Computational Statistics and Data Analysis 75 (2014) 112–123

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Nonnegative bias reduction methods for density estimation using asymmetric kernels

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

[a] Faculty of Economics, Setsunan University, 17-8 Ikeda Nakamachi, Neyagawa, Osaka, 572-8508, Japan

[b] Research Institute of Capital Formation, Development Bank of Japan, 9-7, Otemachi 1-chome, Chiyoda-ku, Tokyo 100-8178, Japan

[c] Institute for Research in Contemporary Political and Economic Affairs, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169-8050, Japan

ARTICLE INFO

Article history:
Received 2 July 2013
Received in revised form 10 December 2013
Accepted 7 January 2014
Available online 31 January 2014

Keywords:
Asymmetric kernel
Bias reduction
Boundary effect
Higher-order bias kernel

ABSTRACT

Two classes of multiplicative bias correction ("MBC") methods are applied to density estimation with support on $[0, \infty)$. It is demonstrated that under sufficient smoothness of the true density, each MBC technique reduces the order of magnitude in bias, whereas the order of magnitude in variance remains unchanged. Accordingly, the mean integrated squared error of each MBC estimator achieves a faster convergence rate of $O\left(n^{-8/9}\right)$ when best implemented, where n is the sample size. Furthermore, MBC estimators always generate nonnegative estimates by construction. Plug-in smoothing parameter choice rules for the estimators are proposed, and their finite sample performance is examined via Monte Carlo simulations.

1. Introduction

Hirukawa (2010) applied two classes of fully nonparametric multiplicative bias correction ("MBC") methods originally proposed for density estimation using symmetric kernels to estimate the density with support on $[0, 1]$ via nonstandard smoothing by the Beta kernel (Chen, 1999). This paper extends the analysis to estimating the density with support on $[0, \infty)$ by asymmetric kernels (Chen, 2000; Jin and Kawczak, 2003; Scaillet, 2004). Let $K_{j(x,b)}(\cdot)$ be the asymmetric kernel indexed by j that depends on a design point $x > 0$ and a smoothing parameter $b > 0$. Given a random sample $\{X_i\}_{i=1}^n$ drawn from a univariate distribution with density f that has support on $[0, \infty)$, the density estimator using asymmetric kernel j can be expressed as

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

Throughout, the kernel j refers to one of the Gamma ("G"), Modified Gamma ("MG"), Inverse Gaussian ("IG"), Reciprocal Inverse Gaussian ("RIG"), Log-Normal ("LN"),[2] and Birnbaum–Saunders ("BS") kernels. Functional forms of these kernels are presented in Table 1. Asymmetric kernels have originally emerged as an alternative to boundary correction methods; see, for instance, Karunamuni and Alberts (2005) for a brief review of the methods. Indeed, because all kernels in Table 1 have

* Corresponding author. Tel.: +81 72 839 8095; fax: +81 72 839 8138.
E-mail addresses: hirukawa@econ.setsunan.ac.jp (M. Hirukawa), marisakudo@gmail.com (M. Sakudo).

1 Tel.: +81 3 3244 1241; fax: +81 3 3270 7084.

2 Our definition of the LN kernel slightly differs from the original one in Jin and Kawczak (2003). This definition ensures that the leading variance of the density estimator (1) becomes $n^{-1}b^{-1/2}f(x) / \left(2\sqrt{\pi}x\right)$ for a design point $x > 0$ so that $x/b \to \infty$ as $b \to 0$.

0167-9473/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2014.01.012

Table 1
Functional forms of asymmetric kernels.

Kernel (j)	Functional form ($u \geq 0$)
G (Chen, 2000)	$K_{G(x/b+1,b)}(u) = u^{x/b}\exp(-u/b) / \left\{b^{x/b+1}\Gamma(x/b+1)\right\}$.
MG (Chen, 2000)	$K_{MG(\rho_b(x),b)}(u) = u^{\rho_b(x)-1}\exp(-u/b) / \left[b^{\rho_b(x)}\Gamma\{\rho_b(x)\}\right]$, where $\rho_b(x) = \begin{cases} x/b & \text{for } x \geq 2b \\ (1/4)(x/b)^2 + 1 & \text{for } x \in [0, 2b) \end{cases}$.
IG (Scaillet, 2004)	$K_{IG(x,1/b)}(u) = \frac{1}{\sqrt{2\pi b u^3}}\exp\left\{-\frac{1}{2bx}\left(\frac{u}{x} - 2 + \frac{x}{u}\right)\right\}$.
RIG (Scaillet, 2004)	$K_{RIG(1/(x-b),1/b)}(u) = \frac{1}{\sqrt{2\pi b u}}\exp\left\{-\frac{x-b}{2b}\left(\frac{u}{x-b} - 2 + \frac{x-b}{u}\right)\right\}$.
LN (Jin and Kawczak, 2003)	$K_{LN(\log x,b)}(u) = \frac{1}{u\sqrt{2\pi b}}\exp\left\{-\frac{(\log u - \log x)^2}{2b}\right\}$.
BS (Jin and Kawczak, 2003)	$K_{BS(b^{1/2},x)}(u) = \frac{1}{2x\sqrt{2\pi b}}\left\{\left(\frac{x}{u}\right)^{1/2} + \left(\frac{x}{u}\right)^{3/2}\right\}\exp\left\{-\frac{1}{2b}\left(\frac{u}{x} - 2 + \frac{x}{u}\right)\right\}$.

