

5-1-2004

Some Improvements in Kernel Estimation Using Line Transect Sampling

Omar M. Eidous

Yarmouk University, Irbid, Jordan, omarm@yu.edu.jo

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Eidous, Omar M. (2004) "Some Improvements in Kernel Estimation Using Line Transect Sampling," *Journal of Modern Applied Statistical Methods*: Vol. 3 : Iss. 1 , Article 15.

DOI: 10.22237/jmasm/1083370500

Available at: <http://digitalcommons.wayne.edu/jmasm/vol3/iss1/15>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Some Improvements in Kernel Estimation Using Line Transect Sampling

Omar M. Eidous

Department of Statistics, Faculty of Science
Yarmouk University

Kernel estimation provides a nonparametric estimate of the probability density function from which a set of data is drawn. This article proposes a method to choose a reference density in bandwidth calculation for kernel estimator using line transect sampling. The method based on testing the shoulder condition, if the shoulder condition seems to be valid using as reference the half normal density, while if the shoulder condition does not seem to be valid, we will use exponential reference density. Accordingly, the performances of the resultant estimator are studied under a wide range of underlying models using simulation techniques. The results demonstrate the improvements that can be obtained by applying this technique.

Key words: Line transect method; kernel density estimation; shoulder condition

Introduction

Line transect sampling is an important technique to estimate population density D of objects in a given region. In line transect sampling an experimenter moves across the region following a specific line with length L looking to the right and to the left of the line and records the perpendicular distance (X_i) from each detected object to the centerline. Assume that n objects has been sighted and the objects on the transect line are seen with probability one. Burnham and Anderson (1976) introduced the fundamental relation for estimating the density of objects D satisfies the following relationship

$$D = \frac{E(n)f(0)}{2L},$$

where $f(x)$ is the conditional density of the line transect distances, given the object is observed. In order to estimate D , one needs to estimate $f(0)$ which is the crucial problem in line transect estimation. When $f(0)$ estimate by an appropriate estimator $\hat{f}(0)$, D can be estimated by

$$\hat{D} = \frac{n\hat{f}(0)}{2L}.$$

Hence, the key aspects in line transect sampling turns out to be modeling $f(x)$ as well as the estimation of $f(0)$.

Various methods have been proposed to estimate $f(0)$ in literature. A parametric approach assuming that $f(x)$ is a member of a family of proper probability density function of known functional form but depend on an unknown parameter(s) θ , where θ may take a vector value and should be estimated by using the perpendicular distances. Estimate θ by $\hat{\theta}$ will lead to $\hat{f}(0) = f(0, \hat{\theta})$, and there is extensive literature on the use of the maximum likelihood techniques for estimation of $f(0)$. See for example, Burnham and Anderson

Dr. Omar Eidous is an Assistant Professor of Applied Statistics. His research interests are in line transect sampling and kernel methods. Email: omarm@yu.edu.jo

(1976); Pollock (1978); Burnham et al. (1980) and Buckland (1985).

To remove the model-dependence of the estimator, nonparametric approaches to estimate $f(0)$ can be implemented. A Fourier series is a nonparametric method has been studied in details by Burnham et al. (1980). Recent works has focused on employing the nonparametric kernel method. Some initial efforts in applying the kernel method to line transect sampling have been made by Buckland (1992); Chen (1996) and Mack and Quang (1998).

It has been widely regarded that the performance of the kernel methods depends largely on the smoothing parameter (bandwidth), and depends very little on the form of the kernel (Silverman, 1986), see also the latest three works mentioned above. In this paper we suggest a new estimator for $f(0)$.

The estimator is developed based on the kernel method itself, while Mach and Quang (1998) recommended using the bandwidth referenced to half normal model; the proposed estimator using the bandwidth reference to half normal or negative exponential models depends on testing the shoulder condition. The bandwidth parameter is selected using the half-normal model as a reference when the shoulder condition is true, that is, $f'(0) = 0$, while the negative exponential model is used when the shoulder condition is not true, that is $f'(0) \neq 0$.

