Statistics and Probability Letters xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

ELSEVIER

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 xxxx

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)	$K_{MB(x,b)}(u)=\frac{u^{\varrho_{b,0}(x)-1}(1-u)^{\varrho_{b,1}(x)-1}}{B\left\{\varrho_{b,0}(x),\varrho_{b,1}(x)\right\}}\mathbf{1}\{u\in[0,1]\}$,
	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 $\operatorname{supp}(X)=[0,1]$ and the gamma ("G"), modified gamma ("MG") and Nakagami-m ("NM") kernels for $\operatorname{supp}(X)=\mathbb{R}_+$. 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 $\tilde{f}_j$ 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 $O\left(b^2\right)$ 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 $O\left(b^2\right)$ 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 $O\left(b^4+n^{-1}b^{-1/2}\right)$. Therefore, when best implemented, the estimator can achieve a $O\left(n^{-8/9}\right)$ MISE-convergence, which is faster than $O\left(n^{-4/5}\right)$, 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.

Please cite this article in press as: Hirukawa M., Sakudo M., Another bias correction for asymmetric kernel density estimation with a parametric start. Statistics and Probability Letters (2018), https://doi.org/10.1016/j.spl.2018.09.002.