support on $[0, \infty)$, they are free of boundary bias near the origin. Besides, the kernels have many other appealing properties, including locally adaptive smoothing via changing their shapes and 'shrinking variance' with the position of x.[3]

Below we formally define two MBC estimators built on the density estimator (1). Throughout (1) is called the bias-uncorrected ("BU") estimator to distinguish it from MBC estimators. In the spirit of Terrell and Scott (1980, abbreviated as "TS" hereafter), the first class of MBC techniques constructs a multiplicative combination of two density estimators employing the same kernel but different smoothing parameters. Let $\hat{f}_{j,b/c}(x)$ be the density estimator using asymmetric kernel j and smoothing parameter b/c, where $c \in (0, 1)$ is some predetermined constant that does not depend on the design point x. Then, the TS-MBC asymmetric kernel density estimator can be defined as

$$\tilde{f}_{TS,j}(x) = \left\{\hat{f}_{j,b}(x)\right\}^{\frac{1}{1-c}}\left\{\hat{f}_{j,b/c}(x)\right\}^{-\frac{c}{1-c}}. \quad (2)$$

On the other hand, the second class of MBC techniques due to Jones et al. (1995, abbreviated as "JLN" hereafter) utilizes a single smoothing parameter b. In light of the identity $f(x) = \hat{f}_{j,b}(x)\left\{f(x)/\hat{f}_{j,b}(x)\right\}$, the JLN-MBC asymmetric kernel density estimator can be defined as

$$\tilde{f}_{JLN,j}(x) = \hat{f}_{j,b}(x)\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{K_{j(x,b)}(X_i)}{\hat{f}_{j,b}(X_i)}\right\}. \quad (3)$$

Recognize that the term inside the bracket is a natural nonparametric estimator of the bias-correction term $f(x)/\hat{f}_{j,b}(x)$. Also observe that both $\tilde{f}_{TS,j}(x)$ and $\tilde{f}_{JLN,j}(x)$ are free of boundary bias and always generate nonnegative density estimates everywhere by construction.

Following the convention, this paper refers to the position of x as "interior x" if $x/b \to \infty$, and "boundary x" if $x/b \to \kappa$ for some constant $\kappa > 0$, as $b \to 0$. As demonstrated shortly, under sufficient differentiability of f, bias convergence of each MBC estimator is accelerated from $O(b)$ to $O(b^2)$, whereas the order of magnitude in variance remains unchanged from the one for (1), i.e. it is still $O\left\{\left(nb^{1/2}\right)^{-1}\right\}$ for interior x. Accordingly, the mean integrated squared error ("MISE") of each MBC estimator for interior x takes the form of $O\left(b^4 + n^{-1}b^{-1/2}\right)$. Therefore, when best implemented, each estimator can achieve the convergence rate of $O\left(n^{-8/9}\right)$ in MISE, which is faster than $O\left(n^{-4/5}\right)$, the MISE-optimal convergence rate within the class of nonnegative kernel estimators (Stone, 1980). Moreover, to implement MBC estimators employing the G and MG kernels, this paper proposes plug-in methods of choosing the smoothing parameter b with gamma density used as a reference.

A few articles other than Hirukawa (2010) have investigated bias reduction methods for density estimation via nonstandard smoothing when the support has a boundary. Hagmann and Scaillet (2007) and Gustafsson et al. (2009) study semi-parametric MBC methods for density estimation with support on $[0, \infty)$. Each method employs asymmetric kernels at the bias correction step after initial parametric density estimation. Unlike MBC methods in this paper, their approaches do not improve the bias convergence in order of magnitude. Moreover, Leblanc (2010) explores a bias reduction method for estimating the density with support on $[0, 1]$ using Bernstein polynomials, and establishes acceleration in bias convergence. However, he adopts an additive bias correction, and thus the bias-corrected estimator does not always generate nonnegative estimates unlike the one in Hirukawa (2010).

The remainder of this paper is organized as follows. Section 2 presents asymptotic properties of two MBC estimators. Section 3 proposes plug-in methods of choosing the smoothing parameter b for MBC estimators using the G and MG kernels, and conducts Monte Carlo simulations to check finite sample properties of the estimators. Section 4 applies two MBC techniques

[3] It is an open question whether the asymmetric kernels studied here may fit with nonparametric analysis of functional or infinite-dimensional data by Ferraty and Vieu (2006). While they consider the asymmetric kernels that take the form of $K((X - x)/b)$ for a data point X, design point x, and smoothing parameter b, none of the kernels in Table 1 can be expressed in this form.