Proc. of the Int. Conference on Mathematics and Applications, Dec 7-8 2021, Blida

MULTIPLICATIVE BIAS CORRECTION FOR INVERSE GAMMA AND BETA PRIME KERNEL DENSITY ESTIMATORS

Lynda HARFOUCHE, Yasmina ZIANE, Nabil ZOUGAB, Smail ADJABI

Research unit LaMOS, Higher School of Computing and Digital Sciences and Technologies (ESTIN), 06008, Amizour, Bejaia, Algeria. 06000 Bejaia, Algeria.
E-mail: lyndaharfouche@hotmail.com
Research unit LaMOS, Faculty of Exact Sciences, Bejaia University, 06000 Bejaia, Algeria.
Research unit LaMOS, Faculty of technology, Bejaia University, 06000 Bejaia, Algeria.
Research unit LaMOS, Faculty of Exact Sciences, Bejaia University, 06000 Bejaia, Algeria.

ABSTRACT

In this paper, we demonstrate that the multiplicative bias correction (MBC) approaches can be extended for both Inverse Gamma (IG) and Beta Prime (BP) kernel density estimators. First, some properties of the MBC-IG and MBC-BP kernel density estimators (bias, variance and mean integrated squared error) are shown. Second, the least square cross validation technique (LSCV) is adapted for the choice of bandwidth.

1. INTRODUCTION

Given a random sample of observations $\{X_i\}_{i=1}^n$ with univariate density f, whitch is supported on $\mathbb{T} = [0,\infty)$. A continuous symmetric or asymmetric kernel estimator $\hat{f}_h(x)$ of $f(x)$ can be defined as follows :

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_{x,h}(X_i),$$

Where $K_{x,h}$ is the continuous symmetric or asymmetric kernel with the target x and $h = h(n) > 0$ is an arbitrary sequence of smoothing parameters (bandwidths).
Recently two classes of multiplicative bias correction ("MBC") techniques were proposed in order to estimate the univariate densities, with compact [Hirukawa(2010)] and with bounded support [Hirukawa and Sakudo(2014)], [Zougab et al.(2015)] for generalized Birnbaum-Saunders kernel density estimators, [Harfouche et al.(2018a)] and [Harfouche et al.(2018b)] in the context of probability mass function with discrete kernels. These two classes are initially investigated in [Terrell and Scott(1980)] and in [Jones et al.(1995)], witch improve bias from $O(h^2)$ as the bandwidth $h \to 0$ to $O(h^4)$ for symmetric kernel functions. It is well known that the symmetric kernel estimator is inappropriate in the context of estimating unknown probability densities whitch are supported on $\mathbb{T} = [0,\infty)$ because it causes boundary bias without producing negative values of the estimate, and we accept that the use of asymmetric kernels functions has many

advantages such as optimal rate of convergence in the mean integrated squared error sense for kernels of order two, witch were originally employed in [Chen(1999)] for nonparametric estimation of compact using Beta as kernel functions, and [Chen(2000)] with (gamma and modified gamma kernels). Two recent papers are respectively proposed by [Mousa et al.(2016)] (Inverse Gamma (IG) kernel) and [Erçelik and Nadar(2018)] (Beta Prime (BP) kernel) in the same context (In the context of unknown probability density estimation). They have shown that the optimal rate of convergence of the mean integrated squared error for this estimator is of order $O(n^{-4/(5)})$ where they assuming that the true density f is twice continuous. [Hirukawa(2010)], [Hirukawa and Sakudo(2014)] and [Zougab et al.(2015)] have demonstrated that the two classes of MBC approaches can be applied in the case of nonparametric density estimation with asymmetric kernel, where they assuming that f is four times continuous with bounded derivatives, the both methods reach an optimal rate of convergence of mean integrated squared error of order $O(n^{-8/(9)})$.

The current work is also motivated by several points. First, The IG and BP kernel density estimators are often used in modeling the hydrological problems and modeling the frequency of certain behavioral acts because of its long tail. Second, they are free of boundary bias, have flexible shape, always nonnegative, and achieve the optimal rate of convergence for the MSE of order of $O(n^{-4/(5)})$ and MISE within the class of nonnegative kernel density estimator. The main aim of this paper is to extend the application of MBC approaches for Inverse Gamma (IG) kernel and Beta Prime (BP) kernel estimator as in [Hirukawa(2010)], [Hirukawa and Sakudo(2014)] and [Zougab et al.(2015)], in order to reach an optimal rate of convergence of mean integrated squared error of order $O(n^{-8/(9)})$. First, we provide the asymptotic properties of these estimators and show that the optimal rate of convergence of the mean integrated squared error is obtained.