In other words, to apply the proposed estimator we need to test whether the dataset at hand satisfies the shoulder condition or not. The bandwidth parameter is chosen by assuming the half normal as the underlying model if the test is accepted and by assuming the negative exponential model if the test is rejected. This method is studied using the simulation technique and the resultant estimator is compared with Mack and Quang (1998)'s estimator.

Methodology

Let X_1, X_2, \dots, X_n be a random sample of perpendicular distances of size n with unknown probability density function $f(x)$. The kernel

estimator $\hat{f}_k(x)$ of $f(x)$ for $x \geq 0$ is (Chen, 1996)

$$\hat{f}_k(x) = \frac{1}{nh} \sum_{i=1}^n \left[K\left(\frac{x - X_i}{h}\right) + K\left(\frac{x + X_i}{h}\right) \right], \quad x \geq 0, \quad (1)$$

where K is a symmetric kernel function and h is the smoothing parameter usually called bandwidth, where both K and h are under the control of the user. Accordingly, the kernel estimator of $f(0)$ is given by

$$\hat{f}_k(0) = \frac{2}{nh} \sum_{i=1}^n K\left(\frac{X_i}{h}\right). \quad (2)$$

As many authors stated, it is very little to choose between the different kernel functions (See for example Silverman, 1986; Wand & Johns, 1995). The crucial problem in kernel density estimation is to select the bandwidth parameter h . The bandwidth controls the smoothness of the fitted density curve. A larger h gives smoother estimate with smaller variance and larger bias. A smaller h produces a rougher estimate with larger variance and smaller bias. One of the most common methods in nonparametric estimation is to find h that minimizing the asymptotic mean integral square error (*AMISE*) or to minimize the asymptotic mean square error (*AMSE*) which compromises between the variance and bias of the estimate. In the remaining of this section we derive the *AMSE* of $\hat{f}_k(0)$. The expected value of $\hat{f}_k(0)$ - which is given by (2)- is

$$\begin{aligned} E(\hat{f}_k(0)) &= \frac{1}{nh} E\left(K\left(\frac{X_1}{h}\right)\right). \\ &= \frac{1}{nh} \int_0^\infty K\left(\frac{x_1}{h}\right) f(x_1) dx_1 \end{aligned}$$

Suppose that the underlying probability density function $f(x)$ has a second-order derivative. Let $u = x_1/h$ and using Taylor's series to expand $f(hu)$ around zero. Then, if $h \rightarrow 0$ as $n \rightarrow \infty$,

$$E(\hat{f}_k(0)) = f(0) + 2hf'(0) \int_0^\infty uK(u)du + h^2 f''(0) \int_0^\infty u^2 K(u)du + O(h^3)$$

Suppose that the shoulder condition is true (i.e. $f'(0) = 0$), then the bias of $\hat{f}_k(0)$ is

$$\text{Bias}(\hat{f}_k(0)) = h^2 f''(0) \int_0^\infty u^2 K(u)du + O(h^3),$$

this indicates that the asymptotic bias of kernel estimator is of order $O(h^2)$ under assumption that the shoulder condition holds. If h is related to n in such a way that $h \rightarrow 0$ and $nh \rightarrow \infty$ as $n \rightarrow \infty$, then the variance of $\hat{f}_k(0)$ is

$$\begin{aligned} \text{Var}(\hat{f}_k(0)) &= \frac{4}{nh^2} \text{Var}\left(K\left(\frac{X_1}{h}\right)\right) \\ &= \frac{4f(0)}{nh} \int_0^\infty K^2(u)du + O(n^{-1}). \end{aligned}$$

It is obvious that as $nh \rightarrow \infty$, a $O(nh)^{-1}$ variance is achieved. Accordingly, the *AMSE* of $\hat{f}_k(0)$ is given by

$$\begin{aligned} \text{AMSE}(\hat{f}_k(0)) &= h^4 f''^2(0) \left(\int_0^\infty u^2 K(u)du \right)^2 \\ &+ \frac{4f(0)}{nh} \int_0^\infty K^2(u)du \end{aligned} \quad (3)$$

where the first term in the right hand side of (3) is the squared bias and the second term is the variance.

