A NEW ESTIMATOR FOR LINE TRANSECT SAMPLING BASED ON GROUPED, PERPENDICULAR DISTANCES.

by

Egerton George Johnson

B.Sc. (Hons.), University of Manchester, 1981

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in the Department

of

Mathematics

© Egerton George Johnson 1984

SIMON FRASER UNIVERSITY

December 1, 1983

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without permission of the author.

APPROVAL

Name: Egerton George Johnson

Degree: Master of science

Title of thesis: A NEW ESTIMATOR FOR LINE TRANSECT SAMPLING BASED ON GROUPED, PERPENDICULAR DISTANCES.

Examining committee:

Chairman: A.R. Freedman

R. Routledge
Senior Supervisor

R. Lockhart

D. Eaves

Dr. T. Chang
External Examiner

Date Approved: December 1, 1983

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

A NEW ESTIMATOR FOR LINE TRANSECT SAMPLING BASED ON GROUPED, PERPENDICULAR DISTANCES

Author:

(signature)

Simeon

(name)

13/03/84

(date)

ABSTRACT

The general theory of the line transect method for estimating wildlife population densities and some of the popular estimators being used are reviewed. Amongst the estimators considered are the following; the half-normal estimator, the Fourier series estimator and the generalised exponential series estimator.

A new estimator for wildlife densities using line transect data and based on shape-restrictions for the detection curve is presented. This estimator uses only grouped, perpendicular distances from the observer's line of travel to the sighted object.

The sampling behaviour of the new estimator is studied through simulation studies. These studies show a marked improvement over existing estimators in stability and efficiency for a variety of detection curves.

ACKNOWLEDGMENT

I would like to thank my supervisor Dr R. Routlegde for suggesting the topic of this thesis and for his invaluable help throughout its preparation; also his financial support has been much appreciated.

Thanks also to Cupido for her assistance in the typing of this thesis.

TABLE OF CONTENTS

Approval ii

Abstract iii

Acknowledgment iv

List of Tables vii

List of Figures viii

CHAPTER 1.

Introduction 1

1.1 Main Results and Recommendations 4

CHAPTER 2.

General Theory of Line Transect Methodology

2.1 Underlying Assumptions for Model 7

2.2 General Model for Perpendicular Distances 9

2.3 Motivation for Shape-Restricted Estimator 12

2.4 Derivation of the Shape-Restricted Estimator 14

2.5 Kuhn-Tucker Theory and The Solution Algorithm. 19

CHAPTER 3.

Examples of Three Models and Associated Methods of Estimation 22

3.1 Maximum Likelihood Estimation 23

3.2 A Parametric Case: The Half-normal Estimator 26

3.3 A Nonparametric Case: The Fourier Series Estimator 29

3.4 A Generalised Parametric Case:

The Exponential Power Series Estimator. 33

CHAPTER 4.

The Estimation Problem 37

4.1 Simulation Study Design 39

4.2 Detailed Findings 43

Dicussion and Recommendations. 45

APPENDIX A 56

BIBLIOGRAPHY 66

LIST OF TABLES

TABLE		PAGE
1	Estimated percent relative root mean squared error of h(0), RMSE[h(0)]/h(0)x100%, for three different estimators using the eight Detection curves in Fig. 2. ..	48
2	Estimated percent relative bias of h(0) for three different estimators using the eight Detection curves in Fig. 2.	49
3	Estimated percent relative standard error of h(0), SE[h(0)]/h(0)x100%, for three different estimators using the eight Detection curves in Fig. 2.	50
4	Estimated percent relative efficiencies for the shape-restricted estimator, relative to the half-normal and Fourier series estimators.	51

LIST OF FIGURES

FIGURE PAGE

1 Diagram showing quantities recorded in the line transect sampling:.............................. 2

2 Three parametric models of detection curves:......... 52

(a) Half Normal, $g(x) = \exp[-x^2/2\sigma^2]$ (----),

(b) Generalized Exponential, $g(x) = \exp[-(x/75)\ \ ]$ (- - - -), and

(c) Reversed Logistic, $g(x) = [11\exp(-x/20)]/[1+10\exp(-x/20)]$ (----).

The dashed vertical lines give the 90th percentiles for each of the corresponding probability density curves.

3 The eight detection curves used in the simulation study.

(a) Half Normal curves, $g(x) = \exp[-x^2/2\sigma^2]$, truncated at σ (----), 2σ (- - - -), and 3σ (----). 53

(b) Cosine curves with $g(x) = [1+\cos(\pi x/100)]/2$ (----), and $g(x) = [1+\cos(\pi x/100) + \cos(2\pi x/100)]/2.25$ (----)..54

(c) Piecewise linear forms with kinks at 20 (----), 50 (- - - -), and 80 (----)..........................55

CHAPTER 1

Introduction

The line transect sampling method is a relatively efficient technique for estimating wildlife densities, which has gained popularity in wildlife studies. The method has been used on birds, land animals, and even on plants.

The technique involves travelling along randomly placed lines (transects), and recording the number of sightings n, and distances (radial (r) and/or perpendicular (y)) from the transects (Fig. 1). These observed distances are then used to estimate the population density. Both the practical and theoretical aspects have been extensively discussed by Eberhardt (1978), Gates (1978, 79) and Burnham, et al., (1980).

In order to model line transect sampling methods, assumptions are made about the distribution, response and method of sighting and counting of the animals in the transect study area. These assumptions are outlined in Section 2.2.1.

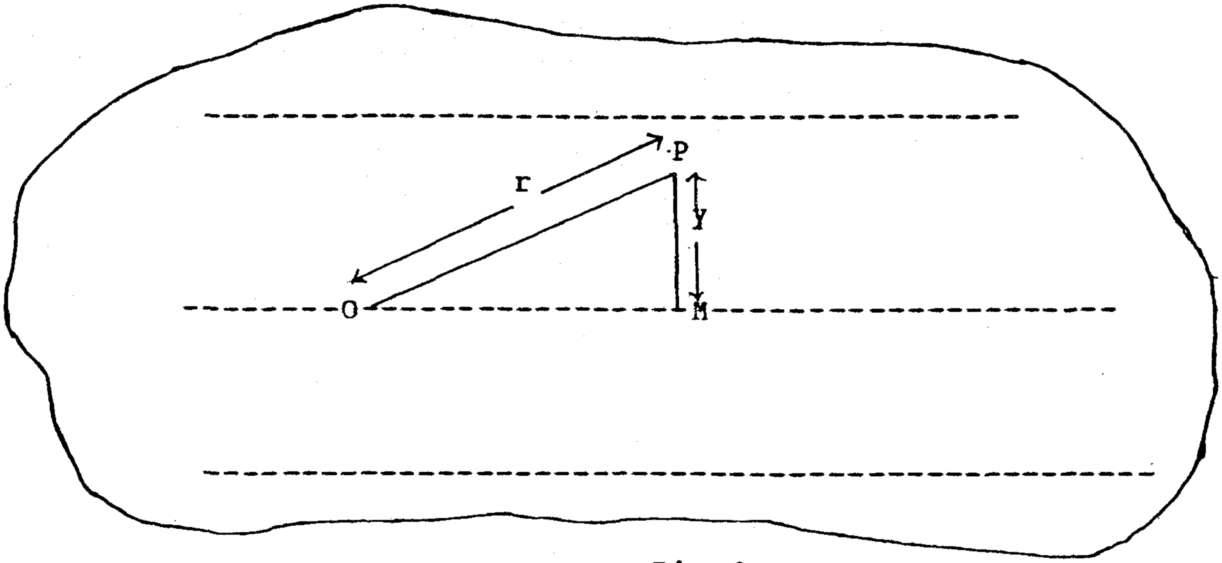

Fig 1
P is the point at which an object is first seen by an observer at O. M is the point on the line perpendicular to the object.

Most development in line transect sampling methods has been concentrated on parametric and nonparametric modelling of the function g(x). This is the probability of a point being seen, given that it is at a perpendicular distance x from the transect line. g(x) is called the detection function. Some of these parametric models are very restrictive and they give reasonable estimates only when the detection function follows closely the form of that used in the model. For example, Gates, et al., (1968) developed the negative exponential, which was later shown to be very restrictive. Any deviation of the detection function from negative exponential produces absurd estimates of f(x). The generalised parametric models (Cain 1974, Pollock 1978) though flexible, are prone to estimation errors. Pollock's (1978) exponential power series model, for example, with

$$g(x) = \exp(-(x/a)^b)$$

encompasses a wide variety of detection curves, including the half-normal (b = 2) and the negative exponential (b = 1) (see Section 3.3.3). However, there is a considerable loss in efficiency when both parameters of the model are estimated. The efficiency can be improved upon by judiciously fixing one of the parameters (Burnham, et al., 1980).

More success has been achieved through nonparametric models. These models, being nonparametric, are designed to perform well over a wide variety of detection curves.
The Fourier series estimator is the most commonly used nonparametric technique for estimating wildlife densities. One possible deficiency of this estimator (and others based on orthogonal series) is that the estimate of f(x) can dip down below zero; i.e., the estimate of f(x) is not a true probability density (Wedgman 1971).

In this thesis, a nonparametric approach based on a constrained least squares technique is developed and its sampling behaviour is investigated.
The main results of a simulation study are given in Section 1.1. In Chapter 2, the general theory of the line transect sampling method based on perpendicular distances is reviewed, and the shape-restricted estimator is derived.

Three models suggested by Burnham and Anderson (1978), Quinn (in a 1977 M.Sc. thesis) and Pollock (1978) are presented in Chapter 3. Two of these are examined in Chapter 4, through an extensive, comparative, simulation study. A FORTRAN list of the simulation program is given in Appendix A.

1.1 Main Results and Recommendations

Since the exact behaviour of estimators subjected to inequality constraints is complex (Barlow, et al., 1972), simulation studies were done to determine the sampling behaviour of the shape-restricted estimator. The simulation results of the shape-restricted estimator will show a general improvement in efficiency and stability over existing estimators.

For comparison purposes, the Fourier series estimator, the shape-restricted estimator and the half-normal estimator were used. Each was tested on eight detection functions shown in Figs. 3(a), 3(b) and 3(c). Statistical quantities such as the average standard error and bias are calculated as percentages of the true density, D.

The shape-restricted estimator behaves well in terms of its efficiency, standard errors, and bias. Unless the detection curve closely resembles the form best suited to one of the other estimators, the shape-restricted estimator has smaller root mean squared error. For example, one of the detection functions used was based on the two-term cosine curve, the root mean squared error of the shaped-restricted estimator based on 100 observation using this curve is 97% and 82% of that for the half-normal and Fourier series estimators respectively. The corresponding efficiency ratios are 106% and 147%. For a wide variety of detection curves, the simulated results show that the shape-restricted estimator has an overall percent relative bias that is comparable to that of the commonly used Fourier series estimator (Table 2).

The results in Table 3 demonstrate the stability of the shape-restricted estimator. This stability and degree of flexibility in fitting detection curves from the set of possible detection curves makes it a fairly robust nonparametric estimator. The half-normal estimator exhibits unstable behaviour. It behaves spectacularly well for detection curves close to the half-normal form, but is prone to large bias for other forms of detection curve (Table 2).

Likewise, the Fourier series estimator exhibits the same unstable behaviour as the half-normal estimator, but with a much more erratic pattern. However, when the detection curve follows closely that of the one-term cosine series (Fig. 3(b)), it has a high degree of accuracy. For other detection curves, it has a much larger relative root mean squared error. This is true even for the two-term cosine (Fig. 3(b)). The Fourier series is unlike the half-normal in that it has much smaller bias, with slight increase in its chance components.

The efficiency and stability of the shape-restricted estimator over a wide variety of detection curves makes it a plausible alternative to existing estimators for grouped data whenever the parametric form of the detection curve is unknown. It has great advantages in allowing the user with prior information about the detection curve, to transfer this information into restrictions on the estimator. One need only to alter the constraint matrix on input. Section 2.4 shows how these restrictions on the detection curve can be transferred to the estimator.