This paper is organized as follows. Section 2 briefly recalls on IG and BP kernels for density estimation. In Section 3, we first introduce the MBC kernel density estimators based on IG and BP kernels. Second, we show some properties of the MBC-IG and MBC-BP kernels density estimators (bias, variance and mean integrated squared error). Third, we adapt the least square cross validation technique (LSCV) for the choice of bandwidth. Conclusion of our paper is given in section 4.

2. A SHORT REVIEW ON IG AND BP KERNELS

This section is dedicated to present a brief recall on both IG and BP kernel density estimators.

2.1. IG and BP kernel estimator

Given a random sample $X_1,\dots,X_n$, the IG and BP estimator of an unknown pdf f with nonnegative support are given by [Mousa et al.(2016)] and [Erçelik and Nadar(2018)] respectively by

$$\widehat{f}_j(x) = \frac{1}{n}\sum_{i=1}^{n} K_{j(x,h)}(X_i) \tag{1}$$

where $x > 0$ is the target (point where the density is estimated), $h > 0$ is a bandwidth (or smoothing parameter), $j = IG, BP$ and the explicit form of $K_{j(x,h)}(X_i)$ is given in table 1.

The expressions of the bias and variance for $\widehat{f}_j(x)$ are derived by [Mousa et al.(2016)] and [Erçelik and Nadar(2018)]. The asymptotic bias when $h \to 0$ is given by

$$\text{bias}(\widehat{f}_j(x)) = q_j(x,f)h + o(h), \tag{2}$$

where the explicit form of $q_j(x,f)$ is given in table 2. Similarly, when $n \to \infty$ and $h \to 0$ the

TABLE 1 – Univariate continuous kernels.

Kernel(j)	Explicit form
IG ([Mousa et al.(2016)])	$K_{IG(\alpha,\beta)}(y)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}(1/y)^{\alpha+1}\exp(-\beta/y)$,
	where $\alpha=\frac{h+x}{h}+1$ *and* $\beta=\frac{x(h+x)}{h}$.
BP ([Erçelik and Nadar(2018)])	$K_{BP(\lambda,\mu)}(\,y)=\frac{y^{\lambda-1}(1+y)^{-\lambda-\mu}}{\beta(\lambda,\mu)}$,
	where $\lambda=\frac{x^2}{h}+x+1$ *and* $\mu=\frac{x}{h}+\frac{1}{x+h}+1$.

TABLE 2 – Explicit form of $q_j(x,f)$.

Kernel(j)	Explicit form
IG ([Mousa et al.(2016)])	$q_{IG}(x,f)=\frac{1}{2}f''(x)x$,
BP ([Erçelik and Nadar(2018)])	$q_{BP}(x,f)=f'(x)+(1+x)\frac{f''(x)}{2}$.

asymptotic variance is

$$\operatorname{Var}(\widehat{f}_j(x))=p_j(x,h)f(x)+o\left(\frac{1}{nh^{1/2}}\right). \tag{3}$$

where the explicit form of $p_j(x,h)$ is given in table 3.
The mean integrated squared error (MISE) is also given in [Mousa et al.(2016)] and [Erçelik and Nadar(2018)]

TABLE 3 – Explicit form of $p_j(x,h)$.

Kernel(j)	Explicit form
IG ([Mousa et al.(2016)])	$p_{IG}(x,h)=\frac{1}{2\sqrt{2\pi}}n^{-1}h^{-1/2}x^{-1/2}$,
BP ([Erçelik and Nadar(2018)])	$p_{BP}(x,h)=\frac{1}{2\sqrt{\pi}}n^{-1}h^{-1/2}(1+x)^{-1/2}$.

and is expressed as

$$\begin{aligned}\operatorname{MISE}(\widehat{f}_j) &= \int_0^\infty \operatorname{bias}^2(\widehat{f}_j(x))dx+\int_0^\infty \operatorname{Var}(\widehat{f}_j(x))dx\\ &= h^2\int q_j^2(x,f)dx\\ &+ \int p_j(x,h)f(x)dx+o\left(h^2+\frac{1}{nh^{\frac{1}{2}}}\right). \end{aligned}\tag{4}$$

Remark
Under sufficient smoothness of the true density, they are shown that the order of magnitude in *MISE* of kernel density estimator (1) is $o\left(h^2+\frac{1}{nh^{\frac{1}{2}}}\right)$. For these raison, two classes of MBC density estimators will be investigated in the next section. The two classes of MBC density estimators by construction nonnegative, and establish a faster convergence rate of $o\left(h^4+\frac{1}{nh^{\frac{1}{2}}}\right)$.

