Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v13n2p562

A non-parametric density estimate adaptation for population abundance when the shoulder condition is violated

By Albadareen et al.

Published: 14 October 2020

Electronic Journal of Applied Statistical Analysis
Vol. 13, Issue 02, October 2020, 562-579
DOI: 10.1285/i20705948v13n2p562

A non-parametric density estimate adaptation for population abundance when the shoulder condition is violated

Baker Albadareen*[a], Noriszura Ismail[a], Omar M. Eidous[b], and Jamil J. Jaber[a,c]

[a] *Universiti Kebangsaan Malaysia, Department of Mathematical Sciences, Bangi, Malaysia*
[b] *Yarmouk University, Department of Statistics, Irbid, Jordan*
[c] *The University of Jordan, Department of Risk Management Insurance, Aqaba, Jordan*

Published: 14 October 2020

The non-parametric kernel density estimation is used in practice to estimate population abundance using the line transect sampling method. Kernel estimator of $f(0)$ usually produces an underestimated value. Assuming the kernel method, this article applies shifted logarithmic transformation to line transect data that violating shoulder condition. Mathematically, the proposed log-transform estimator was shown to be more efficient than the classical kernel estimator. The simulation results present the good properties of the proposed estimator compared to the performance of the classical kernel estimators.

keywords: kernel estimator, line transect, log-transformation, abundance, bandwidth, shoulder condition

1 Introduction

Line transect sampling is a common method to estimate population abundance (density), D. The population in line transect sampling consists of both living or non-living object societies such as plants, animals, birds, and others (Burnham et al., 1980). In the line transect method, the study area has at least a strip of width $2w$ and length L, in which the transect line is randomly placed on the strip. Both sides of the line are observed

*Corresponding author: albadareen.baker@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

before the detected objects are counted, and the perpendicular distances (x) between the transect line and the detected object must be recorded. One of the advantages of the sampling method is that it records the perpendicular distance of the detected objects only.

Assuming that the observer detects a random sample of n objects $x_1, x_2, ..., x_n$. The probability of detecting an object at a perpendicular distance $X = x$ is defined as the conditional probability function P(detected an object given that it has perpendicular distance x), say $g(x)$. The random variable X has a probability density function (pdf) of $f(x)$, $0 \leq x \leq w$, where the distribution shape of $g(x)$ is the same as $f(x)$. However, $f(x)$ is scaled by a constant, making it one unit under the curve because it is a probability density function (pdf). The relationship between $f(x)$ and $g(x)$ is given by $f(x) = \frac{g(x)}{\int_0^w g(t)dt}$. A description of the line transect sampling and its fundamental concepts can be found in Buckland et al. (2001).

The line transect method assumes the probability of detected objects, function $g(x)$, is non-increasing, which means that the closest object to the line has the highest probability to be detected compared to the objects far from the line. Another assumption related to the detection function $g(x)$ is that the observer will never miss any object located on the line $g(0) = 1$ (i.e. at $x = 0$). In fact, the shape of the detection function can be split into two families, according to the distribution shape at $x = 0$; the first is the flat distribution which is known from the literature by the shoulder condition (see Mack and Quang (1998)), and the second is the spike distribution which violates the shoulder condition.

The assumption of shoulder condition is valid if the shape of the detection function $g(x)$ has a shoulder at $x = 0$, or in other words, the probability of detected objects closer to the transect line is still certain. Mathematically, the derivative of $g(0)$ equals to zero, and this is equivalent to $f'(0) = 0$ (Buckland et al., 2001; Eidous, 2015).

Several approaches can be found in the literature to test the shoulder condition (Zhang, 2001). Recent researches present that the shoulder condition is invalid for several line transect of wildlife data such as ruffed grouses, bobwhite quails, scaled quails, white-tailed deer, jackrabbits, capercaillies, and cottontails (Buckland, 1985), and fin whales and striped dolphins (Bauer et al., 2015).

Therefore, this study considers the case when the shoulder condition is violated.

Let D be the population abundance applied in a specific area A. Burnham et al. (1980) has shown that the estimate of the population abundance, D, using the line transect sampling is:

$$\hat{D} = \frac{n\hat{f}(0)}{2L} \tag{1}$$

where $\hat{f}(0)$ is an estimate of $f(0)$ obtained using a robust method based on perpendicular distances $x_1, x_2, ..., x_n$ random sample of size n, while L is the length of the transect line.

If the population area A is known, the population size N is estimated easily using the formula:

$$\hat{N} = \hat{D}A \tag{2}$$

As shown in (1), the estimation value of $f(0)$ is critical to estimate abundance D in the line transect sampling.

Many approaches in the literature have been proposed for estimating $f(0)$. There are two common methods: parametric and non-parametric. For the parametric method, one assumes a specific parametric distribution for the detection function $g(0)$, and the parameters are then estimated using several methods (e.g. maximum likelihood estimation). The negative exponential distribution (Gates et al., 1968) and the half-normal distribution (Burnham et al., 1980) are popular models for fitting the density of perpendicular distances. The parametric models are mentioned in the literature (Barabesi, 2000; Buckland, 1985; Buckland et al., 2001; Miller and Thomas, 2015; Zhang, 2011). The parametric approach is efficient and satisfactory when the data distribution is correctly chosen. Otherwise, it is deemed inappropriate (Buckland et al., 2001).

The second approach is the non-parametric methods such as the Fourier series (Crain et al., 1979), where the series parameters are directly estimated from the perpendicular distances. The kernel estimation is the most common non-parametric approach introduced by Fix and Hodges (1951).

In this paper, the non-parametric method is considered when the shoulder condition is invalid. A new estimator based on the logarithmic transformation is proposed, and the density is obtained via back-transformation. Mathematically, the proposed transformed estimator is shown to be more efficient than the kernel estimator. Simulation is also carried out to study the performance between the proposed transform kernel estimators and the classical one.

