Journal of the Nigerian Association of Mathematical Physics
Volume **11 (November 2007), 497 - 500**

Generalized biases in nonsymmetric univariate kernels

[1]J. E. Osemwenkhae and [2]A. O. Isere
[1]*Department of Mathematics, University of Benin, Nigeria,*
[2]*Department of Mathematics and Statistics, Ambrose Alli University, Nigeria*
[1]**e-mail: josemwenkhae@yahoo.com**

Abstract

This work extends and generalizes biases in nonsymmetric kernels. The practice of obtaining biases of any nonsymmetric kernel when the order of the smoothing parameter, h, is one is seen not to be sufficient as the error size for this case is large. A new scheme for higher order biases in nonsymmetric univariate kernels is proposed. This scheme enjoys not only the possibility of reducing the size of the global error term (MISE), but also generalizes the bias term for any nonsymmetric kernels.

Keywords: nonsymmetric kernel, higher order biases, global error, smoothing parameter.

1.0 Introduction

Kernel estimation came to limelight by the work of Fix and Hodges (1951) [7] and Rosenblatt (1956) [20]. Since then the concept has received a lot of attentions because of its wide applicability to many areas of human endeavour – see Silverman (1986) [23], Wand and Jones (1995) [26], Devroye and Lugosi (1997, 2001 [4,5]), Baxter et al (2000) [2], Dinardo and Tobias (2001) [6], Kim and Heo (2002) [9], Osemwenkhae (2003) [12], Osemwenkhae and Ishiekwene (2006) [13], etc.

Generally, let $X_1, X_2, ..., X_n$ be a random sample from a density function *f*; then the kernel density estimator $\hat{f}$ at *x* is given by:

$$\hat{f}(x,h) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right) \qquad (1.1)$$

where the kernel function *k(.)* is a probability density function (*pdf*) and *h* is the window width or the smoothing parameter. All the analytic properties of *k(.)* is also inherited by $\hat{f}(.)$. The choice of the smoothing parameter has been seen to be crucial – see for examples, Silverman (1986), Hansen (2003), Osemwenkhae and Ogbonmwan (2003), Sheather (2004), etc.

Basically, the most general way of placing a measure on the global accuracy of $\hat{f}$ in (1.1) as an estimate for *f over* all possible data set is the mean integrated squared error (MISE) and is usually defined as:

$$\text{MISE}\,(\hat{f}(x)) = \int bias^2\ \hat{f}(x)dx + \int Var\ \hat{f}(x)\,dx \qquad (1.2)$$

Silverman (1986[23]) showed that for any symmetric kernel of order *2* then (1.2) can be expressed as

$$MISE\hat{f}(x) \approx \frac{1}{4}h^4\ V_2^2 \int f^{//}(x)^2 dx + n^{-1}h^{-1}\int k(t)^2 dt \qquad (1.3)$$

The work of Jones and Signorini (1997) and Osemwenkhae (2003 [12]) extended this to when h is of order *m (for any* even *m)* and obtained that if $k(.)$ is the kernel, then

$$\text{MISE } \hat{f}(x) \approx \left(\frac{1}{(m)!}\right)^2 h^{2m} V_m^2 \int f^{(m)}(x)^2 dx + n^{-1}h^{-1}\int k(t)^2 dt \qquad (1.4a)$$

$$MISE\ \hat{f}(x) \approx \frac{2m+1}{2m}\left[\frac{2m}{(m!)^2} V_m^2 \left\{\int k(t)^2 dt\right\}^{2m} \int f^{(m)}(x)^2 dx\right]^{\frac{1}{2m+1}} \cdot n^{-\frac{2m}{2m+1}} \qquad (1.4b)$$

The simplifications in (1.3) and (1.4b) are necessitated by the assumption of symmetry of $k(.)$ - see Silverman (1986) [23], Simonoff (1996 [24]), Wand and Jones (1995 [26]), Osemwenkhae and Oyegue (2006 [19]) and Osemwenkhae and Odiase (2006a, b [15, 16]).

