Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/100127

Akaike Information Criterion and Fourth-Order Kernel Method for Line Transect Sampling (LTS)

Ali Algarni [1,*] *and Ahmad Almutlg* [2]

[1] Statistics Department, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia
[2] Mathematics Department, Faculty of Science, Qassim University, Qassim, Saudi Arabia

Received: 7 Mar. 2015, Revised: 20 May 2015, Accepted: 12 Aug. 2015
Published online: 1 Jan. 2016

Abstract: Parametric and noparametric approaches were used to fit line transect data. Different parametric detection functions are suggested to compute the smoothing parameter of the nonparametric fourth-order kernel estimator. Among the different candidate parametric detection functions, the researcher suggests to use Akaike Information Criterion (AIC) to select the most appropriate one of them to fit line transect data. More specifically, four different parametric models are considered in this research. Where as two models were taken to satisfy the shoulder condition assumption, the other two do not. Once the appropriate model is determined, it can be used to select the smoothing parameter of the nonparametric fourth-order kernel estimator. As the researcher expected, this technique leads to improve the performances of the fourth-order kernel estimator. For a wide range of target densities, a simulation study is performed to study the properties of the proposed estimators which show the superiority of the resulting proposed fourth-order kernel estimator over the classical kernel estimator in most considered cases.

Keywords: Akaike information criterion, Maximum likelihood estimation, Fourth-order kernel estimator, Exponential model, Half-normal model, Weighted Exponential model, Reversed Logistic model.

1 Introduction

The smoothing parameter h of the fourth-order kernel estimator plays a vital role in its performance. Its performance becomes acceptable compared to the classical kernel estimator when the method of determining the smoothing parameter is chosen correctly. In other words, there is a scope to improve the performances of the fourth-order kernel estimator by considering a suitable parametric method to determine h. mor survey given in Chen [1], Cline and Hart [2], Eidous [3], Eidous and bAl-Masri [4], Cowling and Hall [5], Eidous [6], Eidous and Alshakhatreh [7], Gasser and Muller [8], Gasser et al. [9], Eidous [10], Mack and Quang [11], Wand and Jones [12], Zhang and Karunamuni [13], Eidous [14].

The four parametric models will be again considered in this paper. However, choosing a suitable parametric model for the data is not dependent on guesswork, instead the (AIC) will be used for this purpose in this paper. To investigate the statistical properties of the resultant estimator, a simulation study is performed lately in research

Akaike Information Criterion: Burnham and Anderson [16] illustrated that the (AIC) is a quantitative method to select a suitable model for data. Information theory and an extension of the maximum likelihood principle using AIC given in Akaike [17]. They defined the AIC as:

$$AIC = -2\log L + 2q, \tag{1}$$

where $\log L$ is the natural logarithm of the likelihood function evaluated at the MLE of the model parameters and q is the number of parameters in the model. In this paper, the formulas for AIC will be computed for four parametric models. These parametric models are:

Exponential model

$$f(x) = \frac{1}{\theta} exp\left(\frac{-x}{\theta}\right), \quad x > 0, \tag{2}$$

Half-normal model

$$f(x) = \frac{1}{\sqrt{2\Pi\sigma^2}} exp\left(\frac{-x^2}{2\sigma^2}\right), \quad x > 0, \tag{3}$$

* Corresponding author e-mail: ahalgarni@kau.edu.sa

Weighted Exponential model

$$f(x) = \frac{2\theta}{3\log 3} exp\left(-\theta x\right)\left(2 - exp\left(-\theta x\right)\right), \quad x > 0, \quad (4)$$

Reversed Logistic model

$$\frac{2\theta exp\left(-\theta x\right)}{3\log 3\left(1 + 2exp\left(-\theta x\right)\right)}, \quad x > 0, \quad (5)$$

Also, we derived the smoothing parameter of the fourth-order kernel estimator based on each model separately.