Kernel and Bandwidth Selections

Consider the $\text{AMSE}(\hat{f}_k(0))$ - that is given by (3) - as a function of h (say $g(h)$), then differentiate $g(h)$ with respect to h and equating to zero, we get

$$h = \left\{ \frac{f(0) \int_0^\infty K^2(u)du}{\left(f''(0) \int_0^\infty u^2 K(u)du \right)^2} \right\}^{1/5} n^{-1/5}. \quad (4)$$

If the kernel function is chosen as the standard normal function, that is $K_1(u) = \phi(u)$, it is at once apparent that Equation (4) further simplifies to

$$h_1 = 0.892 \left\{ \frac{f(0)}{f''^2(0)} \right\}^{1/5} n^{-1/5}. \quad (5)$$

The value of h_1 can be substituted back into (3) to give as the minimum achievable *AMSE* for $\hat{f}_k(0)$ given by

$$0.7684 [f''(0)]^{2/5} [f(0)]^{4/5} n^{-4/5}. \quad (6)$$

On the other hand, if the kernel function is chosen as the rectangular function, that is $K_2(u) = 1 - |u|$, if $|u| < 1$ and zero otherwise, (Silverman, 1986) then Equation (4) simplifies to

$$h_2 = 2.169 \left\{ \frac{f(0)}{f''^2(0)} \right\}^{1/5} n^{-1/5} \quad (7)$$

Correspondingly, if the value of h_2 is substituted back into (3) then the minimum achievable *AMSE* for $\hat{f}_k(0)$ is given by

$$0.7908 [f''(0)]^{2/5} [f(0)]^{4/5} n^{-4/5}. \quad (8)$$

Comparing (6) and (8), the two quantities has the same convergence rates as $n \rightarrow \infty$. If $n < \infty$ then (6) is slightly smaller than (8). In other words, the efficiency that can be obtained when K_2 is used instead of K_1 is less than 3% in the basis of the *AMSE*.

This conclusion supports the well known result that says, there is a very little to choose between the different kernel functions as

they all contribute the similar amounts to the *AMSE*. Actually, among the different five kernels which are given in Silverman (1986, pp. 43) if the Epanchnikov kernel is used instead of the standard normal kernel then we obtain the maximum efficiency which is less than 4.1%.

Silverman (1986) presented a table contains the efficiencies of different kernel functions with respect to Epanchnikov kernel. His comparative study is achieved on the basis of the *AMISE* instead of the *AMSE* (that we adopted here) and with data support defined on the entire of the real line, while in this study the data support is defined on the positive half of the real line. Accordingly, in the rest of this paper our derivations and computations are based on the standard normal kernel function (K_1). This kernel is differentiable and has all-order derivatives that are required.

The value of h_1 is based on the parameter $f(0)$ that we aim to estimate it. Buckland (1992) and Mack and Quang (1998) overcame this problem by assuming the half normal model as the underlying model of the data and their formula based on minimizing the *AMISE* of the kernel estimator is given by

$$h = 1.06 \sigma n^{-1/5} \tag{9}$$

where σ is estimated practically by its maximum likelihood estimator $\hat{\sigma} = \sqrt{\sum_{i=1}^n x_i^2 / n}$.

Buckland (1992) used Equation (9) for the deer data and reported very similar results to those obtained by Hermite polynomial method. Mack and Quang (1998) recommended the above formula canceling the constant term, that is, $h = \sigma n^{-1/5}$ which is slightly different from Equation (9). Chen (1996) stated that Equation (9) performs quite well when the underlying distribution is close to the half normal distribution, while when the true $f(x)$ is not close to the half normal, the result can be misleading. He suggested an alternative method called "Least Square Cross-Validation Method" (*LSCVM*).

