Journal of the Korean Statistical Society (2005), **34**: 1, pp 49–60

FREQUENCY HISTOGRAM MODEL FOR LINE TRANSECT DATA WITH AND WITHOUT THE SHOULDER CONDITION

OMAR EIDOUS [1]

ABSTRACT

In this paper we introduce a nonparametric method for estimating the probability density function of detection distances in line transect sampling. The estimator is obtained using a frequency histogram density estimation method. The asymptotic properties of the proposed estimator are derived and compared with those of the kernel estimator under the assumption that the data collected satisfy the shoulder condition. We found that the asymptotic mean square error ($AMSE$) of the two estimators have about the same convergence rate. The formula for the optimal histogram bin width is derived which minimizes $AMSE$. Moreover, the performances of the corresponding k-nearest-neighbor estimators are studied through simulation techniques. In the absence of our knowledge whether the shoulder condition is valid or not a new semi-parametric model is suggested to fit the line transect data. The performances of the proposed two estimators are studied and compared with some existing nonparametric and semiparametric estimators using simulation techniques. The results demonstrate the superiority of the new estimators in most cases considered.

AMS 2000 subject classifications. Primary 62D05.

Keywords. Line transect method; shoulder condition; frequency histogram method; kernel method.

1. INTRODUCTION

Line transect method is commonly used by biologists to estimate population density. In addition to its logical framework with intuitive reasoning, sampling using line transect has been a very convenient, easy and relatively cheaper method

Received March 2004; accepted August 2004.

[1]Department of Statistics, Faculty of Science, Yarmouk University, Irbid, Jordan.(e-mail : omarm@yu.edu.jo)

to obtain density of any living or non-living object in an ecosystem. To achieve the experiment, at least one observer moves across the population following a specific line with length L, looking to the right and to the left of the line. It is not sufficient just to record the number of observed objects, n; instead an observer must take the perpendicular distance (x) from from the centerline to a detected object. When objects are observed from a line transect with a detection function $g(x)$, the distance x to observed object from randomly placed transect will tend to have a probability density $f(x)$ of the same shape as $g(x)$ but scaled so that the area under $f(x)$ equal unity. Buckland *et al.* (1980) constitute the key references for this distance sampling procedure.

Logical considerations deriving from the analysis of the physical sighting process suggest that $g(x)$ may usually be assumed monotonically decreasing and satisfying the shoulder condition (*i.e.* $g'(x) = 0$). Accordingly, $f(x)$ is in turn monotonically decreasing with $f'(0) = 0$. However, recent studies have shown that the shoulder condition may not hold for many wildlife line transect data such as whales, jack rabbits, cotton tails and impalas (Buckland, 1985). The basic model for line transect sampling is introduced in the key paper by Burnham and Anderson (1976) who obtain the fundamental relation for estimating the density of objects in a specific area which can be expressed as

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

Accordingly, D can be estimated by

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

where $\hat{f}(0)$ represents a sample estimator of $f(0)$ based on the n observed perpendicular distances $x_1, x_2, \ldots, x_n$ which is usually supposed to be random sample (Buckland *et al.*, 1993). Hence, the key aspects in line transect sampling turns out to be the modeling of $f(x)$ as well as the estimation of $f(0)$.

In a parametric approach, let $f(x)$ be the unknown probability density function of perpendicular distance. A parametric method assumes a model $f(x, \theta)$ which is a member of a family of proper probability density functions of known functional form but depend on an unknown parameter θ, where θ may take a vector value and should be estimated by using the perpendicular distances. A variety of approaches to estimate θ will lead to $\hat{f}(0) = f(0, \hat{\theta})$. A number of parametric models have been proposed for $f(x)$, and there is extensive literature on the use of the maximum likelihood techniques for estimation of $f(0)$. See for example, Burnham and Anderson (1976); Pollock (1978); Burnham et al.(1980) and Buckland (1985).

