Hacettepe Journal of Mathematics and Statistics
Volume 47 (4) (2018), 1003 – 1021

The bias reduction in density estimation using a geometric extrapolated kernel estimator

Reza Salehi*† and Ali Shadrokh‡

Abstract

One of the nonparametric methods to estimate the probability density is kernel method. In this paper, kernel density estimation methods including the naive kernel(NK) estimator and geometric extrapolation based kernel($GEBK$) method are introduced and discussed. Theoretical properties, including the selection of smoothing parameter, the accuracy of resultant estimators using Monte Carlo simulation are studied. The results show that the amount of bias in the proposed geometric extrapolation based kernel estimator significantly decreases.

Keywords: Kernel density estimation, Bias reduction, Smoothing parameter, Geometric extrapolation.

Mathematics Subject Classification (2010): 62G05, 62G07

Received : 16.03.2016 *Accepted* : 28.09.2016 *Doi* : 10.15672/HJMS.201614922002

1. Introduction

Probability density function is one of the basic and widely used concepts in the statistics by which we can verify random variable behavior. If the form of population is known, the parametric method is usually used; otherwise, nonparametric method may be applied. Several nonparametric methods have been used so far; among them, the easiest and the most popular one is the histogram. Until the early fifties, histogram was the only method raised for density estimation. Although for identifying data distribution process the histogram method seems to be easy to use at the first stage, it can not be the most accurate one due to some of its failures such as roughness. Therefore, it has been replaced by other methods. Over the last twenty years, different nonparametric methods such as kernel method in this field have been expanded and developed. Theoretical studies of this technique and its applications by computers for data processing has caused

*Department of Statistics, Payame Noor University, P.O.Box 19395-4697, Tehran, I.R. of Iran., Email: salehir2003@yahoo.com

†Corresponding Author.

‡Department of Statistics, Payame Noor University, P.O.Box 19395-4697, Tehran, I.R. of Iran., Email: ali_shadrokh@yahoo.com

a revolution in the estimation of the probability density. Rosenblatt [11] and Parzen [10] were the first to introduce the kernel method. Silverman [12], Hardell [4] and many others generalized this method to higher dimensional. Also, Farrell [3] obtained the best asymptotic convergence rate $O(n^{\frac{-2r}{2r+1}})$ of mean squared error(MSE) for orthogonal kernel estimators. Terrell and Scott [14] introduced geometric extrapolation of nonnegative kernels for bias reduction. Kim, et al. [7] developed a skewing method to reduce the bias. Mynbaev and Martins-Filho [9] worked out a bias reduced kernel relative to the classical kernel estimator via Lipschitz condition. Kairat, Saralees Nadarajah, et al.[6] studied improving bias with kernels of Epanechnikov and Gram-Charlier for kernel density estimation. Kim and Kim [8] presented a version of the kernel density estimator that reduces MSE. They showed that MSE reduction is due to bias reduction not variance one. Igarashi and Kakizawa [5] studied asymmetric kernel estimator with support $[0, \infty)$, using nonnegative bias correction techniques with standard normal kernel. They showed that the best mean integrated squared error($MISE$) occurs with $O(n^{\frac{-8}{9}})$.
As we pointed out above, many efforts have been made in the past decade to estimate probability density as accurate as possible. In this paper, we mainly focus on bias reduction of kernel estimators using geometric extrapolation. Indeed, in Section 2, we introduce the naive kernel estimator, and in Section 3 we consider and investigate the geometric extrapolation based kernel method. A Monte Carlo simulation study is described in Section 4 and the results are compared by using standard normal kernel. In Section 5, we present three numerical examples. In Section 6, however, we summerize our results as the finall part of our work.

2. The Naive Kernel Estimator

Let $X_1, X_2, \cdots, X_n$ be independent and identically distributed random variables with an unknown density function f, with size n. For $x \in R$ the naive kernel(NK) density estimator of f can be defined as

$$(2.1) \qquad NK: \qquad \hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - X_i}{h}), \qquad h > 0$$

where $K_h(x) = \frac{1}{h}K(\frac{x}{h})$ denotes the kernel weight function, usually a symmetric density, and h is the bandwidth parameter, which controls the smoothness of the estimator. The kernel weight function is usually required to satisfy the following conditions:

$$\begin{array}{ll} a1) & \int\limits_{-\infty}^{\infty} K(x)dx = 1 \\ a2) & \underset{-\infty<x<\infty}{Sup} |K(x)| < \infty \\ a3) & \lim\limits_{x\to\infty} |xK(x)| = 0 \\ a4) & K(x) = K(-x) \\ a5) & h \to 0 \quad and \quad nh \to \infty \quad as \quad n \to \infty. \end{array}$$

Parzen [10] showed that under the conditions of $a1$ to $a5$, $\hat{f}_h(x)$ estimator is consistent in second-order mean. For the optimal asymptotic choice of h, suppose that an unknown density function f and kernel weight function K, to satisfy the following conditions:

$b1$) $\|f''\|_2^2 = \int_{-\infty}^{\infty} (f''(x))^2 dx < \infty$

$b2$) $\|K\|_2^2 = \int_{-\infty}^{\infty} K^2(x)dx < \infty$

$b3$) $\int_{-\infty}^{\infty} xK(x)dx = 0$

$b4$) $\mu_2(K) = \int_{-\infty}^{\infty} x^2K(x)dx < \infty$

$b5$) f'' is absolutely continuous in a neighborhood of x.

According to Silverman [12] and Wand and Jones [15], we have

$$Bias(\hat{f}_h(x)) = E(\hat{f}_h(x)) - f(x) = \frac{h^2}{2} f''(x) \int x^2K(x)dx + O(h^3)$$

$$Var(\hat{f}_h(x)) = E\Big[\hat{f}_h(x)) - E(\hat{f}_h(x))\Big]^2 = \frac{1}{nh} f(x) \int K^2(x)dx + O(\frac{1}{nh}).$$

Thus, according to the criterion MSE we have

$$\begin{aligned} MSE(\hat{f}_h(x)) &= Bias^2(\hat{f}_h(x)) + Var(\hat{f}_h(x)) \\ &\sim \tfrac{h^4}{4}(f''(x))^2(\textstyle\int x^2K(x)dx)^2 + \tfrac{1}{nh}f(x)\textstyle\int K^2(x)dx + O(\tfrac{1}{n} + h^5). \end{aligned}$$

We can easily see that under the conditions of $b1$ to $b5$, the optimized bandwidth of $\hat{f}_h(x)$ estimator is

$$h_{opt} = \left[\frac{\|K\|_2^2}{\|f''\|_2^2 \, \mu_2{}^2(K)}\right]^{\frac{1}{5}} n^{\frac{-1}{5}}$$

d then the optimal MSE is of the order $O(n^{\frac{-4}{5}})$.