3. MBC FOR IG AND BP DENSITY ESTIMATORS

In this section, we apply two classes of MBC techniques for IG and BP kernel density estimator. The two classes are originally proposed by [Terrell and Scott(1980)] and [Jones et al.(1995)]

for symmetric kernel density estimator. Note that these MBC techniques have also extended recently by [Hirukawa(2010)] and [Hirukawa and Sakudo(2014)] for kernel density estimation on the unit interval using beta and modified beta kernels and for density estimation using asymmetric kernels (gamma, modified gamma, inverse gaussian, reciprocal inverse gaussian, log-Normal and Birnbaum-Saunders kernels), respectively.

3.1. Estimators

Based on the same idea of [Terrell and Scott(1980)] and [Hirukawa(2010)], the MBC kernel density estimator using the IG and BP kernels, which we simply denote TS-j kernel density estimators, can be adapted as follows :

$$\widehat{f}_{TS-j}(x) = \left\{\widehat{f}_{j,h}(x)\right\}^{\frac{1}{1-a}} \left\{\widehat{f}_{j,h/a}(x)\right\}^{-\frac{a}{1-a}}, \tag{5}$$

where $\widehat{f}_{j,h}$ and $\widehat{f}_{j,h/a}$ denote the IG or BP kernel density estimators given by (1) with bandwidths h and h/a, respectively, with $a \in (0,1)$ is a constant that does not depend on the target x; see, e.g., [Hirukawa(2010)].

The second class of MBC techniques for symmetric kernel density estimators is attributed to [Jones et al.(1995)] (see also [Hirukawa(2010)] and [Hirukawa and Sakudo(2014)] for asymmetric kernel density estimators). The analogue of their estimators using IG and BP kernels, which we denote JLN-j kernel density estimators is given by

$$\widehat{f}_{JLN-j}(x) = \widehat{f}_j(x) \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{K_{j(x,h)}(X_i)}{\widehat{f}_j(X_i)} \right\}, \tag{6}$$

where $K_{j(x,h)}$ is the IG or BP kernel.

3.2. Asymptotic properties

The following theorems present the asymptotic bias and variance of the TS-j and JLN-j kernel estimators. We assume that

A1. f has four continuous and bounded derivatives.

A2. The sequence of bandwidths $h = h(n)$ satisfies $\lim_{n\to\infty} h = 0$.

Note that these assumptions have been discussed in [Hirukawa(2010)] and [Hirukawa and Sakudo(2014)].

Theorem 1 *Let $\widehat{f}_{TS-j}$ be the TS-j kernel estimator defined by (5). For a given $x > 0$ and under assumptions A1 and A2, then :*

(i) the bias of the TS-j kernel estimators admit the following expansion

$$\text{bias}(\widehat{f}_{TS-j}) = \frac{1}{a}\left[\frac{1}{2}\left\{\frac{\psi_{1,j}^2(x)}{f(x)} - \psi_{2,j}(x)\right\}\right] h^2 + o(h^2),$$

where $\psi_{1,j}(x)$ and $\psi_{2,j}(x)$ are given in Table 4 and 5 respectively.

(ii) the variance of the TS-j kernel estimators is given by

$$\text{Var}(\widehat{f}_{TS-j}) = \gamma(a) p_j(x,h) f(x) + o\left(\frac{1}{nh^{1/2}}\right).$$

where $\gamma(a) = \frac{(1+a^{5/2})(1+a)^{1/2} - 2\sqrt{2}a^{3/2}}{(1+a)^{1/2}(1-a)^2}$,

TABLE 4 – Explicit form of $\psi_{1,j}(x)$.

Kernel(j)	Explicit form
IG	$\frac{1}{2}xf''(x)$,
BP	$f'(x)+\frac{(x+1)}{2}f''(x)+x^2f'''(x)-\frac{x^4}{12}f''''(x)$.

TABLE 5 – Explicit form of $\psi_{2,j}(x)$.

Kernel(j)	Explicit form
IG	$\frac{1}{4}x^2f''''(x)-\frac{4}{3}xf'''(x)$,
BP	$-\frac{x^3}{8}f''''(x)$.

Theorem 2 *Let $\widehat{f}_{JLN-j}$ be the JLN-j kernel estimator defined by (6). For a given $x>0$ and under assumptions A1 and A2, then :*

(i) the bias of the JLN-j kernel estimator is given by

$$\mathrm{bias}(\widehat{f}_{JLN-j}) = -f(x)\psi_{1,j}(x,g)h^2+o(h^2),$$

where $g(x)=\psi_1(x,j)/f(x)$ and $\psi_{1,j}(x)$ is the same as given in Theorem 1.

