# Toward Estimating the Variance in Acoustic Surveys Based on Sampling Design 

Magnar Aksland<br>Department of Biology, University of Bergen<br>E-mail: magnar.aksland@bio.uib.no<br>Received March 8, 2011; revised March 19, 2011, accepted March 22, 2011


#### Abstract

This paper develops a sampling method to estimate the integral of a function of the area with a strategy to cover the area with parallel lines of observation. This sampling strategy is special in that lines very close to each other are selected much more seldom than under a uniformly random design for the positions of the parallel lines. It is also special in that the positions of some of the lines are deterministic. Two different variance estimators are derived and investigated by sampling different man made signal functions. They show different properties in that the estimator that estimate the biggest variance gives an error interval that, in some situations, may be more than ten times the error interval computed from the other estimator. It is also obvious that the second estimator underestimates the variance. The author has not succeeded to derive an expression for the expectation of this estimator. This work is motivated towards finding the variance of acoustic abundance estimates.


Keywords: Acoustic Abundance Estimation, Line Surveys, Sampling Design, Sampling in the Area, Variance Estimation

## 1. Introduction

This paper is about to find the variance in acoustic surveys to estimate the abundance of fish stocks. Here, usually a ship with a downward looking echo sounder that pings sound pulses into the sea and receive sound reflections from fish, covers an area where the fish stock under estimation is supposed to be located.

Estimating the variance of acoustic abundance estimates of marine resources is an old problem, but up to now not much has been done based on sampling design.

The author of this paper has worked with acoustic estimation of fish populations. Then it is natural that the methods will be related to the observations of the acoustic signal generated by fish echoes, although such methods will always have the potential to be used in other applications.

The echo signal received from a modern echo sounder as a function of the area position of the echo sounder is an example of a function of which the area integral is necessary to estimate. This is explained in the next section.

The method presented here represents the use of unequal probability sampling design, and can be used in
many sampling problems where a resource to be estimated are distributed over an area.

Sampling design in the area or volume is a neglected field in the sampling literature, but several applications are demanding appropriated methods. This includes estimating resources that are distributed over the area or volume. Such resources are sometimes observed continuous along lines, and sometimes at distinct positions. Reference to such problems is found in Stevens and Olsen (2004).

An important assumption for the present method is that a quantity associated with each point in an area can be measured continuously with a movable device. Then, the quantity may be observed on a system of lines covering the area of interest.

Here, the unequal probability design has another purpose than is usual. Instead of seeking inclusion probabilities that are positively correlated with the quantity to be observed, as for the Horvitz-Thompson estimator in sampling design, we use a sampling design that has a reduced probability to select object very close to each other relative to a uniform probability design. The present design thus produces samples that are better spatial balanced than that produced by a uniformly random de-
sign.
The present paper is an attempt to give a contribution to the field of line sampling.

It is a hope that more specialists in sampling theory will work in this field.

## 2. Review of Echo Integration in Terms of Energy

This is a short review of a theory of echo-integration to estimate the abundance of fish.

The traditional method of Echo Integration for estimating the abundance of sound scatters (MacLennan, 1990) is based on the integration of the echo-signal. This quantity represents sound energy, but the conversion of integrator values to abundance, or density, of scatters is based on "back scattering cross-section", which represents sound intensity, or power.

In 1982 R. E. Craig proposed to rewrite sonar theory that is based on power, to a theory based on energy (Craig, 1983). Before this proposal, however, a method of echo integration that bases the conversion to scatter abundance on the energy of single target echoes were in use (Aksland, 1983), although the theory of the method was published in 1986 (Aksland, 1986), and published when the split beam and dual beam system were in regular use (Aksland, 2005, Aksland, 2006). This alternative method of echo integration does not use the concept "back scattering cross-section" or "Target Strength". The method is in fact based on much fewer concepts than the traditional echo integration method. The basis of the method is given below.

Let $I_{2}\left(z ; t_{1}(z), t_{2}(z)\right)$ be the depth integrated 20 $\log r$ TVG (Time Varied Gain) echo signal (echo intensity) between times $t_{1}$ and $t_{2}$ after sound transmission. The parameter $r$ is the distance from the transducer. In general, the times $t_{i}=2 r_{i} / c, i=1,2$ depend on the transducer position $z=(x, y)$. Here $c$ is the sound speed. The downward looking transducer is free to be moved within a horizontal plane with Cartesian coordinates $(x, y)$, here called the transducer plan. Let $Q$ be a given region of the transducer plan.

$$
\begin{equation*}
E A=\iint_{Q} I_{2}(z) \mathrm{d} A \tag{1}
\end{equation*}
$$

where $\mathrm{d} A$ is the area differential, is defined as the Echo Abundance within $Q$ subject to the depth channel. $\left[t_{1}(z), t_{1}(z)\right]$ In applications, the dept channel may also consist of several disjoint time intervals. This may be necessary when integrating echoes from a special type of scatters while excluding others.

Let $I_{p 2}(z)$ be the integrated echo pulse intensity from a single scatter at $20 \log r$ TVG, as it is when the
echo is resolved.

$$
\begin{equation*}
E V C=\iint_{Q} I_{p 2} \mathrm{~d} A \tag{2}
\end{equation*}
$$

where the integral is over an area where the integrand is significant is called the Echo Value Constant of the corresponding scatter. As for the Echo Abundance, this integral is also over the different transducer positions in the transducer plane.