If $X_1, X_2, \ldots, X_n$ is a random sample of size n representing the perpendicular distances and following one of the above pdf models $f(x)$ The corresponding smoothing parameters for each model are,

$$h_{\text{exp}} \cong 1.4408(\theta_1)(n^{\frac{1}{3}}), \text{ where } (\hat{\theta}_1) = \overline{x} \quad (6)$$

$$h_{\text{half}} \cong 1.0066(\sigma)(n^{\frac{1}{9}}), \text{where } \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \quad (7)$$

$$h_{\text{weight}} \cong 0.8689(\frac{1}{\theta_3})(n^{\frac{1}{7}})), \text{where } \hat{\theta}_3 = \frac{7}{6\overline{x}} \quad (8)$$

$$h_{\text{Rever}} \cong 3.5400(\frac{1}{\theta_4})(n^{\frac{1}{3}}), \text{where } \hat{\theta}_4 = \frac{1.3078}{\overline{x}} \quad (9)$$

In the following section, we derived the AIC that corresponds to each model.

2 Akaike Information Criterion for Some Parametric Models

For local likelihood density estimation in line transect sampling see Barabesi [18] and Barabesi [19] Let X_1, X_2, $\ldots$, X_n be a random sample of size n perpendicular distances following the exponential model. The likelihood function is

$$L\left(\theta_1\right) = \left(\frac{1}{\theta_1}\right)^n \exp\left(\frac{-1}{\theta_1}\sum_{i=1}^{n} x_i\right) \quad (10)$$

By taking the natural logarithm of both sides, we obtain

$$\log L\left(\theta_1\right) = \frac{-1}{\theta_1}\sum_{i=1}^{n} x_i - n\log\theta_1 \quad (11)$$

Multiply both sides of the last equation by (-2) we obtain

$$-2\log L\left(\theta_1\right) = \frac{2}{\theta_1}\sum_{i=1}^{n} x_i + 2n\log\theta_1 \quad (12)$$

Since there is one parameter that needs to be estimated, q=1. Therefore, the AIC for the exponential model is

$$\begin{aligned} \text{AIC}_{\text{exp}} &= \tfrac{2}{\hat{\theta}_1}\textstyle\sum_{i=1}^{n} x_i + 2n\log\hat{\theta}_1 + 2 \\ &= 2n\left(1 + \log\overline{x}\right) + 2. \end{aligned} \quad (13)$$

Note that $\overline{x}$is the ML estimator for θ_1.

In the same way we can derive the AIC for the other three models, which are given below:

* For the Half-normal model, the likelihood function and the natural logarithm of the likelihood function are$\hat{\sigma}$

$$L\left(\sigma^2\right) = \left(\frac{1}{\sqrt{2\Pi\sigma^2}}\right)^n \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right). \quad (14)$$

$$\log L\left(\sigma^2\right) = -n\log\sqrt{2\Pi} - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 \quad (15)$$

This gives the AIC by

$$\begin{aligned} \text{AIC}_{\text{half}} &= -2n\log 2 + n\log\left[2\pi\hat{\sigma}^2\right] + n + 2 \\ &= -2n\log 2 + n\log\left[2\pi\tfrac{\sum_{i=1}^{n} x_i^2}{n}\right] + n + 2. \end{aligned} \quad (16)$$

where $\frac{\sum_{i=1}^{n} x_i^2}{n}$ is the ML estimator of σ^2.

* For the weighted exponential model, the likelihood function and the natural logarithm of the likelihood function are

$$L\left(\theta_3\right) = \left(\frac{2\theta_3}{3}\right)^n \exp\left(-\theta_3\sum_{i=1}^{n} x_i\right)\prod_{i=1}^{n}\left(2 - \exp\left(-\theta_3 x\right)\right), \quad (17)$$

$$\begin{aligned} \log L\left(\theta_3\right) &= n\log\left(\tfrac{2\theta_3}{3}\right) - \theta_3\textstyle\sum_{i=1}^{n} x_i \\ &\quad + \log\left[\prod_{i=1}^{n}\left(2 - \exp\left(-\theta_3 x\right)\right)\right], \end{aligned} \quad (18)$$

