Generalized Efficiencies for Higher Order Symmetric Univariate Kernels

J. E. Osemwenkhae and S. M Ogbonmwan
Department of Mathematics
University of Benin, Benin City

Abstract

Much work has been done in kernel density estimation especially in obtaining approximate/exact window widths, h. The Mean Integrated Square Errors (MISE) and the plotting and interpretation of contours in both univariate and multivariate kernel density estimates are now realizable through the use of fast and cheap computational devices. The problem of obtaining the efficiency of symmetric kernels, especially for cases where the order of the smoothing parameter is greater than two has not received much attention. In this work, we have provided a formula for obtaining the generalized efficiency for any higher order symmetric kernel. We hope this will generate interests in researchers.

Key words: *Kernel density estimation, Generalized efficiencies, Higher order kernels, Rectangular, Biweight and Gaussian kernels.*

pp 89 - 96

1.0 Introduction

In the last few decades, the field of nonparametric statistical analysis has broadened its appeal by the adoption of more sophisticated alternatives to traditional parametric models for exploring large amounts of univariate data without making specific distributional assumptions. One of these *all important* tools is the nonparametric density estimation.

Let $X_1, X_2, \ldots X_n$ be a random sample from a density $f(x)$; then the kernel density estimator 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 probability density function *(pdf)* which integrates to one and h is the window width or the smoothing parameter.

There seems to be much consensus that the choice of the window width could be crucial - see Taylor (1989), Silverman (1981, 1986), Ogbonmwan and Osemwenkhae (1997, 2000) since it determines the smoothness or roughness of the density. But we have observed that the efficiency of such symmetric kernel would be crucial as well. – see Silverman (1986), Ogbonmwan and Osemwenkhae (1997), and Jones and Signorini (1997). The work of Bartlett (1963) suggested the use of fourth – order kernels as a way of reducing the size of the global error term – the mean integrated square error. More recently, the work of Jones and Foster (1993), Hall and Murison (1993), Jones and Signorini (1997), etc provided more concrete reasons for this approach – amongst which are its tendency to reduce the value of the bias when estimating the size of error committed at a given order. The need to study the efficiencies of these kernels at such higher order of the window width h, is therefore imperative. The problem of obtaining the efficiency of symmetric kernels has been examined in the literature for optimal window width of order 2. The efficiencies of symmetric kernels for the cases where the order of the smoothing parameter is greater than two has not been examined. This is the main thrust of the work of this paper. The main objectives of this paper are two-folds: (i) We seek to obtain, analytically, a generalized formula for the efficiency of any symmetric kernel. (ii) Examine (with some examples), the efficiencies of some higher order symmetric kernels.

In section 2 of this paper, we discussed the existing method of measuring the efficiency of any symmetric kernel if the order of the window width is 2. We obtained in section 3, a generalized formula for the efficiency of any symmetric kernel, while section 4 examines the efficiencies of some higher order symmetric kernels. The efficiencies of some symmetric kernels for higher orders were computed and the results discussed in section 5. The graph of the results obtained is presented in figure 1. Section 6 focuses on the conclusion of this work.

2.0 The Efficiency of Symmetric Kernels When *h* is of Order 2

Recalling (see Silverman 1986), the value of *h* in (1.1) and simplifying, we have the optimal window width for when *h* is of order 2 as

$$h_{opt} \cong V_2^{-2/5}\left\{\int k(t)^2\,dt\right\}^{-1/5}\left\{\int f^{//}(x)dx\right\}^{-1/5} n^{-1/5} \tag{2.1}$$

and the corresponding Mean Integrated Square Error (MISE) for this case is

$$MISE_x\ \hat{f}(x) \cong \frac{5}{4}V_2^{2/5}\left\{\int k(t)^2\,dt\right\}^{4/5}\left\{\int f^{//}(x)^2\,dx\right\}^{1/5} n^{-4/5} = \frac{5}{4}J(k)\left\{\int f^{//}(x)^2\right\}^{1/5} n^{-4/5} \tag{2.2}$$

where

$$J(k) = V_2^{\frac{2}{5}}\left\{\int k(t)^2\,dt\right\}^{4/5},\ \text{and } V_2(k) = \int t^2 k(t)dt \text{ (see Silverman 1986)} \tag{2.3}$$

provided that the kernel function k satisfies the symmetric conditions:

$$\left.\begin{array}{ll} \text{(i)} & \int k(t)dt = 1 \\ \text{(ii)} & \int t\,k(t)dt = 0 \\ \text{(iii)} & \int t^2 k(t)dt = V_2 \neq 0 \end{array}\right\} \tag{2.4}$$