By transforming the Echo Value Constant (2) to polar coordinates $(\theta, \varphi)$ in the object reference system (a polar reference system with origin in the scatter and $\theta=0$ pointing vertically upward), and using that $\mathrm{d} A=r^{2} \tan \theta \mathrm{~d} \theta \mathrm{~d} \varphi$, it may be expressed as

$$
\begin{equation*}
E V C=\int_{0}^{\theta_{\varepsilon}} \int_{0}^{2 \pi} \varsigma(\theta, \varphi) \mathrm{d} \varphi b_{t r}(\theta) \tan \theta \mathrm{d} \theta \tag{3}
\end{equation*}
$$

where $b_{t r}(\theta)$ is the transmit-receive beam function (here circular symmetric), $\theta_{c}$ is an angle where the integrand outside $\theta_{c}$ is negligible, and $\varsigma(\theta, \varphi)=I_{2 p}((\theta, \varphi)) / b_{t r}(\theta)$ is called the back-scattering energy of the scatter. The back-scattering energy is the same as the beam compensated integrated $40 \log r$ TVG echo pulse received from the scatter when the transducer is in direction $(\theta, \varphi)$ in the object reference system. A necessary condition for this is that the ratio between the $40 \log r$ TVG and $20 \log r$ TVG functions satisfies exactly the relation

$$
\frac{40 \log r \mathrm{TVG}}{20 \log r \mathrm{TVG}}=r^{2}=\frac{c^{2}}{4} t^{2}
$$

Note that the Echo Value Constant of a scatter depends on both the level of the echo signal and the pulse length used, as well as on the beam function. For a fixed acoustic system it is a quantity that varies with the time since movement and other kinds of behaviour of the scatters will affect their Echo Value Constants.

If $\theta_{c}$ in (3) is replaced with a variable angle $\alpha$, the corresponding function is called the Echo Value, that is,

$$
\begin{equation*}
E V(\alpha)=\int_{0}^{\alpha} \int_{0}^{2 \pi} \varsigma(\theta, \varphi) \mathrm{d} \varphi b_{t r}(\theta) \tan \theta \mathrm{d} \theta \tag{4}
\end{equation*}
$$

If $\alpha$ increases from zero, it turns out that the Echo Value increases at first, but flattens out when $\alpha$ approaches the outer part of the main lobe. When $\alpha$ is equal to or bigger than the angle $\theta_{20}$ where the trans-mit-receive beam has fallen 20 dB , the Echo Value is approximately constant due to the strong beam damping. It is the value of this flat region that is called the Echo Value Constant, and this property justifies its name.

It can be proved that both the Echo Abundance and the Echo Value Constants of single scatters are independent of the distance between the transducer and scatters when
$I_{2}$ does not include noise. These properties are due to the $20 \underline{\log r} r$ TVG.

Let $\overline{E V C}_{E A}$ be the average, or mean Echo Value Constant over the scatters that contribute to the Echo Abundance (EA). It can be proved (Aksland, 1986) that

$$
\begin{equation*}
E A=\sum_{i=1}^{N} E V C_{i}=N \overline{E V C}_{E A} \tag{5}
\end{equation*}
$$

where $N$ is the number of scatters and $E V C_{i}$ is the Echo Value Constant of the $i$-th scatter that contributes to the Echo Abundance. Relation (5) is true if the depth integral of overlapping echoes is the same as the sum of the individual integrated resolved echo pulses over all scatters. A necessary condition for this is that the phases of the echoes from each single scatter are uncorrelated, and that shadow effects are negligible (Zhao and Ona, 2003).
By writing (5) as

$$
N=\frac{E A}{\overline{E V C_{E A}}}
$$

we see that the mean Echo Value Constant corresponding to an Echo Abundance is a constant that converts the Echo Abundance to the number of scatters.
The mean Echo Value Constant may be estimated from integrated single target echoes at $40 \log r$ TVG and corresponding detection angles provided that the echoes are representative for all echoes that contributes to the Echo Abundance. The detection angle of an echo is the angle between the beam axis and the direction between the transducer and the scatter, and can be detected within the main lobe with split beam and dual beam echo sounder systems. See Aksland, (2005), (2006) and (2010) for details.

## 3. Estimating the Echo Abundance

The Echo Abundance is the area integral of the echo intensity as a function of the transducer position over an area. The alternative echo integrator-method defines and uses the Echo Abundance, so the sampling methods to be developed here refer to this method. A review of the theory behind the alternative echo integrator-method is given in Section 2.
Short reviews of other methods to estimate the Echo Abundance are given below.

### 3.1. Review of Estimating the Echo Abundance

A non statistical way of estimating the Echo Abundance within a given sea area $Q$ is to observe the echo intensities $I_{2}\left(z ; t_{1}(z), t_{2}(z)\right)$, (see (1)), on a system of lines covering $Q$. Next fit some parametric class of surfaces over $Q$ to the observed values of $I_{2}$. Then, the volume
under the fitted surface within $Q$ will be an estimator of the Echo Abundance. An example of this way to estimate the Echo Abundance is given in Aksland (1983).

This method does not estimate the precision of the estimated Echo Abundance.

The method can, nevertheless be recommended if the precision of the abundance estimate is not very important, or if it is believed that the precision of the Echo Abundance is small compared to other factors of uncertainty affecting the abundance estimate.

To estimate the Echo Abundance with precision is a challenge. Since acoustic data are observed along lines, application of probabilistic survey design methods are not appropriate except for special cases where each strata are covered with parallel lines of observations. This case is equivalent with sampling points on a line, where the integrated echo intensities are projected onto a line orthogonal to the parallel lines. The parallel lines are then projected to points of observation onto the orthogonal line.

This restriction of probability sampling methods may be reduced through a generalization of the foundation of probability sampling theory. Then randomization of other types of lines covering an area may be used. In acoustic abundance estimation, coverings with zigzag lines are common.