The AIC is

$$\begin{aligned} \text{AIC}_{\text{weight}} &= -2n\log\left(\tfrac{2\hat{\theta}_3}{3}\right) + 2n\hat{\theta}_3 \\ &\quad -2\textstyle\sum_{i=1}^{n}\log\left(2 - \exp\left(-\hat{\theta}_3 x\right)\right) + 2. \end{aligned} \quad (19)$$

where $\hat{\theta}_3$ is the ML estimator of θ_3, which does not exist in closed form and a numerical method such as the Newton-Raphson method is needed to obtain the value of θ_3.

* Finally, for the reversed logistic model, the likelihood function and the natural logarithm of the likelihood function are

$$L\left(\theta_4\right) = \left(\frac{2\theta_4}{\log 3}\right)^n \frac{\exp\left(-\theta_4\sum_{i=1}^{n} x_i\right)}{\prod_{i=1}^{n}\left(1 + 2\exp\left(-\theta_4 x\right)\right)} \quad (20)$$

$$\begin{aligned} \log L\left(\theta_4\right) &= n\log\left(\tfrac{2\theta_4}{\log 3}\right) - \theta_4\textstyle\sum_{i=1}^{n} x_i \\ &= -\log\left[\prod_{i=1}^{n}\left(1 + 2\exp\left(-\theta_4 x\right)\right)\right]. \end{aligned} \quad (21)$$

The AIC is

$$\begin{aligned} \text{AIC}_{\text{Revr}} &= -2n\log\left(2\hat{\theta}_4\right) + 2n\log\left(3\right) + 2n\hat{\theta}_4\overline{x} \\ &\quad -2\textstyle\sum_{i=1}^{n}\log\left(1 + 2\exp\left(-\hat{\theta}_4 x\right)\right) + 2. \end{aligned} \quad (22)$$

where $\hat{\theta}_4$ is the ML estimator of θ_4. Again, a closed form for the value of $\hat{\theta}_4$ does not exist in closed form and a numerical method is needed to obtain it.

3 Simulation Study Design

we performed a simulation study to investigate the properties of the proposed estimator. This estimator is the fourth-order kernel whose parameter h can be computed by considering any reasonable parametric model. One of the four parametric models mentioned in Section (3.2) of this chapter is now used to compute h. The AIC is a criterion for choosing the most appropriate model based on the data. The resulting estimator is denoted by $\hat{f}(0)$. The proposed estimator $\hat{f}(0)$ is computed as follows:

1. Simulate the n perpendicular distances $X_1, X_2, \ldots, X_n$ from one of the target models mentioned in Section (2.6).
2. For the n simulated perpendicular distances, compute the value of h that corresponds to the exponential, half-normal, weighted exponential, and reversed logistic models. The values of these smoothing parameters are denoted by h_{exp}, h_{half}, h_{weight}, and h_{Rever}, respectively.
3. Compute the AIC for the exponential, half-normal, weighted exponential, and reversed logistic models.
4. Select the model with the smallest AIC as a reference model to compute the value of h based on the selected model. The formula for computing h for each model are given in Section (3.2).
5. Compute the value of the fourth-order kernel estimator based on the selected h of step (4) and based on the perpendicular distances $X_1, X_2, \ldots, X_n$.

The data are simulated from the 16 models that were given in Section (2.6) with the same values of n, β, and ϖ. For a simple comparison, the results of a classical kernel estimator $\hat{f}_k(0)$ are also presented. The relative bias (RB), relative mean error (RMX), and the efficiency (EFF) of $f_k^*(0)$ with respect to $\hat{f}_k(0)$ are given in Tables (3.1-3.4). Note that,

$$\mathrm{EFF} = \frac{\mathrm{MSE}(\hat{f}_k(0))}{\mathrm{MSE}(f_k^*(0))} \quad (23)$$

as illustrated in Section (2).