CHAPTER 2

General Theory of Line Transect Methodology

2.1 Underlying Assumptions for Model

In order to develop any mathematical model for line transect sampling, several assumptions about both the spatial distribution of the birds, land animals or plants in the study area and about the methodology of sampling must be made. For the purpose of modelling line transect sampling methods, birds, land animals and plants are referred to, in abstract terms, as points or objects. The following are the basic assumptions underlying various parametric and nonparametric models.

1. Points are randomly and independently distributed over the study area, A, with density, D per unit area; i.e., $P(\text{point is in } (x, x+dx) \mid \text{it is in } A) = 2Ldx/A$; where L is the length of strip traversed.
2. Sightings are independent events.
3. No points are counted more than once.
4. Points are stationary, at least until detection.
5. The detection function is constant throughout the study. (In particular the response behaviour of the population ought to remain fairly stable in the course of running

the transect.)

6. Points exactly on the transects are seen with probability one.

Let the probability of a point being seen, given that it is at a perpendicular distance x from the transect line be called the detection function, $g(x)$;
i.e., $g(x) = P(\text{point is seen} \mid \text{perpendicular distance}, x)$.
Then assumption six becomes $g(0) = 1$.

Eberhardt (1978a), discussed these underlying assumptions, and made suggestions for estimating the density (D) when some of these assumptions are violated.
For example, assumption 2 is violated when objects are being observed in groups (birds flushing in groups). This problem can be avoided by estimating the number of groups, and then adjusting this estimate by multiplying it by the average group size. The final estimate may be biased, since the average group size is a biased estimate of the population mean if the group sizes and radial distances are correlated.

It has been observed that in practice, objects may move away from the observer, thus violating the 'no movement' assumption. This problem can be alleviated by using monotonic nonincreasing estimators for estimating the density of the objects (Burnham, et al., 1980). The isotonic regression estimator (Barlow, et al., 1972) and shape-restricted estimator (Section 4.2) satisfy this property.

Given that the detection function, $g(x)$, is continuous at $x = 0$ and $g'(0) = 0$, then it is sensible to assume that in the neighbourhood of the transect, the observer would see objects with probability close to 1; meaning that, the detection function would have a shoulder near $x = 0$. Indeed observers have generally found that whenever $g(0) = 1$, it remains very close to 1 for small x. That is; the graph of $g(x)$ has a shoulder near $x = 0$.

2.2 General Model for Perpendicular Distances

Given the above assumptions, we derive the distribution of the observed distances $x_1, x_2, \ldots\ldots, x_n$ conditional on n, the number of sightings. We define the intensity function, $h(x)$ (through which we shall obtain the density of objects), and establish its relationship to the detection function and the probability density function of observed distances. The density, D, of objects, is derived in terms of the above probability density function and hence the intensity function.

Consider a strip centred on the transect with half-width, T, Let its total area be given by A. Then from assumption 1,

$P(\text{point is in } (x, x+dx) \mid \text{it is in } A) = 2L\,dx/A.$

(Formal limit definitions are omitted for clarity of exposition.) From the definition of $g(x)$,

$P(\text{point is seen} \mid \text{point is in } (x, x+dx)) = g(x),$

so

$P(\text{point is seen in } (x, x+dx) \mid \text{point is in } A)$

$$= [2L\,dx/A]\, g(x),$$

and the probability, P_T of seeing a point at distances $\le T$ from the transect is

$$= [2L/A] \int_0^T g(x)\, dx$$

$$= [1/T] \int_0^T g(x)\, dx\ .$$

Thus $P_T = [1/T]\mu_T$. Where $\mu_T = \int_0^T g(x)\, dx$.

(for the unbounded strip $\mu = \int_0^\infty g(x)\, dx$ assuming $g(x)$ is such that $\lim_{T\to\infty} \mu_T = \mu = \lim_{T\to\infty} \int_0^T g(x)\, dx. < \infty$).

Now $P(\text{point is in } (x, x+dx) \mid \text{point is seen})$

$$= \frac{P(\text{point is seen in } (x, x+dx))}{P(\text{point is seen})}$$

$$= \frac{g(x)\, dx\, /T}{P_T}$$

$$= \frac{g(x)\, dx}{TP_T}$$

$$= \frac{g(x)\, dx}{\mu_T}\ .$$

so for any $T > 0$ and conditional on n, $x_1, x_2, \ldots, x_n$ are i.i.d with density

$$f(x) = g(x)/\mu_T. \qquad (2.2.1)$$

Now define a function, $h(x)$, called the <u>intensity</u> function as the fraction of sightings seen in x to $x + dx$ multiplied by the expected number of sightings seen in the study area. This is just the probability density for the distance of a single observation multiplied by the expected number of sightings; that is, $h(x) = f(x)\, E(n)$.

The relationship between $h(x)$ and $g(x)$ can be shown by considering the number of observations at distances between x and $x+dx$ from the transect. Since only a fraction, $g(x)$, of these points in this range is expected to be detected, the expected number of sightings is $2DLg(x)\,dx$. The intensity function is the limit of the expected number of sightings in x to $x + dx$ divided by dx as $dx \rightarrow 0$. Therefore $h(x) = 2DLg(x)$. This implies that the expected number of sightings, $E(n)$, for the study area is given by;

$$E(n) = \int_0^T h(x)\, dx = 2LD\int_0^T g(x)\, dx.$$

Thus

$$E(n) = 2LD\mu_T. \qquad (2.2.2)$$

In order to estimate the density D, of points per unit area, we only need to estimate $h(0)$. (Since $h(0)= 2LD$, which implies $D = h(0)/2L$.)

From (2.2.2), let $D = n/2L\mu_T$.

From assumption 6, $g(0) = 1$.

Hence from (2.2.1)

$$f(0) = g(0)/\mu_T = 1/\mu_T,$$

which implies

$$f(0) = 1/\mu_T.$$

So to estimate $1/\mu_T$, we could also use an estimate of the probability density function $f(x)$ at zero; i.e., $f(0)$. Since $h(x) = E(n)f(x)$, it is sensible to require that $\hat{h}(x) = n\hat{f}(x)$. We have then an estimate of $h(0)$ given by $n\hat{f}(0)$.

Thus

$$\hat{D} = n\hat{f}(0)/2L = \hat{h}(0)/2L . \qquad (2.2.3)$$

2.3 Motivation for Shape-Restricted Estimator

Histogram-like-estimates are maximum likelihood estimates and have been in use since 1895 (Tapia and Thompson, 1978). As was observed by Wedgman (1972b), they have large errors compared to orthogonal series, but the rate of convergence of their mean integrated squared error, (MISE) seems to be better than those of the orthogonal series or kernel estimates. An estimate of the rate of convergence of the MISE for the classical histogram is $O(n^{-1})$ and the rate for both the orthogonal series and kernel estimates is $O(n^{-4/5})$ (Wedgman 1978b). Thus, if some appropriate smoothing could be introduced, it is conceivable that the large standard errors for histogram estimators, based on small samples, could be improved upon without a concomitant degradation of convergence. Hence, small standard errors might be available for all sample sizes.

One form of smoothing which has been considered by Barlow, et al., (1972), and also by Hayes (unpublished thesis) is that obtained by insisting that the density function be monotonically nonincreasing. This (least squares) estimator known as the isotonic regression estimator is also the maximum likelihood estimator of f(x), under the sole restriction that f(x) be monotonic nonincreasing. Its major weakness is its tendency to produce spiked estimate of f(x) at x = 0.

What is being presented here is a similar form of smoothing, but one which also incorporates some knowledge of plausible detection curves encountered in line transect sampling.

From (2.2.3), the estimate of density (D) of points is given by

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

Where f(0) and h(0) are the underlying p.d.f. and intensity function evaluated at x = 0.

A naive estimation technique for $\hat{h}(0)$ would be to use only observations close to the transect line. But this completely ignores information available from observations away from the transect line and produces inefficient estimates. One way around this is to estimate the entire h(x), incorporating all observations.

This estimation is done through the observed intensity within reasonably small intervals of distance; i.e., through the frequency histogram of observed distances, imposing the following shape-restrictions on g(x). (Restrictions on g(x) transfer directly to h(x)).

(i) Monotonicity (g(x) is assumed monotonic nonincreasing)

(ii) Concavity (g(x) is assumed to curve downwards over the range containing about 90% of the observations), and

(iii) Positivity (g(x) is non-negative in domain of g).

2.4 Derivation of the Shape-Restricted Estimator

The estimator is based on grouped data. Hence consider a partition of the interval [0,T), say

$$0 = t_1 < t_2 <, \ldots, < t_{m+1} = T.$$

Let T_i denote the half-open interval $[t_i, t_{i+1})$ for $i = 1, \ldots, m$. For random samples $x_1, x_2, \ldots, x_n$ in (0,T), let n denote the number of these samples falling in the interval T_i

$$\text{Then } \sum_{i=1}^{m} n_i = n .$$

Let $f = (n_1, n_2, \ldots\ldots, n_m)$, be the vector of frequencies. We now have to transfer the above restrictions (i) through

(iii) on $g(x)$ to restrictions on the estimator. Before developing these restrictions, we define the following quantities.

(a) <u>Convex Set</u>: A set X is convex if for any two points x_1, and x_2 in X, all points on the line segment joint x_1, and x_2 are also in X. That is; x_1, x_2 in X implies $\theta x_1 + (1-\theta)x_2$, is also in X.

(b) <u>Monotone function</u>: A function g is monotone on some interval I if either g is nondecreasing ($x_1 < x_2 \Longrightarrow g(x_1) \le g(x_2)$) on I, or g is nonincreasing ($x_1 < x_2 \Longrightarrow g(x_1) \ge g(x_2)$) on I.

(c) <u>Convex (Concave) function</u>: A function g is convex (concave) over a set X if for any two points x_1 and x_2 in X and for all θ, $0 \le \theta \le 1$;

$$g(\theta x_1 + (1-\theta)x_2) \underset{(\ge)}{\le} \theta g(x_1) + (1-\theta)g(x_2).$$

A function is <u>strictly convex</u> if '≤' is replaced by '<' whenever $x_1 \neq x_2$, and $0<\theta<1$ or alternatively, if the Hessian of $g(x)$ is defined and is positive definite for all x.

To formulate the problem as one of constrained minimisation, we first observe that the raw estimate of $h(x)$ (i.e., the estimate of the intensity function through the frequency histogram of the observed distances) is given by,

$$\hat{h}(x) = \sum_{i=1}^{M} n_i / (t_{i+1} - t_i) \; I(T_i)(x),$$

where $I(T_i)(x)$ denotes the indicator function of the interval T_i. If the smoothed estimate in the i'th class $= y_i$, then the y -values minimise $\sum_{i=1}^{M} (y_i - \hat{y}_i)$, with $\hat{y}_i = \hat{h}(x)$, the raw estimate of the intensity in the i'th class, subject to the constraints that the y-values be non-negative and nonincreasing and their graph be concave out to a prespecified point. That is; we are to

minimise

$$\sum_{i=1}^{M} (y_i - y_i)^2$$

subject to

$$\begin{aligned}
y_1 &\ge y_2 \\
y_1 - y_2 &\le y_2 - y_3 \\
y_2 - y_3 &\le y_3 - y_4 \\
\vdots \quad & \quad \vdots \\
y_{k-2} - y_{k-1} &\le y_{k-1} - y_k \\
y_{k-1} &\ge y_k \\
&\vdots \\
y_{n-1} &\ge y_n \\
y_n &\ge 0.
\end{aligned}$$

where k is that interval containing the 90th percentile. (2.4.1) i.e.; the problem is to minimise

$$\sum_{i=1}^{m} (y_i - \hat{y}_i)^2$$

subject to

$$\begin{aligned} a_1(y) &= y_1 - y_2 && \geq 0 \\ a_i(y) &= -y_{i-1} + 2y_i - y_{i+1} && \geq 0 \quad i=2,\ldots,k \qquad (2.4.2) \\ a_i(y) &= y_i - y_{i+1} && \geq 0 \quad i=k+1,\ldots,n \\ a_n(y) &= y_n && \geq 0. \end{aligned}$$

We can write (2.4.2) in a more compact form as follows;

minimise

$$f_0 = \|y - \hat{y}\|^2 \qquad (2.4.3)$$

subject to

$$Ay \geq 0.$$

Where $\|y\|$ is the Euclidean norm of y defined as

$$\|y\|^2 = \sum_{i=1}^{m} y_i^2.$$

The i'th row of the matrix A is given by a , i=1,...n.