Estimation of the Echo Abundance belongs to the more general problem of estimating an integral

$$
\begin{equation*}
\int_{Z} f(z) \mathrm{d} z \tag{6}
\end{equation*}
$$

where $Z$ is a region of the Euclidean space, and $f(z)$ is observed on a subset of $Z$ with dimension commonly less than the dimension of $Z$. In acoustic abundance estimation $Z$ is of dimension 2 , while the set observed has dimension 1 (a system of lines).

Foote and Stefánson (1993) have described and discussed different methods for estimating fish abundance over an area from line-transects. They recommend kriging methods, but have only one reference to probability sampling (Cochran).

There are mainly two classes of methods that are available for estimating (6) together with an estimate of the precision of the estimate. This is methods within random field models (geostatisics, kriging) and probability survey sampling methods, respectively.

### 3.2. Random Field Models

Here $f(z)$ is considered as a realization of a random process (field) $F(z), z \in Z$.

Methods based on parametric models with stationary increments are elaborated and known as "Geostatistics", (Matheron 1963). Estimators are based on predictors for
the values attained by the process outside the sample (kriging), (Krige 1951). The sample may be selected subjectively. The predictors have certain minimum variance properties for given model parameters (usually trend and auto covariance functions), and these have to be estimated in applications.

Objectives against the application of random field models are that it is difficult to judge whether the models are well related to the spatial distribution they are supposed to describe. Also the bias caused by spatial distributions that cannot adequately be taken to be a normal realization of the used random field model is difficult to judge.

However, the fact that these methods allow subjective selected samples has lead many scientists to choose these methods. See the book by Rivoirard et al. (2000) and references therein.

### 3.3. Probability Survey Methods

The methods within probability survey sampling are mainly concerned with the estimation of the mean $\bar{y}=\frac{1}{N} \sum_{j=1}^{N} y_{j}$ or the total $N \bar{y}$ of an unknown finite vector $Y=\left\{y_{1}, y_{2}, \cdots, y_{N}\right\}$ The foundation given below follows Cassel, Särndal and Wretman (1977).

The estimation is based on a sample from the set of labels $J=\{1,2, \cdots, N\}$, where $J$ must be known.

A sampling design, $P(s)$, is a probability measure on the set $\Upsilon$ of all subsets of $J$. Selection of a sample is done in accordance with the selected sampling design. The data associated with a sample $s$ is denoted by $D=\left\{s, Y_{s}\right\}$, where $Y_{s}$ is the restriction of $Y$ to $s$. In general, an estimator is a real function $t(D), s \in \Upsilon$.

The combination $\{t(D), P(s)\}$ is called a strategy.
The art of probability sampling is to make use of all known information about $Y$ to construct a strategy that will give the most precise and accurate estimates for a given sampling budget.

An estimator that depends only on the values of $Y$ on the selected sample and is otherwise independent on the labels of the sample is denoted label-independent. Most of the traditional and well-known estimators in probability sampling are label-independent.

A sampling design is non-informative if $P(s)$ is independent of values of $Y$. Otherwise the design is informative.

Many methods within probability sampling may easily be generalized to sampling from infinite populations. Sampling from Euclidean space, and from an area in particular, are good examples, where the infinite set of points are the sampling objects. If the population from which samples are selected are a subset of a Euclidean
space, each object has a position, and there is a unique distance between each pair of points. Unfortunately, such quantities associated with sample objects are very seldom and poorly treated in the probability sampling theory. In particular, when sampling a subset of the Euclidean space successively to estimate an integral, estimators that are label-dependent and sampling designs that are informative are likely to be better than the traditional estimators in probability sampling.

Fortunately Thompson (1990) and Thompson and Seber (1996) has worked with adaptive sampling methods that has informative sampling designs that give better precision than conventional estimates of populations having aggregation tendencies in their area distribution. These estimation methods have been used in biological sampling as well as in acoustic surveys (Harbitz, Ona and Pennington, 2009), (Conners and Schwager, 2002), (McQuinn et al., 2005).

When sampling values of some function defined on a given subset of a Euclidean space, the observed values give information about the spatial structure of the function values. Information about this structure should be used when selecting the rest of the sample, as well as in constructing a good estimator. When sampling in the area, estimators that are the volume under some fitted surface to the observed data are likely to have good properties relative to other, more traditional, estimators in the sampling literature. The adapted sampling design methods mentioned above have not been developed far enough to be appropriate for all situations when estimating an integral over Euclidean space.

The case of spatial label sets impose some general demands on the sampling design.
a) The design should reduce the selection of "spatial unbalanced" samples
b) For successively selected samples, the design should be informative

This indicates the need for developing new methods within the field of probability sampling. Bertil Matern (1969) has pointed out this and other problems related to the application of probability sampling methods, but not much has happened with the foundation of sampling design since then.

Case a) in (7) above means that samples that do not cover the space properly, or are too patchy, are "spatial unbalanced". A way to avoid this is to stratify the sample population into $n$ strata, where $2 n$ is the sample size, and chose a probability sample of size 2 in each stratum. When estimating the abundance of animals with aggregation tendencies, which are rather usual, there is another advantage of stratifying the sampling population as fine as possible. The population variance of small strata are likely to be smaller than the population variance of big-
ger strata. Then, under the assumption of equal total sample sizes, the use of many small strata is likely to give more precise estimates than an estimator based on fewer and bigger strata. See Des Raj (1968), ch. 4.

Case b) in (7) holds because when starting to collect a successive sample, information on the spatial structure of the population variable is gained. To increase the precision, this knowledge should be used when selecting the sampling design for the position of the next observation.

## 4. A Strategy for Parallel Line Survey

Assume that we know approximately the area where a pelagic resource is located. This resource can be observed with a downward looking transducer that is moved within the actual area. Cover first the area along parallel lines that runs completely through the locations occupied by the resource so that sailing between the different lines are outside this region. The first covering has the role of a pilot survey that is selected subjectively.

