Journal of the Korean Statistical Society 46 (2017) 194–207

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Inverse gamma kernel density estimation for nonnegative data

Yoshihide Kakizawa [a,*], Gaku Igarashi [b]

[a] Faculty of Economics, Hokkaido University, Nishi 7, Kita 9, Kita-ku, Sapporo 060–0809, Japan

[b] Division of Policy and Planning Sciences, Faculty of Engineering, Information and Systems, University of Tsukuba, 1–1–1 Tennodai, Tsukuba, Ibaraki 305–8573, Japan

ARTICLE INFO

Article history:
Received 15 March 2016
Accepted 29 September 2016
Available online 19 October 2016

AMS 2000 subject classifications:
primary 62G07
secondary 62G20

Keywords:
Asymmetric kernel estimation
Boundary-bias-free
Gamma kernel
Inverse gamma kernel

ABSTRACT

This paper considers a varying asymmetric kernel estimation of the density f for nonnegative data. Regardless of $f(0) = 0$ or $f(0) > 0$, it is important to give a good varying shape/scale parameter for the inverse gamma (IGam) kernel, due to the problem of $\widehat{f}(0) = 0$ in some existing literature. After reformulating the IGam kernel density estimator, asymptotic properties like mean integrated squared error, mean integrated absolute error, strong consistency, and asymptotic normality are investigated in detail, under some conditions on the target density f. Simulation studies are conducted to compare the proposed IGam kernel density estimators with the existing gamma kernel density estimators.

1. Introduction

This paper considers estimation of a density f that has support $[0, \infty) = \mathrm{R}_+$ (say). In this setting, the standard kernel density estimator (e.g., Wand & Jones, 1995); $\widehat{f}(x) = (nh)^{-1}\sum_{i=1}^{n} k((x - X_i)/h)$, where $k(\cdot)$ is a symmetric kernel and $h = h_n > 0$ is a bandwidth tending to 0 as $n \to \infty$, suffers from the so-called boundary bias or edge effect. To solve this problem, many remedies have been suggested in the literature. See Hall and Park (2002), Jones (1993), Karunamuni and Alberts (2005), Marron and Ruppert (1994), and Zhang, Karunamuni, and Jones (1999).

As an alternative device to these boundary correction methods, Chen (2000) proposed to replace the usual (location-scale) kernel $\frac{1}{h}k(\frac{x-s}{h})$ by a gamma (Gam) kernel

$$k_{p,\sigma}^{\mathrm{Gam}}(s) = \frac{s^{p-1}}{\sigma^p \Gamma(p)} \exp\Big(-\frac{s}{\sigma}\Big),$$

in such a way that the shape parameter is suitably parameterized by x and b, where the scale parameter plays a role of a smoothing parameter $b = b_n > 0$ that tends to 0 as $n \to \infty$. It is a "varying kernel", whose support matches the support of the density f to be estimated. More precisely, the Gam kernel density estimators were defined by

$$\widehat{f}_b^{\mathrm{Gam}_1}(x) = \frac{1}{n}\sum_{i=1}^{n} k_{x/b+1,b}^{\mathrm{Gam}}(X_i), \qquad \widehat{f}_b^{\mathrm{Gam}_2}(x) = \frac{1}{n}\sum_{i=1}^{n} k_{\varrho(x/b),b}^{\mathrm{Gam}}(X_i), \quad x \in \mathrm{R}_+, \tag{1}$$

* Corresponding author.
E-mail address: kakizawa@econ.hokudai.ac.jp (Y. Kakizawa).

http://dx.doi.org/10.1016/j.jkss.2016.09.002

where

$$\varrho(t) = \begin{cases} t, & t \geq 2, \\ \dfrac{t^2}{4} + 1, & 0 \leq t \leq 2. \end{cases}$$

Chen (2000) showed that both $\widehat{f}_b^{\mathrm{Gam}_1}$ and $\widehat{f}_b^{\mathrm{Gam}_2}$ are boundary-bias-free density estimators and achieve the rate of convergence of the mean integrated squared error (MISE) of order $O(n^{-4/5})$, and that the second estimator $\widehat{f}_b^{\mathrm{Gam}_2}$ is superior to the first one $\widehat{f}_b^{\mathrm{Gam}_1}$ in the sense of the optimal asymptotic MISE. It should be emphasized that the last fact was further extended by Igarashi and Kakizawa (2014), as follows: In a class of the Gam kernel density estimators

$$\widehat{f}_{b,d}^{\mathrm{Gam}_2}(x) = \frac{1}{n}\sum_{i=1}^{n} k^{\mathrm{Gam}}_{\varrho_d(x/b),b}(X_i), \quad x \in \mathrm{R}_+,$$

where

$$\varrho_d(t) = \begin{cases} t + d, & t \geq 2, \\ (d+1)\left(\dfrac{t}{2}\right)^{2/(d+1)} + 1, & 0 \leq t \leq 2 \end{cases}$$

($d > -1$ is a constant), the estimator $\widehat{f}_{b,1/4}^{\mathrm{Gam}_2}$ has the best performance in terms of the optimal asymptotic MISE; note that $\widehat{f}_{b,1}^{\mathrm{Gam}_2} = \widehat{f}_b^{\mathrm{Gam}_1}$ and $\widehat{f}_{b,0}^{\mathrm{Gam}_2} = \widehat{f}_b^{\mathrm{Gam}_2}$, since $\varrho_1(t) = t + 1$ and $\varrho_0(t) = \varrho(t)$.