For the above quadratic minimisation problem, if we can show that the objective function is strictly convex and the admissible y-values constitute a convex set, then then we are guaranteed that any local minimum of (2.4.3) is also the unique global minimum. In what follows, we show that the restrictions together do indeed constitute a convex set and that f_0 is strictly convex.

Because the functions $a_i(y)$ are concave functions for all i, then the set of points Y satisfying $a_i(y) \geq 0$ $i = 1,\dots,n$ is a convex set. The proof is as follows:

Take any two points, y_1 and y_2 in Y. Then $a_i(y_1) \geq 0$, and $a_i(y_2) \geq 0$ for $i = 1,\dots,n$.
Then for any i, $i = 1,\dots,n$
we have

$$a_i(\theta y_1 + (1-\theta)y_2) \geq \theta a_i(y_1) + (1-\theta)a_i(y_2) \quad \text{by concavity}$$
$$\geq 0 \quad \text{since } \theta, 1-\theta, \geq 0.$$

Hence $\theta y_1 + (1-\theta)y_2$ is in the set Y whenever y_1 and y_2 are, which from the definition of a convex set implies that the inequality constraints together constitute a convex domain.

Similarly, the proof that f is strictly convex is as follows;
We have, $f_0 = \sum_{i=1}^{M} (y_i - \hat{y}_i)^2 = \| y - \hat{y} \|^2$.
and the gradient vector, ∇f, is given by

$$\nabla f = [2(y_1 - \hat{y}_1), 2(y_2 - \hat{y}_2), \dots, 2(y_m - \hat{y}_M)].$$

and the Hessian $H = 2I$ where I is the identity matrix, which is positive definite. Hence f_0 is strictly convex, and there is a unique minimum to (2.4.3).

2.5 Kuhn-Tucker Theory and The Solution Algorithm.

We are faced with the constrained minimisation problem:

Minimise

$$\| y - y \|^2 = f_0(y),$$

subject to

$$-Ay \leq 0.$$

Solutions to such constrained minimisation problems can sometimes be deduced through the Kuhn-Tucker conditions. These conditions are essential in characterising the solutions and also in defining the vector of Lagrange multipliers $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. They form the foundation for development of some algorithms used in solving constrained minimisation problems.

The Kuhn-Tucker necessary conditions for y and λ to be stationary points for the above minimisation problem can be summarised as follows:

(1) $\lambda_1, \lambda_2, \ldots, \lambda_n \leq 0.$

(2) $\nabla f_0(y) - \lambda \nabla Ay = 0.$

i.e., $2(y - \hat{y}) + \lambda A = 0,$

where $(y - \hat{y}) = ((y_1 - \hat{y}_1), (y_2 - \hat{y}_2), \ldots\ldots, (y_m - \hat{y}_m)).$

(3) $\lambda_i(-a_i(y)) = 0 \qquad i = 1, \ldots, n.$

(4) $-Ay \leq 0.$

These necessary conditions are also sufficient conditions since f(y) is strictly convex and the minimisation is over a convex domain (Rockafellar, 1970, Thm. 27.2. and preceding materials). Thus, the local minimum is a global minimum. It is however difficult to solve these resulting conditions explicitly for y and . Consequently, the minimisation problem was transformed to one of LEAST DISTANCE PROBLEM (LDP). (Hanson and Lawson, 1974, Ch. 22). The Kuhn-Tucker conditions form the foundation upon which the algorithm for solving the LDP is based. The LDP, was then solved via a NON-NEGATIVE LEAST SQUARES PROBLEM (NNLS) (see Hanson and Lawson, 1974, Ch. 22 for details).

To transform to LDP format, the problem of

minimising

$$\| y - \hat{y} \|^2 \qquad (2.5.1)$$

subject to

$$Ay \geq 0,$$

let $r = y - \hat{y}$.

Then (2.5.1) becomes,

minimise

$$\| r \|^2$$

subject to

$$Ar \geq v.$$

Where $v = -A\hat{y}$.

We now seek the point r* of the polyhedral set

$\{r \mid Ar \geq v\}$ at least distance from the origin $r = 0$. Lawson and Hanson (1974) NNLS algorithm finds r^* by iteratively searching for the solution to the Kuhn-Tucker conditions.

CHAPTER 3

Examples of Three Models and Associated Methods of Estimation

Most estimation methods in line transect sampling involve parametric or nonparametric estimation of $f(x)$. Parametric estimators are typically maximum likelihood estimators. Two examples of parametric models are the half-normal with $f(x) = (2/\sigma\sqrt{2\pi}) \exp(-x^2/2\sigma^2)$ and parameter σ, and the negative exponential with $f(x) = a \exp(-ax)$ and parameter a (see Sections 3.2 and 3.4).

Nonparametric estimates are usually obtained by the method of expectations or least squares techniques.
The former technique lends itself readily to the Fourier series with ungrouped data, with great ease of computation. One of the problems encountered in using this technique is that, estimates of $f(x)$ may not be true probability densities. In fact, when applied to the Fourier series estimator, the estimate $\hat{f}(x) = (1/T) + \sum_{j=1}^{p} a_j \cos(j\pi x/T)$ may dip below zero.

The general estimation techniques for three popular methods namely, the Fourier series, the half-normal and the exponential power series models are reviewed.

3.1 Maximum Likelihood Estimation

Let $f(x,\theta)$ $0<x<T$ be a pdf with parameter θ, which may be vector valued, $\theta = (\theta_1, \theta_2, \ldots \theta_r)$. As was noted earlier, T can be finite or infinite. Let $x_1, x_2, \ldots, x_n$ be random perpendicular distances from $f(x,\theta)$.

Define the conditional likelihood function given n as

$$L(\theta) = f(x_1,\theta)\ f(x_2,\theta), \ldots, f(x_n,\theta)$$

$$= \prod_{i=1}^{n} f(x_i,\theta),$$

and $l(\theta) = \ln(L(\theta))$ as the log likelihood function. The maximum likelihood (ML) estimate(s) of θ, θ for a well behaved model can be found by solving the following equations for 0.

$$g_j(\theta) = \frac{\partial l(\theta)}{\partial \theta_j} = \theta. \qquad j=1,\ldots,r$$

The asymptotic variance - covariance matrix of θ is $I^{-1}(\theta)/n$ where $I(\theta)$ is the Fisher information matrix and has elements

$$I_{jk}(\theta) = -E\left[\frac{\partial^2 l(\theta)}{(\partial\theta_j)(\partial\theta_k)}\right]. \qquad j,k=1,\ldots,r$$

The former equations often require numerical solutions since the analytic equations involved are sometimes very complex. A hybridisation of the Newton-Raphson and Marquardt procedure is generally employed. An empirical estimate of the information matrix $I(\theta)$ is derived through H, with elements,

$$H_{jk}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 l(x_i, \theta)}{(\partial\theta_j)(\partial\theta_k)}. \qquad j,k = 1,\ldots,r$$

The ML estimate is $\hat{f}(0) = f(0,\hat{\theta})$ and its asymptotic sampling variance is

$$\mathrm{Var}(f(0,\hat{\theta})) = \frac{[\nabla f(0,\hat{\theta})]' \, I^{-1}(\hat{\theta}) \, [\nabla f(0,\hat{\theta})]}{n}.$$

Where $\nabla f(x)$ is the gradient vector of $f(x)$ with respect to θ.

Suppose the data were grouped into m classes with intervals defined by

$$0 = t_1 < t_2 < \ldots\ldots < t_{m+1} = T$$

Since detections are independent events, the n are multinomial random variables, define by

$$P(n_1, n_2, \ldots, n_m) = \frac{n!}{n_1!\, n_2! \ldots n_m!} \, p_1^{n_1} \, p_2^{n_2} \ldots p_m^{n_m}$$

where

$$p_i = \int_{t_i}^{t_{i+1}} f(x,\theta) \, dx \qquad i = 1,\ldots,m$$

$$= F(t_{i+1}) - F(t_i),$$

and

$$\frac{\partial p_i}{\partial\theta_j} = \frac{\partial F(t_{i+1})}{\partial\theta_j} - \frac{\partial F(t_i)}{\partial\theta_j} \qquad \begin{array}{l} i = 1,\ldots,m. \\ j = 1,\ldots,r \end{array}$$

The log likelihood function $l(\theta)$ is given by

$$l(\theta) = \ln K(n) + \sum_{i=1}^{n} n_i \ln p_i,$$

where $K(n) = \dfrac{n!}{n_1!n_2!\ldots n_M!}$.

Under suitable regularity conditions (Rao, 1973), the ML estimates are the solutions to the log likelihood equations

$$\frac{\partial l(\theta)}{\partial \theta_j} = \sum_{i=1}^{M} \frac{n_i}{p_i} \frac{\partial p_i}{\partial \theta_j} = 0. \qquad j = 1,\ldots,r$$

The Fisher information matrix I(0) has elements given by

$$I_{jk}(\theta) = \sum_{i=1}^{M} \frac{1}{p} \left\{\frac{\partial p_i}{\partial \theta_j}\right\} \left\{\frac{\partial p_i}{\partial \theta_k}\right\}. \qquad j,k=1,\ldots,r$$

Fisher's method of scoring can be used with Marquardt modification for an iterative numerical method in solving the above equations.

It is not always possible to use ML techniques in estimation. In cases where the proposed density estimator (f(x)) is not a bona-fide density function (density dipping below zero), a method of expectations is used in avoiding taking logarithms of negative numbers. Estimation for the Fourier series is based on this method for ungrouped data.

3.2 A Parametric Case: The Half-normal Estimator

Consider the untruncated half-normal model defined by

$$f(x) = \frac{1}{\sigma}\left(\frac{2}{\pi}\right)^{1/2} \exp\left(-x^2/2\sigma^2\right) \qquad 0<x<\infty \quad 0<\sigma . \quad (3.2.1)$$

The truncated model is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \frac{\exp\left(-x^2/2\sigma^2\right)}{[F(T/\sigma) - 1/2]} \qquad 0<x<T \quad 0<\sigma . \quad (3.2.2)$$

Where F(.) is the standard normal cumulative distribution function (cdf).

The following example considers ML estimation with ungrouped, truncated data.

Using (3.2.2), the log likelihood l(0) is given by

$$l(\sigma) = -n\ln\sigma - \frac{n}{2}\ln(2\pi) - n\ln\left[F(T/\sigma) - 1/2\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2 .$$

The ML estimate of σ is the solution to the equation

$\partial l/\partial\sigma = 0$. i.e.,

$$\frac{dl}{d\sigma} = \frac{-n}{\hat{\sigma}} - \frac{n}{[F(T/\hat{\sigma}) - 1/2]} F'(T/\hat{\sigma}) + \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{n} x_i^2 = 0.$$

Furthermore,

$$\frac{\partial^2 l}{\partial\sigma} = \frac{n}{\sigma} - n\frac{[F(T/\sigma) - 1/2) F''(T/\sigma] - [F(T/\sigma)]}{[F(T/\sigma) - 1/2]} - \frac{3}{\sigma^4}\sum_{i=1}^{n} x_i^2 ,$$

and the information matrix given by

$$I(\sigma) = -E\left[\frac{\partial^2 l}{\partial\sigma}\right].$$

where $F(u) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-x^2/2\sigma^2)\, dx,$

$$F'(T/\sigma) = \frac{dF(T/\sigma)}{d\sigma} = \frac{1}{\sqrt{2\pi}} \exp(-T^2/2\sigma^2)(-T/\sigma^2),$$

$$F''(T/\sigma) = \frac{1}{\sqrt{2\pi}}\frac{T}{\sigma} \exp(-T^2/2\sigma^2)(2 - T^2/\sigma^2),$$

and

$$E(x^3) = \frac{J(\sigma, T)}{2}\left[2\sigma^4 - \sigma^2 \exp(-T^2/2\sigma^2)(T^2 + 2\sigma^2)\right]. \qquad (3.2.3)$$

Where $J(\sigma, T)$ is defined as in (3.2.6),

and $Var(\hat{f}(0)) = [\nabla f(0)]^2 I^{-1}(\hat{\sigma}),$ (3.2.4)

with $\hat{f}(0) = \frac{1}{\sqrt{2\pi}}\frac{1}{\hat{\sigma}}\frac{1}{(F(T/\hat{\sigma}) - 1/2)},$ (3.2.5)

and

$$\frac{d\hat{f}(0)}{d\sigma} = \frac{f(0)}{\sigma}\left[-1 + Tf(0)\exp(-T^2/2\sigma^2)\right].$$

The ML estimate will have to be found by iterative techniques such as the Newton-Raphson method (Seber 1973).

or the method of scoring (Rao 1973). The ML estimator of f(0) can then be obtained from (3.2.5), and approximate variance from (3.2.4).