Next return and go two parallel legs with random locations between each neighbour lines of the pilot survey. The randomized legs should also go completely through the resource. Figure 1 illustrates this survey.

In this survey, the pilot survey, which is deterministic, defines strata boundaries. Two random lines are selected within each stratum, stochastically independent of the selection within other stratums. To reduce the likelihood of selecting lines very close to each other, or to close to one or both lines of the pilot survey, the positions for the lines are selected with unequal probabilities.

In a strata let the positions of the two selected random
lines be $x_{l}$ and $1-x_{h}$, respectively, where $x_{l}+x_{h}<1$, and the values observed are $y_{l}$ and $y_{h}$. The strata width is set equal to 1 since this will not represent any loss of generality. The function $y(x)$ is then observed at the values $0, x_{l}, 1-x_{h}$ and 1 for $x$. In the following, $y(x)$ will be called the signal function.

We choose the following sampling strategy for the stratum:

The estimator is given by
$T=\frac{1}{2}\left[x_{l} y_{0}+x_{h} y_{1}+\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right]$
This is the same as the area under the step function shown in Figure 2, and defines a label-dependent estimator.

The sampling design is given by the following probability density:

$$
\begin{align*}
& f\left(x_{l}, x_{h}\right)=120 x_{l} x_{h}\left(1-x_{l}-x_{h}\right)  \tag{9}\\
& \text { for } x_{l}>0, x_{h}>0, x_{l}+x_{h}<1
\end{align*}
$$

A plot of this density in terms of $x$ for $x_{l}$ and $y$ for $1-x_{h}$ is shown in Figure 3.

It is seen that this sampling design reduces the likelyhood considerable that the lines will be selected close together, or close to some strata boundary relative to a uniform sampling design, where the probability density is constant for all $x$ and $y$.

This sampling strategy fulfils to a certain degree demand (7a) for a strategy for estimating an integral over a subset of the Euclidean space. The estimator (8) represents the area under a simple interpolation to the observed


Figure 1. The pilot survey: thin lines. The probability survey: thick lines. The isolines illustrate the distribution of the echo intensity that is not known during observation. The whole survey is projected onto the line perpendicular to the parallel lines.


Figure 2. Graphical illustration of the estimator (area under step function). This is the same as the area under the straight-line curve through the observed values.


Figure 3. The joint probability density for the positions of the left $(x)$ and right $(y)$ random lines in a stratum.
values while the sampling design reduces the likelihood for selecting lines close together. However, the present sampling design is not informative, as it is independent of any observed values from the survey. Unfortunately, this strategy does not make use of the observed variation along the parallel lines.

A generalization of this method where the sampling design depends on the observations on the strata boundaries, may appear in the future. The following are some statistical properties of the present strategy, which are
derived in the Appendix I:
The random variables $x_{l}$ and $x_{h}$ have identical marginal distributions with probability density $20 x(1-x)^{3}, \quad 0<x<1$, and the conditional density of one of the variables, $z$, given the other, $x$, is $6 \frac{z(1-x-z)}{(1-x)^{3}}, 0 \leq z \leq 1-x$, . These distributions are used when generating random values for $x_{l}$ and $x_{h}$. Further

$$
\begin{align*}
& \quad \mathrm{E}\left(x_{l}\right)=\frac{1}{3}, \mathrm{E}\left(x_{l}^{2}\right)=\frac{1}{7}, \mathrm{E}\left(x_{l} x_{h}\right)=\frac{2}{21},  \tag{10}\\
& \\
& \mathrm{E}\left(x_{l}^{2} x_{h}\right)=\frac{1}{28}, \mathrm{E}\left(x_{h} \mid x_{l}\right)=\frac{1-x_{l}}{2} \\
& \mathrm{E}(T)=\frac{1}{6}\left(y_{0}+y_{1}\right)+5 \int_{0}^{1} x(1-x)\{1-x(1-x)\} y(x) d x \\
& \text { or }  \tag{11}\\
& \mathrm{E}(T)=\frac{1}{6}\left(y_{0}+y_{1}\right)+5 \int_{0}^{1}\left(x-2 x^{2}+2 x^{3}-x^{4}\right) y(x) d x
\end{align*}
$$

In general $T$ is biased, but $T$ is unbiased if $y(x)$ is linear. Moreover, when the graph of $y(x)$ is a straight line over the strata, $T$ is equal to $\int_{0}^{1} y(x) \mathrm{d} x$ for every selected positions of $x_{l}$ and $x_{h}$. This follows from Figure 2, and implies further that $\operatorname{Var} T=0$ for the class of linear functions $y(x)=a+b x$ for all real constants $a$ and $b$.

It is possible to specify a bigger class of functions $\{y(x)\}$ for which $T$ is unbiased. However, $T$ is in general biased, and although most estimators given in the sampling literature are unbiased, the lack of unbiasedness for $T$ is caused by the fact that $T$ was selected without this in mind. It is seen from Figure 2 that $T$ has positive bias if $y(x)$ is convex and negative bias if $y(x)$ is concave on $[0,1]$, because then $y(x)$ will be below (above) the straight-line function in Figure 2. However, $T$ is supposed to be summed over several strata, and it is the total bias over all strata that are important here. This is likely to be small because the individual strata biases will be both positive and negative. An in depth theoretical analysis of the bias of the sum of $T$ over the strata seems to be difficult. Some considerations are given in the Discussion.