The problem of minimizing $J(k)$ then reduces to that of minimizing $\int k(t)^2\,dt$ subject to the constraint that $\int k(t)dt,\ and\ \int t^2 k(t)dt$ are both equal to one and $\int k(t)^2\,dt \neq 0$ –since $k(t)$ is acclaimed to be symmetric. This problem is solved by evoking the popular Epanechnikov kernel density denoted by

$$k_e(t) = \begin{cases} \frac{3}{4\sqrt{5}}\left(1-\frac{1}{5}t^2\right), |t| \leq 5 \\ 0 \qquad otherwise \end{cases} \tag{2.5}$$

Silverman (1986) pointed out that we can consider the efficiency of any symmetric kernel *k* by comparing it with the Epanechnikov kernel, hence the efficiency of any symmetric kernel *k* is

$$Eff(k) = \{J(k_e)/J(k)\}^{5/4} \tag{2.6a}$$

$$= \frac{3}{5\sqrt{5}}\left\{\int t^2 k(t)dt\right\}^{-\frac{1}{2}}\left\{\int k(t)^2\,dt\right\}^{-1} \tag{2.6b}$$

For the proof of (2.1) to (2.5), see Osemwenkhae (1994). Equation (2.5b) is true for only the case where *h* is of order 2. However, this result does not hold for cases where the order is 4, 6, 8 … 2m for m = 1, 2, 3,…,< ∞.

3.0 Generalized Efficiencies for Symmetric Kernels

Recalling (2.6a), the efficiency of any symmetric kernel is obtained from the expression

$$Eff(k) = \left\{\frac{J(k_e)}{J(k)}\right\}^{5/4}$$

where

$J(k) = V_2^{2/5}\left\{\int k(t)^2\,dt\right\}^{4/5}$, and $J(k_e) = V_2\left\{\int k_e(t)^2\,dt\right\}^{4/5}$, $V_2 = \int t^2 k(t)dt$, and $k_e(t)$ is as defined in (2.5) and $k(t)$ is the symmetric kernel of our choice.

When the conditions given below:

(i) $\int k(t)dt = 1$

(ii) $\int t^{2m-1}k(t)dt = 0$

(iii) $\int t^{2m}k(t)dt = V_{2m} \neq 0$, m = 1, 2, ………, < ∞

are satisfied, Ogbonmwan and Osemwenkhae (2000) showed that the Mean Integrated Square Error can be expressed as

$$MISE_{2m}\hat{f}(x) \cong \frac{4m+1}{4m}\left\{\frac{4m}{((2m)!)^2}\right\}^{\frac{1}{4m+1}} V_{2m}^{\frac{2}{4m+1}}\left\{\int k(t)^2\,dt\right\}^{\frac{4m}{4m+1}}\left\{\int f^{(2m)}(x)^2\,dt\right\}^{\frac{1}{4m+1}} n^{-\frac{4m}{4m+1}} \tag{3.1}$$

Proposition

The efficiency os any symmetric kernel k (denoted here as Eff(k$_{2m}$) is given as

$$Eff(k_{2m}) = \frac{1}{5}\left[\frac{3^{(1+2m)}}{(2m+1)(2m+3)}\right]^{\frac{1}{2m}}\left\{\int t^{2m}k(t)dt\right\}^{-\frac{1}{2m}}\left\{\int k(t)^2\,dt\right\}^{-1}$$

Proof

Recall equation (3.1) that;

$$MISE_{2m}\hat{f}(x) = \frac{4m+1}{4m}\left\{\frac{4m}{((2m)!)^2}\right\}^{\frac{1}{4m+1}} V_{2m}^{\frac{2}{4m+1}}\left\{\int k(t)^2\,dt\right\}^{\frac{4m}{4m+1}}\left\{\int f^{(2m)}(x)^2\,dt\right\}^{\frac{1}{4m+1}} n^{-\frac{4m}{4m+1}}$$

where $V_{2m} = \int t^{2m}k(t)dt$ and m = 1, 2, 3,…, < ∞ or

$$MISE_{2m}\hat{f}(x) = \frac{4m+1}{4m}\left\{\frac{4m}{((2m)!)^2}\right\}^{\frac{1}{4m+1}} J_{2m}(k)\left\{\int f^{(2m)}(x)^2\,dx\right\}^{\frac{1}{4m+1}} n^{-\frac{4m}{4m+1}}$$