2 Kernel Density Estimation

Kernel density estimation is the popular non-parametric method that smooth the data in histogram. For this method, inferences of a specific population value are made based on the random sample. The advantage of this method compared to the parametric one is that it requires no assumption regarding the shape of $f(x)$, which is useful for non-statistician researchers. Moreover, non-parametric methods allow the data to illustrate on its own. Silverman (1986) stated the full description of the kernel Rosenblatt–Parzen estimation method (KDE). The classical KDE is given by

$$\hat{f}_X(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x - x_i}{h}), \quad -\infty < x < \infty \tag{3}$$

where h is the bandwidth (smoothing parameter), and $K(.)$ is the kernel function; it's assumed to be satisfied the conditions $\int_{-\infty}^{\infty} K(u)du = 1$, $\int_{-\infty}^{\infty} uK(u)du = 0$ and $\int_{-\infty}^{\infty} u^2K(u)du \neq 0$ (Silverman, 1986).

Assume that a line transect method, it has a non-negative random sample $x_1, x_2, ..., x_n$ of size n of perpendicular distances, and has a continuous probability density function

$f(x),\ x \geq 0$. To reduce the bias around the boundary for the Rosenblatt–Parzen kernel density estimation (KDE). Schuster (1985) and Silverman (1986) suggested to use a so-called "re-normalization" method applying a reflection technique. The method replaces each sampled data value x_i with x_i and it's reflection $-x_i$, if $K(u)$ is considered as a symmetric density function around zero; the obtained reflected estimator is:

$$\hat{f}_X(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{x-x_i}{h}) + K(\frac{x+x_i}{h}), \quad x \geq 0 \tag{4}$$

The usual reflection kernel estimator of $f_X(0)$ is obtained simply by substituting $x = 0$ in (4), which gives (Chen, 1996)

$$\hat{f}_X(0) = \frac{2}{nh}\sum_{i=1}^{n} K(\frac{x_i}{h}) \tag{5}$$

Thus, the bias and variance of $\hat{f}_X(0)$ are:

$$Bias[\hat{f}_X(0)] = 2hf'_X(0)\int_0^{\infty} uK(u)du + O(h^2) \tag{6}$$

$$Var[\hat{f}_X(0)] = \frac{4}{nh} f_X(0)\int_0^{\infty} K^2(u)du + o(\frac{1}{nh}) \tag{7}$$

If the small terms $o(.)$ and $O(.)$ are ignored, then the asymptotic mean squared error (AMSE) of $\hat{f}_X(0)$ is:

$$AMSE[\hat{f}_X(0)] = \frac{4}{nh} f_X(0)\int_0^{\infty} K^2(u)du + (2hf'_X(0)\int_0^{\infty} uK(u)du)^2 \tag{8}$$

The AMSE is sensitive to the choice of bandwidth h, and thus, several approaches to finding the 'best' one are proposed in the literature. Therefore, our study suggests several methods for the optimal bandwidth, and the description is provided in the subsection "Optimal Bandwidth".

Kernel method is commonly used in practice to find a good estimate of $f_X(0)$ for the line transect data. As examples, Chen (1996) suggested and investigated the characteristics of the estimator in equation (5), Mack (2002) suggested several bias reduction techniques for $f_X(0)$, Eidous (2005) suggested some improvements of the kernel estimator of $f_X(0)$, and Eidous (2012) proposed several simple kernel estimator for $f_X(0)$ when the shoulder condition is violated. Albadareen and Ismail (2018) proposed an adaptive kernel estimator based on a generalized form of Epanechnikov kernel function. Recently, Albadareen and Ismail (2019, 2020) applied shifted power transformation to reduce boundary effect of the kernel estimator using the line transect sampling, and Zhang et al. (2020) estimated distribution function at the boundary based on the kernel estimator.

3 Log-Transformation Method for Kernel Density Estimation

The non-parametric density estimation based on transformed data has existed for many years. From the literature, the log-transformation is a transformation case stated by Box and Cox (1964) to make the random variable X less restrictive from some assumptions such as normal, homoscedastic and linear model and to transform it to be more appropriate for inference. An application of the kernel density using the logarithmic transformation was also suggested by Devroye and Györfi (1985), and studied by Marron and Ruppert (1994). Other studies can also be found in Silverman (1986) who stated that the transformation $Y = \log(X)$ can be used for dealing with the rising of spike. Charpentier and Flachaire (2015) who showed that the logarithmic transformation of $Y = \log(X)$ applied to the classical kernel density provided better fit for estimating the density of heavy-tailed shape such as income distribution.

In this study, the line transect data are positive and the estimation of interest is $f(x)$ at $x = 0$. Since the logarithmic transformation $Y = \log(X)$ is unsatisfactory when $x = 0$, we propose the transformation $Y = \log(X+1)$ in which Y is strictly an increasing function, and the underlying density functions are $f_X(x)$ and $f_Y(y)$ respectively. Using the changing variable equation, the density function is

$$f_Y(y) = f_X(x) \mid \frac{dx}{dy} \mid = f_X(e^y - 1)e^y, \quad y \geq 0 \tag{9}$$

where $\mid \frac{dx}{dy} \mid$ is the Jacobian for the transformation of random variables from X to Y. At $x = 0$, $y = \log(0+1) = 0$, and thus,

$$f_Y(0) = f_X(0)e^0 = f_X(0) \tag{10}$$

which shows that $f_X(0) = f_Y(0)$.

