GLOBAL JOURNAL OF MATHEMATICAL SCIENNCES VOL. 5, NO. 2, 2006: 175-178
COPYRIGHT (C) BACHUDO SCIENCE CO. LTD. PRINTED IN NIGERIA. ISSN 1596-6208

THE NON-SYMMETRIC UNIVARIATE KERNELS.

J. E. OSEMWENKHAE and C. B. ORHIONKPAIYO

(Received 28 August, 2006; Revision Accepted 20 October, 2006)

ABSTRACT

This paper is necessitated by the presence of non-symmetric densities in most disciplines of engineering and the sciences. It examines basic defining properties of the univariate kernels when the kernel in question need not be symmetric. The paper concludes that based on these defining properties, construction of densities corresponding to any set of data drawn from a non-symmetric population is possible.

KEYWORDS: Non-symmetric kernels, univariate kernels, optimal window width.

1. INTRODUCTION

The concept of density estimation dates back to the first published work by Rosenblatt (1956) and since then this concept has received a lot of attention. The Rosenblatt – Parzen kernel estimator (see Wand and Jones, 1995) is defined as:

$$\hat{f}(x)=\frac{1}{nh}\sum k\left(\frac{x-x_i}{h}\right) \tag{1.1}$$

where k(.) is a kernel function and h is the window width, smoothing parameter or simply the kernel size. Traditionally, $\int k(x)dx=1$ and k is assumed to be symmetric with $k(u)=1-k(-u)$. A huge literature exists in this area of research. For detailed reference on the modifications and enlargements of (1.1) see Wand and Jones (1995), Simonoff (1996), Devroye and Lugosi (1997, 2001), Osemwenkhae (2003), Sheather (2004) and Osemwenkhae and Odiase (2006a,b). All these references assume *k(.)* is symmetric. Most of their findings and useful results are based on this symmetric property. Nevertheless, we have seen in real practice that some useful probability functions do not really exhibit this symmetric character. The exponential, χ^2 and the Snedecor F-distributions are some examples of non-symmetric kernels (see Rohatgi, 1984). The above named families of distributions provide probability models that are very useful in engineering and science disciplines, see Devore (1991) and Mugdadi and Lahrech (2004). For instance, the memoryless property of most life length of organisms or pieces of equipment as well as the waiting time a customer spends in a restaurant before he is served are some distinctive uses of the exponential distributions, see Rohatgi (1984, pp405).

In this work, our attempt will be to obtain some desirable properties of (1.1) when the kernel is non-symmetric. To examine this some desirable properties that would be of interest to us are:

(i) the optimal window width for this case.
(ii) the bias for this case,
(iii) the variance for this case,
(iv) the mean integrated square error (MISE) based on (i) – (iii) and
(v) expressions for (i) & (iv) when the kernel function is the exponential.

2 THE NON-SYMMETRIC KERNEL

Suppose we define for (1.1) the following non-symmetric conditions:

$$\left.\begin{aligned}&\text{(i)}\quad \int k(t)dt=1\\ &\text{(ii)}\quad V_1^2=\int t^2k^2(t)dt<\infty\end{aligned}\right\} \tag{2.1}$$

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, we shall find the bias and the variance of $\hat{f}(x)$ defined in (1.1)

J. E. Osemwenkhae, Department of Mathematics, University of Benin, Benin City.
C. B. Orhionkpaiyo, Dept. of Computer Sci. and Information Tech./Mathematics, Igbinedion Univ. Okada, Benin City.

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

$$= \int k(t)\{f(x-ht)dt - f(x)\}dt$$

$$= -hf'(x)\int tk(t)dt + \text{higher order term of h}$$
$$\approx -hf'(x)V_1 \qquad (2.2)$$

Similarly,

$$\text{var}\ \hat{f}(x) = n^{-1}h^{-1}\int f(x-ht)k(t)^2\,dt - n^{-1}\left(f(x) + 0(h^2)\right)^2$$

$$= n^{-1}h^{-1}f(x)\int k(t)^2\,dt + 0(n^{-1})$$

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

From (2.2) and (2.3) and with $MISE\ \hat{f}(x) = \int Bias^2\ \hat{f}(x)dx + \int Var\ \hat{f}(x)dx$ **(see Silverman, 1986),**

$$MISE\ \hat{f}(x) \approx h^2 f'(x)^2 V_1^2 + n^{-1}h^{-1}\int k(t)^2\,dt \qquad (2.4)$$

Differentiating (2.4) and solving for h, we get;

$$h_{opt} \approx \left(\frac{\int k(t)^2\,dt}{2nV_1^2\int f'(x)^2\,dx}\right)^{1/3} \qquad (2.5)$$

(2.5) is the equation of the optimal window width h for any non-symmetric kernel k.