The primary simulation results indicated that the advantage of using the *LSCVM* over

using Formula (9) is not significant despite the computer-intensive procedures that need to apply it. By interesting the last three works mentioned above and the work introduced by Zhang (2001) which was concerned the testing of the shoulder condition, we suggest to use two reference models to choose the bandwidth h . One of these two models is the negative exponential model which does not satisfy the shoulder condition at the origin, and the other is the half normal model which satisfies the shoulder condition at the origin. The criterion to choose between them is by testing the shoulder condition as illustrates in the following section.

Testing the Shoulder Condition

A motivation to assume the half normal or the negative exponential as the underlying model to apply Formula (5) is that, the first model has a shoulder at the origin, while the second one does not. In other words, we expect the reference model that should be used to choose h is the half normal model when the data have a shoulder at the origin, whereas the negative exponential should be used when the data do not have a shoulder at the origin. Accordingly, assume that we are not sure whether the data have the shoulder at the origin or not, in this case and before we decide which model should be used we need to test the shoulder condition.

Zhang (2001) proposed a procedure for testing the shoulder condition of a model based on line transect sampling. Assume that a random sample x_1, x_2, \dots, x_n of perpendicular distances is drawn from a distribution with probability density function $f(x)$. Consider the test $H_0 : f'(0) = 0$ vs $H_1 : f'(0) \neq 0$, according to Zhang (2001), we reject H_0 for

large value of
$$Z = \frac{\sqrt{\sum_{i=1}^n x_i^2}}{\sum_{i=1}^n x_i}.$$
 Zhang

constructed a table of critical values of the sampling distribution for Z with respect to different sample sizes by Monte Carlo simulation. For example, at level of significant $\alpha = 0.5$ we reject H_0 in favor of H_1 if

$Z > q$, where $q = 0.1880, 0.1308, 0.0914$ for $n = 50, 100, 200$ respectively.

Accordingly, to choose the bandwidth h we consider the following two steps

1) If H_0 is not rejected, use the half-normal model as a reference model and then compute the value of h by using the following formula

$$h = 0.933 \hat{\sigma} n^{-1/5} \quad (10)$$

2) If H_0 is rejected, use the negative exponential model as a reference model and then compute the value of h by using

$$h = 0.892 \hat{\lambda} n^{-1/5} \quad (11)$$

where $\hat{\sigma}$ is as defined in Section (3) and $\hat{\lambda}$ is the maximum likelihood estimator for the scale parameter λ in the case of the negative exponential density, which is given by

$$\hat{\lambda} = \sum_{i=1}^n x_i / n.$$

Results

To assess the practical impact of our technique, we undertook some numerical investigations in which we compared our proposed estimator with an ordinary estimator given by Mack and Quang (1998). In this numerical study we considered several parent densities. These densities are those considered by Barabesi (2001) which are commonly used in line transect studies. The exponential power family (Pollock, 1978)

$$f(x) = \frac{1}{\Gamma(1+1/\beta)} e^{-x^\beta}, \quad x \geq 0, \quad \beta \geq 1,$$

The hazard-rate family (Hayes and Buckland, 1983)

$$f(x) = \frac{1}{\Gamma(1-1/\beta)} \left(1 - e^{-x^{-\beta}}\right), \quad x \geq 0, \quad \beta > 1$$

and the beta model (Eberhardt, 1968)

$$f(x) = (1 + \beta)(1 - x)^\beta, \quad x \geq 0, \quad \beta \geq 0.$$

In our simulation design, these three families were truncated at some distance w . Four models were selected from the exponential power family with parameter values $\beta = 1.0, 1.5, 2.0, 2.5$ and corresponding truncation points given by $w = 5.0, 3.0, 2.5, 2.0$ (Figure 1a). Four models were selected from the hazard-rate family with parameter values $\beta = 1.5, 2.0, 2.5, 3.0$ and corresponding truncation points given by $w = 20, 12, 8, 6$ (Figure 1b).