The estimation of $f_X(0)$ in equation (1) requires $f_Y(0)$ to be substituted with $\hat{f}_Y(0)$. In our study, we use the kernel method, which is the common non-parametric method to derive $\hat{f}_Y(0)$:

$$\hat{f}_X(0) = \hat{f}_Y(0) = \frac{2}{nh}\sum_{i=1}^{n} K(\frac{y_i}{h}) \ , \qquad y_i = \log(x_i + 1) \tag{11}$$

The bias and variance of $\hat{f}_Y(0)$ are:

$$Bias[\hat{f}_Y(0)] = 2hf'_Y(0)\int_0^\infty uK(u)du + h^2 f''_Y(0)\int_0^\infty u^2K(u)du + o(h^2) \tag{12}$$

$$= 2h(f_X(0) + f'_X(0))\int_0^\infty uK(u)du + O(h^2) \tag{13}$$

$$Var[\hat{f}_Y(0)] = \frac{4}{nh} f_Y(0)\int_0^\infty K^2(u)du + o(\frac{1}{nh}) \tag{14}$$

$$= \frac{4}{nh} f_X(0) \int_0^\infty K^2(u)du + o(\frac{1}{nh}) \quad (15)$$

If the small terms $o(.)$ and $O(.)$ are ignored, the asymptotic mean squared error of $\hat{f}_Y(0)$ is:

$$AMSE[\hat{f}_Y(0)] = \frac{4}{nh} f_X(0) \int_0^\infty K^2(u)du + (2h(f_X(0) + f'_X(0)) \int_0^\infty uK(u)du)^2 \quad (16)$$

By comparing (15) and (7), it was found that $Var[\hat{f}_Y(0)]$ equals to $Var[\hat{f}_X(0)]$.

It should be noted that the transformed estimator (11) is applied when violated the shoulder condition.

If the Gaussian kernel function $K(u)$ is assumed, it is also important to note here that the value of the reflection estimators at (4) and (5) decays to zero for $x > w$ when w value is sufficiently large such that $w \geq \max(x_i) + 4h$. Because, when $\mid x \mp x_i \mid > 4h$; the value of $K(\frac{x \mp x_i}{h})$ is vanishing.

4 Simulation Study

The theoretical characteristics of the proposed estimator in terms of bias, variance and MSE are asymptotic, implying that the estimator requires the assumption $n \rightarrow \infty$. In this simulation study, the performance of the proposed estimator is compared to the classical kernel estimator using a wide range of sample sizes, which are $n = 50$, 100, and 500. The performance indicators are the relative bias (RB) and the relative mean error (RME) which are computed using $RB = E[\hat{f}(0) - f(0)]/f(0)$ and $RME = \sqrt{MSE[\hat{f}(0)]}/f(0)$.

The simulation study is focused on generating random samples from three common density families under the line transect sampling that violate the shoulder condition. Four models are considered for each density group, so that altogether there are 12 detection functions. The detection functions cover several potential models for the perpendicular distances in the case of invalid shoulder condition, as demonstrated in Figure 1.

The three density families are:

a) Reserved logistic model (Burnham et al., 1980)

   The detection function is $g(x) = \frac{(1+b)e^{-\beta x}}{1+be^{-\beta x}}$ and $f(x) = \frac{\beta b e^{-\beta x}}{\log(1+b)(1+be^{-\beta x})}, \beta, b > 0$, $0 \leq x \leq w$, where we use the following parameter values for the four models; $\beta = 1.5$, 2.0, 2.5 and 3.0 with $b = 2$, and truncation point $w = 5.0$.

b) Beta (BE) model (Eberhardt, 1968)

   The detection function is $g(x) = (1 - x)^\beta$ and $f(x) = (1 + \beta)(1 - x)^\beta$, $0 \leq x \leq w, \beta \geq 1$, where parameter values are $\beta = 1.5$, 2.0, 2.5 and 3.0 and truncation point $w = 1$ for the four models.

c) Negative exponential model (Gates et al., 1968)

The detection function is $g(x) = e^{-\beta x}$ and $f(x) = \beta e^{-\beta x}$, $\beta > 0 \leq x \leq w$. We use parameter values $\beta = 1.0$, 1.5, 2.0 and 2.5 and truncation point $w = 3.0$ for the four models.

4.1 Optimal Bandwidth

The estimator in equation (5) gives convergent results when the symmetric kernel functions are chosen, such as Gaussian, biweight, and Epanechnikov, which are based on the mean squared error (Wand and Jones, 1995). Therefore, we consider the Gaussian kernel function in this paper.

The kernel density estimator is highly sensitive to the smoothing parameter (bandwidth h) because the performance of $\hat{f}(0)$ depends on the value of h (Gerard and Schucany, 1999). The approaches to choose an efficient bandwidth can be found from past studies. In general, Ghosh (2018) recommended the use of a sequence of bandwidths for the kernel estimators and their performance were compared to understand the underlying structure of the unknown pdf because there is no "best" method in choosing the bandwidth. The same approach is also used by Silverman (1981). Thus, different bandwidths are used in this simulation study, along with some comparisons between the estimators. We consider the following bandwidths and estimators for this simulation study:

a) Rule of thumb (Silverman, 1986): $h = 1.06\hat{\sigma}n^{(-1/5)}$

The bandwidth is computed by minimizing the approximate mean squared error (AMSE) of $\hat{f}_X(x)$. In the line transect method, the half-normal distribution is assumed as the reference density of $f(x)$ to estimate σ via the maximum likelihood estimator $\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$. The bandwidth is applied to the original data which resulted in the produced estimator, Est1.

b) Modified rule of thumb (Silverman, 1986): $h = 0.9\ min(\frac{IQR(X)}{1.349}, \hat{\sigma})n^{(-1/5)}$

The bandwidth is similar to the previous bandwidth but is modified to the long-tailed and skewed distribution, which better fits the density that has no shoulder in the line transect data, where $\hat{\sigma}$ is computed as illustrated in part (a) above, and $IQR(x)$ is the interquartile range of the perpendicular distances X. The bandwidth is applied to the original data which resulted in the produced estimator, Est2.

c) Similar to the h value in part (a). The bandwidth applied to the transformed data will result in the produced estimator, Est3.

d) Similar to the h value in part (b). The bandwidth applied to the transformed data will result in the produced estimator, Est4.

Another efficient bandwidth can be obtained by minimizing the $AMSE[\hat{f}_Y(0)]$ in equation (16). If the $AMSE[\hat{f}_Y(0)]$ is derived with respect to h. Then, the smoothing parameter is: $h = (\frac{f_X(0)\int_0^\infty [K(u)]^2 du}{2n(f_X(0)+f'_X(0))^2(\int_0^\infty uK(u)du)^2})^{\frac{1}{3}}$.