While parametric methods are very powerful, they are highly dependent on the specification of the model. As an alternative method to parametric approach, recent works has focused on employing the nonparametric kernel method to estimate $f(0)$ (Chen, 1996; Mack and Quang, 1998). Eidous (2004a) introduced the frequency histogram and the frequency polygon methods using line transect data. He showed that the two methods give equivalent estimators for $f(0)$. Eidous (2004b) proposed a simple approach to reduce the bias of the frequency histogram estimator from $O(h^2)$ to $O(h^3)$. Beside that the frequency histogram method is a nonparametric method which removes the model-dependence of the estimator; it employs only the first bin (which contains $x = 0$) in final estimate. That is, apart from the choice of the bin width, the remaining data are superfluous. This tends to highlight the rather strong assumption that the distribution of distances needs to be uniform for line transect methods to work correctly (Buckland et al, 1993).

On one hand, this paper suggested a new estimator for $f(0)$ based on the frequency histogram method in the case that the shoulder condition is true, that is, $f'(0) = 0$. The mathematical derivations showed that the new estimator is very competitor for the kernel estimator in the sense that the new estimator achieves the same orders of bias and variance as the kernel estimator. Moreover, the simulation results -which based on the k-nearest-neighbor approach to choose the bin width- showed that the new estimator performed better that the kernel estimator in some cases considered. On the other hand, if we are not sure about the existing of the shoulder condition, another estimator is suggested. The proposed estimator combines the negative exponential and the frequency histogram models. Its properties are studied through simulation technique, and the results showed that some improvements are attained over the first estimator.

2. THE FREQUENCY HISTOGRAM ESTIMATOR

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a probability density function $f(x)$. Suppose $f(x)$ has support $\Omega = [a, b]$, where a and b are usually taken to encompass the observed data. Partition $[a, b]$ into k non-overlapping bins $T_i = [t_i, t_{i+1})(i = 1, 2, \ldots, k)$ where $a = t_1 < t_2 < \ldots < t_k = b$. Let h be a common bin width, $h = t_{i+1} - t_i$ for all i, and ν_i be the number of sample values falling in bin T_i, where $\sum_{i=1}^{k} \nu_i = n$. Then the frequency histogram estimate,

$\hat{f}(x)$ of $f(x)$ is defined by (Eidous, 2004a)

$$\hat{f}(x) = \frac{\nu_i}{nh} = \frac{1}{nh}\sum_{j=1}^{n} I_{T_i}(x_j), \ x \in T_i, \ i = 1, 2, \ldots, k$$

where I_{T_i} is the indicator function of the ith bin. To implement the above estimator in line transect sampling we concern with the bin that contains zero, that is, $x = 0$, because our aim is to estimate the probability density function at zero perpendicular distance. Take $a = t_1 = 0$ and $t_2 = h$, then we concern with the first bin and the frequency histogram estimate, $\hat{f}(0)$ of $f(0)$ is given by

$$\hat{f}(0) = \frac{\nu_1}{nh} = \frac{1}{nh}\sum_{j=1}^{n} I_{[0,h)}(x_j). \tag{2.1}$$

We note here that, while the frequency histogram estimates of $f(x)$ is a step function the frequency polygon estimates of $f(x)$ (Scott, 1985) is a continuous linear function connecting the bin centers of a frequency histogram. As an estimators for $f(0)$, Eidous (2004a) showed that the frequency histogram and the frequency polygon methods give the same estimator.

3. Asymptotic Properties

In this section we derived the asymptotic mean square error ($AMSE$) of the proposed estimator $\hat{f}(0)$ and compared with the $AMSE$ for the kernel estimator $\hat{f}_k(0)$ given by (5.1). The number of perpendicular distances ν_1 that falls into bin $[0, h)$ is a binomial random variable with parameters n and p, where p is the cell probability given by

$$p = \int_0^h f(u)du.$$

Thus, the expected value of $\hat{f}(0)$ for given the sample size n is

$$E\left(\hat{f}(0)\right) = \frac{1}{nh}E(\nu_1) = \frac{1}{h}\int_0^h f(u)du.$$

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

$$E\left(\hat{f}(0)\right) = f(0) + \frac{h}{2}f'(0) + \frac{h^2}{6}f''(0) + O(h^3).$$