Moreover, four models were selected from beta model with parameter values $\beta = 1.5, 2.0, 2.5, 3.0$ and $w = 1$ for all cases (Figure 1c). The considered models cover a wide range of perpendicular distance probability density functions which vary near zero from spike to flat. It should be remarked that the truncated exponential power model with $\beta = 1$ and the beta model do not satisfy the shoulder condition. This choice was made in order to assess the robustness of the considered estimators with respect to the shoulder condition.

For each model and for sample sizes $n = 50, 100, 200$ one thousand samples of distances were randomly drawn. For each model and for each sample size, Table (1) reports the simulated value of the relative bias (RB):

$$RB_i = \frac{E(\hat{f}(0)) - f(0)}{f(0)}, \quad i = 1, 2;$$

the relative mean error (RME)

$$RME_i = \frac{\sqrt{MSE(\hat{f}(0))}}{f(0)}, \quad i = 1, 2,$$

for each considered estimator, and the efficiency (EFF) of the proposed estimator with respect to Mack and Quang (1998)'s estimator,

$$EFF = \frac{MSE_2}{MSE_1},$$

Table (1). Relative Biases (RB) and Relative Mean Error (RME) for the proposed estimator and for the kernel estimator using bandwidth rule based on half normal model.

n	β	w	$RB1$	$RME1$	$RB2$	$RME2$	EFF
Exponential Power Model							
50			-0.245	0.283	-0.333	0.347	1.226
100	1.0	5.0	-0.217	0.238	-0.305	0.313	1.312
200			-0.191	0.207	-0.276	0.282	1.362
50			-0.134	0.201	-0.168	0.205	1.021
100	1.5	3.0	-0.101	0.159	-0.136	0.164	1.032
200			-0.079	0.127	-0.114	0.135	1.065
50			-0.067	0.160	-0.084	0.150	0.942
100	2.0	2.5	-0.059	0.124	-0.071	0.119	0.961
200			-0.044	0.099	-0.053	0.096	0.965
50			-0.047	0.144	-0.055	0.137	0.949
100	2.5	2.0	-0.022	0.119	-0.029	0.112	0.944
200			-0.023	0.091	-0.027	0.088	0.965
Hazard Rate Model							
50			-0.174	0.236	-0.402	0.417	1.765
100	1.5	20.0	-0.118	0.167	-0.354	0.363	2.174
200			-0.072	0.114	-0.306	0.311	2.730
50			-0.063	0.166	-0.247	0.276	1.658
100	2.0	12.0	-0.034	0.119	-0.206	0.225	1.890
200			-0.012	0.086	-0.159	0.172	2.007
50			-0.016	0.161	-0.119	0.173	1.077
118	2.5	8.0	-0.001	0.121	-0.083	0.124	1.032
119			0.009	0.094	-0.053	0.087	0.934
50			0.001	0.156	-0.049	0.132	0.845
100	3.0	6.0	0.000	0.118	-0.034	0.095	0.807
200			0.012	0.099	-0.011	0.073	0.737
Beta Model							
50			-0.167	0.218	-0.183	0.219	1.005
100	1.5	1.0	-0.150	0.182	-0.163	0.186	1.023
200			-0.128	0.149	-0.139	0.155	1.039
50			-0.186	0.235	-0.208	0.239	1.018
100	2.0	1.0	-0.158	0.193	-0.177	0.199	1.030
200			-0.140	0.169	-0.159	0.176	1.037
50			-0.205	0.247	-0.231	0.256	1.033
100	2.5	1.0	-0.170	0.208	-0.198	0.218	1.045
200			-0.152	0.178	-0.179	0.191	1.073
50			-0.201	0.249	-0.237	0.261	1.048
100	3.0	1.0	-0.176	0.213	-0.212	0.228	1.074
200			-0.149	0.178	-0.187	0.199	1.118

where $RB1$, $RME1$ are the RB and RME of the proposed estimator $\hat{f}_k(0)$ and $RB2$, $RME2$ are the RB and RME for the Mack and Quang's estimator $\hat{f}_M(0)$ say.