Assuming that $K(u)$ follows Gaussian distribution, the unknown values ($f(0)$ and $f'(0)$) are estimated using a suitable density as the reference distribution family (see Al-Bassam and Eidous (2018)). This bandwidth value will be used in parts (e) and (f).

The following bandwidths and estimators are also considered in this simulation study:

e) Assuming that the reference distribution model; $f_X(x)$ is the negative exponential model in the case of violated shoulder condition. Applying the maximum likelihood estimators ($\hat{f}_X(0) = \frac{1}{\bar{x}}$) and ($\hat{f}'_X(0) = \frac{-1}{\bar{x}^2}$); then: $h = (\frac{(\frac{1}{\bar{x}})(\frac{1}{4\sqrt{\pi}})}{2n(\frac{1}{\bar{x}}-\frac{1}{\bar{x}^2})^2(\frac{1}{\sqrt{2\pi}})^2})^{(1/3)} = (\frac{(\sqrt{\pi})(\bar{x})^3}{4n(\bar{x}-1)^2})^{(1/3)}$. The bandwidth applied to the transformed data will result in the produced estimator, Est5.

f) Assuming that the reference distribution model; $f_X(x)$ is the half-normal model in the case of validity of shoulder condition. Applying the maximum likelihood estimators ($\hat{f}_X(0) = \frac{2}{\hat{\sigma}\sqrt{2\pi}}$) and ($\hat{f}'_X(0) = 0$); then $h = (\frac{(\frac{2}{\hat{\sigma}\sqrt{2\pi}})(\frac{1}{4\sqrt{\pi}})}{2n(\frac{2}{\hat{\sigma}\sqrt{2\pi}}+0)^2(\frac{1}{\sqrt{2\pi}})^2})^{\frac{1}{3}}$ $= (\frac{\hat{\sigma}\pi}{4\sqrt{2n}})^{\frac{1}{3}}$, where $\hat{\sigma}$ is computed as illustrated in part (a) above. The bandwidth applied to the transformed data will result in the produced estimator, Est6.

Our simulation study focused on six estimation values of $f(0)$, where Est1 and Est2 are applied to the original data as suggested in the literature, while Est3, Est4, Est5 and Est6 are the proposed estimators using the transformed data.

Tables 1-3 provide the relative bias (RB) and the relative mean error (RME) for the estimators under the reserved logistic model, beta model, and negative exponential model based on the simulation study.

The simulation results show that the proposed estimator (Est4) (based on the modified rule of thumb (Silverman, 1986)) is superior, in which it has smaller absolute relative bias and relative mean error compared to the original kernel estimators (Est1 and Est2) for each density family and all cases. The RME of (Est4) decreases as the sample size increases, hence it is concluded that (Est4) is asymptotically consistent. The results are depicted in Figures 2-4.

The proposed estimator (Est3) in the reserved logistic model is also more efficient than other estimators at $\beta = 1.5$, 2.0 and 2.5 for all sample sizes. On the other hand, the proposed estimators (Est4, Est5 and Est6) are generally performing well in the case of negative exponential model compared to other estimators (Est1, Est2 and Est3) when $\beta = 1.5$, 2.0 and 2.5 for all sample sizes. The proposed estimators (Est4, Est5 and Est6) also have smaller RME than the kernel estimator when applied to the original data (Est1) for all studied densities with large samples ($n = 500$).

Based on RB and RME, comparison between two bandwidths which are the modified bandwidth introduced by Silverman (1986) $h = 0.9\ min(\frac{IQR(X)}{1.349}, \hat{\sigma})n^{(-1/5)}$ and the original bandwidth introduced by Silverman (1986) $h = 1.06\hat{\sigma}n^{(-1/5)}$ can be carried out. The modified bandwidth (Est2 and Est4) produces superior efficiency than the original bandwidths (Est1 and Est3) whether the data is log-transformed or not. Therefore, the modified bandwidth introduced by Silverman (1986) where $h = 0.9\ min(\frac{IQR(X)}{1.349}, \hat{\sigma})n^{(-1/5)}$, is more appropriate for the line transect sampling when the shoulder condition is violated.

5 Conclusion

This study introduces an efficient and consistent estimator for population abundance using the line transect sampling. The Log-transformation method applied to the kernel estimation is proven to perform well compared to the classical kernel estimator in the case of violated shoulder condition. The asymptotic bias, variance and mean squared error (AMSE) of the proposed estimator were obtained. The proposed estimator was shown to have smaller AMSE if the detection function has no shoulder. A simulation study was conducted to compare the performance of the proposed estimators and the classical kernel estimators. The simulation results showed that the proposed transformed estimators have smaller relative mean error than not-transformed kernel estimators for each density family. In addition, the proposed estimators have better performance than the not-transformed kernel estimators in terms of absolute relative error with fixed variance in all cases.

Acknowledgement

The authors would like to thank the editor and the reviewer for their valuable comments and constructive reviews to further improve this paper. The authors are also grateful and would like to acknowledge the financial support granted by the Ministry of Higher Education (MOHE) and Universiti Kebangsaan Malaysia (UKM) in the form of research grants (FRGS/1/2019/STG06/UKM/01/5 and GUP-2019-031).

References

Al-Bassam, M. and Eidous, O. M. (2018). Combination of parametric and nonparametric estimators for population abundance using line transect sampling. *Journal of Information and Optimization Sciences*, 39(7):1449–1462.

Albadareen, B. and Ismail, N. (2018). Adaptive kernel function using line transect sampling. *AIP Conference Proceedings*, 1940:020112.

Albadareen, B. and Ismail, N. (2019). An Adaptation of Kernel Density Estimation for Population Abundance using Line Transect Sampling When the Shoulder Condition is Violated. *International Journal of Innovative Technology and Exploring Engineering*, 9(2):3494–3498.