3. Bias Reduction by Geometric Extrapolation

Geometric extrapolation was introduced for kernel density estimation by Terrell and Scott [14]. Consider the naive kernel density estimator with two different bandwidths h and $2h$:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - X_i}{h})$$

and

$$\hat{f}_{2h}(x) = \frac{1}{2nh} \sum_{i=1}^{n} K(\frac{x - X_i}{2h}).$$

Suppose the kernel weight function K above is symmetric so that all the odd moments of K are zero. Terrell and Scott [14] proposed the following estimator:

(3.1) $GEBK_{TS}$: $$f_h^*(x) = \Big(\hat{f}_h(x)\Big)^{\frac{4}{3}} \Big(\hat{f}_{2h}(x)\Big)^{\frac{-1}{3}}.$$

Note that f^* doesn't have integral one. We can easily see that the optimized $MSE(f_h^*(x))$ is of the order $n^{\frac{-8}{9}}$. We modify the estimator so that the bias can be considerably reduced comparing to the estimator 3.1. Indeed, we propose the estimator 2.1 with two different bandwidths h and ah.

Now, assume also that the second derivative f'' of the underlying density f is absolutely continuous and square integrable. Then, using Taylor expansion of $f(x+hz)$ around x, we have

$$\begin{aligned} E\left(\hat{f}_h(x)\right) &= n\int \tfrac{1}{nh}K(\tfrac{x-t}{h})f(t)dt \\ &= \int K(z)f(x-hz)dz \\ &= \int K(z)\left[f(x)-hzf'(x)+\tfrac{(hz)^2}{2!}f''(x)+\cdots\right]dz \ , \\ &= f(x)+\sum_{i=1}^{r}\tfrac{(-1)^i}{i!}f^{(i)}(x)m_ih^i+O\left(h^{r+1}\right) \end{aligned}$$

where

$$m_i(K)=\int z^iK(z)dz=0 \quad , \quad i=1,2,\cdots,r-1 \quad \mathit{and} \quad m_r(K)\neq 0$$

means that K is a kernel of order r. Let

$$I_h(x)=E(\hat{f}_h(x)).$$

Then, for symmetric kernels for which all the odd moments of K are zero, we have

$$I_h(x)=f(x)+\frac{h^2}{2!}f''(x)\int z^2K(z)dz+\frac{h^4}{4!}f^{(4)}(x)\int z^4K(z)dz+\cdots .$$

By taking $c_i=\frac{(-1)^i}{i!}f^{(i)}(x)m_i(K), \quad i=1,2,3,\cdots$, then we have

$$I_h(x)=f(x)\left[1+\frac{c_2}{f(x)}h^2+\frac{c_4}{f(x)}h^4+\cdots\right].$$

Accordingly,

$$\log(I_h(x))=\log(f(x))+\log\left(\left[1+\frac{c_2}{f(x)}h^2+\frac{c_4}{f(x)}h^4+\cdots\right]\right).$$

By Taylor expansion of $log(1+x)$, we also have

$$\begin{aligned} \log(I_h(x)) &= \log(f(x))+\tfrac{c_2}{f(x)}h^2+\tfrac{c_4}{f(x)}h^4+\cdots \\ &\quad +(\tfrac{-1}{2})\left[\tfrac{c_2^2}{f^2(x)}h^4+\tfrac{c_4^2}{f^2(x)}h^8+\cdots\right] \\ &\quad +\tfrac{1}{3}\left[\tfrac{c_2^3}{f^3(x)}h^6+\tfrac{c_4^3}{f^3(x)}h^{12}+\cdots\right]+\cdots \\ &= \log(f(x))+\tfrac{c_2}{f(x)}h^2+\tfrac{c_4f(x)-\frac{1}{2}c_2^2}{f^2(x)}h^4+\cdots . \end{aligned}$$

Similarly, for the parameter ah we have

$$\log(I_{ah}(x))=\log(f(x))+\frac{c_2}{f(x)}(ah)^2+\frac{c_4f(x)-\frac{1}{2}c_2^2}{f^2(x)}(ah)^4+\cdots .$$

Thus, for bias reduction, we propose the following geometric interpolation estimator:

$$\tilde{f}_h(x)=\left(\hat{f}_h(x)\right)^{w_1}\left(\hat{f}_{ah}(x)\right)^{w_2}. \tag{3.2}$$

Therefore,

$$\log(\tilde{f}_h(x))=w_1\log\left(\hat{f}_h(x)\right)+w_2\log\left(\hat{f}_{ah}(x)\right)$$

$$\begin{aligned} w_1\log\left(I_h(x)\right)+w_2\log\left(I_{ah}(x)\right) &= w_1\log(f(x))+w_1\tfrac{c_2}{f(x)}h^2 \\ &\quad +w_1\tfrac{c_4f(x)-\frac{1}{2}c_2^2}{f^2(x)}h^4+\cdots+w_2\log(f(x)) \\ &\quad +w_2\tfrac{c_2}{f(x)}(ah)^2+w_2\tfrac{c_4f(x)-\frac{1}{2}c_2^2}{f^2(x)}(ah)^4+\cdots \\ &= (w_1+w_2)\log(f(x))+(w_1+a^2w_2)\tfrac{c_2}{f(x)}h^2 \\ &\quad +(w_1+a^4w_2)\tfrac{c_4f(x)-\frac{1}{2}c_2^2}{f^2(x)}h^4+\cdots . \end{aligned}$$

To reach our goal, we need $w_1 \log(I_h(x)) + w_2 \log(I_{ah}(x))$ has term $log(f(x))$ but has h^2 term disappeared. Thus, w_1 and w_2 have to satisfy

$$\begin{cases} w_1 + w_2 = 1 \\ w_1 + a^2 w_2 = 0 \end{cases} \Rightarrow \begin{cases} w_1 = \frac{a^2}{a^2-1} \\ w_2 = \frac{-1}{a^2-1} \end{cases}, \qquad a > 0, a \neq 1$$

Hence,

$$\frac{a^2}{a^2-1}\log(I_h(x)) + \frac{-1}{a^2-1}\log(I_{ah}(x)) = \log(f(x)) + \frac{\frac{1}{2}a^2c_2^2 - a^2c_4 f(x)}{f^2(x)}h^4 + \cdots.$$

Now, using a series expansion for the exponential functions, we have

$$(I_h(x))^{\frac{a^2}{a^2-1}}(I_{ah}(x))^{\frac{-1}{a^2-1}} = f(x) + \frac{\frac{1}{2}a^2c_2^2 - a^2c_4 f(x)}{f(x)}h^4 + \cdots.$$