The MISE in (2.4) will be explicitly obtained by substituting (2.5) into (2.4) to obtain

$$MISE\ \hat{f}(x) \approx \left\{\int k(t)^2\,dt\right\}^{2/3}\left\{V_1^2\int f'(x)^2\,dx\right\}^{\frac{1}{3}}\left\{(2n)^{-\frac{2}{3}} + \frac{2}{n^{2/3}}\right\}$$

$$= \frac{1+2^{\frac{5}{3}}}{(2n)^{2/3}}\left\{\int k(t)^2\,dt\right\}^{\frac{2}{3}}\left\{V_1^2 f'(x)^2\,dx\right\} \qquad (2.6)$$

Equations (2.2), (2.3), (2.5) and (2.6) are very vital in non-symmetric kernels of order one and would be used in assessing the general behaviour of kernels that are not symmetric.

3. TEST KERNEL

If we take the exponential kernel, $k(t) = e^{-t}$, $0 < t < \infty$ as our non-symmertic kernel, then,

$$V_1^2 = \int_{-\infty}^{\infty} t^2 k(t)^2\,dt = -\frac{1}{2}\left[t^2 e^{-2t} + te^{-2t} + \frac{1}{2}e^{-2t}\right] = \frac{1}{4} \qquad (3.1)$$

Also, $\int_0^{\infty} k(t)^2\,dt = \frac{1}{2}$ (3.2)

From (3.1) and (3.2), the h_{opt} corresponding to (2.5) would be

$$h_{opt} \approx \left(\frac{1}{4n\int f'(x)^2\,dx}\right)^{1/3} \qquad (3.3)$$

The MISE in (2.6) with $k(t) = e^{-t}$ can be expressed as:

$$MISE\ \hat{f}(x) \approx \frac{1+2^{5/3}}{(2n)^{2/3}}\left\{\left(\frac{1}{4}\right)^{2/3}\left(\int f'(x)^2\,dx\right)^{1/3}\right\}$$

$$=\frac{1+2^{5/3}}{(8n)^{2/3}}\left\{\int\left(f'(x)^2\right)^{1/3}dx\right\} \qquad (3.4)$$

If the X_i's are sufficiently large and independent then $f(x)$ could be taken to be the normal distribution via the central limit theorem – Towers (2002). If we proceed as in Silverman (1986) and define

$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, then

$$I=f'(x)=-\frac{(x-\mu)}{\sigma^3\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\therefore\ \int I^2dx=\frac{1}{2\pi\sigma^3}\Gamma\left(\frac{3}{2}\right)=\frac{\sqrt{\pi}}{4\pi\sigma^3} \qquad (3.5)$$

Therefore, the optimal smoothing parameter (3.3) becomes,

$$h_{opt}\approx\left\{\frac{\pi\sigma^3}{n\sqrt{\pi}}\right\}^{1/3}=\pi^{1/6}\sigma n^{-1/3} = 1.21\sigma\ n^{-\frac{1}{3}} \qquad (3.6)$$

So, the 'ideal' window width corresponding to the exponential kernel if *h* is of order 1, is (3.6). The symmetric counterpart of order 2 is given as $h_{opt}\approx 1.06\sigma n^{-\frac{1}{6}}$ (Silverman 1986, pp45).

To obtain the corresponding MISE, substitute (3.5) into (3.4) to get:

$$MISE\ \hat{f}(x)\approx\frac{\left(1+2^{5/3}\right)\pi^{2/3}}{4n^{2/3}\cdot(4\pi)^{1/3}}\sigma^{-1}=\left(1+2^{5/3}\right)4^{-4/3}\pi^{1/3}\cdot n^{-2/3}\sigma^{-1} \qquad (3.7)$$

From equations (3.6) and (3.7), the optimal smoothing parameter, h_{opt}, and $MISE\ \hat{f}(x)$ are obtained from $x_1, x_2, ..., x_n$ by substituting for *n* (sample size) and σ (the sample standard deviation).