To be able to find a formula for $\operatorname{Var}(T)$, the expected value of $T^{2}$ is given.

$$
\begin{align*}
\mathrm{E}\left(T^{2}\right)= & \frac{1}{84}\left\{3\left(y_{0}^{2}+y_{1}^{2}\right)+4 y_{0} y_{1}\right\}+\frac{1}{2} \int_{0}^{1}\left(3 x-5 x^{2}+10 x^{3}-20 x^{4}+18 x^{5}-6 x^{6}\right) y^{2}(x) \mathrm{d} x \\
& +y_{0} \int_{0}^{1}\left(5 x^{2}-10 x^{3}+5 x^{4}+2 x^{5}-2 x^{6}\right) y(x) \mathrm{d} x+y_{1} \int_{0}^{1}\left(2 x-5 x^{2}+10 x^{3}-15 x^{4}+10 x^{5}-2 x^{6}\right) y(x) \mathrm{d} x  \tag{12}\\
& +60 \int_{0}^{1}(1-z) z y(z) \int_{0}^{z}\left\{\left(x-x^{2}\right) z-x^{2}+x^{3}\right\} y(x) \mathrm{d} x \mathrm{~d} z
\end{align*}
$$

Formula (12) may also be expressed with the polynomials on factorial form.

$$
\begin{aligned}
\mathrm{E}\left(T^{2}\right)= & \frac{1}{84}\left\{3\left(y_{0}^{2}+y_{1}^{2}\right)+4 y_{0} y_{1}\right\}+\frac{1}{2} \int_{0}^{1} x(1-x)\left(3-2 x+8 x^{2}-12 x^{3}+6 x^{4}\right) y^{2}(x) \mathrm{d} x \\
& +y_{0} \int_{0}^{1} x^{2}(1-x)\left(5(1-x)+2 x^{3}\right) y(x) \mathrm{d} x+y_{1} \int_{0}^{1} x(1-x)^{2}\left\{5 x+2(1-x)^{3}\right\} y(x) \mathrm{d} x \\
& +60 \iint_{\substack{x>0, z>0 \\
x+z<1}}(1-x)(1-z) x z(1-x-z) y(x) y(1-z) \mathrm{d} x \mathrm{~d} z
\end{aligned}
$$

With (11) and (12), expressions for the variance and mean square error of $T$ may be derived. The variance expression given here is obtained by squaring the last
expression of (11) and subtracting it from the first expression of (12).

$$
\begin{align*}
\operatorname{Var}(T) & =\frac{1}{126}\left(y_{0}^{2}+y_{1}^{2}-y_{0} y_{1}\right)+\frac{y_{0}}{3} \int_{0}^{1}\left(-5 x+25 x^{2}-40 x^{3}+20 x^{4}+6 x^{5}-6 x^{6}\right) y(x) \mathrm{d} x \\
& +\frac{y_{1}}{3} \int_{0}^{1}\left(x-5 x^{2}+20 x^{3}-40 x^{4}+30 x^{5}-6 x^{6}\right) y(x) \mathrm{d} x+\frac{1}{2} \int_{0}^{1}\left(3 x-5 x^{2}+10 x^{3}-20 x^{4}+18 x^{5}-6 x^{6}\right) y^{2}(x) \mathrm{d} x  \tag{13}\\
& +60 \int_{0}^{1}(1-z) z y(z) \int_{0}^{z}\left\{\left(x-x^{2}\right) z-x^{2}+x^{3}\right\} y(x) \mathrm{d} x \mathrm{~d} z-25\left\{\int_{0}^{1}\left(x-2 x^{2}+2 x^{3}-x^{4}\right) y(x) \mathrm{d} x\right\}^{2}
\end{align*}
$$

An estimator for $\operatorname{Var}(T)$ may be obtained from (13) by computing it for the signal function $y(x)$ given by the piecewise straight-line function through the observed values (see Figure 2). However, it is not easy to evaluate the expectation of this estimator. Since the integrals in (13) are tractable for piecewise linear functions for

$$
\begin{align*}
\widehat{\operatorname{Var}}(T)= & \frac{1}{126}\left(y_{0}^{2}+y_{1}^{2}-13 y_{0} y_{1}\right)+\frac{1}{4}\left\{\left(1-x_{h}\right)^{2} y^{2}\left(x_{l}\right)+\left(1-x_{l}\right)^{2} y^{2}\left(1-x_{h}\right)\right\} \\
& +\left\{\frac{y_{0}}{2}\left(x_{l}-\frac{1}{3}\right)+\frac{y_{1}}{2}\left(x_{h}-\frac{1}{3}\right)\right\}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}  \tag{14}\\
& +\frac{1}{2}\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)-\frac{1}{4}\left[\hat{\mathrm{E}}^{2}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}\right]
\end{align*}
$$

where $\hat{\mathrm{E}}^{2}$ is an estimator of $\mathrm{E}^{2}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}$. Formula (14) is not in a form that can be directly applied because an estimator for the last line has to be inserted. The formula is derived in Appendix I.

Another reason why the proposed variance estimators cannot be directly used is that it is derived for strata of with 1 . This was done to simplify mathematical derivations, but it may easily be generalized to the case with different strata widths.

Assume that we have $n$ strata with widths $x_{i}, i=1, \mathrm{~L}, n$. Let (8) be denoted by $T_{i}$ in strata number $i$. Then, $T_{i}$ is an estimator of $\int_{0}^{1} y(x) \mathrm{d} x$, where $x=x^{\prime} / x_{i}$, and $x^{\prime}$ runs from 0 to $x_{i}$ over strata number $i$. Since the Echo Abundance in strata number $i$ is equal to $\int_{0}^{x_{i}} y\left(x^{\prime}\right) \mathrm{d} x$, we have that $x_{i} T_{i}$ is an estimator of the Echo Abundance in strata number $i$, while $\sum_{i=1}^{n} x_{i} T_{i}$ is the corresponding estimator for the total Echo Abundance.