If the shoulder condition is true then the bias of $\hat{f}(0)$ is

$$Bias\left(\hat{f}(0)\right) = \frac{h^2}{6}f''(0) + O(h^3),$$

which indicates that the asymptotic bias of histogram estimator is of order $O(h^2)$ under assumption that the shoulder condition holds. Notice that, we use notation $O(s)$ to represent term as the same order of magnitude of s as explained by Serfling (1980). If h is related to n in such a way that $h \to 0$ and $nh \to \infty$ as $n \to \infty$, then the variance of $\hat{f}(0)$ is

$$\begin{aligned} Var\left(\hat{f}(0)\right) &= \frac{p(1-p)}{nh^2} \\ &= \frac{f(0)}{nh} + O(n^{-1}). \end{aligned}$$

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

$$AMSE\left(\hat{f}(0)\right) = \frac{h^4}{36}[f''(0)]^2 + \frac{f(0)}{nh}, \tag{3.1}$$

where the first term in the right hand side of (3.1) is the square bias and the second term is the variance. Under the assumption that $f'(0) = 0$, Chen (1996) showed that (as the normal kernel has been used) the asymptotic bias and variance of the kernel estimator, $\hat{f}_k(0)$ -which is given by (5.1)- are $h^2 f''(0)/2 + O(h^3)$ and $f(0)/(nh\sqrt{\pi}) + O(n^{-1})$ respectively, and the $AMSE$ of is $h^4[f''(0)]^2/4 + f(0)/(nh\sqrt{\pi})$. Comparing our results with the results obtained by Chen, we find that the frequency histogram has the same convergence rates as the kernel estimator. In which, for frequency histogram estimator the bias is proportional to $f''(0)/6$ and the variance is proportional to $f(0)$. While for the kernel estimator, the bias is proportional to $f''(0)/2$ and the variance is proportional to $f(0)/\sqrt{\pi}$. In other words, if the two approaches use the same value of then the bias of the histogram estimator is less than that of the kernel estimator, and the converse is true when we talk about the variances. From viewpoint of the approximation used, it is easy to show that the $AMSE$ for the histogram estimator is less than that of the kernel estimator if

$$nh^5[f''(0)]^2 > 1.96 f(0).$$

4. BIN WIDTH SELECTION

To implement the new estimator in practice we need to choose the value of the bin width, h. One of the most common methods in nonparametric estimation is to find h that minimizing the $AMSE$. Consider the $AMSE$ as a function of h (say $z(h)$), then differentiate $z(h)$ with respect to h and equating to zero to get,

$$h = \left(\frac{9f(0)}{f''^2(0)}\right)^{1/5} n^{-1/5}. \tag{4.1}$$

The value of h given by (4.1) can be substituted back into (3.1) to give as the minimum achievable MSE for $\hat{f}(0)$ given by

$$\frac{1.25}{3^{0.4}}[f''(0)]^{2/5}[f(0)]^{4/5}n^{-4/5}. \tag{4.2}$$

Correspondingly, if one chose the bandwidth parameter h for the kernel estimator based on minimizing the $AMSE$, then the optimal value of h is h_k given by

$$h_k = \left(\frac{f(0)}{\sqrt{\pi}f''^2(0)}\right)^{1/5} n^{-1/5}. \tag{4.3}$$

Substitute the value of h_k into the $AMSE$ of $\hat{f}_k(0)$ then we get the minimum achievable MSE for $\hat{f}_k(0)$ given by

$$\frac{1.25}{\pi^{0.4}}[f''(0)]^{2/5}[f(0)]^{4/5}n^{-4/5}, \tag{4.4}$$

Comparing (4.2) with (4.4), the two quantities has the same convergence rates as $n \to \infty$. If $n < \infty$ then (4.4) is slightly smaller than (4.2).

