	(0.01263)	0.00259
	Weibull	G	0.02666	(0.00957)	0.03970	(0.01386)	0.01830
		JLN	0.02480	(0.01009)	0.03859	(0.01539)	0.01788
		HG	0.02262	(0.00914)	0.03275	(0.01322)	0.00548
		JSH	0.02464	(0.00945)	0.03520	(0.01349)	0.00495
	Half-Normal	G	0.02397	(0.00859)	0.03504	(0.01223)	0.01349
		JLN	0.02185	(0.00926)	0.03175	(0.01383)	0.01165
		HG	0.02418	(0.00963)	0.04022	(0.01552)	0.00991
		JSH	0.02574	(0.00971)	0.04143	(0.01551)	0.00813
	Generalized Gamma	G	0.03168	(0.01121)	0.04148	(0.01471)	0.02611
		JLN	0.03096	(0.01087)	0.04206	(0.01447)	0.02612
		HG	0.03818	(0.04458)	0.05704	(0.07169)	0.02181
		JSH	0.03149	(0.03271)	0.04676	(0.06469)	0.01581
200	Gamma	G	0.02035	(0.00662)	0.03191	(0.00943)	0.01397
		JLN	0.01840	(0.00634)	0.03284	(0.00973)	0.01286
		HG	0.01646	(0.00610)	0.02334	(0.00878)	0.00039
		JSH	0.01797	(0.00645)	0.02521	(0.00909)	0.00153
	Weibull	G	0.02100	(0.00688)	0.03191	(0.01026)	0.01461
		JLN	0.01952	(0.00711)	0.03123	(0.01112)	0.01450
		HG	0.01730	(0.00612)	0.02518	(0.00905)	0.00437
		JSH	0.01868	(0.00638)	0.02687	(0.00933)	0.00330
	Half-Normal	G	0.01895	(0.00625)	0.02777	(0.00921)	0.01049
		JLN	0.01685	(0.00694)	0.02444	(0.01037)	0.00925
		HG	0.01880	(0.00684)	0.03153	(0.01109)	0.00790
		JSH	0.01964	(0.00699)	0.03190	(0.01109)	0.00568
	Generalized Gamma	G	0.02503	(0.00823)	0.03291	(0.01090)	0.02072
		JLN	0.02490	(0.00790)	0.03374	(0.01046)	0.02145
		HG	0.03559	(0.03404)	0.05467	(0.05543)	0.02314
		JSH	0.03213	(0.03047)	0.05121	(0.05663)	0.01944
500	Gamma	G	0.01500	(0.00436)	0.02467	(0.00665)	0.01033
		JLN	0.01346	(0.00435)	0.02665	(0.00707)	0.00983
		HG	0.01151	(0.00384)	0.01653	(0.00584)	0.00026
		JSH	0.01245	(0.00409)	0.01774	(0.00609)	0.00070
	Weibull	G	0.01541	(0.00483)	0.02381	(0.00716)	0.01081
		JLN	0.01431	(0.00509)	0.02388	(0.00767)	0.01106
		HG	0.01236	(0.00416)	0.01807	(0.00604)	0.00329
		JSH	0.01314	(0.00429)	0.01907	(0.00617)	0.00190
	Half-Normal	G	0.01358	(0.00409)	0.02024	(0.00601)	0.00745
		JLN	0.01179	(0.00437)	0.01709	(0.00649)	0.00656
		HG	0.01339	(0.00436)	0.02306	(0.00733)	0.00562
		JSH	0.01370	(0.00446)	0.02275	(0.00725)	0.00331
	Generalized Gamma	G	0.01784	(0.00549)	0.02373	(0.00723)	0.01463
		JLN	0.01860	(0.00547)	0.02526	(0.00718)	0.01620
		HG	0.03006	(0.02062)	0.04717	(0.03393)	0.02236
		JSH	0.02937	(0.01961)	0.04706	(0.03474)	0.02122

Note: "Ave" and "SD" are simulation averages and standard deviations, respectively.

(i.e., G and JLN) and semiparametric (i.e., HG and JSH) estimators. Four curves in Fig. 1 are the plots of

$$Z(x) = \begin{cases} \dfrac{\log \tilde{r}_G(x) + n^{-1}b^{-1/2}/\left\{4\sqrt{\pi}\sqrt{x}f\left(x;\hat{\theta}\right)\right\}}{n^{-1/2}b^{-1/4}/\sqrt{2\sqrt{\pi}\sqrt{x}f\left(x;\hat{\theta}\right)}} & \text{for interior } x \\ \dfrac{\log \tilde{r}_G(x) + n^{-1}b^{-1}\Gamma(2\kappa+1)/\left\{2^{2(\kappa+1)}\Gamma^2(\kappa+1)f\left(x;\hat{\theta}\right)\right\}}{n^{-1/2}b^{-1/2}\sqrt{\Gamma(2\kappa+1)/\left\{2^{2\kappa+1}\Gamma^2(\kappa+1)f\left(x;\hat{\theta}\right)\right\}}} & \text{for boundary } x \end{cases}$$