We rewrite $\hat{f}_h$ and $\hat{f}_{ah}$ as

$$\hat{f}_h(x) = I_h(x) + Z$$

and

$$\hat{f}_{ah}(x) = I_{ah}(x) + W,$$

respectively. We have shown in $E\left(\hat{f}_h(x)\right)$ and $Var\left(\hat{f}_h(x)\right)$ that the random variables Z and W have expectation zero while the variances and covariances of Z and W are of order $O(\frac{1}{nh})$. So, our proposed estimator becomes

$$\text{(3.3)} \qquad GEBK_P: \quad \tilde{f}_h(x)) = \left(\hat{f}_h(x)\right)^{\frac{a^2}{a^2-1}}\left(\hat{f}_{ah}(x)\right)^{\frac{-1}{a^2-1}}.$$

Now, using the expansion of $(1+k)^x = 1 + kx + O(k^2)$, we have

$$\begin{aligned}
\tilde{f}_h(x)) &= (I_h(x))^{\frac{a^2}{a^2-1}}\left[1 + \tfrac{Z}{I_h(x)}\right]^{\frac{a^2}{a^2-1}}(I_{ah}(x))^{\frac{-1}{a^2-1}}\left[1 + \tfrac{W}{I_{ah}(x)}\right]^{\frac{-1}{a^2-1}} \\
&= (I_h(x))^{\frac{a^2}{a^2-1}}(I_{ah}(x))^{\frac{-1}{a^2-1}}\left[1 + \tfrac{a^2}{a^2-1}\tfrac{Z}{I_h(x)} + O\left((\tfrac{Z}{I_h(x)})^2\right)\right] \\
&\times\left[1 + \tfrac{-1}{a^2-1}\tfrac{W}{I_{ah}(x)} + O\left((\tfrac{W}{I_{ah}(x)})^2\right)\right] \\
&= (I_h(x))^{\frac{a^2}{a^2-1}}(I_{ah}(x))^{\frac{-1}{a^2-1}}[1 + \tfrac{a^2}{a^2-1}\tfrac{Z}{I_h(x)} + O\left((\tfrac{Z}{I_h(x)})^2\right) \\
&+ \tfrac{-1}{a^2-1}\tfrac{W}{I_{ah}(x)} + \tfrac{-a^2}{(a^2-1)^2}\tfrac{Z}{I_h(x)}\tfrac{W}{I_{ah}(x)} + \tfrac{-1}{a^2-1}\tfrac{W}{I_{ah}(x)}O\left((\tfrac{W}{I_{ah}(x)})^2\right) \\
&+ O\left((\tfrac{W}{I_{ah}(x)})^2\right) + \tfrac{a^2}{a^2-1}\tfrac{Z}{I_h(x)}O\left((\tfrac{W}{I_{ah}(x)})^2\right) \\
&+ O\left((\tfrac{Z}{I_h(x)})^2\right)O\left((\tfrac{W}{I_{ah}(x)})^2\right)] \\
&= (I_h(x))^{\frac{a^2}{a^2-1}}(I_{ah}(x))^{\frac{-1}{a^2-1}} + \tfrac{a^2}{a^2-1}Z\left(\tfrac{I_h(x)}{I_{ah}(x)}\right)^{\frac{1}{a^2-1}} \\
&+ \tfrac{-1}{a^2-1}W\left(\tfrac{I_h(x)}{I_{ah}(x)}\right)^{\frac{a^2}{a^2-1}} + O\left((Z+W)^2\right).
\end{aligned}$$

From definition $I_h(x)$, we also have

$$\frac{I_h(x)}{I_{ah}(x)} = 1 + O\left(h^2\right).$$

Then,

$$Bias\left(\tilde{f}_h(x)\right) = E\left(\left(\hat{f}_h(x)\right)^{\frac{a^2}{a^2-1}}\left(\hat{f}_{ah}(x)\right)^{\frac{-1}{a^2-1}}\right) - f(x)$$
$$= E[(I_h(x))^{\frac{a^2}{a^2-1}}(I_{ah}(x))^{\frac{-1}{a^2-1}} + \tfrac{a^2}{a^2-1}Z\big(1+O(h^2)\big)^{\frac{1}{a^2-1}}$$
$$+ \tfrac{-1}{a^2-1}W\big(1+O(h^2)\big)^{\frac{a^2}{a^2-1}} + O\left((Z+W)^2\right)] - f(x)$$
$$= f(x) + \tfrac{1}{2}a^2c_2^2h^4 + O\left(h^4\right) + 0 + 0 + O\left(\tfrac{1}{nh}\right) - f(x)$$
$$= \tfrac{1}{2}a^2c_2^2h^4 + O\left(h^4 + \tfrac{1}{nh}\right)$$

and

$$Var\left(\tilde{f}_h(x)\right) = Var\left(\left(\hat{f}_h(x)\right)^{\frac{a^2}{a^2-1}}\left(\hat{f}_{ah}(x)\right)^{\frac{-1}{a^2-1}}\right)$$
$$= E\left[\tfrac{a^2}{a^2-1}Z + \tfrac{-1}{a^2-1}W\right]^2 + O\left(\tfrac{1}{n}\right)$$
$$= Var\left(\tfrac{a^2}{a^2-1}\hat{f}_h(x) + \tfrac{-1}{a^2-1}\hat{f}_{ah}(x)\right) + O\left(\tfrac{1}{n}\right) = O\left(\tfrac{1}{nh}\right).$$

Thus, the mean integrated squared error($MISE$) for the estimator $\tilde{f}_h(x)$ is

$$MISE(\tilde{f}_h(x)) = \int_{-\infty}^{\infty} E\left[\tilde{f}_h(x) - f(x)\right]^2 dx$$
$$= \int_{-\infty}^{\infty} Var\left(\tilde{f}_h(x)\right)dx + \int_{-\infty}^{\infty} Bias^2\left(\tilde{f}_h(x)\right)dx$$
$$= \int_{-\infty}^{\infty} (\tfrac{a^2}{a^2-1})^2 Var\left(\hat{f}_h(x)\right)dx + \int_{-\infty}^{\infty} (\tfrac{-1}{a^2-1})^2 Var\left(\hat{f}_{ah}(x)\right)dx$$
$$+ \int_{-\infty}^{\infty} \tfrac{1}{4}a^4c_2^4h^8dx + O\left(h^8\right)$$
$$= (\tfrac{a^2}{a^2-1})^2\tfrac{1}{nh}\|K\|_2^2 + O\left(\tfrac{1}{n}\right) + (\tfrac{-1}{a^2-1})^2\tfrac{1}{anh}\|K\|_2^2 + O\left(\tfrac{1}{n}\right)$$
$$+ \tfrac{1}{4}a^4\left(\tfrac{1}{2}\right)^4\|f''\|_2^4\left(\int z^2K(z)dz\right)^4h^8 + O\left(h^8\right)$$
$$= \tfrac{a^5+1}{a(a^2-1)^2nh}\|K\|_2^2 + \tfrac{a^4}{64}\|f''\|_2^4\mu_2^4(K)h^8 + O\left(h^8 + \tfrac{1}{nh}\right).$$

Under the conditions of $b1$ to $b5$, the optimal choice of h that minimizes $MISE(\tilde{f}_h(x))$ satisfies

$$h_{opt} = \left[\frac{8(a^5+1)\|K\|_2^2}{a^5(a^2-1)^2\|f''\|_2^4\mu_2^4(K)}\right]^{\frac{1}{9}} n^{\frac{-1}{9}}.$$

We can easily see that the optimized bandwidth is $h = O\left(n^{\frac{-1}{9}}\right)$ and, therefore, the optimal MSE is of the order $O(n^{\frac{-8}{9}})$. These results are summarized into the following theorem.