For grouped data estimation, integrating f(x) with respect to x gives

$$B(x) = J(\sigma, T) \int_0^x \exp(-u^2/2\sigma^2) \; du$$

$$= 2 J(\sigma, T) \; (F(x/\sigma) - 1/2).$$

Where $J(\sigma, T) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{1}{(F(T/\sigma) - 1/2)}$. (3.2.6)

The cell probabilities are given by

$$p_i = B(t_{i+1}) - B(t_i)$$

$$= 2 J(\sigma, T) \; [F(t_{i+1}/\sigma) - F(t_i/\sigma)],$$

then $$\frac{\partial p_i}{\partial \sigma} = 2 J(\sigma, T) [F'(t_{i+1}/\sigma) - F'(t_i/\sigma)]$$

$$+ J'(\sigma, T) [F(t_{i+1}/\sigma) - F(t_i/\sigma)].$$

where $J'(\sigma, T)$ and $F'(t/\sigma)$ are calculated using equations (3.2.3). PROGRAM TRANSECT has these equations coded in FORTRAN routines. This program was used in the comparative study.

3.3 A Nonparametric Case: The Fourier Series Estimator

The Fourier series estimator is based on the Fourier series expansion of a function over a finite interval, (Tarter and Kronmal 1968,1976).

Define the even extension $\phi(x)$ of $f(x)$ from $[0,T]$ to $[-T,T]$ as

$$\phi(x) = \begin{cases} f(x) & 0 \leq x < T \\ f(-x) & -T \leq x \leq 0. \end{cases}$$

The Fourier expansion of $\phi(x)$ over $[-T,T]$ is

$$\phi(x) = \frac{1}{T} + \sum_{j=1}^{\infty} \left[a_j \cos\left(\frac{j\pi x}{T}\right) + b_j \sin\left(\frac{j\pi x}{T}\right) \right] \quad -T \leq x \leq T.$$

where $$a_j = \frac{1}{T}\int_{-T}^{T} \phi(x) \cos\left(\frac{j\pi x}{T}\right) dx$$

$$= \frac{2}{T}\int_{0}^{T} f(x) \cos\left(\frac{j\pi x}{T}\right) dx$$

$$b_j = \frac{1}{T}\int_{-T}^{T} \phi(x) \sin\left(\frac{j\pi x}{T}\right) dx = 0.$$

It follows then that,

$$f(x) = \frac{1}{T} + \sum_{j=1}^{\infty} a_j \cos\left(\frac{j\pi x}{T}\right) \quad 0 \leq x \leq T.$$

We want to estimate $f(0)$. Noting that $\cos(0) = 1$. we have

$$\hat{f}(0) = \frac{1}{T} + \sum_{j=1}^{\infty} \hat{a}_j \doteq \frac{1}{T} + \sum_{j=1}^{p} a_j \quad 0 \le x \le T.$$

Although f(x) integrates to 1, f(x) may not be a bona-fide density function. The method of expectation is used to estimate the a_j's.

We first observe that

$$a_j = \frac{2}{T} \int_0^T f(x) \cos\left(\frac{j\pi x}{T}\right) dx$$

$$= \frac{2}{T} E\left[\cos\left(\frac{j\pi x}{T}\right)\right].$$

and so an estimator for a_j is

$$\hat{a}_j = \frac{2}{T}\frac{1}{n} \sum_{i=1}^{n} \cos\left(\frac{j\pi x_i}{T}\right). \qquad j = 1,\ldots,p$$

This estimate, $\hat{a}_j$, is unbiased for a . The Fourier coefficients minimise the Mean Integrated Squared Error (MISE), defined by

$$MISE = E\left[\int_0^T (\hat{f}(x) - f(x))^2 \right] dx,$$

$$\text{where} \quad f(x) = \frac{1}{T} + \sum_{j=1}^{p} a_j \cos\left(\frac{j\pi x}{T}\right).$$

Kronmal and Tarter (1968) suggested the following stopping rule based on the MISE; which could be expressed in terms of the variance of the estimate of a_j :

Add terms to the series until $\mathrm{Var}(\hat{a}_{p+1}) \geq (a_{p+1})^2$;

i.e., $$\frac{1}{T}\left(\frac{(T\, a_{2p+2} + 2)}{n+1}\right)^{1/2} \geq |a_{p+1}|.$$

It was observed by Burnham, et al., (1980), that typically $a_{2p+2} << a_{p+1}$, and they suggested computing their stopping rule by setting $\hat{a}_{2p+2} = 0$. This leads to computing $\hat{a}_p$ sequentially until the first estimate $\hat{a}_{p+1}$; such that,

$$\frac{1}{T}\left[\frac{2}{n+1}\right]^{1/2} \geq |\hat{a}_{p+1}|.$$

The stopping rule used by Burnham, et al., (1980) relies on the fact that $a_{2p+2} << a_{p+1}$. Although, this fact was noted by Kronmal and Tarter (1968), they did not make any assumptions about $\hat{a}_{2p+2}$ relative to $\hat{a}_{p+1}$.

The Burnham et al. (1980) stopping rule value $[2/(T^2(n+1))]^{1/2}$ does not depend on the data set. Even if some part of the data is changed, the value of the stopping rule is not affected. One can find no theoretical justification for assuming $\hat{a}_{2p+2}$ to be zero.

Simulation studies show that $\hat{a}_{2p+2}$ is not always much less than $\hat{a}_{p+1}$ and that in practice, very rarely are more than one or at most two cosine terms included in the series. Consequently, the Fourier series estimate has a tendency to retain the half-bell-shaped curve even when it is inappropriate (e.g. the piecewise linear detection curve with a = 20 as in Fig. 3).

We now consider estimation with grouped data. Though the cell probabilities, p , may in theory be negative, however it has been found in practice that they are always positive (Burnham, et al., 1980, p.134).

Integrating the Fourier series model yields

$$F(x) = \frac{x}{T} + \sum_{j=1}^{p} a_j \frac{T}{j\pi} \sin\left(\frac{j\pi x}{T}\right),$$

and since

$$p_i = F(t_{i+1}) - F(t_i)$$

$$= \frac{t_{i+1} - t_i}{T} + \sum_{j=1}^{p} a_j \frac{T}{j\pi} \sin\left(\frac{j\pi t_{i+1}}{T}\right) - \sin\left(\frac{j\pi t_i}{T}\right).$$

Then

$$\frac{\partial p_i}{\partial a_j} = \frac{T}{j\pi}\left[\sin\left(\frac{j\pi t_{i+1}}{T}\right) - \sin\left(\frac{j\pi t_i}{T}\right)\right] \qquad \begin{array}{l} j=1,\ldots,p. \\ i=1,\ldots,m \end{array}$$

Burnham, et al., (1980 p 70) argued that the stopping rule for ungrouped data is not applicable here, since the estimates of a_j may vary slighty with the number of terms to be included in f(x) They used a likelihood ratio test in deciding the appropriate model in conjunction with the usual goodness of fit test.

The likelihood ratio test of the following hypotheses;

$$H_0 : E(\hat{a}_{p+1}) = 0$$

vs.

$$H_1 : E(\hat{a}_{p+1}) \neq 0$$

is based on the difference between $l_p(\hat{a})$ and $l_{p+1}(\hat{a})$.

Where $l_p(a) = \text{constant} + \sum_{i=1}^{p} n_i \ln p_i(a)$.

Given that $-2(l_p(\hat{a}) - l_{p+1}(\hat{a})) \sim \chi_1^2$,

we reject H_0 if this difference exceeds 3.84 and conclude that $\hat{a}_{p+1}$ is a significant term in the series expansion of f(x).

3.4 A Generalised Parametric Case: The Exponential Power Series Estimator.

The exponential power series estimator was proposed by Pollock (1978). The detection function g(x) is given by

$g(x) = \exp(-(x/a)^b)$.

with the p.d.f f(x) given by

$$f(x) = \frac{\exp(-(x/a)^b)}{a\,(1 + b^{-1})} \qquad x>0,\ a>0,\ b>0. \qquad (3.4.1)$$

Therefore $\hat{f}(0) = (\hat{a}\,\Gamma(1 + b^{-1}))^{-1}$. (3.4.2)

Maximum likelihood estimates of a and b and hence f(0) are found by maximising the log likelihood l(a,b).

$$l(a,b) = \exp\left(-\sum_{i=1}^{n}(x_i/a)^b\right)/(a\,\Gamma(1 + b^{-1}))^n, \qquad (3.4.3)$$

with

$$\frac{\partial l}{\partial a} = \frac{\hat{b}}{\hat{a}}\sum_{i=1}^{n}\left(\frac{x}{\hat{a}}\right)^{\hat{b}} - \frac{n}{\hat{a}} = 0, \qquad (3.4.4)$$

and

$$\frac{dl}{db} = - \sum_{i=1}^{n} \left(\frac{x_i}{\hat{a}}\right)^{\hat{b}} \ln \frac{(x_i)}{\hat{a}} - n\,\psi \frac{(1 + \hat{b}^{-1})}{\hat{b}} = 0. \qquad (3.4.5)$$

The maximum likelihood estimator of a and b are the solutions to equations (3.4.4) and (3.4.5). The function ψ is the digamma function.

Assuming now that sufficient prior information is available on the form of the detection function, g(x), to make b known, then from (3.4.4),

$$\hat{a} = [\,(b \sum_{i=1}^{n} (x_i)/n)\,]^{1/b}. \qquad (3.4.6)$$

We now consider two special cases:

Case 1: b= 1

$$f(x) = \exp\,(-(x/a))/a\Gamma(1 + 1)$$

$$= \exp\,(-(x/a))/a$$

which is the negative exponential distribution. Maximum likelihood estimate of a, using (3.4.6), is

$$\hat{a} = \sum_{i=1}^{n} x_i/n = \bar{x}$$

and

$$\hat{f}(0) = (\hat{a}\,\Gamma(1 + b^{-1}))^{-1}$$

$$= (1/\bar{x})\Gamma(2) = 1/\bar{x}\ .$$

This estimator was first proposed by Leopold et al. (1951), based on the negative exponential distribution. It differs from Gates' (1968) estimator by a factor of

$(n-1)/n$ which is Gates' correction for bias.

The density estimate for ungrouped, untruncated data is $\hat{D} = n\hat{f}(0)/2L = (n-1)/2L\bar{x}$.

Case 2: b=2

$$f(x) = \exp(-(x/a)^2)/a\,\Gamma(1+1/2) \qquad (3.4.7)$$

$$= 2\exp(-(x/a)^2)/a\sqrt{\pi},$$

which is the half-normal distribution. The maximum likelihood estimate of a is given by

$$\hat{a} = \left[2\sum_{i=1}^{n} x_i^2/n\right]^{1/2},$$

and

$$\hat{f}(0) = \frac{1}{\Gamma(3/2)\left(2\sum_{i=1}^{n} x_i^2/n\right)^{1/2}} = \left(\frac{2}{\pi\sum_{i=1}^{n} x_i^2/n}\right)^{1/2}. \qquad (3.4.8)$$

This estimator was proposed by Hemingway (1971) and differs from Quinn's (unpublished thesis) by a factor of $(n-0.8)/n$, which is Quinn's correction for bias.