Depending on the simulation results given in Table (1), several conclusions can be drawn from examining the results in regard to model robustness (RB) and (RME). The estimator $\hat{f}_M(0)$ is with large $|RB2|$ for the exponential power model with $\beta = 1$ and for the hazard rate model with $\beta = 1.5, 2.0$, the maximum $|RB2|$ value turns out to be 0.402 for the hazard rate model with $\beta = 1.5$ (). For the exponential power and the hazard rate models, the $|RB2|$ for $\hat{f}_M(0)$ increases as the shape parameter β decreases, while it decreases as β increases for the beta model. On the other hand, the $RME2$ ranges in $[0.132, 0.417]$ if $n = 50$, in $[0.095, 0.363]$ if $n = 100$ and in $[0.073, 0.311]$ if $n = 200$.

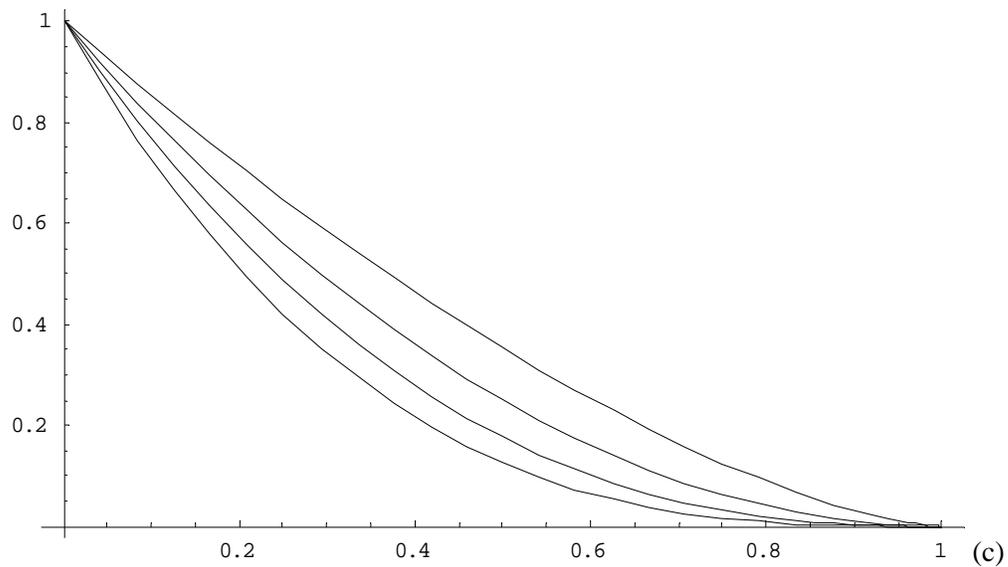
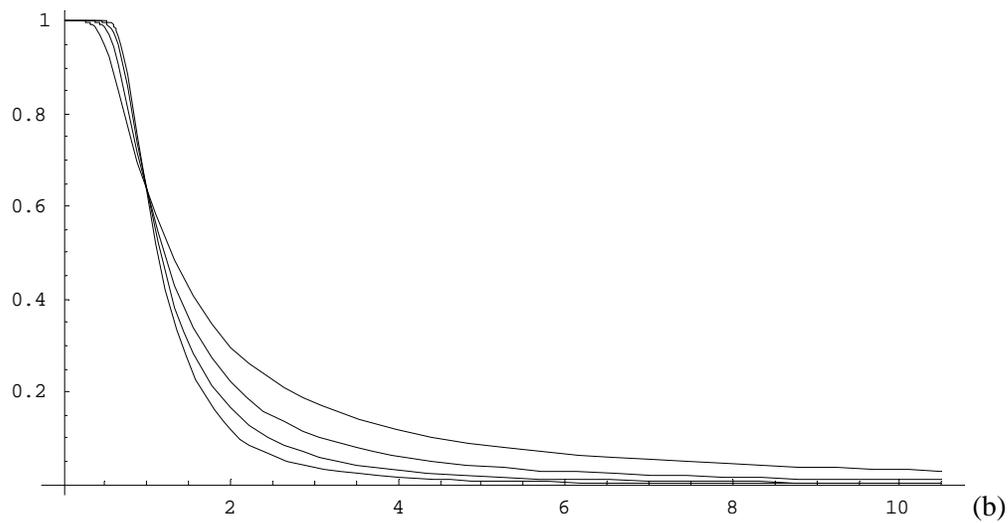
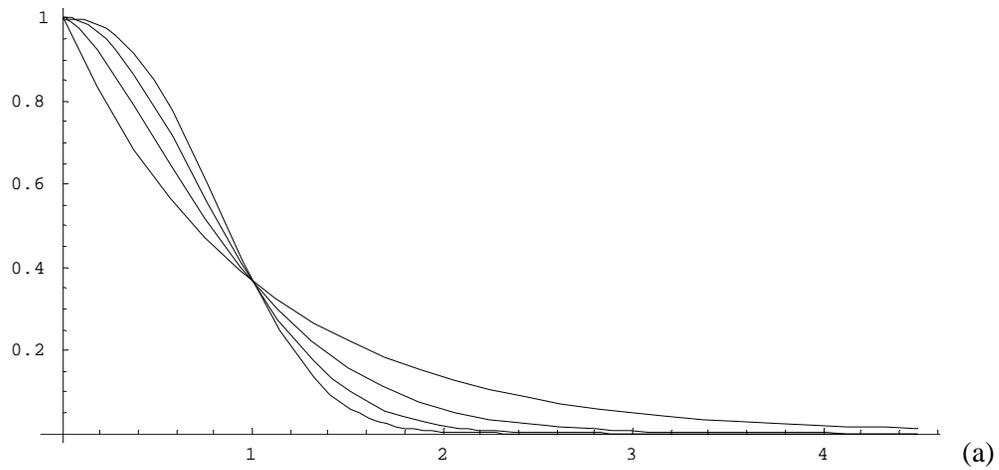
As to our proposed estimator $\hat{f}_k(0)$, it generally produces rather small $|RB1|$ s, the maximum $|RB1|$ value turns out to be 0.283 for the exponential power model with $\beta = 1.0$. Comparing the $|RB2|$ s of $\hat{f}_M(0)$ with that of $\hat{f}_h(0)$, the simulation results demonstrated clearly that the $|RB1|$ s of $\hat{f}_k(0)$ are smaller than the corresponding $|RB2|$ s of $\hat{f}_M(0)$, especially for the exponential power model with $\beta = 1.0, 1.5$ (in which the shapes are spike and has a moderate shoulder respectively); the