Since $\operatorname{Var}\left(x_{i} T_{i}\right)=x_{i}^{2} \operatorname{Var}\left(T_{i}\right)$, (7) or an alternative multiplied with $x_{i}^{2}$, is an estimator for $\operatorname{Var}\left(x_{i} T_{i}\right)$, and these may be summed over all strata to obtain an estimator for the variance of $\sum_{i=1}^{n} x_{i} T_{i}$.

Note that in strata number $i, x_{l}=x_{l}^{\prime} / x_{i}$, and $x_{h}=x_{h}^{\prime} / x_{i}$, where $x^{\prime}$ and $x_{i}-x_{h}^{\prime}$ are the absolute positions of the two randomized legs in the strata.

Results from sampling a set of artificial functions are given in Appendix II.

## 5. Discussion

Estimating the Echo Abundance may also be useful in the classic echo integrating method. However, the present methods are independent of what quantity that is estimated.

### 5.1. Estimating the Echo Abundance

Many acoustic surveys are carried out in a way where the
covering of the resource is done subjectively using common sense decisions. Among such decisions there can be found quite wise ideas to generalize the theory of sampling design. Although most coverings are done in some systematic ways using the available knowledge of how the actual resource distributes over the area, the cruise leaders will always also want to use observation data from the finished part of the survey to decide upon the remaining covering. When covering a resource subjectively, legs are never selected very close to each other, but when using sampling design, this may happen. Many cruise leaders will consider close legs as a waist of money. But within sampling design from a finite population, every object must have a positive probability to be selected in the sample. Moreover, to be able to estimate the variance of the used estimator, every pair of objects should have a positive probability to be selected. This is a strict requirement for the Horvitz-Thompson estimator based on unequal probabilities given in RAJ (1968). This is the reason why close legs should have a positive probability to be selected when using sampling design to select parallel lines. There are certain problems with unequal probability designs. See Tillé (1996), but these do not affect the present methods.

There is another way to avoid the possible selection of close legs. This is when the echo intensity as a function of the area position is modelled stochastically. In Cassel, Särndal and Wretman (1977) it is shown how super population models reduce the importance of a sampling design and may even allow estimation of variances from subjectively selected samples. However, when variances are estimated from subjective samples based on some stochastic model for the area echo intensity, all probabilities comes from this model, and the variance estimates cannot be expected to be more reliable than the underlying stochastic model.

An advantage with probability sampling methods to estimate populations with difficult area distribution, is that the estimators do not depend on some underlying super population model, but on man made probabilities expressed by the chosen sampling design. As long as samples are selected in accordance with the sampling design, the estimates are objective. Subjective knowledge is used with probability sampling methods in various ways, for instance when forming strata, and in general when deciding upon the sampling design. Whenever it is felt that an estimate based on probability sampling design is biased, or are subject to other kinds of errors, this feeling comes from knowledge that is not used when deciding about the sampling strategy. All supplementary knowledge should be used when deciding the sampling strategy.

### 5.2. Miscellaneous

This paper shows that it is possible to reduce the prob-
ability to select legs close to each other, but the probability should not be zero. An estimator that is biased within strata in general may bother someone, but since it is the bias of the estimator summed over all strata that is of importance, the bias within strata is not serious. The bias for every signal function $y(x)$ for which (8) can be integrated can be computed analytically. If $y(x)=x^{2}$, which is convex, the bias is $+10.7 \%$. When $y(x)=1-x^{2}$, the bias has the same absolute value, but is negative. If the likelihood to observe signal functions with biases that cancel each other is similar, then the bias of the estimated sum over many strata will tend to zero when the number of strata increases. However, this is not likely to occur exactly. An example that may occur is a function that is zero everywhere except for a very high value over a short distance caused by a fish school. If the school happens to stay close to a strata boundary, the estimator within the actual strata may have a considerable negative bias. The opposite situation to this, where the function has a high value everywhere except for a short distance where it is zero, is very unlikely to occur. However, to find out more about the bias of the estimator (8) summed over strata, the best study would be to compute by programming the true value as well as the expectation for a lot of different functions $y(x)$ defined over many strata. The study in Appendix II throws some light on this problem, but more signal functions and strata numbers are needed. However, the figures in Appendix II show that the bias usually shrinks when the number of strata is increased from 5 to 10 . Otherwise, the figures in Appendix II indicate that the difference between (13) and (14) is less when the sampling density is small and when sampling signal functions with sharp and big variations.

Some may have noted that the estimator $T_{i}$ and $T_{i+1}$ corresponding to strata no $i$ and $i+1$ both contain the observed value on the common strata line. But this does not violate the stochastic independence of $T_{i}$ and $T_{i+1}$ because observed values from the pilot survey (on strata boundaries) are not stochastic.

### 5.3. The Estimators of Variance

The two estimators developed in this paper may not be the best. The estimator based on (13) is obviously too small. This is indicated in the Appendix II. Estimator (14) is almost unbiased in theory, but Appendix II indicates that it is too big.

The interval, Estimate $\pm S D(T)$, based on (14), contains the true value in almost all cases in Appendix II, as well as in other cases. However, the square root of an unbiased estimator of the variance of an estimator has an expectation that is less than the expectation of the Stan-
dard Deviation of the estimator. Therefore, if $\operatorname{Var}(T)$ is unbiased, Estimate $\pm \sqrt{\operatorname{Var}(T)}$, should not be grater than the Estimate $\pm S D(T)$ in average.