where

$$J_{2m}(k) = V_{2m}^{\frac{2}{4m+1}}\left\{\int k(t)^2\,dt\right\}^{\frac{4m}{4m+1}} \tag{3.2}$$

and

$$J_{2m}(k_e) = V_{2m}^{\frac{2}{4m+1}}\left\{\int k_e(t)^2\,dt\right\}^{\frac{4m}{4m+1}} \tag{3.3}$$

with

$$V_{2m}(e) = \int t^{2m}k_e(t)dt \text{, and } k_e(t) = \begin{cases}\frac{3}{4\sqrt{5}}\left(1-\frac{1}{5}t^2\right), & |t|\le 5\\ 0 & otherwise\end{cases}$$

(where k_e is the Epanechnikov version of k_{2m}).

So that $V_{2m}(e) = \frac{3}{4\sqrt{5}}\int_{-\sqrt{5}}^{\sqrt{5}} t^{2m}\left(1-\frac{1}{5}t^2\right)dt = \left.\frac{3t^{2m+1}}{4\sqrt{5}(2m+1)} - \frac{3t^{2m+3}}{20\sqrt{5}(2m+3)}\right|_{-\sqrt{5}}^{\sqrt{5}}$

$$= \frac{6\left(\sqrt{5}\right)^{2m+1}}{4\sqrt{5}\,(2m+1)} - \frac{6\left(\sqrt{5}\right)^{2m+3}}{20\sqrt{5}(2m+3)} = \frac{6\left(\sqrt{5}\right)^{2m+1}}{4\sqrt{5}}\left\{\frac{1}{2m+1}-\frac{1}{2m+3}\right\} = \frac{3\left(\sqrt{5}\right)^{2m}}{(2m+1)(2m+3)} \tag{3.4}$$

Also,

$$\int_{-\sqrt{5}}^{\sqrt{5}} k_e(t)^2\,dt = \frac{3}{4\sqrt{5}}\int_{-\sqrt{5}}^{\sqrt{5}}\left(1-\frac{1}{5}t^2\right)dt = \frac{3\sqrt{5}}{25} \tag{3.5}$$

From (3.3), (3.4) and (3.5) we obtain

$$J_{2m}(k_e)=\left\{\frac{3\left(\sqrt{5}\right)^{2m}}{(2m+1)(2m+3)}\right\}^{\frac{2}{4m+1}}\left(\frac{3}{5\sqrt{5}}\right)^{\frac{4m}{4m+1}}$$

Therefore from 3.2 and 3.3

$$Eff\left(k_{2m}\right)=\left\{\frac{J\left(k_{2m}{}^{(e)}\right)}{J\left(k_{2m}\right)}\right\}^{\frac{4m+1}{4m}} \text{ or } Eff\left(k_{2m}\right)=\left\{\frac{\left[\frac{3\left(\sqrt{5}\right)^{2m}}{(2m+1)(2m+3)}\right]^{\frac{2}{4m+1}}\left[\frac{3}{5\sqrt{5}}\right]^{\frac{4m+1}{4m}}}{\left[V_{2m}^{\frac{2}{4m+1}}\int k(t)^2\,dt\right]^{\frac{4m}{4m+1}}}\right\}^{\frac{4m+1}{4m}}$$

$$=\frac{\left\{\left[\frac{3\left(\sqrt{5}\right)^{2m}}{(2m+1)(2m+3)}\right]^{\frac{2}{4m+1}}\left[\frac{3}{5\sqrt{5}}\right]^{\frac{4m}{4m+1}}\right\}^{\frac{4m+1}{4m}}}{V_{2m}^{\frac{1}{2m}}\left[\int k(t)^2\,dt\right]}=\frac{3^{\frac{1}{2m}}.\sqrt{5}}{(2m+1)(2m+3)}\,\frac{3}{5\sqrt{5}}\;\Big]\Big/\,V_{2m}^{\frac{1}{2m}}\left[k(t)^2\right]dt$$

$$=\frac{1}{5}\left\{\frac{3^{(1+2m)}}{(2m+1)(2m+3)}\right\}^{\frac{1}{2m}}\left\{\int t^{2m}k(t)dt\right\}^{-\frac{1}{2m}}\left\{\int k(t)^2\,dt\right\}^{-1} \tag{3.6}$$

Equation (3.6) presents a generalized formula for the efficiency of any symmetric kernels for a given order m of the window width h. Thus, the result (3.6) of this proposition is very important as the burden of considering the efficiencies for any specified even order of h has been removed.