hazard rate model with $\beta = 1.5, 2.0$ (in which, the two shapes have the shoulder at $x = 0$ but when $\beta = 2.0$ the curve drops sharply –but less than that of $\beta = 1.5$ – when we move far from $x = 0$) and for the beta model with different values of β (the shapes do not have the shoulder at $x = 0$). The different shapes of these models are depicted in Figure 1.

On the other hand, the $RME1$ ranges in $[0.144, 0.283]$ if $n = 50$, in $[0.118, 0.238]$ if $n = 100$ and in $[0.086, 0.207]$ if $n = 200$. Regarding the stability of the accurate of the two estimators, the performance of $\hat{f}_k(0)$ is more stable and hence its performance is better than that of $\hat{f}_M(0)$. The efficiency (EFF) values in Table 1 show that, for some of the models investigated, a considerable gain in the accuracy of the proposed estimator is achieved. The efficiency values increase as the sample size n increases for the exponential power model with $\beta = 1.0, 1.5, 2.0$; the hazard rate model with $\beta = 1.5, 2.0$ and for the beta model with different values of β , in the cases where the proposed estimator performs better than $\hat{f}_M(0)$. In the other cases where the shoulder condition is –in some sense– large the efficiency is less than one which indicates the performance of $\hat{f}_M(0)$ is better than $\hat{f}_k(0)$ but the efficiency remains acceptable in these cases.