If the estimator based on (13) has much less expectation than $\operatorname{Var}(T)$, it is dangerous to increase it with a constant factor, because the right factor may vary with the shape of the signal function, and as well on the sampling density. One possible reason why this estimator estimates a too small variance, is that in all strata where the observed numbers $y_{0}, y_{l}, y_{h}$ and $y_{1}$ all happen to have values close to some straight line, the estimated variance in these strata are close to zero.
A difference between the used man made signal functions in the Appendix II and real signal functions based on the Echo Abundance generated by a fish population is that the latter is not static. This means that observed values are hardly equal if they are observed twice at different times. As the man made signal functions used are continuous with continuous derivative, a function like this will converge to linear within strata when the strata widths goes to zero. Therefore, the variance estimator based on (13) may produce estimates that tend too fast to zero when the number of strata increases. It is not sure that the estimator has this property if it used on real non-static signal functions.

Another way to find an estimator for $\operatorname{Var}(T)$ is to express it as a multiple polynomial

$$
\operatorname{Var}(T)=\sum c_{i j m n} y_{0}^{i} y_{1}^{j} y_{l}^{m} y_{h}^{n}
$$

where $c_{i j m n}$ are constants, or polynomials in $x_{l}$ and $x_{h}$. The problem is to find coefficients $c_{i j m n}$ so that $\mathrm{E}\{\operatorname{Var}(T)\}=\operatorname{Var}(T)$. This is a generalization of the method given in Des Raj (1968) to find an unbiased variance estimator when sampling with unequal probabilities. The present author has not yet succeeded to find an estimator by this method.

### 5.4. Sampling with Unequal Probabilities

In the literature about sampling design, interesting results about sampling with unequal probabilities have been derived (see Des Raj (1968)). It may be tempting to try the Horvitz-Thompson estimator and variance estimator here, but in the present case this estimator is not appropriate. This is because the estimator does not make use of the observed values on the strata boundaries.

Also, there are different reasons to apply unequal probabilities in the present case and in the HorvitzThompson's case. The reason for choosing a not uniform random sampling strategy when sampling an area is to spread the sample better than is obtained with a uniform random sampling strategy. Living resources' have social
behaviour, and then it is believed that the echo intensities at positions close to each other are seldom very different. This is the reason to spread the sample in the present case.

The general reason for selecting unequal probabilities in the Horvitz-Thompson's case is that probabilities can be chosen that are positive correlated with the variable to be observed.

### 5.5. The Future

The author of this paper does not look upon the present results as a finite solution. It is more a start of using sampling design to estimate the variance. Hopefully, the results of this paper build on some principles that are new in sampling design. The combination of a subjective and a randomized covering where the estimator is label dependent and depends on both deterministic and randomized observations is not common. It is a hope that generalized methods building on this principle can be developed. A real challenge is to combine the common subjective coverings with additional randomized observations for estimating the variance.
Use of adaptive sampling strategies in sampling design is difficult. But if sampling designs can be based on the non-random observations from a deterministic part of the survey, the same variance reductions may be obtained with less statistical difficulties.

## 6. Conclusions

A special sampling strategy is proposed for covering an area with parallel lines of observation. The strategy consists of a deterministic covering followed by a randomized covering between the deterministic covering.

A label dependent estimator is proposed that depends on both deterministic and randomized observations. The sampling design is with unequal probabilities with the purpose to produce better spatial balanced samples.

The theoretical Expectation and Variance of this estimator are derived, and two estimators of the variance have been found. Further properties of the estimators of variance are studied by sampling man made functions. This study showed that one of the estimators is likely to underestimate the variance.
The two variance estimators may not be the best for the proposed sampling strategy. However, if the proposed strategy is generalized and based on similar principles, the results in this paper is important as a special case.

Estimating the variance of an estimated integral based on a line sample requires a generalization of the foundation of sampling design. This is a big job.

## 7. References

[1] Stevens D. L. and Olsen A. R. Spatially Balanced Sampling of Natural Resources. Journal of the Americal Statistical Association 99, No. 465, 2004, 262-278.
[2] D. N. MacLennan. Acoustical measurement of fish abundance. J. Acoust. Soc. Am.; vol. 87(1): 1990, 1-15.
[3] Craig R. E. Re-definition of sonar theory in terms of energy. In: Selected papers of the ICES/FAO Symposium on fisheries acoustics. FAO Fish. Rep; (300), 1983, 331 p.
[4] Aksland M. Acoustic abundance estimation of the spawning component of the local herring stock in Lindaaspollene, western Norway. FiskDir. Skr.Ser. HavUnders. 1983, 297-334.
[5] Aksland M. Estimating numbers of pelagic fish by echo integration. J. Cons. int. Explor. Mer. 43, 1986, 7-25.
[6] Aksland M. An alternative echo-integrating method. ICES Journal of Marine Science 62, 2005, 226-235.
[7] Aksland M. Applying an alternative method of echo-integration. ICES Journal of Marine Science 63, 2006, 1438-1452.
[8] Zhao X. and Ona E. Estimation and compensation models for the shadowing effect in dense fish aggregations. ICES Journal of Marine Science 60, 2003, 155-163.
[9] Aksland M. Analysing Estimators in the Alternative Echo Integrator Method. The Open Ocean Engineering Journal 3, 2010, 116-128.
[10] Foote K. G., and Stefánson G. Definition of the problem of estimating fish abundance over an area from acoustic line-transect measurements of density. ICES Journal of Marine Science 50, 1993, 369-381.
[11] Matheron G. Principles of geostatistics. Economic Geology 58, 1963, 1246-1266.
[12] Krige D. G. A statistical approach to some mine valuations and allied problems at the Witwatersrand. Master's thesis of the University of Witwatersrand. 1951.
[13] Rivoirard J. Simmonds K. G., Foote P., Fernandes and N. Bez. Geostatistics for estimating fish stock abundance Blackwell Science, Osney Mead, Oxford, 2000, 206 pp.
[14] Cassel C. M., Särndal C. E. and Wretman J. H. Foundations of Inference in Survey Sampling. John Wiley \& Sons, 1977.
[15] Thompson, S. K. Sampling, 2nd edn. John Wiley and Sons, New York, 2002.
[16] Thompson, S. K., and Seber, G. A. F. Adaptive Sampling. John Wiley and Sons, New York, 1996.
[17] Harbitz A., Ona E. and Pennington M. The use of an adaptive acoustic-survey design to estimate the abundance of highly skewed fish populations. ICES Journal of Marine Science 66, 2009, 1-6.
[18] Conners M. E. and Schwager S. J. The use of adaptive cluster sampling for hydroacoustic surveys. ICES Journal of Marine Science 59, 2002, 1314-1325.
[19] McQuinn I. H., Simard Y., Stroud W. F., Beaulieu J. L., and Walsh S. J. An adaptive, integrated "acoustic-trawl" survey design for Atlantic cod (Gadus morhua) with estimation of the acoustic and trawl dead zones. ICES Journal of Marine Science 62, 2005, 93-106.
[20] Matern B. Sample Survey Problems. Bulletin of the International Statistical Institute Vol. XLII, Book 1, 1969, 143-154.
[21] Des Raj. SAMPLING THEORY. McGraw-Hill Publishing Company LTD, 1968.
[22] Tillé Y. Some remarks on unequal probability designs without replacement. Annales D'Économie et de Statistique No 44, 1996, 177-189.