In this work the general assumption of symmetry of the kernel function that has simplified most of the available literature would be dropped. Hence, *k(.)* can take any of the following nonsymmetric kernels: the exponential, the weibull, the gamma, the chi-square and the Snedecor F-distributions. These densities have been seen to be very useful in the field of engineering and the sciences – see Devore (1991 [4]), Mugdadi and Lahrech (2004 [11]), Rohatgi (1984 [20]) and Barlow and Proschan (1981). So, the need to investigate how the global error (MISE) size could behave at higher order biases for nonsymmetric kernels needed be examined. Fundamentally, the aim of this paper is to

(*i*) examine the bias term when the order of h is one (the fundamental case) and its consequence on the *MISE*

(*ii*) propose a generalized bias reduction technique theoretically that would help in reducing the problem(s) in (*i*) above,

(*iii*) highlight the implication (*ii*) above of this in nonsymmetric univariate kernels.

2.0 The bias of nonsymmetric kernel (when the order of *k*(.) is one)

We shall first briefly consider the situation when the smoothing parameter h, is order one as contained in Osemwenkhae and Orhionkpaiyo (2006 [18]). Suppose we define for (1.1) the following nonsymmetric conditions:

$$\left.\begin{array}{l} \text{(i)}\ \int k(t)dt = 1 \\ \\ \text{(ii)}\ V_1^2 = \int t^2 k^2(t)dt < \infty \end{array}\right\} \qquad (2.1)$$

Also assume that f' and f'' are not only continuous but also square integrable as well as $\lim_{n\to\infty} h = 0$ and $\lim_{n\to\infty} nh = \infty$. With the definition of (2.1) above, the bias and integrated variance of $\hat{f}(x)$ defined in (1.1) are respectively,

$$Bias\ \hat{f}(x) = E\hat{f}(x) - f(x) = -hf'(x)\int tk(t)dt + \text{higher order term of h}$$

$$\approx -hf'(x)V_1 \qquad (2.2)$$

and

$$\int \text{var}\ \hat{f}(x)dx \approx n^{-1}h^{-1}\int k(t)^2 dt \qquad (2.3)$$

From (2.2) the order of *the bias* term is order h^1 and invariably the optimal window width and the $MISE\hat{f}(x)$ for (2.1) are respectively;

$$h_{opt} \approx \left(\frac{\int k(t)^2 dt}{2nV_1^2 \int f'(x)^2 dx}\right)^{1/3} \qquad (2.4)$$

$$MISE\ \hat{f}(x) \approx \frac{3}{(2n)^{\frac{2}{3}}}\left\{\int k(t)^2 dt\right\}^{2/3}\left\{V_1^2 f'(x)^2 dx\right\} \tag{2.5}$$

Equations (2.2), (2.3), (2.4) and (2.5) are fundamental equations in non-symmetric kernels of order one. Clearly, the rate of convergence of (2.5) is $n^{-2/3}$. This would converge but at a very slow rate. This has been the existing pattern/style in literature. For detailed proofs of the above equations see Mugdadi and Lahrech (2004 [11]), Osemwenkhae and Izevbizua (2005 [14]) and Osemwenkhae and Orhionkpaiyo (2006 [18]). So, in the next section we shall propose a method of reducing the bias term and examine the implication of this for the global error term in any non-symmetric kernel.

2.0 Bias reduction technique (Proposed Scheme).

Suppose we modify (2.1) by imposing the following regularity conditions:

$$\left.\begin{aligned} &(i)\ \int k(t)dt = 1 \\ &(ii)\ \int tk(t)dt = \int t^2 k(t)dt = \ldots = \int t^{2m-3}k(t)dt = \int t^{2m-2}k(t)dt = 0 \\ &(iii)\ \int t^{2m-1}k(t)dt = J_{2m-1} \neq 0 \text{ for m} = 1(1)\text{r} \end{aligned}\right\} \tag{3.1}$$