3.1. Theorem. *Let K be symmetric and all the odd moments of K be zero. Assume also that conditions b_1 to b_5 hold. Then,*

$$Bias\left(\tilde{f}_h(x)\right) = \frac{a^2}{8}h^4(f''(x))^2\left(\int z^2K(z)dz\right)^2 + O\left(h^4 + \frac{1}{nh}\right)$$

and

$$Var\left(\tilde{f}_h(x)\right) = O\left(\frac{1}{nh}\right).$$

Consequently,

$$MSE(\tilde{f}_h(x)) = O(h^8 + \frac{1}{nh})$$

and the optimal MSE *is of the order* $O(n^{\frac{-8}{9}})$ *with* $h = O\left(n^{\frac{-1}{9}}\right)$.

4. Simulation Studies

4.1. Optimal Choice for Parameter a. In this section we report on a simulation study which is designed to show that the minimizer of $MSE(\tilde{f}_h(x))$ yields a reliable choice of a. Therefore, to introduce an optimal estimator reducing the bias, we should look for an a that minimizes the MSE of the proposed estimator. Since bias is the area occupied by true density and estimated density, we apply numerical methods such as rectangular and composite Simpson(Burden and Faires [2]) ones to optimize a by minimizing the MSE. In each case the data come from a standard normal density f. Also the kernel K is a standard normal density and in each case we set $x = 0$. In a simulation study, when we know the true f, we may compute the true bias and variance and hence MSE. Considering Figure 1, the optimal values for a are 0.6272 and 0.6429 in the rectangular and Simpson methods, respectively, using R software. Having noted that the Simpson method considers to be more accurate, we prepose 0.6429 as an optimal a. Thus, our final proposed estimator will be as

$$\tilde{f}_h(x) = \left(\hat{f}_h(x)\right)^{-0.7045}\left(\hat{f}_{0.6429h}(x)\right)^{1.7045}. \tag{4.1}$$

This estimator reduces the bias considerably comparing to the $GEBK_{TS}$ one.

4.2. Monte Carlo Simulation Study. In this section, we carry out some simulation studies for comparing the three proposed estimators: the naive kernel(NK) $\hat{f}_h(x)$ given in 2.1, the geometric extrapolation based kernel($GEBK_{TS}$) $f^*_h(x)$ given in 3.1, and the geometric extrapolation based kernel($GEBK_P$) $\tilde{f}_h(x)$ given in 4.1. We randomly select 1000 independent samples of size $n=$ 50, 100, 200 and 400. We arbitrarily choose $x = 0, 0.5, 1, 1.5, 2, 2.5, 3$ and 3.5 at which the kernel estimators are calculated and compared. Also we choose the standard normal as the kernel function K, that is

$$K(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}I_{\{-\infty<x<\infty\}}.$$

Having selected optimal bandwidth parameters, the bias, variance and MSE are estimated, respectively, as

$$Bias(\hat{f}) = \tfrac{1}{1000}\sum_{i=1}^{1000}\left(\hat{f}_i - f\right),$$
$$Var(\hat{f}) = \tfrac{1}{1000}\sum_{i=1}^{1000}\left(\hat{f}_i - \bar{\hat{f}}\right)^2 \quad with \quad \bar{\hat{f}} = \tfrac{1}{1000}\sum_{i=1}^{1000}\hat{f}_i$$

and

$$MSE(\hat{f}) = \frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{f}_i - f\right)^2,$$

where f is the true density $f(x)$ and $\hat{f}$ is either $\hat{f}_h(x)$, $f^*_h(x)$ or $\tilde{f}_h(x)$. It should be noted that f can be chosen as the standard normal density(density 1), gamma($\alpha = 2, \beta = 1$) density(density 2) and Weibull($\alpha = 2, \beta = 1$) density(density 3). The simulation results are presented in Tables 1−8 and Figures 2 and 3.

Table 1. Bias, variance and MSE of different kernel density estimators evaluated at $x = 0$

n	$estimator$	$criterion$	$density$ 1	$density$ 2	$density$ 3
50	NK	$Bias$	−0.0380254	0.1237169	0.1621021
		$Variance$	0.0022895	0.0006136	0.0031785
		MSE	0.0037355	0.0159195	0.0294556
	$GEBK_{TS}$	$Bias$	−0.0226931	0.1356625	0.1813864
		$Variance$	0.0024247	0.0005772	0.0031123
		MSE	0.0029396	0.0189815	0.0360133
	$GEBK_P$	$Bias$	−0.0046226	0.0985899	0.0954033
		$Variance$	0.0050519	0.0011061	0.0030946
		MSE	0.0050733	0.0108261	0.0121964
100	NK	$Bias$	−0.0306791	0.1202835	0.1435720
		$Variance$	0.0014492	0.0003704	0.0016117
		MSE	0.0023905	0.0148386	0.022225
	$GEBK_{TS}$	$Bias$	−0.0211265	0.1342227	0.1699952
		$Variance$	0.0011991	0.0003425	0.0016123
		MSE	0.0016454	0.0183582	0.0305107
	$GEBK_P$	$Bias$	−0.0026483	0.0937691	0.0913473
		$Variance$	0.0026414	0.0005457	0.0015521
		MSE	0.0026484	0.0093383	0.0098964
200	NK	$Bias$	−0.0241311	0.1137246	0.1282513
		$Variance$	0.0008785	0.0001933	0.0008019
		MSE	0.0014608	0.0131266	0.0172503
	$GEBK_{TS}$	$Bias$	−0.0163765	0.1311195	0.1584170
		$Variance$	0.0007303	0.0001951	0.0007596
		MSE	0.0009985	0.01738749	0.0258556
	$GEBK_P$	$Bias$	−0.0022946	0.0892354	0.0828551
		$Variance$	0.0017981	0.0002937	0.0008015
		MSE	0.0018033	0.0082567	0.0076665
400	NK	$Bias$	−0.0191010	0.1063954	0.1133275
		$Variance$	0.0005286	0.0001267	0.0003952
		MSE	0.0008935	0.0114467	0.0132383
	$GEBK_{TS}$	$Bias$	−0.0136772	0.1260452	0.1476991
		$Variance$	0.0003946	0.0001056	0.0004228
		MSE	0.0005817	0.0159930	0.0222378
	$GEBK_P$	$Bias$	−0.0013672	0.0849751	0.0756382
		$Variance$	0.0009379	0.0001502	0.0003869
		MSE	0.0009398	0.0073709	0.0061080