(ii) the variance of the JLN-j kernel estimators has asymptotic form

$$\mathrm{Var}(\widehat{f}_{JLN-j}) = p_j(x,h)f(x)+o\left(\frac{1}{nh^{1/2}}\right).$$

3.3. Global propriety

The criterion to use for the global propriety is the mean integrated squared error (MISE) defined as

$$\mathrm{MISE}(\widehat{f}_{MBC-j}) = \int_0^\infty \mathrm{bias}^2(\widehat{f}_{MBC-j}(x))dx+\int_0^\infty \mathrm{Var}(\widehat{f}_{MBC-j}(x))dx, \tag{7}$$

where $\widehat{f}_{MBC-j}$ is the TS-j or the JLN-j kernel density estimators.

The mean integrated squared error (MISE) of the TS-j kernel estimators given in (5) is expressed as

$$\mathrm{MISE}(\widehat{f}_{TS-j}) = \frac{h^4}{4a^2}\int_0^\infty\left\{\frac{\psi_{1,j}^2(x)}{f(x)}-\psi_{2,j}(x)\right\}^2dx+\gamma(a)\int_0^\infty p_j(x,h)f(x)dx+o\left(\frac{1}{nh^{1/2}}+h^4\right). \tag{8}$$

The optimal bandwidth minimizing the corresponding MISE (8) is such that

$$h_{TS-IG}^{opt} = \left\{\frac{a^2\gamma(a)\int_0^\infty \frac{f(x)}{x^{1/2}2\sqrt{2\pi}}dx}{8\int_0^\infty\left\{\frac{\psi_{1,IG}^2(x)}{f(x)}-\psi_{2,IG}(x)\right\}^2dx}\right\}^{2/9} n^{-2/9}. \tag{9}$$

$$h_{TS-BP}^{opt} = \left\{\frac{a^2\gamma(a)\int_0^\infty \frac{f(x)}{(1+x)^{1/2}2\sqrt{\pi}}dx}{8\int_0^\infty\left\{\frac{\psi_{1,BP}^2(x)}{f(x)}-\psi_{2,BP}(x)\right\}^2dx}\right\}^{2/9} n^{-2/9}. \tag{10}$$

Similarly, the MISE of the JLN-j kernel estimators given in (6) is given by

$$\mathrm{MISE}(\widehat{f}_{JLN-j}) = h^4 \int_0^\infty f^2(x)\psi_{1,j}^2(x,g)dx + \int_0^\infty p_j(x,h)f(x)dx + o\left(\frac{1}{nh^{1/2}} + h^4\right). \tag{11}$$

By minimizing (11) in the bandwidth h, we obtain the optimal value

$$h_{JLN-IG}^{opt} = \left\{ \frac{\int_0^\infty \frac{f(x)}{x^{1/2}2\sqrt{2\pi}}dx}{8\int_0^\infty f^2(x)\left[\frac{1}{2}xg''(x)\right]^2 dx} \right\}^{2/9} n^{-2/9}. \tag{12}$$

$$h_{JLN-BP}^{opt} = \left\{ \frac{\int_0^\infty \frac{f(x)}{(1+x)^{1/2}2\sqrt{\pi}}dx}{8\int_0^\infty f^2(x)\left[g'(x) + \frac{(x+1)}{2}g''(x) + x^2g'''(x) - \frac{x^4}{12}g''''(x)\right]^2 dx} \right\}^{2/9} n^{-2/9}. \tag{13}$$

Note that the bandwidths (9), (10), (12) and (13) can not be employed in practice. Then, next subsection presents a practical procedure to bandwidth selection.

3.4. Choice of bandwidth for MBC-j kernel estimators

The optimal bandwidths given by (9), (10), (12) and (13) depend on the unknown density f and on its derivatives f', f'', f''' and f'''', for these raison, they can not be exploited in practice. In this paper, we adapt the least square cross validation (LSCV) method. In the case of the LSCV technique, for a given estimator $\widehat{f}_{MBC-j}$, which denote the TS-j or JLN-j kernel estimators, the optimal bandwidth h_{ucv} of h is obtained by

$$h_{LSCV} = \arg\min_h \mathrm{LSCV}(h),$$

where

$$\mathrm{LSCV}(h) = \int \widehat{f}_{MBC-j}^2(x)dx - \frac{2}{n}\sum_{i=1}^n \widehat{f}_{MBC-j}^{(-i)}(X_i),$$