Albadareen, B. and Ismail, N. (2020). A General Base of Power Transformation to Improve the Boundary Effect in Kernel Density without Shoulder Condition. *Pakistan Journal of Statistics and Operation Research*, 16(2):279–285.

Barabesi, L. (2000). Local likelihood density estimation in line transect sampling. *Environmetrics*, 11(4):413–422.

Bauer, R. K., Fromentin, J.-M., Demarcq, H., Brisset, B., and Bonhommeau, S. (2015). Co-Occurrence and Habitat Use of Fin Whales, Striped Dolphins and Atlantic Bluefin Tuna in the Northwestern Mediterranean Sea. *PLOS ONE*, 10(10):e0139218.

Box, G. E. P. and Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2):211–243.

Buckland, S. T. (1985). Perpendicular Distance Models for Line Transect Sampling. *Biometrics*, 41(1):177–195.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2001). *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, London, 1 edition.

Burnham, K. P., Anderson, D. R., and Laake, J. L. (1980). Estimation of Density from Line Transect Sampling of Biological Populations. *Wildlife Monographs*, (72):3–202.

Charpentier, A. and Flachaire, E. (2015). Log-Transform Kernel Density Estimation of Income Distribution. *L'Actualité économique*, 91(1-2):141–159.

Chen, S. X. (1996). A Kernel Estimate for the Density of a Biological Population by Using Line Transect Sampling. *Applied Statistics*, 45(2):135–150.

Crain, B. R., Burnham, K. P., Anderson, D. R., and Lake, J. L. (1979). Nonparametric estimation of population density for line transect sampling using Fourier series. *Biometrical Journal*, 21(8):731–748.

Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation*. John Wiley and Sons, Montreal - Canada, 1 edition.

Eberhardt, L. L. (1968). A Preliminary Appraisal of Line Transects. *The Journal of Wildlife Management*, 32(1):82–88.

Eidous, O. M. (2005). On improving kernel estimators using line transect sampling. *Communications in Statistics - Theory and Methods*, 34(4):931–941.

Eidous, O. M. (2012). A new kernel estimator for abundance using line transect sampling without the shoulder condition. *Journal of the Korean Statistical Society*, 41(2):267–275.

Eidous, O. M. (2015). Nonparametric Estimation of f(0) Applying Line Transect Data with and without the Shoulder Condition. *Journal of Information and Optimization Sciences*, 36(4):301–315.

Fix, E. and Hodges, J. (1951). Nonparametric discrimination: consistency properties. *DTIC Document*, (May).

Gates, C. E., Marshall, W. H., and Olson, D. P. (1968). Line Transect Method of Estimating Grouse Population Densities. *Biometrics*, 24(1):135–145.

Gerard, P. D. and Schucany, W. R. (1999). Local bandwidth selection for Kernel estimation of population densities with line transect sampling. *Biometrics*, 55(3):769–773.

Ghosh, S. (2018). *Kernel Smoothing: Principles, Methods and Applications*. Chapman & Hall, New York, 1 edition.

Mack, Y. P. (2002). Bias-corrected confidence intervals for wildlife abundance estimation. *Communications in Statistics - Theory and Methods*, 31(7):1107–1122.

Mack, Y. P. and Quang, P. X. (1998). Kernel Methods in Line and Point Transect Sampling. *Biometrics*, 54(2):606–619.

Marron, J. S. and Ruppert, D. (1994). Transformations to Reduce Boundary Bias in Kernel Density Estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(4):653–671.

Miller, D. L. and Thomas, L. (2015). Mixture models for distance sampling detection functions. *PLoS ONE*, 10(3):1–19.

Schuster, E. F. (1985). Incorporating Support Constraints into Nonparametric Estimators of Densities. *Communications in Statistics - Theory and Methods*, 14(5):1123–1136.

Silverman, B. W. (1981). Using Kernel Density Estimates to Investigate Multimodality. *Journal of the Royal Statistical Society. Series B (Methodological)*, 43(1):97–99.

Silverman, B. W. (1986). *Density estimation: For statistics and data analysis.* Number 1. Chapman & Hall, London, 1 edition.

Wand, M. P. and Jones, M. C. (1995). *Kernel Smoothing.* Number 1. Chapman & Hall, New York, 1 edition.

Zhang, S. (2001). Generalized likelihood ratio test for the shoulder condition in line transect sampling. *Communications in Statistics - Theory and Methods*, 30(11):2343–2354.

Zhang, S. (2011). On parametric estimation of population abundance for line transect sampling. *Environmental and Ecological Statistics*, 18(1):79–92.

Zhang, S., Li, Z., and Zhang, Z. (2020). Estimating a Distribution Function at the Boundary. *Austrian Journal of Statistics*, 49(1):1–23.