4 Results

Depending on the simulation results of Tables (3.1-3.4), several conclusions can be drawn by inspecting the results with regard to , , and EFF.

1. It is obvious that the RBs that are associated with the proposed estimators $f_k^*(0)$ are smaller than (in their magnitude) the corresponding RBs that are associated with the classical kernel estimator $\hat{f}_k(0)$.
2. The RMEs for the two estimators $f_k^*(0)$ and $\hat{f}_k(0)$ decrease when the sample size increases. This is a strong indication of the consistency of these estimators.
3. Compared to $f_k^*(0)$, the performance of the classical kernel estimator seems to be reasonable for the EP model with β=2 at which the smoothing parameter of $\hat{f}_k(0)$ is computed under this model (i.e., under the half- normal model). For the two cases, the HR model with (β,ω)=(2.5,10) and for the BE model with large value of β, the classical kernel estimator performs better than $f_k^*(0)$. However, the efficiencies in these cases are around 1, which indicates that the performances of the two estimators are similar.
4. With the exception of the three cases mentioned in (3), the performance of $f_k^*(0)$ is better than that of $\hat{f}_k(0)$. In fact, out of the 16 target models, there are 12 cases in which $f_k^*(0)$ beats $\hat{f}_k(0)$.
5. In general, the performance of the fourth-order kernel estimator $f_k^*(0)$ is very good for almost all considered target models. In addition, for the cases in which $\hat{f}_k(0)$ seems to be better than $f_k^*(0)$, the gain of efficiency is not significant compared with the large efficiencies of $f_k^*(0)$ over $\hat{f}_k(0)$ in certain cases. For example, the efficiency of $f_k^*(0)$ for the model HR with (β,ω,n)=(1,35,200) is 4.086, which means that the performance of $\hat{f}_k(0)$ becomes the same as that of $f_k^*(0)$ for a sample size of approximately 4.086$\times$200 $\cong$800.
6. Finally, we can say that Tables (3.1-3.4) demonstrate clearly that there is a significant improvement when applying the estimator $f_k^*(0)$ instead of the classical kernel estimator $\hat{f}_k(0)$.

References

[1] Chen, S. X. (1996). A kernel estimate for the density of a biological population by using line transect sampling. Applied Statistics, 45, 135-150.

[2] Cline, D. B. H., & Hart, J. D. (1991). Kernel estimation of densities of discontinuous derivatives. Statistics, 22, 69-84.

[3] Eidous, O. M., (2009). Kernel method starting with half-normal detection function for line transect density estimation. Commun. in Statist.-Theory and Methods. 38, 2366-2378.

[4] Eidous, O. M.,& Al-Masri, A. (2010). Fourth-order kernel method using line transect sampling. Advances and Applications in Statistics, 19, 65-79.

[5] Cowling, A., & Hall, P. (1996). On pseudodata method for removing boundary effects in kernel density estimation. Journal of the Royal Statistical Society, Ser. B, 58, 551-563.

[6] Eidous, O. M., (2012). A new kernel estimator for abundance using line transect sampling without the shoulder condition. Journal of the Korian Statistical Society, 41, 267-275.

[7] Eidous, O. M., & Alshakhatreh, M. K. (2011). Asymptotic unbiased kernel estimator for line transect sampling.

Table 3.1: RB and RME of the classical and fourth-order kernel estimators when the perpendicular distances are simulated from exponential power (EP) detection function.