Table 2. Bias, variance and MSE of different kernel density estimators evaluated at $x = 0.5$

n	*estimator*	*criterion*	*density* 1	*density* 2	*density* 3
50	NK	$Bias$	−0.0269493	−0.0817046	−0.0864387
		$Variance$	0.0021085	0.0013812	0.0096969
		MSE	0.0028347	0.0080568	0.0171686
	$GEBK_{TS}$	$Bias$	−0.0142055	−0.0783841	−0.0681626
		$Variance$	0.0018904	0.0011125	0.0097117
		MSE	0.0020922	0.0072566	0.0143579
	$GEBK_P$	$Bias$	−0.0022709	−0.0421594	0.0004914
		$Variance$	0.0049798	0.0036793	0.0260016
		MSE	0.0049849	0.0053567	0.0260019
100	NK	$Bias$	−0.0217940	−0.0731437	−0.0648438
		$Variance$	0.0012769	0.0007589	0.0060570
		MSE	0.0017518	0.0061089	0.0102617
	$GEBK_{TS}$	$Bias$	−0.0117034	−0.0731111	−0.0579405
		$Variance$	0.0010599	0.0006599	0.0055246
		MSE	0.0011968	0.0060051	0.0088817
	$GEBK_P$	$Bias$	−0.0035742	−0.0388310	−0.0047309
		$Variance$	0.0032213	0.0016868	0.0146176
		MSE	0.0032341	0.0031947	0.0146399
200	NK	$Bias$	−0.0160644	−0.0664662	−0.0499959
		$Variance$	0.0007851	0.0004387	0.0038702
		MSE	0.0010431	0.0048565	0.0063697
	$GEBK_{TS}$	$Bias$	−0.0110948	−0.0700265	−0.0459248
		$Variance$	0.0006330	0.0003746	0.0028464
		MSE	0.0007561	0.0052783	0.0049555
	$GEBK_P$	$Bias$	0.0018186	−0.0346332	−0.0057377
		$Variance$	0.0013759	0.0009856	0.0078136
		MSE	0.0013792	0.0021850	0.0078465
400	NK	$Bias$	−0.0117784	−0.0573178	−0.0407104
		$Variance$	0.0005382	0.0002612	0.0021712
		MSE	0.0006770	0.0035466	0.0038285
	$GEBK_{TS}$	$Bias$	−0.0085058	−0.0670166	−0.0382721
		$Variance$	0.0003492	0.0001810	0.0016659
		MSE	0.0004215	0.0046722	0.0031307
	$GEBK_P$	$Bias$	0.0010195	−0.0310965	−0.0012154
		$Variance$	0.0008117	0.0005196	0.0042133
		MSE	0.0008127	0.0014866	0.0042148

Table 3. Bias, variance and MSE of different kernel density estimators evaluated at $x = 1$

n	*estimator*	*criterion*	*density* 1	*density* 2	*density* 3
50	NK	$Bias$	−0.0044831	−0.0776301	−0.0307330
		$Variance$	0.0017487	0.0018225	0.0097852
		MSE	0.0017688	0.0078490	0.0107297
	$GEBK_{TS}$	$Bias$	0.0054640	−0.0739304	0.0003583
		$Variance$	0.0015922	0.0017133	0.0083076
		MSE	0.0016220	0.0071790	0.0083077
	$GEBK_P$	$Bias$	0.0023854	−0.0150179	0.0014841
		$Variance$	0.0037588	0.0037797	0.0255406
		MSE	0.0037645	0.0040053	0.0255428
100	NK	$Bias$	−0.0020890	−0.0675426	−0.0286426
		$Variance$	0.0011101	0.0009061	0.0057692
		MSE	0.0011145	0.0054681	0.0065896
	$GEBK_{TS}$	$Bias$	0.0071681	−0.0688615	0.0026775
		$Variance$	0.0008745	0.0009346	0.0045265
		MSE	0.0009259	0.0056765	0.0045336
	$GEBK_P$	$Bias$	0.0011245	−0.0102334	0.0014844
		$Variance$	0.0023634	0.0022433	0.0146129
		MSE	0.0023647	0.0023481	0.0146151
200	NK	$Bias$	−0.0014821	−0.0517373	−0.0192061
		$Variance$	0.0006081	0.0005720	0.0033769
		MSE	0.0006103	0.0032488	0.0037458
	$GEBK_{TS}$	$Bias$	0.0044963	−0.0603400	−0.0012688
		$Variance$	0.0005027	0.0004392	0.0025029
		MSE	0.0005229	0.0040801	0.0025046
	$GEBK_P$	$Bias$	0.0002443	−0.0038863	−0.0004044
		$Variance$	0.0012744	0.0010577	0.0076728
		MSE	0.0012745	0.0010728	0.0076730
400	NK	$Bias$	−0.0000334	−0.0411430	−0.0172528
		$Variance$	0.0003878	0.0003302	0.0022073
		MSE	0.0003878	0.0020230	0.0025050
	$GEBK_{TS}$	$Bias$	0.0048745	−0.0514574	0.0012382
		$Variance$	0.0002815	0.0002584	0.0015127
		MSE	0.0003052	0.0029062	0.0015142
	$GEBK_P$	$Bias$	0.0006123	−0.0035961	0.0012974
		$Variance$	0.0006972	0.0007113	0.0043657
		MSE	0.0006976	0.0007242	0.0043674

Table 4. Bias, variance and MSE of different kernel density estimators evaluated at $x = 1.5$