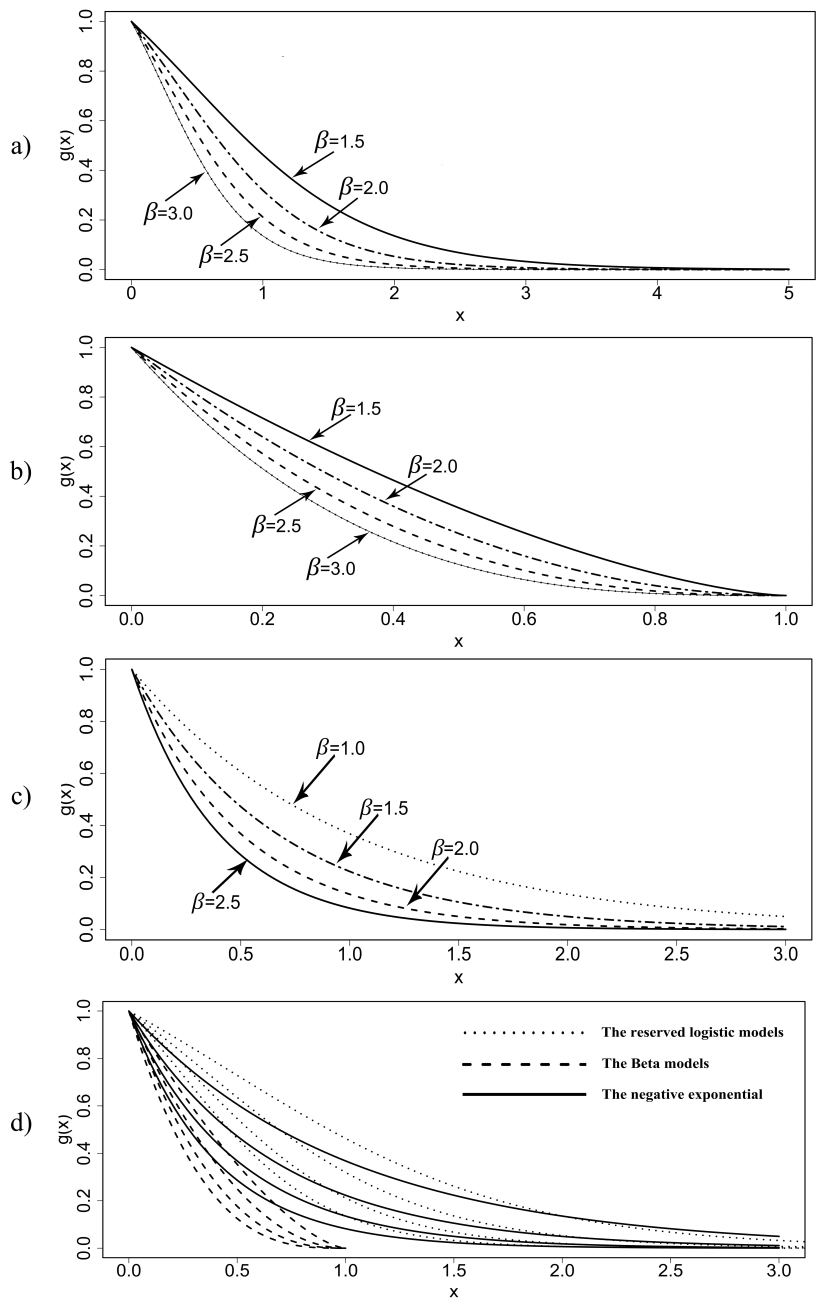

Figure 1: The detection functions of a) Reserved logistic models at $b = 2$, b) Beta models, c) Negative exponential models and d) All models

Table 1: Simulation results for reserved logistic model

		$n = 50$		$n = 100$		$n = 500$	
	Estimator	RB	RME	RB	RME	RB	RME
$\beta = 1.5$	Est1	-0.232	0.257	-0.202	0.218	-0.146	0.153
	Est2	-0.127	0.216	-0.106	0.165	-0.072	0.100
	Est3	-0.030	0.109	0.004	0.075	0.050	0.065
	Est4	0.042	0.179	0.052	0.137	0.055	0.090
	Est5	-0.168	0.348	-0.081	0.246	0.052	0.082
	Est6	0.056	0.197	0.055	0.163	0.041	0.105
$\beta = 2$	Est1	-0.233	0.255	-0.203	0.219	-0.151	0.157
	Est2	-0.126	0.209	-0.101	0.163	-0.076	0.101
	Est3	-0.079	0.127	-0.049	0.092	-0.006	0.041
	Est4	-0.002	0.165	0.016	0.129	0.018	0.071
	Est5	-0.003	0.157	0.019	0.123	0.020	0.074
	Est6	0.012	0.158	0.025	0.136	0.018	0.086
$\beta = 2.5$	Est1	-0.240	0.260	-0.207	0.222	-0.149	0.156
	Est2	-0.136	0.209	-0.108	0.168	-0.072	0.101
	Est3	-0.117	0.151	-0.083	0.112	-0.034	0.056
	Est4	-0.039	0.163	-0.017	0.129	0.001	0.072
	Est5	-0.025	0.160	-0.005	0.136	0.006	0.086
	Est6	-0.037	0.137	-0.014	0.121	0.005	0.080
$\beta = 3$	Est1	-0.238	0.261	-0.210	0.224	-0.151	0.157
	Est2	-0.134	0.217	-0.111	0.167	-0.075	0.101
	Est3	-0.134	0.169	-0.107	0.131	-0.056	0.070
	Est4	-0.054	0.179	-0.036	0.130	-0.014	0.069
	Est5	-0.031	0.185	-0.019	0.146	-0.006	0.088
	Est6	-0.066	0.143	-0.041	0.114	-0.012	0.071

Table 2: Simulation results for beta model

		$n = 50$		$n = 100$		$n = 500$	
	Estimator	RB	RME	RB	RME	RB	RME
$\beta = 1.5$	Est1	-0.193	0.223	-0.169	0.190	-0.125	0.135
	Est2	-0.125	0.214	-0.100	0.166	-0.073	0.104
	Est3	-0.113	0.154	-0.092	0.125	-0.060	0.078
	Est4	-0.069	0.187	-0.048	0.141	-0.031	0.081
	Est5	-0.041	0.215	-0.022	0.176	-0.018	0.109
	Est6	-0.103	0.147	-0.073	0.119	-0.037	0.076
$\beta = 2$	Est1	-0.217	0.243	-0.193	0.209	-0.145	0.152
	Est2	-0.136	0.216	-0.116	0.169	-0.083	0.108
	Est3	-0.152	0.184	-0.130	0.152	-0.091	0.102
	Est4	-0.090	0.191	-0.072	0.143	-0.048	0.084
	Est5	-0.053	0.216	-0.044	0.170	-0.024	0.106
	Est6	-0.154	0.180	-0.117	0.142	-0.064	0.086
$\beta = 2.5$	Est1	-0.241	0.261	-0.209	0.224	-0.160	0.166
	Est2	-0.150	0.217	-0.120	0.172	-0.092	0.114
	Est3	-0.187	0.210	-0.157	0.176	-0.114	0.122
	Est4	-0.109	0.191	-0.082	0.148	-0.061	0.091
	Est5	-0.069	0.213	-0.042	0.175	-0.030	0.107
	Est6	-0.201	0.217	-0.153	0.170	-0.088	0.102
$\beta = 3$	Est1	-0.245	0.267	-0.224	0.239	-0.173	0.179
	Est2	-0.146	0.221	-0.134	0.182	-0.101	0.122
	Est3	-0.198	0.224	-0.178	0.195	-0.133	0.140
	Est4	-0.111	0.199	-0.101	0.159	-0.074	0.101
	Est5	-0.066	0.219	-0.061	0.181	-0.039	0.111
	Est6	-0.231	0.243	-0.187	0.199	-0.111	0.122