where $\widehat{f}_{MBC-j}^{(-i)}(y)$ is the leave-one-out estimator computed as $\widehat{f}_{MBC-j}(y)$ by excluding the observation X_i. For the TS-j kernel estimators, the LSCV function is given

$$\begin{aligned}\mathrm{LSCV}_{TS-j}(h) &= \int_0^\infty \left\{\widehat{f}_{j,h}(x)\right\}^{\frac{2}{1-a}} \left\{\widehat{f}_{j,h/a}(x)\right\}^{-\frac{2a}{1-a}} dx - \frac{2}{n(n-1)} \\ &\times \sum_i \left[\left\{\sum_{k\neq i} K_{j(X_i,h)}(X_k)\right\}^{\frac{1}{1-a}} \left\{\sum_{k\neq i} K_{j(X_i,h/a)}(X_k)\right\}^{-\frac{a}{1-a}} \right].\end{aligned} \tag{14}$$

In the case of JLN-j kernel estimators, the expression of LSCV is

$$\begin{aligned}\mathrm{LSCV}_{JLN-j}(h) &= \frac{1}{n^2}\int_0^\infty \widehat{f}_j(x)^2 \left\{\sum_{i=1}^n \frac{K_{j(x,h)}(X_i)}{\widehat{f}_j(X_i)}\right\}^2 dx - \frac{2}{n(n-1)} \\ &\times \sum_i \sum_{k\neq i} K_{j(X_i,h)}(X_k)\frac{\widehat{f}_j(X_i)}{\widehat{f}_j(X_k)}.\end{aligned} \tag{15}$$

4. CONCLUSIONS

This paper has extended the application of the multiplicative bias correction (MBC) approaches for Inverse Gamma (IG) and Beta Prime (BP) kernels in order to estimate densities with nonnegative data. As in several papers, [Hirukawa(2010)] and [Hirukawa and Sakudo(2014)], we have demonstrated that these two approaches of MBC improve the order of magnitude in bias from $O(h)$ to $O(h^2)$. The performances of the MBC-IG and MBC-BP kernel estimators (TS-IG, JLN-IG, TS-BP and JLN-BP kernel estimators) with the least square cross validation (LSCV) bandwidth selectors are investigated.

5. REFERENCES

[Chen(1999)] Chen, S.X., 1999. Beta kernel estimators for density functions. Computational Statistics and Data Analysis. 31, 131–145.

[Chen(2000)] Chen, S.X., 2000. Gamma kernel estimators for density functions. Annals of the Institute of Statistical Mathematics. 52, 471–480.

[Erçelik and Nadar(2018)] Erçelik, E., Nadar, M., 2018. Nonparametric density estimation based on beta prime kernel. Communications in Statistics - Theory and Methods, DOI : 10.1080/03610926.2018.1538458.

[Harfouche et al.(2018a)] Harfouche L, Adjabi S, Zougab N, Funke B., 2018. Multiplicative bias correction for discrete kernels. Statistical Methods and Applications. 27, 253-276.

[Harfouche et al.(2018b)] Harfouche L, Adjabi S, Zougab N., 2018. Bayesian bandwidth selection with two multiplicative bias correction estimators for discrete kernel. Communication in statistics-Case studies, Data Analysis and Applications. DOI : 10.1080/23737484.2018.1512909.

[Hirukawa(2010)] Hirukawa, M., 2010. Nonparametric multiplicative bias correction for kernel-type density estimation on the unit interval. Computational Statistics and Data Analysis. 54, 473–495.

[Hirukawa and Sakudo(2014)] Hirukawa, M., Sakudo, M., 2014. Nonnegative bias reduction methods for density estimation using asymmetric kernels. Computational Statistics and Data Analysis. 75, 112–123.

[Jones et al.(1995)] Jones, M.C., Linton, O., Nielsen, J.P., 1995. A simple bias reduction method for density estimation. Biometrika. 82, 327–338.

[Mousa et al.(2016)] Mousa, A. M., Hassan, M. Kh., Fathi, A., 2016. A New Nonparametric Estimator For Pdf Based On Inverse Gamma Distribution. Communications in Statistics-Theory and Methods. 45, 7002–7010 .

[Terrell and Scott(1980)] Terrell, G.R., Scott, D.W., 1980. On improving convergence rates for nonnegative kernel density estimators. Annals of Statistics. 8, 1160–1163.

[Zougab et al.(2015)] Zougab, N., Adjabi, S., 2015. Multiplicative bias correction for generalized Birnbaum-Saunders kernel density estimators and application to nonnegative heavy tailed data. Journal of the Korean Statistical Society.