The illustration below shall be based on the memorlyless property of the exponential distributions (see Rohatgi, 1984). We shall apply this concept to the duration (in months) of 200 fairly used car reams on our roads, before they failed some specific gauge test. The data was obtained from a mechanic workshop and spare part station in an urban city in Nigeria. The data gave a mean of 20 months and a standard deviation of 5 months. See Appendix 1 for this data. Applying (3.6) and (3.7), we obtained that for this set of data that $h_{opt}\approx 1.035$ and $MISE\ \hat{f}(x)\approx 5.631\times10^{-3}$.

These result suggest that to construct the density corresponding to this set of data, the optimal window width, *h*, would approximately be 1.035 and for this *h*, the size of the global error between *f* and $\hat{f}$ would approximately be 5.631E-03.

4. CONCLUDING REMARKS

Some defining properties of any univariate non-symmetric kernels were considered and found that these basic defining properties of any kernel (bias, variance, MISE and h_{opt}) as well as their mathematical expressions were not only peculiar to symmetric kernels but could also be obtained for these distributions. These defining properties were also considered specifically for the exponential kernel and found to be useful in constructing the shape of this kernel.

REFERENCES

Devore, L. J., 1991. Probability and Statistics for Engineering and the Sciences. California: Brooks/Cole

Devroye, L. and Lugosi, G., 1997. Nonasymptotic universal smoothing factors, kernel complexity and yatracos classes. Annals of Stat. 25: 2626 – 2637.

Devroye, L. and Lugosi, G., 2001. Combinational methods in density estimation. Spinger – Verleg, New York.

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

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

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 (Ghana)

Osemwenkhae J. E. and Odiase, J. I., 2006b. Derivation of asymptotic biases for some symmetric kernels. Accepted for publication in 'Science Focus' (the official journ. of the Faculty of Pure and Applied Sciences, Ladoke Univ. of Tech.,Ogbomoso, Nigeria

Rohatgi, V. J., 1984. Statistical Inference. New York: John Wiley and Sons.

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

Sheather, S. J., 2004. Density Estimation. Statistical Science, 18, 588 – 597.

Silverman, B. W., 1986. Density estimation for statistics and data analysis. London: Chapman and Hall.

Simonoff, J. S., 1996. Smoothing methods in Statistics. New York: Springer-Verlag.

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

Wand, M. P. and Jones, M. C., 1995. Kernel smoothing. London: Chapman and Hall.

APPENDIX 1

DURATION OF 200 FAIRLY USED CAR REAMS ON OUR ROADS.

The duration (in months) of 200 fairly used car reams on our roads before they failed some specific gauge test is displayed below. The data was obtained from a mechanic workshop and spare part station in an urban city in Nigeria.

18.5	18.4	12.8	23.5	31.0	17.4	25.2	20.9	26.8	21.3
13.6	18.1	15.8	21.6	27.2	24.2	20.7	22.7	28.4	27.4
21.2	26.7	12.4	15.3	26.5	22.6	25.7	18.8	21.2	20.4
26.4	19.6	18.2	18.8	20.6	17.0	19.3	13.9	20.7	14.4
26.0	19.1	19.8	20.7	20.0	26.5	16.1	26.3	29.3	13.1
28.7	17.4	20.1	22.8	22.3	11.2	25.4	18.6	19.8	21.0
9.1	29.9	18.4	20.7	19.9	22.8	17.1	13.5	16.7	24.8
18.8	24.3	31.0	15.4	14.7	19.4	22.7	23.8	24.5	12.5
25.5	31.9	11.3	29.4	11.1	20.2	22.7	23.9	20.0	17.0
14.6	16.7	16.3	22.4	24.1	16.7	18.4	22.1	25.2	22.6
16.5	28.3	7.1	20.4	22.2	17.3	17.8	22.0	22.2	18.9
11.5	11.9	27.2	24.1	23.1	24.2	13.2	16.8	27.3	18.3
10.8	22.7	13.6	24.3	21.1	24.0	30.0	23.6	19.4	19.2
15.1	24.5	16.7	16.8	14.9	22.3	17.2	22.6	15.2	17.8
16.1	29.6	23.8	15.4	26.2	23.5	20.4	25.4	12.3	16.0
9.4	19.6	22.3	25.6	18.4	28.2	18.8	6.2	7.2	23.2
17.2	17.4	24.4	14.0	15.8	21.5	34.2	22.3	20.5	20.3
18.0	23.4	23.0	12.2	15.9	22.9	26.3	27.3	21.5	26.3
20.7	18.1	13.1	23.6	17.9	29.3	24.4	11.4	19.9	21.3
18.2	23.8	14.4	23.2	17.7	18.3	26.7	20.2	21.0	20.3