Assume that the underlying probability density function $f(x)$ to be half-normal with scale parameter σ^2 then from (4.1) we find $h \cong 1.624\ \hat{\sigma}\ n^{-1/5}$ and from (4.3) we find $h \cong 0.934\ \hat{\sigma}\ n^{-1/5}$, where $\hat{\sigma}$ is the maximum likelihood estimator for σ. A simulation study is performed (which is not stated in this paper) based on the above two values of h and h_k. The simulation results indicated that the performances of the histogram and the kernel estimators are very similar to each other for all models considered in Section (4.4). The simulation study -given in Section (4.4)- is achieved by adopting the k-nearest-neighbor selector, which does not require any assumption about the shape of $f(x)$. Loftsgaarden and Quesenberry (1965) introduced the k-nearest-neighbor selector given by $h = x_{(k)}$, where $x_{(k)}$ represents the kth order statistic in the observed sample. As to the selection of k, a common choice is given by $k = [[n^{\varepsilon}]]$, where $0 < \varepsilon < 1$ and $[[.]]$ denotes the greatest integer function. In this setting, we used $\varepsilon = 4/5$ (See for example, Mack and Rosenblatt, 1979).

5. TESTING THE SHOULDER CONDITION

In the above discussion we assumed that the shoulder condition is true and then achieve a $O(h^2)$ bias. Without the shoulder condition assumption, we get a $O(h)$ bias which is significantly greater than $O(h^2)$. If we are not sure about the validity of the shoulder condition, we propose to use a new estimator for $f(0)$. The basic idea is to use a semiparametric estimator that combines the negative exponential model -which does not satisfy the shape criterion- and the frequency histogram model. Gates *et al.* (1968) suggested the negative exponential model to fit the perpendicular distances. The basic model is

$$f(x, \lambda) = \frac{1}{\lambda}\, e^{-x/\lambda}, \quad x \geq 0.$$

Thus, the proposed semiparametric estimator in this case is of the form

$$\hat{f}^*(0) = (1 - m) f(0, \hat{\lambda}) + m\hat{f}(0).$$

The parameter m is estimated from the data and its estimate $\hat{m}$ is then used in $\hat{f}^*(0)$ as the proposed estimate for $f(0)$. The parameter $f(0, \lambda) = 1/\lambda$ can be estimated by the maximum likelihood estimator $f(0, \hat{\lambda}) = 1/\bar{x}$, where $\bar{x}$ represents the mean for the observed perpendicular distances. In this setting $f(0, \lambda)$ is estimated by the unbiased estimator $(n-1)/n\bar{x}$.

What is less clear in the above semiparametric model is how m should be choosen in the estimator $\hat{f}^*(0)$. The main point is that we need to force $\hat{m}$ to be close to unity when the shoulder condition for the underlying model of the data at hand holds and to be far from unity toward zero when the shoulder condition fails to hold. In other words, a good $\hat{f}^*(0)$ is expected to give high weight for the histogram estimator when the shoulder condition holds and less weight when it does not.

Mack (1998) proposed a procedure for testing the shoulder condition of a model based on line transect sampling. Assume that a random sample $x_1, \ldots, x_n$ of perpendicular distances is drawn from a distribution with probability density function $f(x)$. Then the kernel estimate of $f(0)$ is (Chen, 1996),

$$\hat{f}_k(0) = \frac{2}{nh} \sum_{i=1}^{n} K\left(\frac{x_i}{h}\right). \tag{5.1}$$

The most widely used kernel is the standard normal kernel,

$$K(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}}.$$

Thus as the standard normal kernel is used, $\hat{f}_k(0)$ becomes

$$\hat{f}_k(0) = \frac{2}{nh\sqrt{2\pi}} \sum_{i=1}^{n} e^{-x_i^2/2h^2}.$$

Now to perform the test concerns the shoulder condition we follow Mack (1998)'s testing method. According to him, consider the test $H_0 : f'(0) = 0$ vs $H_1 : f'(0) \neq 0$ then we reject H_0 in favor of H_1 if $|T| > Z_{\alpha/2}$, where $Z_{\alpha/2}$ represents the $\alpha/2$ quantile of the standard normal distribution. The following quantities should be computed to find the value of the test statistics T:

$$T = f'(0)\sqrt{\frac{nb^3}{2\hat{f}_k(0)}},$$

where $f'(0) = [F_n(2b) - 2F_n(b)]/b^2$, $b = \hat{\sigma}\, n^{-1/4}$. $F_n(u)$ is the empirical cumulative distribution function defined by $F_n(u) = \sharp x_i \in [0, u]/n$. The idea to choose the weight parameter m is based on the p-value for the above test. The p-value for the above test is given by

$$\begin{aligned} p &= 2P(Z < -|T|) \\ &= 2\Phi(-|T|), \end{aligned}$$

where $\Phi(x)$ is the standard normal distribution function. The p-value indicates how strongly H_0 is supported by the data. A large p-value indicates in some sense that $f(0)$ is close to the histogram estimator $\hat{f}(0)$. Thus we can use this p-value to estimate m. Based on our preliminary simulations we suggest taking m as $\hat{m} = p^d$ where $0 < d < 1$. In this paper we take $d = 0.1$ as an estimator for d, which performed generally quite well in line transect data (several values of d were tried and we found $d = 0.1$ satisfactory). Thus, the proposed semiparametric estimator is

$$\hat{f}^*(0) = (1 - p^{0.1})\hat{f}(0, \hat{\lambda}) + p^{0.1}\hat{f}(0).$$

6. Simulation Study

Because the exact behavior of the proposed semiparametric estimator $\hat{f}^*(0)$ is complex, we chose to study the sample properties of $\hat{f}^*(0)$ in addition to the first estimator $\hat{f}(0)$ through simulation techniques. The proposed estimators were compared with the nonparametric kernel estimator $\hat{f}_k(0)$ by adopting the k-nearest-neighbor selector method to choose the bandwidth parameter, h. The reason to adopt the k-nearest-neighbor method is that, the method uses the same values of the parameter h for the two estimators, histogram and kernel. As stated earlier, when the value of h taken to be different based on the minimizing the $AMSE$ for each estimator then the performances of the two estimators are similar, which coincide with our discussion in Section (4). The Buckland (1992) semiparametric estimator based on a key half-normal model with Hermite polynomial correction, $\hat{f}_H(0)$ is also considered. Our simulation design is similar to that of Barabesi, (2001), in which three families of models which are commonly used as references in line transect studies were considered in the simulation. The exponential power family (Pollock, 1978)

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

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

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

and the beta model (Eberhardt, 1968)

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

In our simulation design, these three families were truncated at some distance w which required in computing of $\hat{f}_H(0)$. 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 $w = 5.0,\ 3.0,\ 2.5,\ 2.0$. 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 $w = 20,\ 12,\ 8,\ 6$. 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. 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 shape criterion. This choice was made in order to assess the robustness of the considered estimators with respect to the shape criterion. 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 6.1 reports the simulated value of the relative bias (RB)

$$RB = \frac{E\left(\hat{f}(0)\right) - f(0)}{f(0)} \ ,$$

TABLE 6.1 *RB and RME(in parentheses) for the different four estimators of $f(0)$.*