The density for ungrouped, untruncated data is;

$$\hat{D} = n\hat{f}(0)/2L = \frac{(n-0.8)}{2L}\left[\frac{2}{\pi\sum_{i=1}^{n} x_i^2/n}\right]^{1/2}$$

The exponential power series estimator is plagued with numerical problems. It is often necessary to restrict the value of the shape parameter to be greater than or equal to one; i.e., $b \geq 1$, since otherwise the estimate of the density function will have a spike at $x = 0$. As was noted in Section 1.1, estimation of both parameters is met with a considerable loss in efficiency.

Theoretically, this estimator can take a wide variety of plausible shapes for the detection functions, but estimation problems have limited its usage. It was one of the estimators that was originally planned to be used in the simulation study (Chapter 4) but was abandoned because of severe numerical convergence problems. For a more detailed treatment, see Pollock (1978) and Burnham, et al., (1980).

CHAPTER 4

The Estimation Problem

Seber (1973) and Burnham and Anderson (1976) supplied the general framework for estimation based on perpendicular distances. This framework is reviewed in Chapter 2. Suppose that we require an estimator of density, D, of objects based on the set $x_1, x_2, \ldots, x_n$ of observed perpendicular distances. Under the assumptions in Section (2.2.1) these distances are identically and independently distributed with density, $f(x) = g(x)/\mu_T$, over $0 \leq x < T$ where g(x) is the conditional probability of an object been seen, given that it is at a perpendicular distance x from the transect and $\mu_T = \int_0^T g(x)\,dx$. The intensity h(0) can be estimated by $n\hat{f}(0)$, and the estimate of Density, $\hat{D}$, is given by $\hat{D} = n\hat{f}(0)/2L = \hat{h}(0)/2L$.

The fundamental problem is thus reduced to estimating h(0), given a sample of independent observations. In order to compare estimators of h(x) in line transect sampling methods, and also to classify these estimators as acceptable in estimating wildlife densities, the following criteria must be taken into consideration:

1. Measure of Robustness, Resistance and Stability: Robustness in nonparametric estimation is hard to define. In line transect theory, Burnham, et al., (1980) introduced the concept of model Robust estimators, which refers to estimators that are sufficiently flexible in fitting different types of detection functions. They associated robustness of an estimator to small standard error and small bias of the estimator. If the ratio of the bias to the standard error of that estimator is small, say 0.5, they concluded that the estimator is robust. It is true that small size of estimation errors is an important criterion. However, one should look at the resistance and stability of these estimation errors.

   a. Resistance: A statistic is said to be resistant if any changes in a small part of the data, no matter how substantial, fails to produce any substantial changes in the statistic (Mosteller and Tukey, 1977).

   b. Stability: An estimator said to be stable if changing the underlying detection function does not substantially change the behaviour of the estimation errors.

   We shall also consider resistance and stability of the estimation errors of different estimators, as being a necessary measure for robustness as used in the above context.

2. Estimator Efficiency: We would like our estimators to be statistically efficient; i.e., in the class of all

acceptable estimators we want those with very small sampling mean squared error.

3. Shape Criterion: We shall further constrain our possible choice of estimators by insisting that they have a shoulder near x = 0. This restriction is consistent with the notion that in some small neighbourhood near the transect, g(x) will indeed be equal to 1.

Resistance and Stability of estimation errors point to reliable estimation techniques. One would normally be skeptical of a technique, if the estimation errors were to depend critically on a small portion of the data set, or were the errors tended to be large for certain, plausible detection curves. Furthermore, estimation errors are serious only if they form a substantial percentage of the actual quantity being estimated. For this reason, the average errors are given as percentages of h(0).

4.1 Simulation Study Design

In the simulation study that was done, perpendicular distances were generated from the following detection functions:

(i) PIECEWISE-LINEAR with shape parameter a

$$f(x) = \begin{cases} 0 & x \le 0 \\ b & 0 < x \le a \\ (T-x)\,b/h & a < x \le T \\ 0 & x > T. \end{cases}$$

Where $b = 2/(a + T)$ and $T > 0$.

(ii) ONE-TERM COSINE

$$f(x) = 1/T\,[\,1 + \cos(\pi x/T)\,] \qquad 0 \le x \le T$$

(iii) TWO-TERM COSINE

$$f(x) = 1/T\,[\,1 + \cos(\pi x/T) + 0.25\cos(2\pi x/T)\,] \qquad 0 \le x \le T$$

(iv) HALF-NORMAL with parameter σ

$$f(x) = 2/\sqrt{2\pi}\,\sigma \exp(-x^2/2\sigma^2) \qquad 0 \le x < \infty.$$

In the simulations, T was set equal to 100. The different detection functions were simulated from the piecewise linear with a =20 , 50 , and 80. and, T for the half-normal detection

function was set at σ, 2σ, and 3σ. The data were grouped into 10 classes of equal intervals. 250 replicates were generated from sample sizes 60 and 100 for each of the 8 detection functions. A list of the program for estimating f(0) using the shape-restricted estimator is given in Appendix A.
An average f(0) was calculated from the estimate of each sample.

The following statistical quantities were used in assessing efficiency, stability of estimation errors, robustness and robustness of efficiency. $\hat{h}_i(0)$ = estimate from replicate # i, and R is the number of replicates.

1. The Percent Relative Bias (PRB), given by

$$PRB = \frac{(\bar{h}(0) - h(0))}{h(0)} \times 100\%,$$

where $\bar{h}(0) = \frac{1}{R} \sum_{i=1}^{R} h_i(0)$,

and R is the number of replicate lines.

2. The estimate of the sampling variances of h(0), Var(h(0))

$$\hat{Var}(\hat{h}(0)) = \frac{1}{R-1} \sum_{i=1}^{R} (\hat{h}_i(0) - \bar{h}(0))^2.$$

3. The Mean Squared Error (MSE), of $\hat{h}(0)$, is given by

$$\hat{MSE}(\hat{h}(0)) = \hat{Var}(\hat{h}(0)) - (h(0) - \bar{h}(0))^2.$$

4. The estimate of percent relative Standard Error of $\hat{h}(0)$

given by

$$\frac{\widehat{SE}(\hat{h}(0))}{h(0)} \times 100\%.$$

5. The estimate of the percent relative Root Mean Square Error of $\hat{h}(0)$, given by

$$\frac{\widehat{RMSE}(\hat{h}(0))}{h(0)} \times 100\%.$$

Where $RMSE(.) = (MSE(.))^{1/2}$.

6. The Relative Efficiency (RE), of the Shape-Restricted estimator, given by

$$RE(\hat{h}_{SRE}(0)) = \frac{\widehat{MSE}(\hat{h}.(0))}{\widehat{MSE}(\hat{h}_{SRE}(0))}.$$

Where $\hat{h}(0)$ is some other estimator and $\hat{h}_{SRE}(0)$ is the shape-restricted estimate of $h(0)$.

Program TRANSECT was used in calculating estimates of $h(0)$ for the Fourier (cosine) series and half-normal estimators. These estimates of $h(0)$ were then used in calculating the above quantities.

4.2 Detailed Findings

On examining the bias (Table 2) of different estimators, one immediately notices the magnitude of the bias of the half-normal estimator. This manifests itself when we simulate from detection functions that have fairly broad shoulders. An example of this is the piecewise linear with a = 50 for which the bias is 20%. Both the Fourier series and shape-restricted estimators have moderate sizes for their corresponding biases; having a range of -0.2% to 7.3% for sample size 100. As was previously observed in Chapter 2, the half-normal and Fourier (cosine) series estimators had only negligible bias when the underlying detection functions are the same form as the estimator used.

It is interesting to note that, the Fourier (cosine) series estimator does poorly when estimating h(0) for the two-term cosine detection function, as seen from the estimate of its relative standard error and relative root mean squared error. This anomaly could be explained as follows: More terms of the cosine series are needed to fit the two-term cosine curve with a corresponding loss in precision of the estimates of the coefficients. This loss in precision produces the loss in efficiency.

The problem is even more pronounced with the half-normal estimator, since the two-term cosine curve is not very similar in shape to the half-normal (Figs. 3(a) and 3(b)). The two-term cosine curve has a significant probability of large observations. Both the Fourier series and half-normal estimators attempt to fit these outliers, thereby distorting the estimate of the remaining portion of the density curve.
The shape-restricted estimator does not suffer from this deficiency.

Improved performance of the shape-restricted estimator over the Fourier series and half-normal estimators is seen by observing Tables 1 and 3. The relative standard error and relative root mean squared error are stable with respect to changes in the detection function. This stability is not displayed by any of the other estimators considered in the simulation.

One can also conclude form Tables 1 and 3 that the shape-restricted estimator is resistant to changes in the tail observations: Since changes in this part of the data do not substantially change the statistic. For example, the half-normal curve with $T = 2\sigma$ and 3σ; there is hardly any change in the relative standard error (14.7% and 14.0%) and relative root mean squared error (14.7% and 14.2%) even though the tail has changed substantially (it had been eliminated when $T = 2\sigma$). This resistance is not displayed by the other estimators considered in the simulation.

On examining statistical efficiency (Table 4), one finds not one estimator that has higher efficiency, uniformly. The shape-restricted estimator, though not surpassing the Fourier series and half-normal estimators in efficiency, does show reasonably good efficiency. Apart from cases where the estimator naturally assumes the same form as the detection function used, the shape-restricted estimator has high efficiencies.

Dicussion and Recommendations.

The shape-restricted estimator provides a monotonically nonincreasing concave step function as a representation of the underlying detection function. It meets the shape criterion that in some small neighbourhood near the transect, g(x), will indeed be equal to 1.

With regards to resistance and stability of the estimation errors, it is encouraging to see a marked improvement over existing estimators. However, it would be statistically more efficient to improve upon the size of these errors. It is envisaged that a method for determining exactly where the detection curve changes concavity would help to improve the estimation errors.

The assumptions that the function be monotonically nonincreasing and concave up to at least the 80th percentile have additional advantages: Monotonic nonincreasing estimators are 'robust' to movement of the population away from the line in response to the observer (Burnham, et al., 1980). Since movement of the population is often away from the observer's line of travel, the number of sightings in the neighbourhood of $x = 0$ will decrease the apparent underlying detection curve near $x = 0$. Assuming the detection curve is continuous at $x = 0$, then $g(x) < 1$ and $g'(x) > 0$ for x close to zero. Any estimator that is not constrained to be monotone nonincreasing will estimate this bogus detection curve more closely. The monotonicity and concavity restrictions ought to reduce the magnitude of such estimation errors.

Though the shape restricted estimator is not universally superior to existing estimators, it does have qualities desirable of estimators in line transect work. It is truly nonparametric in that it does not perform markedly better on a limited class of detection functions.

One potential criticism is that its implementation requires the possibly subjective choice of the point of possible inflection. (In the simulation, it was set at 90'th percentile.) However, additional simulations have shown that the location of the point where the underlying detection function changes from concave to possibly convex, is not extremely crucial as long as it is between the 80th and 95th percentiles.

For the estimator to be fully operational, procedures for computing the variance of the estimator and for producing confidence intervals are required. Furthermore, as was noted earlier, techniques for improving upon the estimation errors and handling of ungrouped data are desirable. Work is still being done in these two areas. Research for reducing bias and producing confidence intervals could be directed towards the Bootstrap technique (Efron, 1978, 1983). Further research along these lines looks promising.

Table 1

Estimated percent relative root mean squared error of $\hat{h}(0)$, $\widehat{RMSE}[\hat{h}(0)]/h(0) x100\%$, for three different estimators using the eight detection curves in Fig. 2. For the half-normal curve, T = the truncation point, and σ = the standard deviation; For the piecewise linear form, a = the position of the kink.