Mnatsakanov and Ruymgaart (2012) studied the density estimator

$$\widehat{f}^{\mathrm{MD_{MR}}}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{\alpha}{x(\alpha-1)!}\left(\frac{\alpha X_i}{x}\right)^{\alpha-1}\exp\left(-\frac{\alpha X_i}{x}\right) = \frac{1}{n}\sum_{i=1}^{n} k^{\mathrm{Gam}}_{\alpha,x/\alpha}(X_i), \quad x \in \mathrm{R}_+, \tag{2}$$

which is referred to as a moment density (MD) estimator by construction, where $\alpha = \alpha_n \in \mathrm{N}$ diverges as $n \to \infty$. Mnatsakanov and Sarkisian (2012) (see also Koul & Song, 2013) further developed other MD estimators

$$\widehat{f}^{\mathrm{MD^*_{MS}}}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{X_i(\alpha-1)!}\left(\frac{\alpha x}{X_i}\right)^{\alpha}\exp\left(-\frac{\alpha x}{X_i}\right), \qquad \widehat{f}^{\mathrm{MD_{MS}}}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{X_i\alpha!}\left(\frac{\alpha x}{X_i}\right)^{\alpha+1}\exp\left(-\frac{\alpha x}{X_i}\right), \quad x \in \mathrm{R}_+. \tag{3}$$

The leading term of the variance of the MD estimators (2) and (3) was shown to be $(\sqrt{\alpha}/n)f(x)/(2\sqrt{\pi}x)$ for each $x > 0$; see Mnatsakanov and Ruymgaart (2012) and Mnatsakanov and Sarkisian (2012), whereas, for $x/b \to \infty$, the leading term of the variance of Chen's (2000) density estimator (1) was given by $(n\sqrt{b})^{-1}f(x)/(2\sqrt{\pi x})$. At first inspection, the MD estimator may be more attractive than Chen (2000), since the coefficient function $1/(2\sqrt{\pi}x)$ in the asymptotic variance formula becomes smaller as x increases, with the decay $1/x$, rather than $1/\sqrt{x}$ in Chen's case. However, the condition $\int_{\mathrm{R}_+} f(x)/x\,dx < \infty$ for the MISE; see (16) and (23) in Mnatsakanov and Ruymgaart (2012), excludes any bounded continuous density f on R_+ with $f(0) > 0$. Note that the integrated variance of Chen's density estimators (1) is approximated by $(n\sqrt{b})^{-1}\int_{\mathrm{R}_+} f(x)/(2\sqrt{\pi x})\,dx$, which exists, at least, when the target density f is bounded continuous on R_+. To make matters worse, one observes that the definitions (2) and (3) are somewhat "bad", because the intrinsic constraints $\widehat{f}^{\mathrm{MD_{MR}}}(0) = \widehat{f}^{\mathrm{MD^*_{MS}}}(0) = \widehat{f}^{\mathrm{MD_{MS}}}(0) = 0$ are undesirable when $f(0) > 0$.

Despite of these facts, the MD estimators (3) can be viewed as inverse gamma (IGam) kernel density estimators. That is, using the notation

$$k^{\mathrm{IGam}}_{p,\sigma}(s) = \frac{\sigma^p}{\Gamma(p)s^{p+1}}\exp\left(-\frac{\sigma}{s}\right),$$

we have $\widehat{f}^{\mathrm{MD^*_{MS}}}(x) = n^{-1}\sum_{i=1}^{n} k^{\mathrm{IGam}}_{\alpha,\alpha x}(X_i)$ and $\widehat{f}^{\mathrm{MD_{MS}}}(x) = n^{-1}\sum_{i=1}^{n} k^{\mathrm{IGam}}_{\alpha+1,\alpha x}(X_i)$, $x \in \mathrm{R}_+$, with shape parameter $p = \alpha$ or $\alpha + 1$ and scale parameter $\sigma = \alpha x$. Consequently, the integer parameter $\alpha \in \mathrm{N}$ ($\alpha \to \infty$) of Mnatsakanov and Ruymgaart (2012) and Mnatsakanov and Sarkisian (2012) can be enlarged to be the positive real number parameter (i.e., $\alpha = 1/b \to \infty$). Recently, Mousa, Hassan, and Fathi (2016) considered another parameterization of (p, σ) to suggest a new density estimator[1]

$$\widehat{f}_b^{\mathrm{IGam_{MHF}}}(x) = \frac{1}{n}\sum_{i=1}^{n} k^{\mathrm{IGam}}_{x/b+2,x(x/b+1)}(X_i), \quad x \in \mathrm{R}_+ \tag{4}$$

as an alternative to Chen (2000). This definition, however, is still "bad", due to $\widehat{f}_b^{\mathrm{IGam_{MHF}}}(0) = 0$.

[1] Mousa et al. (2016) used the argument by Scaillet (2004) (see also Jin & Kawczak, 2003), i.e., the Taylor expansion of $s^{-1/2}f(s)$ around $s = x$, to derive the asymptotic variance of their estimator (4). However, such a task would be inconvenient, compared to our derivation, relying only on the Taylor expansion of $f(s)$ around $s = x$, as in Igarashi (2016) and Igarashi and Kakizawa (2014). Anyway, their final result of $V[\widehat{f}_b^{\mathrm{IGam_{MHF}}}(x)]$ is incorrect; we see that, when $x/b \to \infty$, its asymptotic variance should read $(nb^{1/2})^{-1}f(x)/(2\sqrt{\pi x})$, as in Chen's (2000) density estimators (1).
Mnatsakanov and Sarkisian (2012) and Mousa et al. (2016) seemed to mention the MISEs of the estimators (3) and (4), without careful analysis of the error terms (i.e., their derivations were formal).