Exponential Power Model	n	$\hat{f}_k(0)$	$\hat{f}_H(0)$	$\hat{f}(0)$	$\hat{f}^*(0)$
	50	-0.31(0.34)	-0.28(0.31)	-0.24(0.28)	-0.19 (0.26)
$\beta = 1, \omega = 5$	100	-0.27(0.29)	-0.27(0.29)	-0.19(0.24)	-0.16(0.21)
	200	-0.25(0.26)	-0.24(0.26)	-0.18(0.20)	-0.13(0.16)
	50	-0.18(0.22)	-0.08(0.16)	-0.10(0.21)	-0.06 (0.19)
$\beta = 1.5, \omega = 3$	100	-0.15(0.18)	-0.08(0.13)	-0.08(0.16)	-0.03(0.15)
	200	-0.13(0.15	-0.08(0.11)	-0.07(0.13)	-0.02(0.11)
	50	-0.12(0.17)	0.01(0.19)	-0.05(0.18)	-0.00(0.18)
$\beta = 2, \omega = 2.5$	100	-0.10(0.14)	0.02(0.15)	-0.04(0.14)	0.02(0.15)
	200	-0.09(0.11)	0.02(0.11)	-0.04(0.11)	0.02(0.12)
	50	-0.08(0.16)	0.06(0.24)	-0.03(0.18)	0.03 (0.19)
$\beta = 2.5, \omega = 2$	100	-0.06(0.11)	0.06(0.19)	-0.02(0.13)	0.05(0.15)
	200	-0.06(0.09)	0.06(0.17)	-0.02(0.11)	0.05(0.12)
Hazard Rate Model					
	50	-0.11(0.23)	-0.37(0.39)	0.02(0.24)	-0.02 (0.20)
$\beta = 1.5, \omega = 20$	100	-0.06(0.16)	-0.36(0.38)	0.09(0.21)	-0.03(0.13)
	200	-0.01(0.11)	-0.36(0.38)	0.13(0.19)	-0.06(0.10)
	50	-0.08(0.19)	-0.10(0.21)	0.03(0.21)	0.04 (0.18)
$\beta = 2, \omega = 12$	100	-0.07(0.16)	-0.08(0.15)	0.03(0.14)	0.04(0.13)
	200	-0.03(0.09)	-0.08(0.14)	0.05(0.12)	0.05(0.10)
	50	-0.07(0.16)	0.08(0.19)	0.01(0.17)	0.05 (0.18)
$\beta = 2.5, \omega = 8$	100	-0.05(0.11)	0.07(0.15)	0.02(0.13)	0.05(0.14)
	200	-0.03(0.08)	0.07(0.11)	0.02(0.10)	0.05(0.11)
	50	-0.06(0.15)	0.10(0.20)	0.00(0.17)	0.06 (0.18)
$\beta = 3, \omega = 6$	100	-0.03(0.10)	0.11(0.17)	0.02(0.13)	0.05(0.14)
	200	-0.02(0.07)	0.11(0.16)	0.01(0.10)	0.06(0.11)
Beta Model					
	50	-0.21(0.25)	-0.03(0.16)	-0.15(0.23)	-0.08(0.20)
$\beta = 1.5, \omega = 1$	100	-0.19(0.22)	-0.04(0.12)	-0.13(0.19)	-0.07(0.16)
	200	-0.17(0.18)	-0.03(0.09)	-0.11(0.15)	-0.05(0.12)
	50	-0.22(0.26)	-0.08(0.15)	-0.15(0.23)	-0.11 (0.21)
$\beta = 2.0, \omega = 1$	100	-0.21(0.23)	-0.08(0.12)	-0.15(0.20)	-0.08(0.16)
	200	-0.19(0.20)	-0.08(0.11)	-0.13(0.16)	-0.06(0.13)
	50	-0.24(0.28)	-0.13(0.18)	-0.17(0.25)	-0.12(0.20)
$\beta = 2.5, \omega = 1$	100	-0.22(0.24)	-0.12(0.15)	-0.15(0.20)	-0.09(0.17)
	200	-0.19(0.21)	-0.13(0.14)	-0.13(0.16)	-0.07(0.13)
	50	-0.25(0.28)	-0.15(0.19)	-0.18(0.25)	-0.13 (0.22)
$\beta = 3.0, \omega = 1$	100	-0.23(0.25)	-0.16(0.18)	-0.15(0.21)	-0.10(0.17)
	200	-0.20(0.21)	-0.16(0.17)	-0.13(0.17)	-0.09(0.14)

and the relative mean error (RME)

$$RME = \frac{\sqrt{MSE\left(\hat{f}(0)\right)}}{f(0)},$$

for each considered estimator.

7. RESULTS AND CONCLUSIONS

Depending on the simulation results given in Table 6.1, several conclusions can be drawn from examining the results in regard to model robustness (RB) and (RME). The estimators $\hat{f}_H(0)$ introduced by Buckland (1992) is with large $|RB|$ for the exponential power model with $\beta = 1$ and for the hazard rate model with $\beta = 1.5$. However, it is with quite small $|RB|$ for other cases. The estimator turn out to be the best for the exponential power model with $\beta = 1.5$ and $n = 50,\ 100$ and for the beta model with $\beta = 1.5,\ 2.0$, with $\beta = 2.5,\ 3.0$ for small sample size. Table 6.1 shows clearly that the $|RB|$ of the histogram estimator $\hat{f}(0)$ is generally smaller than the $|RB|$ of the kernel estimator $\hat{f}_k(0)$, which coincide with our results in Section (3). We note here that $\hat{f}_k(0)$ tends to have a downward bias for all cases considered.