	Estimators					
	n=60			n=100		
Detection function	Fourier Series	Half Normal	Shape restricted	Fourier Series	Half Normal	Shape restricted
Half-Normal $T = \sigma$	18.1	12.2	12.0	15.7	10.0	9.6
Half-Normal $T = 2\sigma$	13.8	11.7	14.9	10.2	8.4	11.8
Half-Normal $T = 3\sigma$	14.0	11.0	14.2	11.7	7.6	11.4
One-term Cosine	9.5	14.2	14.2	7.1	11.7	11.9
Two-term Cosine	16.8	13.8	15.5	13.7	11.5	11.2
Piecewise Linear a=20	14.8	15.7	14.3	12.2	13.0	10.7
Piecewise Linear a=50	26.8	23.9	13.6	22.0	22.1	11.1
Piecewise Linear a=80	25.0	19.7	13.5	21.0	16.4	10.3

Table 2

Estimated percent relative bias of $\hat{h}(0)$ for three different estimators using the eight detection curves in Fig. 2. For the half-normal curve, T = the truncation point, and σ = the standard deviation; For the piecewise linear form, a = the position of the kink.

	Estimators					
	n=60			n=100		
Detection function	Fourier Series	Half Normal	Shape restricted	Fourier Series	Half Normal	Shape restricted
Half Normal $T = \sigma$	-0.5	-0.1	-0.2	-0.2	-0.0	-0.5
Half Normal $T = 2\sigma$	0.3	0.7	-2.5	1.4	0.2	-2.1
Half Normal $T = 3\sigma$	-6.8	2.2	-2.2	5.5	1.3	-4.0
One-term Cosine	1.0	8.9	-1.9	1.0	8.1	-1.5
Two-term Cosine	-7.2	-7.7	-3.8	-5.1	-8.9	-3.1
Piecewise Linear, a=20	6.9	9.9	2.0	7.3	9.8	2.5
Piecewise Linear, a=50	5.3	19.9	6.6	2.3	20.0	6.0
Piecewise Linear, a=80	3.1	13.6	8.1	1.7	12.7	6.3

Table 3

Estimated percent relative standard error of $\hat{h}(0)$, $\hat{SE}[\hat{h}(0)]/h(0) \times 100\%$, for three different estimators using the eight detection curves in Fig. 2. For the half-normal curve, T = the truncation point, and σ = the standard deviation; For the piecewise linear form, a = the position of the kink.

	Estimators					
	n=60			n=100		
Detection function	Fourier Series	Half Normal	Shape restricted	Fourier Series	Half Normal	Shape restricted
Half-Normal $T = \sigma$	18.0	12.2	12.0	15.7	10.0	9.6
Half-Normal $T = 2\sigma$	13.8	11.7	14.7	10.1	8.4	11.6
Half-Normal $T = 3\sigma$	12.2	10.8	14.0	10.3	7.5	10.7
One-term Cosine	9.5	11.1	14.1	7.0	8.4	11.8
Two-term Cosine	15.2	11.5	15.0	12.7	7.4	10.8
Piecewise Linear a=20	13.1	12.1	14.2	9.8	8.5	10.5
Piecewise Linear a=50	26.3	13.2	11.9	21.8	9.4	9.3
Piecewise Linear a=80	24.8	14.2	10.8	20.9	10.4	8.1

Table 4

Estimated percent relative efficiencies for the Shape-restricted estimator, relative to the half-normal and Fourier series estimators.
$(\widehat{MSE}[\hat{h}_{\bullet}(0)]/\widehat{MSE}[h_{SRE}(0)]) \times 100\%$.
For the half-normal curve, T = the truncation point, and σ = the standard deviation; For the piecewise linear form, a = the position of the kink.

Estimators

Detection function	n=60 Fourier Series	n=60 Half Normal	n=100 Fourier Series	n=100 Half Normal
Half-Normal T = σ	2.3	1.0	2.7	1.0
Half-Normal T = 2σ	0.9	0.6	0.7	0.6
Half-Normal T = 3σ	1.0	0.6	1.1	0.4
One-term Cosine	0.4	1.0	0.4	1.0
Two-term Cosine	1.2	0.8	1.5	1.1
Piecewise Linear a=20	1.1	1.2	1.3	1.5
Piecewise Linear a=50	3.9	3.1	3.9	4.0
Piecewise Linear a=80	3.5	2.1	4.2	2.5

Proper Figures to Follow .

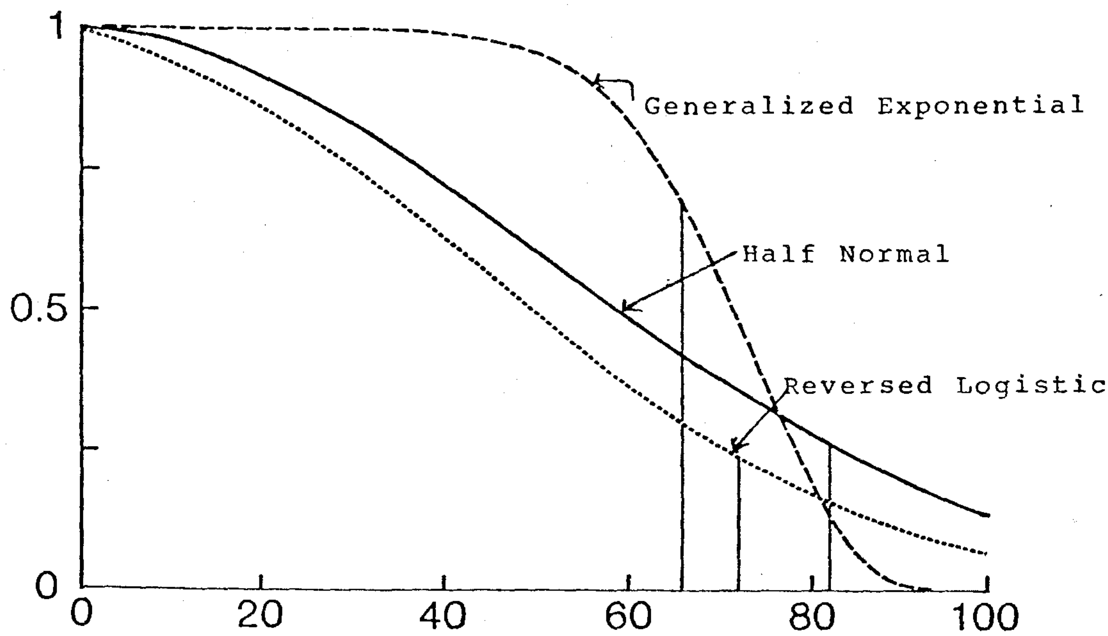

Fig. 2

Three parametric models of detection curves:

(a) Half Normal, $g(x) = \exp(-x^2/2\sigma^2)$ (———),

(b) Generalized Exponential, $g(x) = \exp(-(x/75)^8)$ (-----), and

(c) Reversed Logistic, $g(x) = (11\exp(-x/20))/(1+10\exp(-x/20))$ (••••••).

The dashed vertical lines give the 90th perventiles for each of the corresponding probability density curves.

Proper Figures to Follow

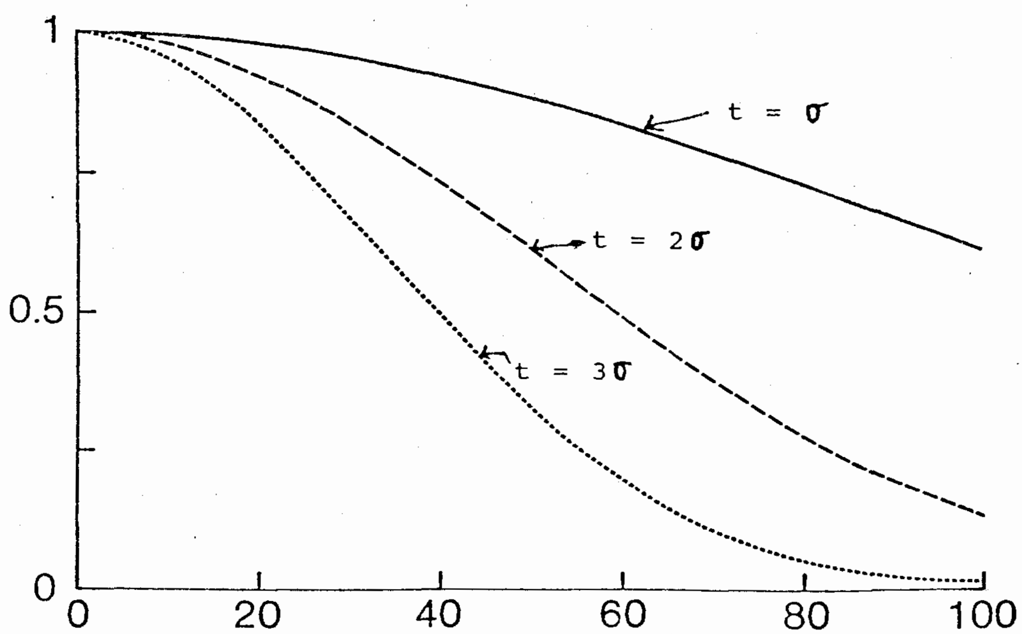

Fig. 3(a)

The six detection curves used in the simulation study. Half Normal curves, $g(x) = \exp(-x^2/2\sigma^2)$, truncated at σ(———), 2σ(-----), and 3σ(··········).

Proper Figures to Follow

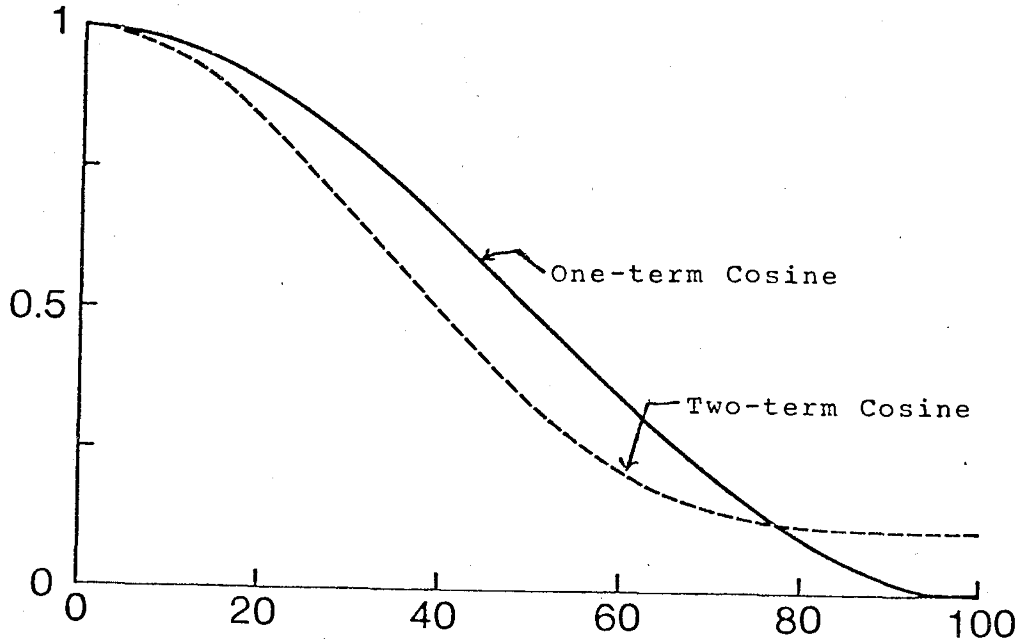

Fig. 3(b)

Cosine curves with g(x) = (1+cos(πx/100))/2 (———),

and g(x) = (1+cos(x/100)+cos(2πx/100))/2.25(-----).