Parameter	Estimator	n=50			n=100			n=200		
		EFF	RME	RB	EFF	RME	RB	EFF	RME	RB
$\hat{f}_k(0)$	$\beta = 1.0$	-0.316	0.331	1.000	-0.292	0.303	1.000	-0.267	0.273	1.000
$f_k^*(0)$	$\omega = 5.0$	-0.124	0.231	1.429	-0.100	0.189	1.600	-0.075	0.144	1.894
$\hat{f}_k(0)$	$\beta = 1.5$	-0.143	-0.069	1.000	-0.121	0.155	1.000	-0.105	0.124	1.000
$f_k^*(0)$	$\omega = 3.0$	-0.069	0.217	0.885	-0.070	0.183	0.850	-0.099	0.139	0.891
$\hat{f}_k(0)$	$\beta = 2.0$	-0.073	0.153	1.000	-0.066	0.123	1.000	-0.042	0.091	1.000
$f_k^*(0)$	$\omega = 2.5$	-0.032	0.165	0.925	-0.039	0.114	1.073	-0.025	0.076	1.197
$\hat{f}_k(0)$	$\beta = 2.5$	-0.050	0.151	1.000	-0.037	0.115	1.000	-0.031	0.088	1.000
$f_k^*(0)$	$\omega = 2.0$	-0.016	0.14	1.076	-0.005	0.112	1.025	-0.004	0.077	1.139

Table 3.2: RB and RME of the classical and fourth-order kernel estimators when the perpendicular distances are simulated from hazard-rate (HR) detection function.

Parameter	Estimator	n=50			n=100			n=200		
		EFF	RME	RB	EFF	RME	RB	EFF	RME	RB
$\hat{f}_k(0)$	$\beta = 1.0$	-0.592	0.651	1.000	-0.623	-0.648	0.625	1.000	0.593	1.000
$f_k^*(0)$	$\omega = 35$	-0.414	0.435	1.496	-0.357	0.372	1.679	-0.292	0.303	1.954
$\hat{f}_k(0)$	$\beta = 1.5$	-0.514	0.524	1.000	-0.490	0.495	1.000	-0.448	0.450	1.000
$f_k^*(0)$	$\omega = 25$	-0.154	0.231	2.258	-0.105	0.168	2.936	-0.058	0.110	4.086
$\hat{f}_k(0)$	$\beta = 2.0$	-0.275	0.299	1.000	-0.235	0.251	1.000	-0.195	0.205	1.000
$f_k^*(0)$	$\omega = 14$	0.007	0.183	1.633	0.032	0.145	1.732	0.031	0.128	1.605
$\hat{f}_k(0)$	$\beta = 2.5$	-0.128	0.186	1.000	-0.100	0.134	1.000	-0.070	0.096	1.000
$f_k^*(0)$	$\omega = 10$	0.045	0.229	0.814	0.020	0.152	0.879	0.016	0.133	0.726

Table 3.3: RB and RME of the classical and fourth-order kernel estimators when the perpendicular distances are simulated from the beta (BE) family detection function.

Parameter	Estimator	n=50			n=100			n=200		
		EFF	RME	RB	EFF	RME	RB	EFF	RME	RB
$\hat{f}_k(0)$	$\beta = 1.5$	-0.169	0.207	1.000	-0.157	0.184	1.000	-0.131	0.152	1.000
$f_k^*(0)$	$\omega = 1$	-0.140	0.196	1.055	-0.139	0.166	1.106	-0.119	0.139	1.094
$\hat{f}_k(0)$	$\beta = 2$	-0.196	0.228	1.000	-0.171	0.196	1.000	-0.150	0.166	1.000
$f_k^*(0)$	$\omega = 1$	-0.172	0.227	1.005	-0.152	0.198	0.991	-0.145	0.160	1.032
$\hat{f}_k(0)$	$\beta = 2.5$	-0.202	0.234	1.000	-0.189	0.210	1.000	-0.168	0.182	1.000
$f_k^*(0)$	$\omega = 1$	-0.143	0.243	0.961	-0.168	0.218	0.965	-0.179	0.189	0.962
$\hat{f}_k(0)$	$\beta = 3$	-0.221	0.252	1.000	-0.203	0.220	1.000	-0.182	0.193	1.000
$f_k^*(0)$	$\omega = 1$	-0.154	0.258	0.977	-0.161	0.231	0.953	-0.179	0.209	0.925

Table 3.4: RB and RME of the classical and fourth-order kernel estimators when the perpendicular distances are simulated from the beta (BE) family detection function.