Regarding the RME, the performance of $\hat{f}(0)$ is better than $\hat{f}_k(0)$ for the exponential power model with $\beta = 1.0,\ 1.5$ and for the beta model for all combinations of β and n. The performance of the semiparametric estimator $\hat{f}^*(0)$ is more attractable, it is with the smallest $|RB|$ among the other estimators for almost all cases considered. In terms of RME, the performance of $\hat{f}^*(0)$ is better than $\hat{f}(0)$ and $\hat{f}_k(0)$ in most cases considered, specially when the shoulder condition of the simulated data model fails to valid, e.g. the exponential power model with $\beta = 1$ and the beta model. Moreover, the performance of $\hat{f}^*(0)$ is better than $\hat{f}_H(0)$ for the exponential power model with $\beta = 1.0,\ 2.5$; for the Hazard rate model for all values of β and n; and for the beta model $\beta = 2.5,\ 3.0$ with and large n. Sometimes its performance is similar to that of $\hat{f}_H(0)$ as in the case of the exponential power model with $\beta = 2.0$.

ACKNOWLEDGEMENTS

The author thanks two referees and the Editor for helpful comments and suggestions which have improved the presentation of the paper. Also thanks for Ayman Baklizi for helpful discussion, and to Abedel-Qader Al-Masri for his assistance in the computing work. This research is supported by a grant from Yarmouk University.

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.

BUCKLAND, S.T., ANDERSON, D.R., BURNHAM, K.P. AND LAAKE, J.L. (1993). *Distance sampling*, Chapman and Hall, London.

BURNHAM, K.P. AND ANDERSON, D.R. (1976). "Mathematical models for nonparametric inferences from line transect data", *Biometrics*, **32**, 325-336.

BURNHAM, K.P., ANDERSON, D.R. AND LAAKE, J.L. (1980). "Estimation of density from line transect sampling of biological populations", *Wildlife Monograph*, number 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", *Journal of Wildlife Management*, **32**, 82-88.

EIDOUS, O. M. (2004a). "Histogram and polygon methods using line transect sampling", To appear in *Journal of Applied Statistical Science.*

EIDOUS, O. M. (2004b). "Bias correction for histogram estimator using line transect sampling", *Environmetrics*, **15**.

GATES, C.E., MARSHALL, W.H. AND OLSON, D.P. (1968). "Line transect method of estimating grouse population densities", *Biometrics*, **24**, 135-145.

HAYES, R.J., AND BUCKLAND, S.T. (1983). "Radial distance models for line-transect method", *Biometrics*, **39**, 29-42.

LOFTSGAARDEN, D.O. AND QUESENBERRY, C.P. (1965). "A nonparametric estimate of multivariate density function", *Annals of Mathematical Statistics*, **36**, 1049-1051.

MACK, Y.P. (1998). "Testing for the shoulder condition in line transect sampling", *Communication in Statistics-Theory and Method*, **27**, 423-432

MACK, Y.P. AND QUANG, P.X. (1998). "Kernel methods in line and point transect sampling", *Biometrics*, **54**, 606-619.

MACK, Y.P. AND ROSENBLATT, M. (1979). "Multivariate k-nearest neighbor density estimates", *Journal of Multivariate Analysis*, **9**, 1-15.

POLLOCK, K.H. (1978). "A family of density estimators for line transect sampling", *Biometrics*, **34**, 475-478.

SCOTT, D.W. (1985). "Frequency polygons: Theory and application", *Journal of the American Statistical Association*, **80**, 348-354.

SERFLING, R.J. (1980). *Approximation theorems of mathematical statistics*, Wiley, New York.