n	*estimator*	*criterion*	*density* 1	*density* 2	*density* 3
50	NK	$Bias$	0.0163337	−0.0361186	0.0175634
		$Variance$	0.0013168	0.0014166	0.0058930
		MSE	0.0015836	0.0027211	0.0062014
	$GEBK_{TS}$	$Bias$	0.0180325	−0.0257172	0.0281780
		$Variance$	0.0011698	0.0014960	0.0060529
		MSE	0.0014950	0.0021574	0.0068469
	$GEBK_P$	$Bias$	0.0067766	0.0026524	0.0075476
		$Variance$	0.0025116	0.0029131	0.0140342
		MSE	0.0025575	0.0029202	0.0140912
100	NK	$Bias$	0.0125088	−0.0274810	0.0164478
		$Variance$	0.0007636	0.0008103	0.0039152
		MSE	0.0009201	0.0015655	0.0041857
	$GEBK_{TS}$	$Bias$	0.0152086	−0.0191209	0.0202111
		$Variance$	0.0006913	0.0008234	0.0032287
		MSE	0.0009226	0.0011890	0.0036372
	$GEBK_P$	$Bias$	0.0026997	0.0044329	−0.0007799
		$Variance$	0.0013676	0.0015985	0.0065855
		MSE	0.0013749	0.0016182	0.0065861
200	NK	$Bias$	0.0088776	−0.0208889	0.0117048
		$Variance$	0.0004531	0.0004437	0.0021572
		MSE	0.0005319	0.0008800	0.0022942
	$GEBK_{TS}$	$Bias$	0.0137086	−0.0141658	0.0184052
		$Variance$	0.0003634	0.0004132	0.0017842
		MSE	0.0005513	0.0006138	0.0021229
	$GEBK_P$	$Bias$	0.0024436	0.0040050	0.0042177
		$Variance$	0.0007157	0.0010144	0.0038191
		MSE	0.0007217	0.0010305	0.0038369
400	NK	$Bias$	0.0076227	−0.0145428	0.0101131
		$Variance$	0.0002394	0.0003073	0.0013722
		MSE	0.0002975	0.0005188	0.0014744
	$GEBK_{TS}$	$Bias$	0.0110717	−0.0089278	0.0013634
		$Variance$	0.0001990	0.0002322	0.0009623
		MSE	0.0003216	0.0003119	0.0011482
	$GEBK_P$	$Bias$	0.0006322	0.0041933	0.0023575
		$Variance$	0.0003941	0.0005149	0.0021371
		MSE	0.0003945	0.0005325	0.0021426

Table 5. Bias, variance and MSE of different kernel density estimators evaluated at $x = 2$

n	*estimator*	*criterion*	*density* 1	*density* 2	*density* 3
50	NK	*Bias*	0.0172091	−0.0075158	0.0179870
		Variance	0.0006947	0.0010850	0.0021138
		MSE	0.0009909	0.0011415	0.0024373
	$GEBK_{TS}$	*Bias*	0.0187117	0.0079709	0.0162472
		Variance	0.0006738	0.0009557	0.0017180
		MSE	0.0010239	0.0010192	0.0019819
	$GEBK_P$	*Bias*	0.0038751	0.0093389	0.0046634
		Variance	0.0011432	0.0028226	0.0031812
		MSE	0.0011582	0.0029098	0.0032030
100	NK	*Bias*	0.0128326	−0.0043542	0.0154606
		Variance	0.0003725	0.0006386	0.0012068
		MSE	0.0005372	0.0006575	0.0014458
	$GEBK_{TS}$	*Bias*	0.0148732	0.0092789	0.0136737
		Variance	0.0003575	0.0004819	0.0009661
		MSE	0.0005787	0.0005679	0.0011531
	$GEBK_P$	*Bias*	0.0025133	0.0043210	0.0015805
		Variance	0.0005249	0.0015332	0.0017512
		MSE	0.0005312	0.0015519	0.0017537
200	NK	*Bias*	0.0104526	−0.0025784	0.0108814
		Variance	0.0002141	0.0004371	0.0006454
		MSE	0.0003233	0.0004437	0.0007638
	$GEBK_{TS}$	*Bias*	0.0123268	0.0111069	0.0106401
		Variance	0.0001775	0.0002840	0.0004962
		MSE	0.0003294	0.0004073	0.0006094
	$GEBK_P$	*Bias*	0.0021000	0.0037071	0.0028888
		Variance	0.0003207	0.0008691	0.0008926
		MSE	0.0003251	0.0008829	0.0009010
400	NK	*Bias*	0.0085324	−0.0009745	0.0083108
		Variance	0.0001193	0.0002692	0.0003651
		MSE	0.0001921	0.0002701	0.0004342
	$GEBK_{TS}$	*Bias*	0.0105904	0.0120981	0.0078759
		Variance	0.0000912	0.0001621	0.0002832
		MSE	0.0002033	0.0003084	0.0003453
	$GEBK_P$	*Bias*	0.0014057	0.0034578	−0.0002835
		Variance	0.0001875	0.0004880	0.0004975
		MSE	0.0001895	0.0004999	0.0004976

Table 6. Bias, variance and MSE of different kernel density estimators evaluated at $x = 2.5$

n	*estimator*	*criterion*	*density* 1	*density* 2	*density* 3
50	NK	$Bias$	0.0114241	0.0056051	0.0049395
		$Variance$	0.0002584	0.0009013	0.0002923
		MSE	0.0003889	0.0009327	0.0003167
	$GEBK_{TS}$	$Bias$	0.0122514	0.0194716	0.0046412
		$Variance$	0.0002638	0.0008259	0.0002026
		MSE	0.0004139	0.0012051	0.0002242
	$GEBK_P$	$Bias$	0.0026980	0.0048414	0.0020619
		$Variance$	0.0003645	0.0025276	0.0004583
		MSE	0.0003717	0.0025511	0.0004625
100	NK	$Bias$	0.0080130	0.0036021	0.0044851
		$Variance$	0.0001495	0.0005750	0.0001739
		MSE	0.0002137	0.0005879	0.0001941
	$GEBK_{TS}$	$Bias$	0.0104656	0.0176385	0.0040371
		$Variance$	0.0001279	0.0003874	0.0001021
		MSE	0.0002375	0.0006985	0.0001183
	$GEBK_P$	$Bias$	0.0018538	0.0031253	0.0016258
		$Variance$	0.0001888	0.0014389	0.0002398
		MSE	0.0001923	0.0014487	0.0002425
200	NK	$Bias$	0.0062307	0.0037391	0.0032572
		$Variance$	0.0000812	0.0003159	0.0000934
		MSE	0.0001201	0.0003298	0.0001041
	$GEBK_{TS}$	$Bias$	0.0080626	0.0161488	0.0025871
		$Variance$	0.0000667	0.0002391	0.0000590
		MSE	0.0001317	0.0004999	0.0000657
	$GEBK_P$	$Bias$	0.0008541	0.0017022	0.0006414
		$Variance$	0.0001047	0.0007238	0.0001225
		MSE	0.0001054	0.0007267	0.0001230
400	NK	$Bias$	0.0048465	0.0031157	0.0022005
		$Variance$	0.0000446	0.0002111	0.0000506
		MSE	0.0000681	0.0002208	0.0000555
	$GEBK_{TS}$	$Bias$	0.0063295	0.0135302	0.0020593
		$Variance$	0.0000340	0.0001412	0.0000332
		MSE	0.0000741	0.0003243	0.0000375
	$GEBK_P$	$Bias$	0.0009277	0.0014529	0.0003791
		$Variance$	0.0000601	0.0004324	0.0000705
		MSE	0.0000609	0.0004345	0.0000706