The conditions stated in (3.1) are modifications and enlargements of the conditions in Osemwenkhae and Orhionkpaiyo (2006 [18]). The bias corresponding to (3.1) is obtained by taking the expansion of $\hat{f}$ (x), using Taylor series, to *(2m – 1)th* term to have

$$Bias\ \hat{f}(x) = E\hat{f}(x) - f(x) = \int h^{-1}k\left(\frac{x-y}{h}\right)f(y)dy - f(x)$$

$$= \int k(t)\{f(x-ht)dt - f(x)\}dt \quad (\ if \quad t = \frac{x-y}{h} \quad)$$

$$= -\frac{1}{(2m-1)!}h^{2m-1}f^{(2m-1)}(x)\int t^{2m-1}k(t)dt + \text{higher order term of h}$$

$$\approx -\frac{1}{(2m-1)!}h^{2m-1}f^{(2m-1)}(x)J_{2m-1} \qquad \text{for } m = 1(1)\text{r} \tag{3.2}$$

All the lower order terms in h drop out because of the nonsymmetric conditions in (3.1) above. Also,

$$\int \text{var}\ \hat{f}(x)dx \approx n^{-1}h^{-1}\int k(t)^2 dt \tag{3.3}$$

Comparing (2.2) with (3.2) shows that for this proposed scheme, the order of the bias has increased from h^1 *to* h^{2m-1}. Invariably, from $MISE\hat{f}(x) = \int Bias^2\ \hat{f}(x)dx + \int Var\hat{f}(x)ds$, and using (3.2) and (3.3) we get

$$MISE\ \hat{f}(x) \approx \frac{1}{((2m-1)!)^2}\int h^{4m-2}J_{2m-1}^2 f^{(2m-1)}(x)^2 dx + n^{-1}h^{-1}\int k(t)^2 dt \tag{3.4}$$

Invoking equation (3.5b) of Osemwenkhae and Izevbizua (2005 [14]) we obtain (3.4) explicitly as

$$MISE\hat{f}(x) \approx J_{2m-1}^{\frac{2}{4m-1}}\left\{k(t)^2 dt\right\}^{\frac{4m-2}{4m-1}}\left\{\int f^{(2m-1)}(x)dx\right\}^{\frac{1}{4m-1}} n^{-\frac{4m-2}{4m-1}}$$

$$\left\{\frac{1}{((2m-1)!)^2}\left[\frac{((2m-1)!)^2}{4m-2}\right]^{\frac{4m-2}{4m-1}} + \left[\frac{((2m-1)!)^2}{4m-2}\right]^{-1}\right\} \tag{3.5}$$

From (2.5) and (3.5), the rate of convergence of *MISE* has increased from $n^{-\frac{2}{3}}$ to $n^{-\frac{4m-2}{4m-1}}$. This

is made possible by the regularity conditions in (3.1). Invariably, the rate of convergence of the *MISE* has been enhanced by this new scheme. Apart from the size of the error term that has been reduced, conditions (3.1) has helped to generalize the bias term (see equation (3.2)) in higher order nonsymmetric kernels.

3.0 Discussion of Findings.

The order of the bias term for the fundamental case is h^1. This is the simplest form of the bias term in nonsymmetric kernels. The corresponding order of the *MISE* term is $n^{-2/3}$. This fundamental case has been the pattern/style in existing literature (see Osemwenkhae and Orhionkpaiyo, 2006 [18]). Nevertheless, for this proposed scheme, the bias is of order h^{2m-1} and the *MISE* term is of order $n^{-\frac{4m-2}{4m-1}}$ ($m = 1, 2, 3, \ldots, < \infty$). The proposed scheme has not only given us the generalized form for the bias, but the *MISE* in nonsymmetric univariate kernels has also been reduced when (3.5) is compared with (2.5).

5.0 Conclusion

The proposed scheme for higher order bias in nonsymmetric univariate kernels has not only reduced the size of the *MISE*, but has also generalized the bias term for any kernel that belongs to this special class. This is definitely a boost to researchers in general and statisticians in particular.

References

[1] Barlow, E. R and Proschan, F. (1981). Statistical theory of reliability and life testing (probability models). TO BEGIN WITH; Silver Spring MD, USA.

[2] Baxter, M. I., Beardah, C. C. and Westwood, S. (2000). Sample size and related issues in the analysis of lead isotope data. Journal of Archaeological Science 27, 973 – 980.