Proper Figures to Follow

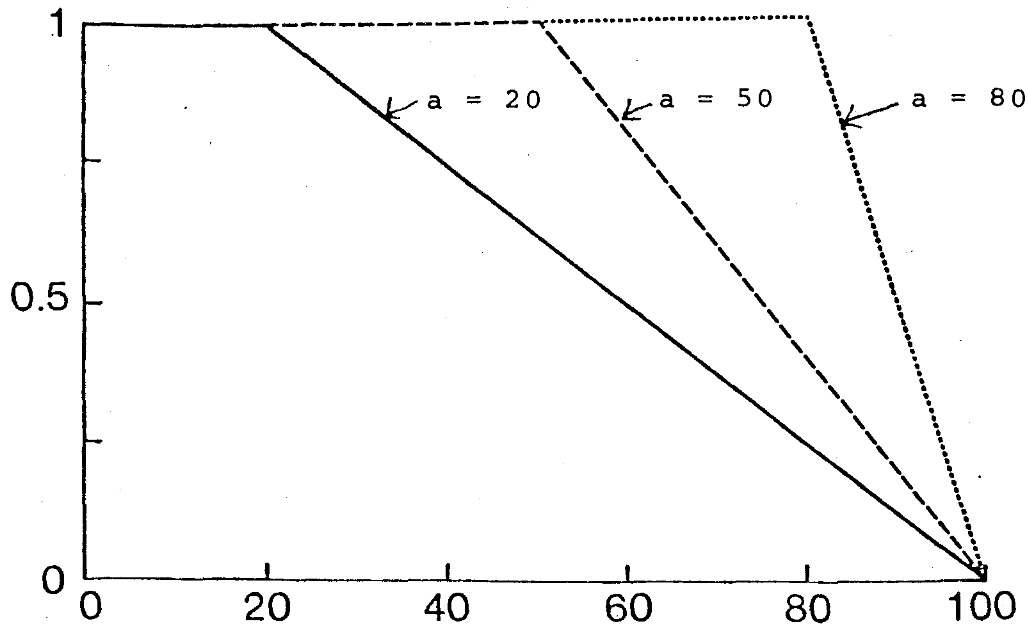

Fig. 3(c)

Piecewise linear forms with kinks at 20(———) 50(- - - - -), and 80(··········).

APPENDIX A

In order to run this program, and do the simulation study as was presented in the paper, the following program packages are needed:

1. Program TRANSECT and
2. Program LLSQ

Program TRANSECT could be obtained from the following address:

SHARE Program Library Agency
P.O. Box 12076
Research Triangle Park
NC 27709 USA

in ANSI FOTRAN 1V at an approximate cost of $40.00

This program was used in generating estimates of f(0) using the Fourier series and Half-normal estimators. One has to surpress most of the output from this program, since, only the raw estimates of f(0) are needed. These estimates are stored in a temporary file for later use.

Subroutine CLSA would generate both control and data file for running TRANSECT, if the CALL PRINT is included in the main program. (see main program list).

The LLSQ Program can be obtained from the following address :

IMSL
Customer Relation
Sixth Floor ,NBC. Building
7500 Bellaire Boulevard
Houston Texas 77036-5085 USA.

Subroutine CLSA calls Subroutine LDP which is in Program LLSQ. It writes output on two files namely, [8] and [9]. [8] contains the control and data file for TRANSECT and [9] contains the raw estimates of f(0) from the Shape-Restricted estimator. To get estimates of f(0) for the Fourier series and Half-Normal estimators, Program TRANSECT is loaded specifying [8] as the file data is to be read from. Output from TRANSECT are written on [6].

In the simulation study, grouped, perpendicular distances were generated from the piecewise linear, half-normal, one-term cosine and two-term cosine detection functions.

```
      REAL *8 DSEED
      REAL *4 CC(300),F0
      INTEGER MS(300)
      DSEED=123457.0D0
      DO 76 ILOOP=1,1
      TP=0.90
      NRR=1
      MSTAT=1
      A1=20
      NR=100
      TSTAR=100.
      CALL PRINT(TSTAR,NRR,NR)
      DO 1 MTRY=1,NRR
      N=10
      CALL CLSA(N,A1,NR,DSEED,MSTAT,MTRY,TSTAR,CC,MS,TP,F0)
 1    CONTINUE
      WRITE(9,19) (CC(I),I=1,NRR)
 19   FORMAT(1X,10E15.7)
      WRITE(8,57)
 57   FORMAT('END.')
 76   CONTINUE
      STOP
      END
C
      SUBROUTINE CLSA(N,A1,NR,DSEED,MSTAT,MTRY,TSTAR,CC,MS,
     -TP,F0)
C
C     SUBROUTINES CALLED- FRQCY ,PISWIS, NCOS, CAUS
C                         GGNPM, LDP
C     CALCULATES ESTIMATES OF F(O) FOR THE SHAPE-RESTRICTED
C     ESTIMATOR AND SETUP DATA FILE FOR USE IN PROGRAM
C     TRANSECT.
C     N    :NUMBER OF CLASSES USED IN GROUPING DATA (INPUT)
C     A1   :DENSITY FUNCTION PARAMETER DEPENDING ON MSTAT (INPUT)
C          MSTAT = 1 :PICECWISE LINEAR A1 IS THE KINK POSITION
C          MSTAT = 2 :ONE-TERM COSINE  A1 IS PARAMETER T=TSTAR
C          MSTAT = 3 :TWO-TERM COSINE  A1 IS PARAMETER T=STAR
C          MSTAT = 4 :HALF-NORMAL      A1 IS PARAMETER (SIGMA)
C     NR   :SAMPLE SIZE OF EACH REPLICATE (INPUT)
C     MSTAT :AS DEFINED ABOVE (INPUT)
C     NRR  :NUMBER OF REPLICATES LINES  (INPUT)
C     TSTAR :TRUNCATION POINT OF DATA (INPUT)
C     CC   :VECTOR CONTAINING THE SHAPE-RESTRICTED ESTIMATES
C          FOR EACH OF THE NRR FRPLICATES (OUTPUT)
C     MS   :VECTOR CONTAINING THE SAMPLE SIZES (OUTPUT)
C     F0   :OUTPUT VALUE OF TRUE F(O) OF DENSITY FUNCTION USED
C     TP   :PERCENTILE AT WHICH CONCAVITY IS ASSUMED UP TO .
C
      DIMENSION    G2(20,20), H2(20), X(20), Z(20), W(20)
      DIMENSION WLDP(500),CI(20)
      INTEGER INDEX(20)
      REAL *4 Y(300),R(300),CC(1),H1(20),D(20)
```

```
      REAL *8 DSEED
      INTEGER *4 FQ(20),MS(1)
      DATA IPOT,PI/0,3.141592/
C
      IF(MSTAT .EQ. 2) GOTO 140
      IF(MSTAT .EQ. 3) GOTO 150
      IF(MSTAT .EQ. 4) GOTO 160
C
C     THIS PORTION IS GENERATES  PIECEWISE DEVIATES AND GROUPS
C     THEN INTO N CLASSES
C
      H=TSTAR - A1
      CI(1)=0.0
      UMP=(A1+H)/FLOAT(N)
      DO 15 I=1,N
 15   CI(I+1)=CI(I) + UMP
      CALL PISWIS(NR,A1,H,R,Y,DSEED)
      CALL FRQCY(NR,IPOT,CI,N,Y,FQ)
      NP1=N + 1
      WRITE(8,87)MTRY, (CI(I),I=2,NP1)
      WRITE(8,89) (FQ(I),I=1,N)
 87   FORMAT('SIMULATED PIECEWISE',1X,I3,/20F6.1)
C
      B=2/(H + (2*A1))
      F0=(TSTAR - A1)*(A1 - TSTAR +(2*H))
      F0=B*F0/(2*H)
      F0=A1*B + F0
      F0=B/F0
C
C     THIS IS THE END OF PIECEWISE GENERATION
C
      GOTO 170
C
C     THIS PORTION GENERATES  GROUPED ONE-TERM COSINE DEVIATES
C
 140  A2=PI/A1
      CALL NCOS(NR,N,A2,R,FQ,D,DSEED)
      CI(1)=0.
      DO 19 I=1,N
  19  CI(I+1)=D(I)
      UMP=D(1)
      NP1=N + 1
      WRITE(8,107)MTRY,(CI(I),I=2,NP1)
      WRITE(8,89)(FQ(I),I=1,N)
 107  FORMAT('SIMULATED ONE-TERM COSINE',1X,I3,/20F7.2)
  89  FORMAT(20I4)
C
      F0=A2*(1+COS(A2*0.0))/3.141592
      F0=F0/CDF(TSTAR,A2)
C
C     END OF THE ONE-TERM COSINE GENERATION
C
      GOTO 170
```

```
C
C     THIS PORTION GENERATES GROUPED TWO-TERM COSINE
C
 150  AP=PI/A1
      CALL CAUS(NR,N,AP,R,FQ,D,DSEED)
      CI(1)=0.
      DO 59 I=1,N
  59  CI(I+1)=D(I)
      UMP=D(1)
      NP1=N + 1
      WRITE(8,117)MTRY,(CI(I),I=2,NP1)
      WRITE(8,88)(FQ(I),I=1,N)
 117  FORMAT('SIMULATED TWO-TERM COSINE ',1X,I3,/20F7.2)
  88  FORMAT(20I4)
C
      F0=(AP + AP*COS(AP*0.) + 0.25*AP*COS(2*AP*0.))/3.141592
      F0=F0/CCF(TSTAR,AP)
C
C     END OF THE TWO-TERM COSINE GENERATION
C
      GOTO 170
C
C     THIS  PORTION GENERATES HALF-NORMAL DEVIATES AND
C     GROUPS THEN INTO N CLASSES
C
 160  SIGMA=A1
      UMP=TSTAR/(FLOAT(N)*SIGMA)
      CI(1)=0.0
      DO 115 I=1,N
 115  CI(I+1)=CI(I) + UMP*SIGMA
C     GENERATE N(0,1) AND TRANSFORM TO N(0,SIGMA)
C
      NTC=IFIX((TSTAR/SIGMA)+0.5)
      NRA=2*NR
      CALL GGNPM(DSEED,NRA,R)
      J=0
      DO 16 I=1,NRA
      R(I)=ABS(SIGMA*R(I))
      IF(R(I) .GT. TSTAR) GOTO 16
      J=J + 1
      Y(J)=R(I)
      IF( J .EQ. NR ) GOTO 26
  16  CONTINUE
  26  CALL FRQCY(NR,IPOT,CI,N,Y,FQ)
      NP1=N + 1
      WRITE(8,111)MTRY,(CI(I),I=2,NP1)
      WRITE(8,112)(FQ(I),I=1,N)
 111  FORMAT('SIMULATED HALF-NOML',1X,I2,/20F6.1)
 112  FORMAT(20I4)
      UMP=UMP*SIGMA
C
      F0=SQRT(2*3.141529)*SIGMA
      XE=TSTAR/SIGMA
```

```
      F=0.5*ERFC(-0.7071068*XE)
      F0=1/((F-0.5)*F0)
C
C     END OF HALF-NORMAL GENERATION
C
C     PROCEED TO FIND CLASS AT WHICH 90TH PERCENTILE IS.
C
 170  NSUM=0
      DO 53 I=1,N
      NSUM=FQ(I)+NSUM
      NTRUN=I
      TMP=FLOAT(NSUM)/FLOAT(NR)
      IF(TMP .GE. TP ) GOTO 135
  53  CONTINUE
 135  IF(NTRUN .EQ. N) GOTO 99
      NTP1=NTRUN + 1
      DO 55 I=1,N
      ICI=IFIX((CI(NTRUN+1)*10) +0.5)
      ISTAR=IFIX(TSTAR*10)
      NT=I
      IF(ICI .EQ. ISTAR) GOTO 56
  55  CONTINUE
  56  N=NT
  98  DO 24 I=NTP1,N
  24  NSUM=NSUM + FQ(I)
  99  MS(MTRY)=NSUM
C
C     SET PARAMETERS FOR SUBROUTINE LDP
C
      NN=N -1
      MDGH=20
      K=NTRUN
C
C     G2 DEFINES THE CONSTRAINT MATRIX IN Z-COORDINATES SYSTEM
C
      DO 1 I=1,N
      DO 1 J=1,N
   1  G2(I,J)=0.0
      IF(K .EQ. 1) GOTO 13
      G2(1,1)=1.0
      DO 2 I=2,K
      G2(I,I)=2.0
      G2(I-1,I)=-1.0
   2  G2(I,I-1)=-1.0
      G2(K,K-1)=0.0
  13  DO 3 I=K,NN
      G2(I,I)=1.0
   3  G2(I,I+1)=-1.0
      G2(N,N)=1.0
C
C     H2 DEFINES CONSTRAINT RT SIDE FOR  Z-COORDINATES SYSTEM
C
      NTM1=K -1
```

```
      H2(1)=FLOAT(FQ(2) - FQ(1))
      IF(NTM1 .EQ. 0) GOTO 23
      DO 20 I=2,NTM1
   20 H2(I)=FLOAT(FQ(I-1) + FQ(I+1) - 2*FQ(I))
   23 DO 22 I=K,NN
   22 H2(I)=FLOAT(FQ(I+1) - FQ(I))
      H2(N)=FLOAT(-FQ(N))
C
      CALL LDP (G2,MDGH,N,N,H2,Z,ZNORM,WLDP,INDEX,MODE)
C
C      TRANSFORM BACK FROM Z-COORDINATES TO X-COORDINATES.
C
          DO 60 J=1,N
   60     X(J)=Z(J)+FLOAT(FQ(J))
C
C      COMPUTE THE RESIDUALS.
C
      RES=ZNORM
   67 FORMAT(1X,I8,F10.5)
C
C     COMPUTE RELATIVE ESTIMATES OF X
C
      X(1)=X(1)/(NSUM*UMP)
      CC(MTRY)=X(1)
      RETURN
      END
       SUBROUTINE PISWIS(NR,A1,H,R,PW,DSEED)
C
C      GENERATES PIECEWISE LINEAR RANDOM DEVIATES
C      SUBROUTINE CALLED- GGUW (FROM IMSL)
C      NR :SAMPLE SIZE (INPUT)
C      A1 :POSITION OF THE KINK (INPUT)
C      H  :LENGTH WHICH PIECEWISE IS TRIANGLAR (INPUT)
C      DSEED :SEED FOR GENERATING UNIFORMS (INPUT)
C      PW :OUTPUT VECTOR CONTAINING PIECEWISE DEVIATES
C      R :OUTPUT VECTOR CONTAINING UNIFORM DEVIATES
C
       REAL *4 R(1),PW(1)
       REAL *8 DSEED
       IOPT=1
       B=1.0/(A1 + H/2.0)
       AREA=A1*B
       CALL GGUW(DSEED,NR,IOPT,R)
       DO 1 I=1,NR
       IF (R(I) .LE. AREA) GOTO 2
C
C       CALCULATE PW(I)=A1+H(1-SQRT((1-R(I)/(1-AREA)))
C
       PWH=SQRT((1 - R(I))/(1 - AREA))
       PWH=1 - PWH
       PW(I)=A1 + H*PWH
       GOTO 1
    2  PW(I)=A1*(R(I)/AREA)
```

```
    1   CONTINUE
        RETURN
        END
        SUBROUTINE NCOS(NR,N,A2,R,FQ,D,DSEED)
C
C       GENERATES GROUPED COSINE RANDOM DEVIATES
C       SUBROUTINE CALLED- CCUW (FROM IMSL)
C       FUNCTION USED- CCF
C       NR :SAMPLE SIZE (INPUT)
C       N  :NUMBER OF CLASSES USED IN GROUPING DATA (INPUT)
C       A2 :COSINE DENSITY PARAMETER (INPUT)
C       R  :OUTPUT VECTOR CONTAINIG UNIFORM DEVIATES
C       FQ :OUTPUT VECTOR CONTAINING FREQUENCY COUNTS
C       D  :OUTPUT VECTOR CONTAINING CLASS BOUNDARIES
C       DSEED :SEED FOR GENERATING UNIFORM DEVIATES (INPUT)
C
        REAL *4 R(1),D(1)
        REAL *8 DSEED
        INTEGER *4 FQ(1)
        IOPT=1
        AN=A2*FLOAT(N)
        P=3.141592
        H=P/AN
        DO 2 J=1,N
        XX=H*FLOAT(J)
        D(J) = CDF(XX,A2)
    2   FQ(J)=0
        CALL GGUW(DSEED,NR,IOPT,R)
        NN=N-1
        DO 3 I=1,NR
        RN=R(I)
        DO 4 J=1,NN
        IF(RN .GE. D(J)) GOTO 4
        FQ(J)=FQ(J) + 1
        GOTO 3
    4   CONTINUE
        FQ(N)=FQ(N) + 1
    3   CONTINUE
        DO 5 J=1,N
    5   D(J)=H*FLOAT(J)
        RETURN
        END
        FUNCTION CDF(XX,A2)
        CDF=(A2*XX +SIN(A2*XX))/3.141592
        RETURN
        END
       SUBROUTINE FRQCY(NR,IPOT,CI,N,Y,FQ)
C
C      GROUPS DATA INTO N CLASSES
C      N  :NUMBER IF CLASSES (INPUT)
C      NR :SAMPLE SIZE (INPUT)
C      CI :VECTOR CONTAINING CLASS BOUNDARIES (INPUT)
C      Y  :RANDON DEVIATES TO BE GROUPED (INPUT)
```

```
C     FQ :OUTPUT VECTOR CONTAINING FREQUENCY COUNTS
C
      REAL *4 CI(1),Y(1)
      INTEGER *4 FQ(1)
      DO 1 I= 1,N
    1 FQ(I)=0
      NN2=N-1
      DO 3 I=1,NR
      RN=Y(I)
      DO 4 J=1,NN2
      IF (RN .GE. CI(J+1) ) GOTO 4
      FQ(J)=FQ(J) + 1
      GOTO 3
    4 CONTINUE
      FQ(N)=FQ(N) + 1
    3 CONTINUE
      RETURN
      END
      SUBROUTINE CAUS(NR,N,AP,R,FQ,D,DSEED)
C
C      GENERATES GROUPED COSINE RANDOM DEVIATES
C      SUBROUTINE CALLED- CCUW (FROM IMSL)
C      FUNCTION USED- CCF
C      NR :SAMPLE SIZE (INPUT)
C      N  :NUMBER OF CLASSES USED IN GROUPING DATA (INPUT)
C      AP :COSINE DENSITY PARAMETER (INPUT)
C      R  :OUTPUT VECTOR CONTAINIG UNIFORM DEVIATES
C      FQ :OUTPUT VECTOR CONTAINING FREQUENCY COUNTS
C      D  :OUTPUT VECTOR CONTAINING CLASS BOUNDARIES
C      DSEED :SEED FOR GENERATING UNIFORM DEVIATES (INPUT)
C
      REAL *4 R(1),D(1)
      REAL *8 DSEED
      INTEGER *4 FQ(1)
      IOPT=1
      AN=AP*FLOAT(N)
      P=3.141592
      H=P/AN
      DO 2 J=1,N
      XX=H*FLOAT(J)
      D(J) = CCF(XX,AP)
    2 FQ(J)=0
      CALL GGUW(DSEED,NR,IOPT,R)
      NN=N-1
      DO 3 I=1,NR
      RN=R(I)
      DO 4 J=1,NN
      IF(RN .GE. D(J)) GOTO 4
      FQ(J)=FQ(J) + 1
      GOTO 3
    4 CONTINUE
      FQ(N)=FQ(N) + 1
    3 CONTINUE
```

```
      DO 5 J=1,N
  5   D(J)=H*FLOAT(J)
      RETURN
      END
      FUNCTION CCF(XX,AP)
      CCF=(AP*XX +SIN(AP*XX) +0.125*SIN(2*AP*XX))/3.141592
      RETURN
      END
      SUBROUTINE PRINT(WSTAR,NRR,NR)
C
C     SETUP CONTROL FILE FOR DATA USED IN RUNNING
C     PROGRAM TRANSECT.
C     OUTPUT APPEARS ON OUTPUT FILE 8
C     WSTAR :TRUNCATION POINT OF DATA (INPUT)
C     NRR   :NUMBER OF REPLICATES (INPUT)
C     NR    :SAMPLE SIZE (INPUT)
C
      REAL *4 PL(300)
      INTEGER *4 MS(300)
      DO 1 I=1,NRR
      MS(I)=NR
  1   PL(I)=40.0
      WRITE(8,10)
 10   FORMAT('*ANALYSIS OF EXAMPLE FROM SIMULATION*'/'*DISTANCE
     $MEASURED IN INCHES*',/,'*LENGTH MEASURED IN INCHES*',
     $/,'*AREA IS IN SQUARED INCHES*')
      WRITE(8,20)(PL(I),I=1,NRR)
 20   FORMAT(13(F5.1,1H,),1X,1H$)
      WRITE(8,30)(MS(I),I=1,NRR)
 30   FORMAT(13(I6),1X,1H$)
      WRITE(8,40)WSTAR
 40   FORMAT('*PEST,GRPD,NPOL*',1X,F5.1,/'2.',/,'*FSER*',
     $/,'*HNOR*')
      RETURN
      END
```