Numerical Example

We apply the proposed estimator to the classical wooden stakes data set, given in Burnham *et. al.* (1980, p:61). The data are collected from line transect survey to estimate the density of stakes in a given area. The stakes data are the perpendicular distances (in meters) of detected a stake to the transect line, in which 150 stakes were placed at random in an area of 1000 meters long. Out of 150 stakes, 68 stakes are detected using line transect technique. The true form of $f(x)$ is unknown, but the true value of $f(0)$ is known which equals $f(0) = 0.110294$, thus the actual density of stakes was 37.5 stakes/ha. Calculation shows that $Z = 0.1624$, the empirical critical value for $\alpha = 0.05$ and $n = 68$ is 0.1605 (Zhang , 2001), so the shoulder condition is rejected. In this case, the formula $h = 0.892 \hat{\lambda} n^{-1/5}$ should be used, where the computed value of $\hat{\lambda} = 6.115$, so $h = 2.346$. In turn, the resulting

Figure (1). (a) Exponential power model for $\beta = 1, 1.5, 2, 2.5$. (b) Hazard-rate model for $\beta = 1.5, 2, 2.5, 3$ and (c) Beta model for $\beta = 1.5, 2, 2.5, 3$.



estimate is $\hat{f}_k(0) = 0.10463$ and $\hat{D} = 35.6$ stakes/ha. By adopting the Mack and Quang's estimator, $h = 0.933 \hat{\sigma} n^{-1/5}$, computation gives $\hat{\sigma} = 8.19$ and $h = 3.522$. In turn the resulting estimate is $\hat{f}_M(0) = 0.10005$ and $\hat{D} = 34.01$ stakes/ha. Burnham *et. al.* (1980) analyze the same data by using a cosine series estimator, and they obtain an estimate for $f(0)$ given by 0.1148 with corresponding density estimate $\hat{D} = 39.00$ stakes/ha. It should be remarked that the cosine series estimator employs an exact value for the maximum perpendicular distance (take to be 20 meters for this example), that is, more information is used in this case.

References

- Barabesi, L. (2001). Local parametric density estimation methods in line transect sampling. *Metron LIX*, 21-37.
- Buckland, S. T. (1985). Perpendicular distance models for line transect sampling. *Biometrics*, 41, 177-195.
- Buckland, S. T. (1992). Fitting density functions with polynomials. *Applied Statistics*, 41, 63-76.
- Burnham, K. P. & Anderson, D. R. (1976). Mathematical models for nonparametric inferences from line transect data. *Biometrics*, 32, 325-336.
- Burnham, K. P., Anderson, D. R. & Laake, J. L. (1980). Estimation of density from line transect sampling of biological populations. *Wildlife Monograph*, 72.
- Chen, S. X. (1996). A kernel estimate for the density of a biological population by using line-transect sampling. *Applied Statistics*, 45, 135-150.
- Eberhardt, L. L. (1968). A preliminary appraisal of line transects. *J. Wildlife. Manage.*, 32, 82-88.
- Hayes, R. J., & Buckland, S. T. (1983). Radial distance models for line-transect method. *Biometrics*, 39, 29-42.
- Mack, Y. P. and Quang, P. X. (1998). Kernel methods in line and point transect sampling. *Biometrics*, 54, 606-619.
- Pollock, K. H. (1978). A family of density estimators for line transect sampling. *Biometrics*, 34, 475-478.
- Silverman, B. W. (1986). *Density Estimation*. Chapman & Hall, London.
- Wand, M., & Jones, M. C. (1995). *Kernel smoothing*. Chapman & Hall, London.
- Zhang, S. (2001). Generalized likelihood ratio test for the shoulder condition in line transect sampling. *Communication. in Statistics Theory and Methods*, 30, 2343-2354.