[3] Devore, L. J (1991). Probability and Statistics for Engineering and the Sciences. California: Brooks/Cole

[4] Devroye, L. and Lugosi, G. (1997). Nonasymptotic universal smoothing factors, kernel complexity and yatracos classes. Annals of Stat. 25, 2626 – 2637.

[5] Devroye, L. and Lugosi, G. (2001). Combinational methods in density estimation. Spinger – Verleg, New York.

[6] Dinardo, J. and Tobias, J. L. (2001). Nonparametric density and regression estimation. Journal of Economic Perspectives, 15, 153 – 158.

[7] Fix, E. and Hodges, J. L.(1951). Discriminatory analysis, nonparametric estimation: consistency properties report no 4, project no 21-49-004' USAF School of Aviation Medicine, Randoiph Field, Texas.

[8] Hansen, (2003). Exact Mean Integrated Square Error of higher – order kernel estimators. Joun. of Amer. Stat. Ass., 87, 1 – 16.

[9] Kim, K. D. and Heo, J. H. (2002). Comparative Study of flood quartiles estimation by nonparametric models. Journals of Hydrology, 260, 176 – 193.

[10] Jones, M. C. and Signorini, D. F. (1997). A comparison of higher – order bias kernel density estimates. Journ. Of American Stat. Ass; 92, 1063 – 1073.

[11] Mugdadi, R. A. and Lahrech, M. A. (2004). The exponential kernel in density estimation. Available at http://www.math.siu.edu/mugdadi/expk

[12] Osemwenkhae, J. E. (2003). Higher order forms in kernel density estimation method. Ph.D thesis, Department of Mathematics, Univ. of Benin, Nigeria.

[13] Osemwenkhae, J. E. and Ishiekwene, C. C. (2006). Application of adjustable kernel in test evaluation. Global journal of Pure and Applied Sciences, 13, 107 – 116.

[14] Osemwenkhae, J. E. and Izevbizua, O. (2005). Generalized forms in nonsymmetric univariate kernels. Nigerian Journal of Applied Science, 23, 226 – 230.

[15] Osemwenkhe J. E. and Odiase J. I. (2006a). Improving the choice of higher-order univariate kernels through bias reduction technique. Accepted for publication in the Jour. of Science and Technology, 26, 19 – 23.

[16] Osemwenkhae J. E. and Odiase, J. I (2006b). Derivation of asymptotic biases for some symmetric kernels. Accepted for publication in 'Science Focus' (the official journal. of the Faculty of Pure and Applied Sciences, Ladoke Univ. of Tech.,Ogbomoso, Nigeria), 11, 15 – 18.

[17] Osemwenkhae J. E. and Ogbonmwan (2003). Generalized efficiencies for higher order symmetric univariate kernels. Journ. of the Nig. Ass. of Math. Physics, 7, 89 – 96.

[18] Osemwenkhae J. E. and Orhionkpaiyo,B. C. (2006). The nonsymmetric univariate kernels. Global Journal of Mathematical Sciences, 5, 175 – 178.

[19] Osemwenkhae J. E. and Oyegue, F. O. (2006). Comparing bias and mean integrated square error (*MISE*) in higher order univariate kernel. Accepted for publication in the Journ. of Res. in Phy. Sci., Uyo.

[20] Rohatgi, V. J. (1984). Statistical Inference. New York: John Wiley and Sons.

[21] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. Berlin, Springer-Verlag, 181- 190.

[22] Sheather, S. J. (2004). Density Estimation. Statistical Science, 18, 588 – 597.

[23] Silverman, B. W. (1986). Density estimation for statistics and data analysis. London: Chapman and Hall.

[24] Simonoff, J. S. (1996). Smoothing methods in Statistics. New York: Springer-Verlag.

[25] Towers, S. (2002). Kernel probability density estimation. Seminar reports in particle physics, Stat. Univ. of New York at Stony Brook.

[26] Wand, M. P. and Jones, M. C. (1995). Kernel smoothing. London: Chapman and Hall.