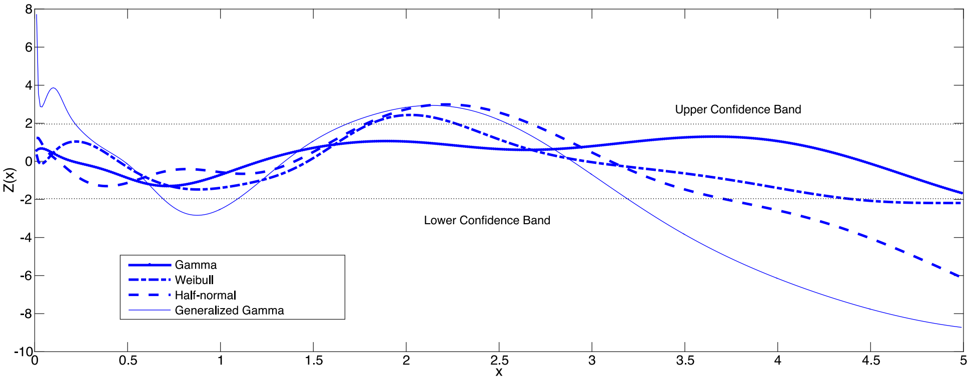

Fig. 1. Plots of $Z(x)$ for JSH-MBC estimation.

constructed from 500 random draws of the four distributions we consider. When the parametric start coincides with the truth, $Z(x)$ is approximately distributed as $N(0,1)$ for each x, and the curve should move within ± 1.96 about 95% of the time. Based on a visual inspection of the figure, we suggest the followings:

1. If the curve is within ± 1.96 (Gamma), then we do not need to go beyond ML.
2. If some parts of the curve lie outside ± 1.96 but it always stays near ± 1.96 (Weibull), then we rely on a semiparametric estimator (i.e., either HG or its bias-corrected version JSH).
3. If the curve considerably deviates from ± 1.96 (Half-Normal and Generalized Gamma), then we resort to a fully nonparametric estimator (i.e., either G or its bias-corrected version JLN).

4. Conclusion

This paper has extended yet another semiparametric MBC technique originally proposed by JSH to asymmetric kernel density estimation. It is demonstrated that the JSH-MBC estimator smoothed by asymmetric kernels accelerates the bias from $O(b)$ to $O(b^2)$ in general under sufficient smoothness of the true density, whereas the order of magnitude in variance remains unchanged. It follows that when best implemented, the MISE of the estimator achieves a faster convergence rate of $O(n^{-8/9})$. Monte Carlo simulations confirm nice bias properties of the JSH-MBC estimator.

Acknowledgments

We would like to thank an anonymous referee for his/her constructive comments and suggestions, which substantially improve the quality of the paper. This research was supported, in part, by the grant from Japan Society for the Promotion of Science (Grant Number 15K03405). The views expressed herein and those of the authors do not necessarily reflect the views of the Development Bank of Japan.

References

Chen, S.X., 1999. Beta kernel estimators for density functions. Comput. Statist. Data Anal. 31, 131–145.
Chen, S.X., 2000. Probability density function estimation using gamma kernels. Ann. Inst. Statist. Math. 52, 471–480.
Gustafsson, J., Hagmann, M., Nielsen, J.P., Scaillet, O., 2009. Local transformation kernel density estimation of loss distributions. J. Bus. Econom. Statist. 27, 161–175.
Hagmann, M., Scaillet, O., 2007. Local multiplicative bias correction for asymmetric kernel density estimators. J. Econometrics 141, 213–249.
Hirukawa, M., 2010. Nonparametric multiplicative bias correction for kernel-type density estimation on the unit interval. Comput. Statist. Data Anal. 54, 473–495.
Hirukawa, M., Sakudo, M., 2014. Nonnegative bias reduction methods for density estimation using asymmetric kernels. Comput. Statist. Data Anal. 75, 112–123.
Hirukawa, M., Sakudo, M., 2015. Family of the generalised gamma kernels: a generator of asymmetric kernels for nonnegative data. J. Nonparametr. Stat. 27, 41–63.
Hjort, N.L., Glad, I.K., 1995. Nonparametric density estimation with a parametric start. Ann. Statist. 23, 882–904.
Jin, X., Kawczak, J., 2003. Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data. Ann. Econ. Finance 4, 103–124.
Jones, M.C., Linton, O., Nielsen, J.P., 1995. A simple bias reduction method for density estimation. Biometrika 82, 327–338.
Jones, M.C., Signorini, D.F., Hjort, N.L., 1999. On multiplicative bias correction in kernel density estimation. Sankhyā A 61, 422–430.
Scaillet, O., 2004. Density estimation using inverse and reciprocal inverse Gaussian kernels. J. Nonparametr. Stat. 16, 217–226.

Stone, C.J., 1980. Optimal rates of convergence for nonparametric estimators. Ann. Statist. 8, 1348–1360.
Stuetzle, W., Mittal, Y., 1979. Some comments on the asymptotic behavior of robust smoothers. In: Gasser, T., Rosenblatt, M. (Eds.), Smoothing Techniques for Curve Estimation: Proceedings of a Workshop Held in Heidelberg, April 2–4, 1979. Springer-Verlag, Berlin, pp. 191–195.
White, H., 1982. Maximum likelihood estimation of misspecified models. Econometrica 50, 1–25.