Table 7. Bias, variance and MSE of different kernel density estimators evaluated at $x = 3$

n	*estimator*	*criterion*	*density* 1	*density* 2	*density* 3
50	NK	$Bias$	0.0053992	0.0099611	0.0008735
		$Variance$	0.0000893	0.0008149	0.0000249
		MSE	0.0001185	0.0009142	0.0000257
	$GEBK_{TS}$	$Bias$	0.0061383	0.0186149	0.0011123
		$Variance$	0.0000722	0.0006981	0.0000190
		MSE	0.0001099	0.0010446	0.0000203
	$GEBK_P$	$Bias$	0.0011763	0.0029935	0.0000649
		$Variance$	0.0000875	0.0019080	0.0000242
		MSE	0.0000889	0.0019170	0.0000242
100	NK	$Bias$	0.0041763	0.0083558	0.0005643
		$Variance$	0.0000443	0.0004638	0.0000136
		MSE	0.0000617	0.0005336	0.0000139
	$GEBK_{TS}$	$Bias$	0.0042406	0.0160118	0.0005289
		$Variance$	0.0000314	0.0004171	0.0000073
		MSE	0.0000494	0.0006734	0.0000076
	$GEBK_P$	$Bias$	0.0011112	0.0012068	0.0004638
		$Variance$	0.0000480	0.0010476	0.0000251
		MSE	0.0000492	0.0010490	0.0000253
200	NK	$Bias$	0.0026957	0.0060232	0.0004045
		$Variance$	0.0000213	0.0002936	0.0000078
		MSE	0.0000286	0.0003299	0.0000080
	$GEBK_{TS}$	$Bias$	0.0032553	0.0134671	0.0005427
		$Variance$	0.0000154	0.0002236	0.0000052
		MSE	0.0000261	0.0004050	0.0000055
	$GEBK_P$	$Bias$	0.0008229	0.0003962	0.0002968
		$Variance$	0.0000261	0.0005555	0.0000113
		MSE	0.0000267	0.0005557	0.0000113
400	NK	$Bias$	0.0018985	0.0051464	0.0003744
		$Variance$	0.0000108	0.0001770	0.0000046
		MSE	0.0000144	0.0002035	0.0000047
	$GEBK_{TS}$	$Bias$	0.0026215	0.0103112	0.0004018
		$Variance$	0.0000081	0.0001243	0.0000022
		MSE	0.0000149	0.0002307	0.0000024
	$GEBK_P$	$Bias$	0.0005308	0.0005353	0.0002322
		$Variance$	0.0000136	0.0003065	0.0000065
		MSE	0.0000139	0.0003068	0.0000066

Table 8. Bias, variance and MSE of different kernel density estimators evaluated at $x = 3.5$

n	*estimator*	*criterion*	*density* 1	*density* 2	*density* 3
50	NK	$Bias$	0.0020862	0.0076653	0.0000608
		$Variance$	0.0000215	0.0007181	0.0000008
		MSE	0.0000258	0.0007769	0.0000008
	$GEBK_{TS}$	$Bias$	0.0022672	0.0125833	0.0001709
		$Variance$	0.0000147	0.0006783	0.0000015
		MSE	0.0000198	0.0008366	0.0000015
	$GEBK_P$	$Bias$	0.0007112	0.0019440	−0.0000186
		$Variance$	0.0000235	0.0014798	0.0000001
		MSE	0.0000240	0.0014836	0.0000001
100	NK	$Bias$	0.0014651	0.0068481	0.0000295
		$Variance$	0.0000108	0.0003958	0.0000004
		MSE	0.0000130	0.0004427	0.0000004
	$GEBK_{TS}$	$Bias$	0.0016563	0.0122393	0.0000989
		$Variance$	0.0000074	0.0003578	0.0000007
		MSE	0.0000101	0.0005076	0.0000007
	$GEBK_P$	$Bias$	0.0003899	0.0008883	0.0000056
		$Variance$	0.0000105	0.0007592	0.0000005
		MSE	0.0000106	0.0007600	0.0000005
200	NK	$Bias$	0.0008463	0.0057157	0.0000212
		$Variance$	0.0000047	0.0002519	0.0000003
		MSE	0.0000054	0.0002845	0.0000003
	$GEBK_{TS}$	$Bias$	0.0012135	0.0089317	0.0000708
		$Variance$	0.0000029	0.0001878	0.0000003
		MSE	0.0000044	0.0002676	0.0000003
	$GEBK_P$	$Bias$	0.0003514	0.0010664	0.0000525
		$Variance$	0.0000065	0.0004216	0.0000011
		MSE	0.0000066	0.0004228	0.0000011
400	NK	$Bias$	0.0006275	0.0040279	0.0000323
		$Variance$	0.0000025	0.0001433	0.0000002
		MSE	0.0000029	0.0001595	0.0000002
	$GEBK_{TS}$	$Bias$	0.0008837	0.0068232	0.0000546
		$Variance$	0.0000014	0.0000944	0.0000001
		MSE	0.0000022	0.0001410	0.0000001
	$GEBK_P$	$Bias$	0.0003165	−0.0005211	0.0000101
		$Variance$	0.0000033	0.0002150	0.0000002
		MSE	0.0000034	0.0002153	0.0000002

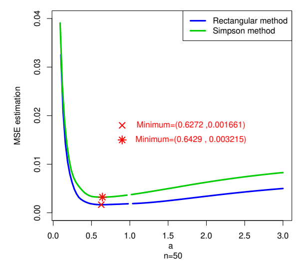

Figure 1. $\widehat{MSE}(\tilde{f}_h(0))$ as a function of a.

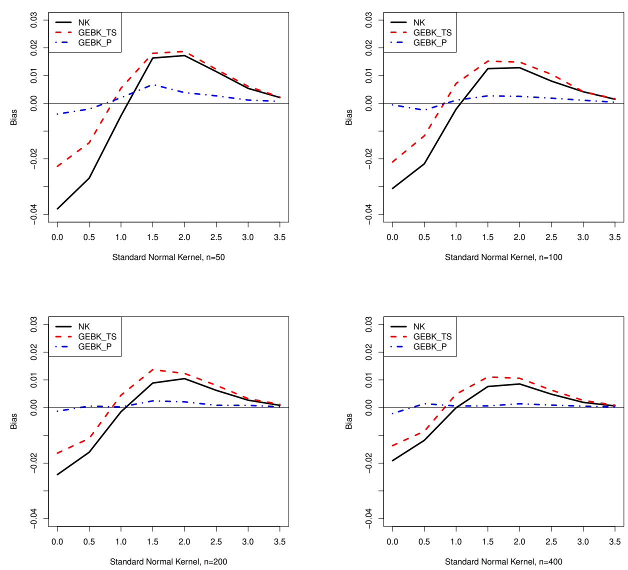

Figure 2. Bias of different kernel estimators at points $x=$ 0, 0.5, 1, 1.5, 2, 2.5, 3 and 3.5 for standard normal density.