Some specific cases of equation (3.6) are as follows: For $m = 1$, that is, if h is of order 2:

$$Eff\left(k_2\right)=\frac{3\sqrt{5}}{25}\left\{\int t^2k(t)dt\right\}^{-\frac{1}{2}}\left\{\int k(t)^2\,dt\right\}^{-1} \tag{3.7}$$

Equation (3.7) agrees with equation (3.26) of Silverman (1986 p 42).

For $m = 2$, i.e. if h is of order 4,

$$Eff\left(k_4\right)=\left(\frac{15}{7}\right)^{1/4}\left(\frac{3\sqrt{5}}{25}\right)\left\{\int t^4k(t)dt\right\}^{-\frac{1}{4}}\left\{\int k(t)^2\,dt\right\}^{-1}$$

For $m = 3$, i.e. if h is of order 6

$$Eff\left(k_6\right)=\left(\tfrac{125}{21}\right)^{1/6}\left(\tfrac{3\sqrt{5}}{25}\right)\left\{\int t^8k(t)dt\right\}^{-\frac{1}{6}}\left\{\int k(t)^2\,dt\right\}^{-1} \text{ and so on}$$

4.0 Efficiencies of Some Specific Univariate Kernels

In this section we used (3.6) to obtain the efficiencies of some specific univariate kernels. Some common univariate kernels of interest to us are the rectangular, the biweight and the Gaussian kernels. The rectangular kernel is given by $k(t) = \frac{1}{2}, \ \forall |t| \le 1$, with

$$\int_{-1}^{1} k(t)^2 \, dt = \frac{1}{2} \text{ and } \int_{-1}^{1} t^{2m} k(t) dt = \frac{1}{2m+1} \tag{4.1}$$

Equation (4.1) into (3.6) gives

$$Eff_R(k_{2m}) = \frac{2}{5}\left\{\frac{3^{(1+2m)}}{2m+3}\right\}^{\frac{1}{2m}} \tag{4.2}$$

To test (4.2), if $m = 1$,

$$Eff(k_2) = 0.9295. \tag{4.3}$$

Similarly, $m = 2$,

$$Eff(k_4) = 0.9709$$

$m = 3$

$$Eff(k_6) = 0.9992, \text{ etc.}$$

If the kernel is the biweight given as $k(t) = \frac{15}{16}\left(1-t^2\right), \ \forall |t| \le 1$

$$\int_{-1}^{1} k(t)^2 \, dt = \frac{5}{7} \text{ and } \int_{-1}^{1} t^{2m} k(t) dt = \frac{15}{(2m+1)(2m+3)(2m+5)} \tag{4.4}$$

Substituting (4.4) into (3.6) we get;

$$Eff_B(k_{2m}) = \frac{7}{5}\left\{\frac{(2m+5)\left(3^{(1+2m)}\right)}{15}\right\}^{\frac{1}{2m}} \tag{4.5}$$

In a similar way as above; (4.6)

$m = 1, \ Eff_B(k_2) = 0.9939$

$m = 2, \ Eff_B(k_4) = 0.9730$

$m = 3, \ Eff_B(k_6) = 0.9580$, etc.

For Gaussian kernel given as $k(t) = \frac{1}{\sqrt{2\pi}}, e^{-\frac{t^2}{2}}, \ \forall |t| < \infty$

$$\int_{-\infty}^{\infty} k(t)^2 \, dt = \frac{1}{2\sqrt{\pi}} \text{ and } \int_{-\infty}^{\infty} t^{2m} k(t) dt = \frac{2m}{\sqrt{\pi}} \Gamma\left(\frac{2m+1}{2}\right), m = 1, 2, \ldots, < \infty \tag{4.7}$$

∴ (4.7) into (3.6) gives

$$Eff_G(k_{2m}) = \frac{2}{5}\sqrt{\pi}\left\{\frac{3^{(1+2m)}\sqrt{\pi}}{2^m (2m+1)(2m+3)\Gamma\left(\frac{2m+1}{2}\right)}\right\} \tag{4.8}$$

To test (4.8); if (4.9)

$m = 1, \ Eff_G(k_2) = 0.9512$

$m = 2, \ Eff_G(k_4) = 0.8745$

$m = 3, \ Eff_G(k_6) = 0.8174$, etc.