BIBLIOGRAPHY

Barlow, R.E.,D.J. Bartholomew, J.M. Bremner, and H.D. Brunk. 1972. Statistical Inference under Order Restrictions. John Wiley & Sons, New York, N.Y. 388 pp.

Burnham, K.P. and D.R. Anderson. 1976. Mathematical Models for Nonparametric Inferences from Transect Data. Biometrics 32(2):325-336.

-------,------, and J.L. Laake. 1980. Estimation of Density from the Line Transect sampling of Biological Populations. Wildlife Monograph no. 72.

Crain, D.R. 1964. Estimation of Distribution using Orthogonal Expansion. Ann. Statist. 2(3):454-463.

Eberhardt, L.L. 1978a. Transect Methods for Population Studies. J. Wildl. Manage. 42(1): 1-31.

Efron, B. 1979. Bootstrap Methods. Another Look at the Jackknife. Annals of Statistics 7, 1-26.

-----. 1981. Nonparametric Standard Errors and Confidence Intervals. The Canadian Journal of Statistics 9(2):p139-172.

Gates, C.E. 1979. Simulation Study of Estimators for the Line Transect Sampling Method. Biometrics 25(2):317-328.

-------,W.H. Marshall, and D.P. Olson. 1968. Line Transect Method of Estimating Grouse Population Densities. Boimetrics 24(1): 135-145.

Golub, G.H and M.A. Saunders. 1970. Linear Least Squares and Quadratic Programming. Integer and Nonlinear programming. North-Holland Publ. Co., Amsterdam. J .Abadie (ed). p229 - 256.

Kronmal, J.L. and M. Tarter. 1968. The Estimation of Probability Densities and Cumulatives by the Fourier Series Method. J. Amer. Statist. Ass. 63(3):925-952.

Laake, J.L., K.P. Burnham, K.P. and D.R. Anderson. 1997. User's Manual for Program TRANSECT.Utah, St Univ. Press, Logan, Utah. 26 pp.

Lawson, C.L. and R.J. Hanson. 1974. Solving Least Squares Problems. Prentice-Hall,Inc., Englewood Cliffs, N.J. 400pp.

Mosteller, F. and J.W. Tukey. 1977. Data Analysis and Regression. Addison-Wesley Publ. Co., inc., London. 588pp

Pollack, K.H. 1978. A Family of Density Estimators for Line Transect. Biometrics 34(3): 475-478.

Rao, C.R, 1973. Linear Statistical Inference and its Applications. John Wiley & Sons,inc., New York,N.Y. 625 pp.

Rockafellar, R.T. 1970. Convex Analysis . Princeton, N.J. Princeton Univ. Press. 452pp.

Seber, G.A.F. 1973. The Estimation of Animal Abundance Hafner Publ. Co.,inc., New York, N.W. 506 pp.

Sen, A.R., J. Tourigny, and G.E. Smith. 1974. On the Line Transect Sampling Method. Biometrics 30(3):329-340.

Tapia, R.A., and J.R. Thompson. 1978. Nonparametric Probability Density Estimation. The John Hopkins Press, Baltimore, Md. 176 pp.

Tarter, M.E., and R.A. Kronmal. 1976. An Introduction to the Implementation and Theory of Nonparametric Density Estimation. Amer. Statist . 30(8):105-112.

Wegman, E.J. 1972a. Nonparametric Probability Density Estimation: Summary of Available Methods. Technometrics 14(3):533-546.

------. 1972b. Nonparametric Probability Density estimation: A Comparison of Density Estimation Methods. J. Comp. Simul. 1 :225-245.