Table 3: Simulation results for negative exponential model

		$n = 50$		$n = 100$		$n = 500$	
	Estimator	RB	RME	RB	RME	RB	RME
$\beta = 1.0$	Est1	-0.272	0.291	-0.245	0.259	-0.179	0.185
	Est2	-0.156	0.220	-0.138	0.182	-0.088	0.110
	Est3	-0.093	0.132	-0.065	0.099	-0.012	0.043
	Est4	-0.002	0.151	0.008	0.118	0.032	0.074
	Est5	-0.169	0.305	-0.109	0.231	0.009	0.062
	Est6	0.026	0.175	0.032	0.150	0.048	0.108
$\beta = 1.5$	Est1	-0.340	0.353	-0.311	0.319	-0.244	0.247
	Est2	-0.203	0.255	-0.184	0.216	-0.138	0.151
	Est3	-0.227	0.244	-0.195	0.207	-0.133	0.139
	Est4	-0.103	0.183	-0.089	0.143	-0.059	0.085
	Est5	-0.116	0.197	-0.092	0.151	-0.047	0.085
	Est6	-0.094	0.165	-0.074	0.137	-0.037	0.085
$\beta = 2.0$	Est1	-0.359	0.372	-0.328	0.336	-0.263	0.266
	Est2	-0.217	0.263	-0.191	0.221	-0.148	0.159
	Est3	-0.274	0.290	-0.243	0.253	-0.181	0.184
	Est4	-0.144	0.206	-0.121	0.164	-0.089	0.106
	Est5	-0.128	0.205	-0.099	0.161	-0.061	0.098
	Est6	-0.167	0.202	-0.128	0.161	-0.076	0.099
$\beta = 2.5$	Est1	-0.367	0.378	-0.334	0.342	-0.267	0.270
	Est2	-0.226	0.264	-0.192	0.222	-0.149	0.160
	Est3	-0.300	0.313	-0.267	0.276	-0.201	0.204
	Est4	-0.168	0.216	-0.136	0.176	-0.102	0.118
	Est5	-0.137	0.210	-0.101	0.170	-0.064	0.102
	Est6	-0.215	0.235	-0.168	0.190	-0.102	0.117

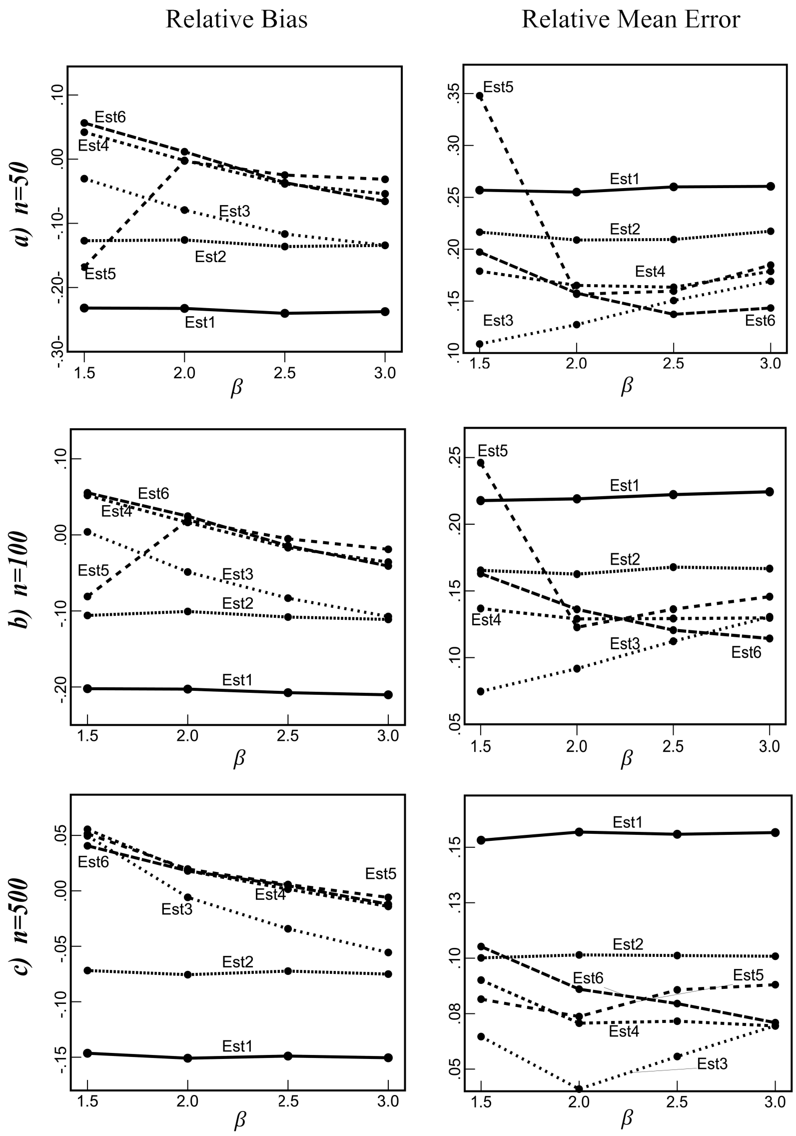

Figure 2: Relative bias and relative mean error of reserved logistic models at $b = 2$