Table 1: Efficiencies of some symmetric kernels of various orders

Kernels	Function K(t)	Efficiency (to 4 d.p)								
		Order 2 (h^2)	Order 4 (h^4)	Order 6 (h^6)	Order 8 (h^8)	Order 10 (h^{10})	Order 12 (h^{12})	Order 14 (h^{14})	Order 16 (h^{16})	Order 18 (h^{18})
Epanec-hnikov	$\frac{3}{4\sqrt{5}}\left(1-\frac{1}{5}t^2\right), \forall \lvert t\rvert \le \sqrt{5}$; 0 otherwise	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Rectangular	$\frac{1}{2}, \forall \lvert t\rvert < 1$; 0 otherwise	0.9295	0.9709	0.9992	1.0201	1.0362	1.0492	1.0601	1.0692	1.0770
Biweight	$\frac{15}{16}\left(1-t^2\right)^2, \forall \lvert t\rvert \le 1$; 0 otherwise	.9939	.9730	.9580	.9466	.9315	0.9251	0.9198	0.9188	0.9143
Gaussian	$\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}, \forall \lvert t\rvert \le \infty$	.9512	.8745	.8174	.7679	0.7284	0.6949	0.6660	0.6408	0.6054

Kernels	Function K(t)	Efficiency (to 4 d.p)											
		Order 20(h^{20})	Order 22(h^{22})	Order 24 (h^{24})	Order 26 (h^{26})	Order 28 (h^{28})	Order 30 (h^{30})	Order 32 (h^{32})	Order 34 (h^{34})	Order 36 (h^{36})	Order 38 (h^{38})	Order 40 (h^{40})	Order 80 (h^{80})
Epanec-hnikov	$\frac{3}{4\sqrt{5}}\left(1-\frac{1}{5}t^2\right), \forall \lvert t\rvert \le \sqrt{5}$; 0 otherwise	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Rectangular	$\frac{1}{2}, \forall \lvert t\rvert < 1$; 0 otherwise	1.0838	1.0897	1.0950	1.0997	1.040	1.1078	1.1113	1.1145	1.1175	1.1202	1.1227	1.1512
Biweight	$\frac{15}{16}\left(1-t^2\right)^2, \forall \lvert t\rvert \le 1$; 0 otherwise	0.9104	0.9069	0.9038	0.9011	0.8986	0.8963	0.8942	0.8923	0.8906	0.8889	0.8874	0.8771
Gaussian	$\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}, \forall \lvert t\rvert \le \infty$	0.5979	0.5797	0.5631	0.5448	0.5340	0.5211	0.5092	0.4980	0.4876	0.4779	0.4687	0.3743