Parameter	Estimator	n=50			n=100			n=200		
	$\beta = 0.6$	EFF	RME	RB	EFF	RME	RB	EFF	RME	RB
$\hat{f}_k(0)$	$b = 1$	-0.252	0.274	1.000	-0.220	0.235	1.000	-0.192	0.204	1.000
$f_k^*(0)$	$\omega = 1$	-0.130	0.235	1.166	-0.094	0.209	1.128	-0.035	0.148	1.372
$\hat{f}_k(0)$	$b = 5$	-0.147	0.193	1.000	-0.114	0.152	1.000	-0.100	0.120	1.000
$f_k^*(0)$	$\omega = 1$	-0.103	0.206	0.937	-0.098	0.153	0.994	-0.095	0.111	1.087
$\hat{f}_k(0)$	$b = 10$	-0.103	0.168	1.	-0.084	0.132	1.	-0.064	0.101	1.
$f_k^*(0)$	$\omega = 1$	-0.060	0.173	0.970	-0.061	0.115	1.143	-0.045	0.085	1.193
$\hat{f}_k(0)$	$b = 30$	-0.049	0.149	1.	-0.041	0.114	1.	-0.039	0.099	1.
$f_k^*(0)$	$\omega = 1$	-0.014	0.142	1.052	-0.011	0.101	1.125	-0.015	0.087	1.136

Commun. in Statist.-Theory and Methods, 40(24), 4353-4363.

[8] Gasser, T., & Muller, H. G. (1979). Kernel estimation of regression functions. In: Gasser, T., & Rosenblatt, M. (Eds), Smoothing techniques for curve estimation. Lecture notes in mathematics, Vol. 757, Spring, Heidelberg, 23-68.

[9] Gasser, T., Muller, H. G., & Mammitzsch, V. (1985). Kernels for nonparametric curve estimation. Journal of the Royal Statistical Society, Ser. B, 47, 238-256.

[10] Eidous, O. M. (2005a). On improving kernel estimator using line transect sampling. Commun. In statistic .- Theory and Methods, 34, 931-941.

[11] Mack, Y. P. & Quang, P. X. (1998). Kernel methods in line and point transect sampling. Biometrics, 54, 606-619.

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

[13] Zhang, S., & Karunamuni, R. J. (1998). On kernel density estimation near endpoints. Journal of Statistical Planning and Inference, 70, 301-316.

[14] Eidous, O. M., (2011a). Variable location kernel method using line transect sampling. Environments, 22, 431-440.

[15] Ababneh, F. and Eidous, O. M. (2012). A weighted exponential detection function model for line transect data. Journal of Modern Applied Statistical Methods, 11(1), 144-151.

[16] Burnham, K. P., & Anderson, D. R. (1976). Mathematical models for nonparametric inferences from line transect data. Biometrics, 32, 325-336.

[17] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. An international Symposium of Information Theory, 2ndedn, Budapest, Hungary, 267-281.

[18] Barabesi, L. (2000). Local likelihood density estimation in line transect sampling. Environmetrics, 11, 413-422.

[19] Barabesi, L. (2001). Local parametric density estimation method in line transect sampling. Metron, LIX, 21-37.

Ali Algarni has obtained his PhD degree at 2013 from School of Mathematics and Applied Statistics, Wollongong University, Australia. He is a Investigator and Data Analyst at center of survey methodology and data analysis,Wollongong. Australia. He is a member, Saudi Association of Mathematics and Applied Statistics. He is presently employed as assistant professor in Statistics Department, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia.

Ahmad Almutlg Taking his Diploma in Risk Management from European Business (Learning Academy) in the UK [October December 2010].In Dec.2014 taking the master in Statistics from King Abdulaziz University, Jeddah. Thesis Topic Akaike Information Criterion and Fourth-Order Kernel Method for Line Transect Sampling.