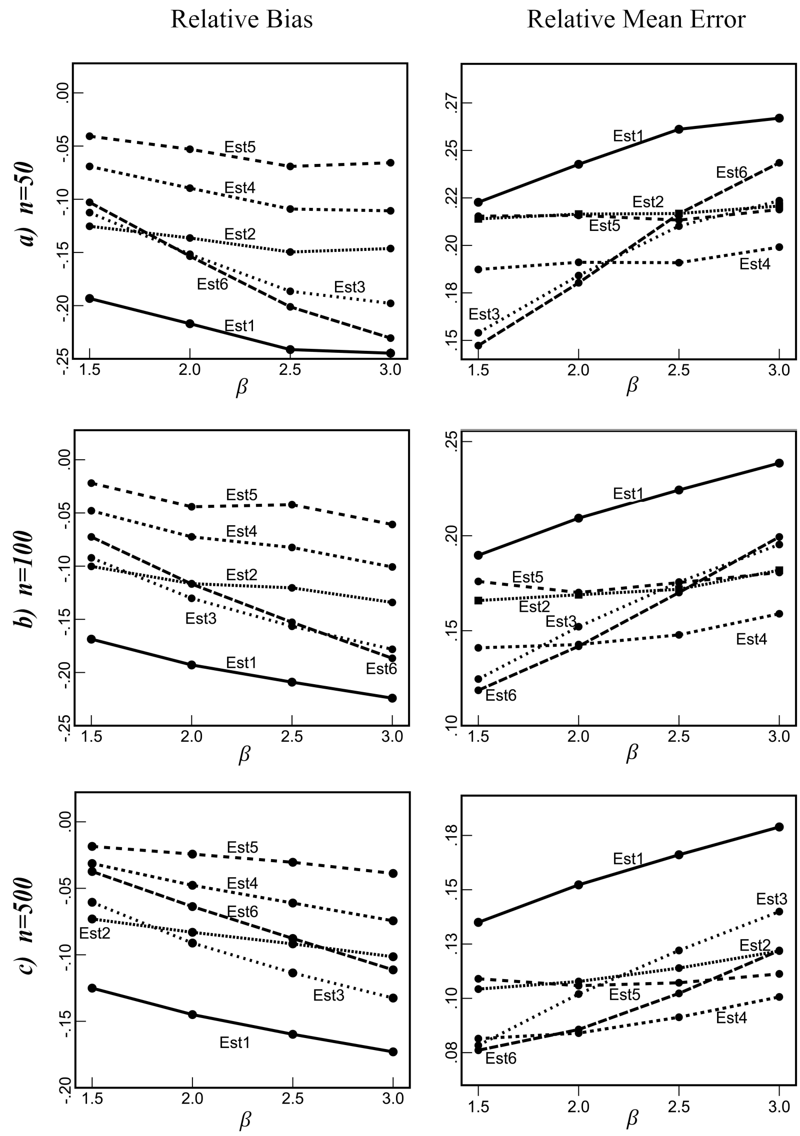

Figure 3: Relative bias and relative mean error of beta model

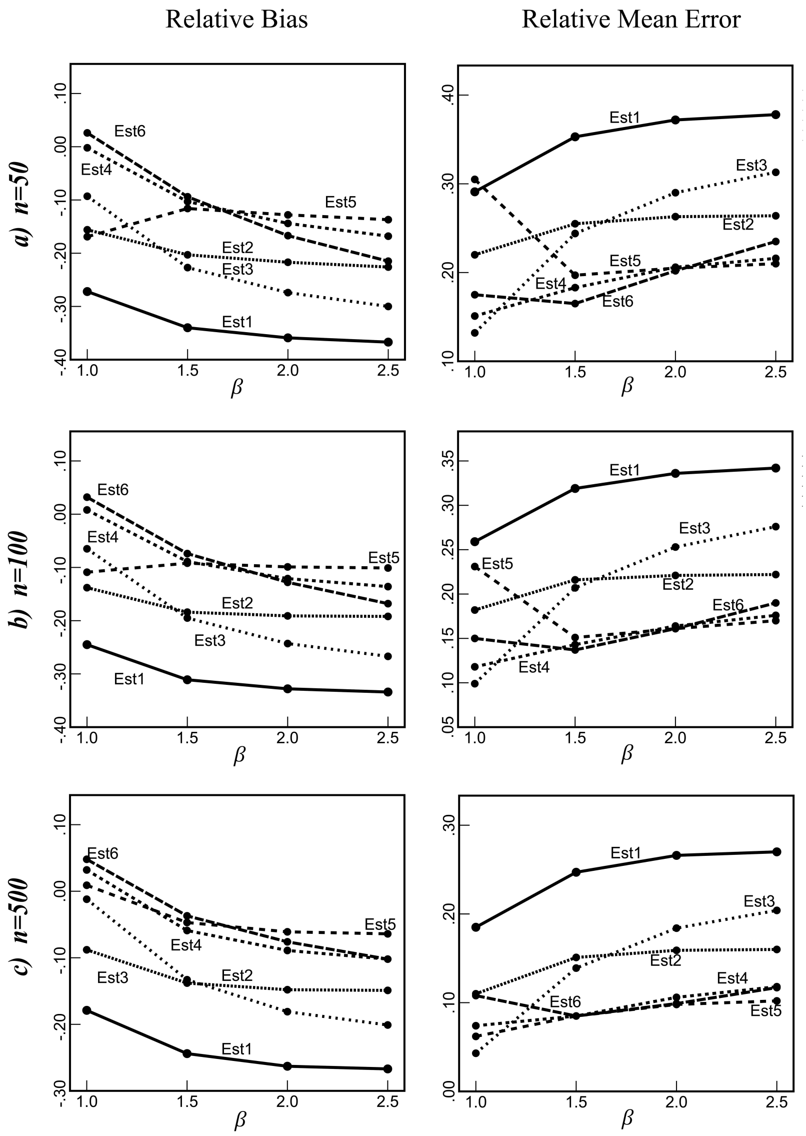

Figure 4: Relative bias and relative mean error of negative exponential model