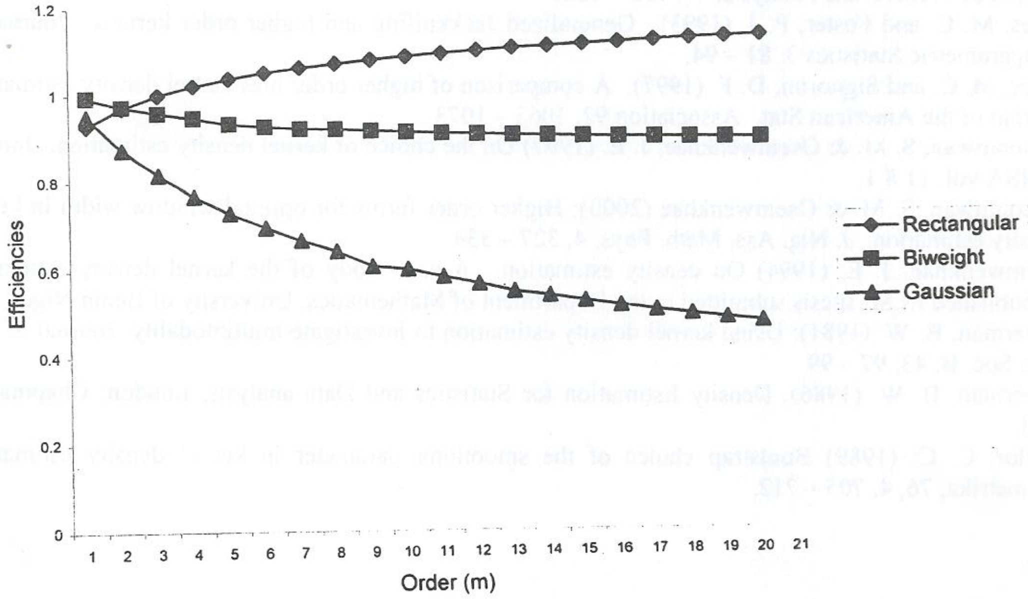

Figure 1: Graph of the efficiencies of the Rectangular, Bi-Weight and Gaussian kernels for higher order of *h*.

As stated in section 2, the pre-existing idea has been on the efficiency of any symmetric kernel if *h* is of order 2. It is worth knowing that equations (4.3), (4.6) and (4.9) agree with the table values on Table (3.1) of Silverman (1986, p. 43) on the efficiencies of those respective kernels when the order of *h* is 2. See Table 1 for higher order values of $Eff(k_{2m})$ for any of these kernels.

5.0 Discussion of Findings

In Table 1, we give the efficiencies of some kernels for higher order of *h*. We find that with higher order of *h*, the efficiencies of the rectangular kernel increased. However, the efficiencies of the Biweight, and the Guassian kernels decreased with higher order of *h*. An overview of most kernels in Table 1 showed that if the order of a kernel is 2 (as in existing literature) there is no much difference between the efficiencies of these kernels. However, if we want to go for higher order of *h* and higher order of the MISE (thereby reducing the rate of error propagation) the choice of kernels should be viewed with caution: the rectangular, and the Biweight kernels seem to perform better as their efficiencies all through the various orders do not show any much significant change. The use of the Guassian kernel should be with serious caution especially if the order of *h* is greater than 6 due to the observed sharp decay. However, if the order of *h* is greater than 6, the performance of the rectangular kernel seem too good and in fact seem better than the Epanechnikor. In contrast, after order 6, the Biweight and the Gaussian kernels perform rather badly. We conjecture that *at order 6, univariate kernels perform best with very good efficiencies and reasonably low MISE.*

6.0 Conclusion

There are lots of gains in the efficiencies of higher order symmetric kernels. Although some kernels, the Gaussian kernels as an example, is differentiable to any higher order, but its efficiency beyond the order of 6 is too discouraging. If better efficiency is required the rectangular and biweight kernels would rather be chosen than the Gaussian kernel. However, this does not mean the Gaussian kernel is not a good kernel of choice as its higher differentiability quality is an outstanding trait.

Thus, the work of this paper has brought a simplification via a generalized formula in obtaining the efficiencies of any higher order symmetric kernel. Researchers can now explore the benefits of higher order kernels based on their efficiencies.

References

[1] Bartlett, M. S. (1963). Statistical estimation of Density functions. Sankhya, Ser. A, 25, 245 254.

[2] Hall, P. and Murison, R. D (1993). Correcting the negativity of high – order kernel density estimators. Journal of Multivariate Analysis, 47, 103 – 122.

[3] Jones, M. C. and Foster, P. J. (1993). Generalized Jackknifing and higher order kernels. Journal of Nonparametric Statistics 3, 81 – 94.

[4] Jones, M. C. and Signorini, D. F. (1997). A comparison of higher order bias kernel density estimators. Journal of the American Stat. Association 92, 1063 – 1073.

[5] Ogbonmwan, S. M. & Osemwenkhae, J. E. (1997) On the choice of kernel density estimation. Journal of NSA vol. 11 # 1.

[6] Ogbonmwan, S. M. & Osemwenkhae (2000); Higher order forms for optimal window width in kernel density estimation. J. Nig. Ass. Math. Phys. 4, 327 – 334.

[7] Osemwenkhae, J. E. (1994) On density estimation. A case study of the kernel density estimation Unpublished M.Sc. thesis submitted to the Department of Mathematics, University of Benin Nigeria.

[8] Silverman, B. W. (1981): Using kernel density estimation to investigate multimodality. Journal of Roy Stat. Soc. B, 43, 97 – 99.

[9] Silverman, B. W. (1986): Density Estimation for Statistics and Data analysis, London: Chapman & Hall.

[10] Taylor, C. C: (1989) Bootstrap choice of the smoothing parameter in kernel density estimation. Biometrika, 76, 4, 705 – 712.