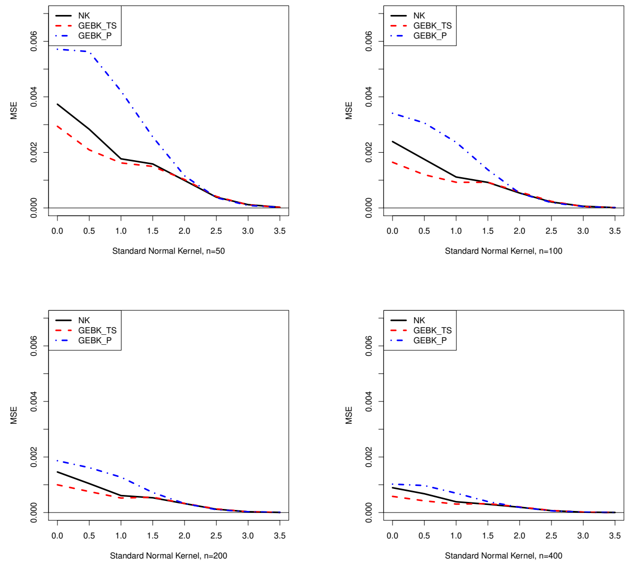

Figure 3. MSE of different kernel estimators at points $x=$ 0, 0.5, 1, 1.5, 2, 2.5, 3 and 3.5 for standard normal density.

5. Empirical illustrations

In this section, we conduct an analysis of a real data on all three old faithful geyser (Hardle [4], Azzalini and Bowman [1]), glass fibres that measured at the National Physical Laboratory, England(Souza, Santos and Cordeiro [13]) and lengths of treatment spells(in days) of control patients in suicide study(Silverman [12]) to demonstrate the use of different kernel estimators proposed in this paper. The old faithful geyser data is a data set with 272 observations and 2 variables: waiting time between eruptions and duration of eruption for the old faithful geyser in Yellowstone National Park. We estimate the density of waiting time for next eruption (in minutes) using the three different kernel estimators including NK, $GEBK_{TS}$ and $GEBK_P$. The results are given in Figure 4(a). From Figure 4(a) we can see that the two kernel estimators NK and $GEBK_{TS}$ perform strikingly close of the curve. While kernel estimator $GEBK_P$ has the best fit to the data among the other two estimators. For Figures 4(b) and 4(c), also, similar discussion may be repeated.

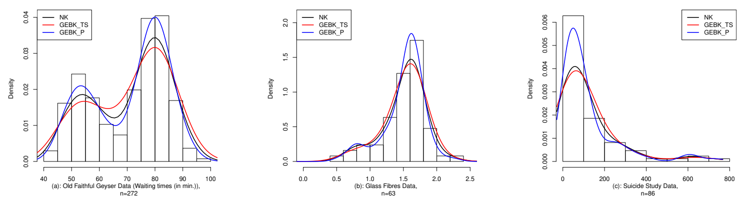

Figure 4. Different kernel density estimates from real data imposed on a histogram of the data.

6. Conclusions

This paper aims to improve the bias reduction by introducing the new kernel density estimator. In section 2, we introduce a very intuitive and feasible kernel density estimator which name naive kernel(NK) density estimator that its bias is of the order $O(h^3)$ and the optimal MSE is of the order $O(n^{\frac{-4}{5}})$. In section 3, we first recall Terrell and Scott estimator($GEBK_{TS}$). Then, we propose a feasible kernel density estimator which reduces the bias significantly, namely the geometric extrapolation based kernel($GEBK_P$) density estimator. We calculate theoretically the bias, variance and MSE and found that the bias is of the order $O(h^4)$ and the optimal MSE is of the order $O(n^{\frac{-8}{9}})$. It seems that $GEBK_P$ estimator doesn't improve over the $GEBK_{TS}$ one. However, considering Figure 2, the optimized bias in $GEBK_P$ is less, practically, than those of NK and $GEBK_{TS}$ estimators. Our simulation study at Tables 1-8 and Figure 2 shows that for finite sample size both estimators NK and $GEBK_{TS}$ perform approximately similar. When the two estimators with geometric extrapolation are compared, $GEBK_P$ estimator generally has smaller bias than $GEBK_{TS}$ estimator. Accordingly, the bias of the $GEBK_P$ estimator is reduced significantly. Therefore, there is no doubt that the $GEBK_P$ estimator is consistently the best estimator comparing with NK and $GEBK_{TS}$ estimators. In addition, MSE of the $GEBK_P$ estimator has insignificant difference when compared with the NK and $GEBK_{TS}$ estimators, especially when sample size is small(Tables 1-8 and Figure 3).

Acknowledgement

The authors would like to thank the Editor and the referees for careful reading and comments which greatly improved the paper.

References

[1] Azzalini, A. and Bowman, A. W. *A look at some data on the Old Faithful geyser*, Applied Statistics **39**, 357-365, 1990.

[2] Burden, Richard L. and Faires, J. Douglas *Numerical analysis.* 9th ed., Brooks/Cole, 2010.

[3] Farrell, R.H. *On the best obtainable asymptotic rates of convergence in estimation of a density function at a point*, The Annals of Mathematics and Statistics, **43**, 170-180, 1972.

[4] Hardle, W. *Smoothing techniques: with implementation in S*, Springer-Verlag Newyork, 1991.

[5] Igarashi G. and Kakizawa Y. *Bias corrections for some asymmetric kernel estimators*, Journal of Statistical Planning and Inference, **159**, 37-63, 2015.

[6] Kairat T. Mynbaev, Saralees Nadarajah, Christopher S. Withers, Aziza S. Aipenova, *Improving bias in kernel density estimation*, Statistics and Probability Letters, **94**, 106-112, 2014.

[7] Kim, C., Kim, W. and Park, B.U. *Skewing and generalized Jackknifing in kernel density estimation*, Communications in Statistics: Theory and Methods, **32**, 2153-2162, 2003.

[8] Kim, J. and Kim, C. *Reducing the mean squared error in kernel density estimation*. Journal of the Korean Statistical Society, **42**, 387-397, 2013.

[9] Mynbaev, K. and Martins-Filho, C. *Bias reduction in kernel density estimation via Lipschitz condition*. Journal of Nonparametric Statistics, **22**, 219-235, 2010.

[10] Parzen E. *On estimation of a probability density function and mode*, Ann. Math. Stat. **33**(3), 1065-1076, 1962.

[11] Rosenblatt M. *Remarks on some nonparametric estimates of a density function*, Ann. Math. Stat. **27**(3), 832–837, 1956.

[12] Silverman B.W. *Density estimation for statistics and data analysis*, first ed., Chapman and Hall, USA. 1986.

[13] Souza, W. B., Santos, A. H. S., and Cordeiro, G. M., *The beta generalized exponential distribution*, Journal of Statistical Computation and Simulation, **80**, 159-172, 2010.

[14] Terrell, G.R. and Scott, D.W. *On improving convergence rates for nonnegative kernel density estimators*. The Annals of Statistics, **8**, 1160-1163, 1980.

[15] Wand, M.P. and Jones, M.C. *Kernel smoothing*. Chapman & Hall, London, 1995.