## Appendix I

## Derivation of Results

Let us write the probability distribution (9) as

$$
\begin{equation*}
f(x, z)=120 x z(1-x-z), x>0, z>0, x+z<1 \tag{15}
\end{equation*}
$$

Where $x$ and $z$ are variables for $x_{l}$ and $x_{h}$, respectively. Note that this distribution is symmetric in the variables $x$ and $z$. This means that the marginal distributions of $x_{l}$ and $x_{h}$ are identical, and

$$
\begin{aligned}
& \mathrm{E}\left(x_{l}^{n} x_{h}^{m}\right)=\mathrm{E}\left(x_{l}^{m} x_{h}^{n}\right), m \geq 0, n \geq 0 \\
& \text { and } \mathrm{E}\left(x_{l} \mid x_{h}\right)=\mathrm{E}\left(x_{h} \mid x_{l}\right)
\end{aligned}
$$

The marginal density of $x_{l}$ is obtained by integrating (15) with respect to $z$.

$$
\begin{aligned}
\int_{0}^{1-x} f(x, z) \mathrm{d} z & =120 x \int_{0}^{1-x} z(1-x-z) \mathrm{d} z \\
& =120 x \int_{0}^{1-x}\left\{(1-x) z-z^{2}\right\} \mathrm{d} z \\
& =20 x(1-x)^{3}
\end{aligned}
$$

Hence, the marginal density is given by

$$
\begin{equation*}
g(x)=20 x(1-x)^{3}, 0<x<1 \tag{16}
\end{equation*}
$$

Now

$$
\begin{align*}
\mathrm{E}\left(x_{l}\right) & =20 \int_{0}^{1} x^{2}(1-x)^{3} \mathrm{~d} x \\
& =20 \int_{0}^{1}\left(x^{2}-3 x^{3}+3 x^{4}-x^{5}\right) \mathrm{d} x  \tag{17}\\
& =\frac{1}{3}=\mathrm{E} x_{h}
\end{align*}
$$

and

$$
\begin{align*}
\mathrm{E}\left(x_{l}^{n} x_{h}^{m}\right) & =120 \iint_{0<x<1-z, 0<z<1} x^{n+1} z^{m+1}(1-x-z) \mathrm{d} x \mathrm{~d} z \\
& =\int_{0}^{1} z^{m+1} \int_{0}^{1-z} x^{n+1}(1-x-z) \mathrm{d} x \mathrm{~d} z \\
& =120 \int_{0}^{1} z^{m+1} \int_{0}^{1-z}\left\{(1-z) x^{n+1}-x^{n+2}\right\} \mathrm{d} x \\
& =120 \int_{0}^{1} z^{m+1}\left\{\frac{(1-z)(1-z)^{n+2}}{n+2}-\frac{(1-z)^{n+3}}{n+3}\right\} \mathrm{d} z  \tag{18}\\
\mathrm{E}\left(x_{l}^{n} x_{h}^{m}\right) & =\frac{120}{(n+2)(n+3)} \int_{0}^{1} z^{m+1}(1-z)^{n+3} \mathrm{~d} z
\end{align*}
$$

By using special cases of (18), the expectation formulas except the conditional in (10) are derived.

The conditional distribution of $x_{h}$ given $x_{l}$ is given by

$$
\begin{align*}
h(z \mid x) & =\frac{f(x, z)}{g(x)}=\frac{120 x z(1-x-z)}{20 x(1-x)^{3}} \\
& =6 \frac{z(1-x-z)}{(1-x)^{3}}, \quad 0 \leq z \leq 1-x \tag{19}
\end{align*}
$$

The conditional expectation follows as:

$$
\begin{equation*}
E\left(x_{h} \mid x_{l}=x\right)=\frac{6 \int_{0}^{1-x} z^{2}(1-x-z) \mathrm{d} z}{(1-x)^{3}}=\frac{1-x}{2} \tag{20}
\end{equation*}
$$

Below, the relation $\int_{0}^{1} f(x) \mathrm{d} x=\int_{0}^{1} f(1-x) \mathrm{d} x$ is used some times. To show the relation, change $x \rightarrow 1-x$ in the integral.

By using (8),

$$
\begin{align*}
\mathrm{E}(T)= & \frac{1}{2}\left[y_{0} \mathrm{E} x_{l}+y_{1} \mathrm{E} x_{h}+\mathrm{E}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)\right\}\right. \\
& \left.+\mathrm{E}\left\{\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}\right] \\
= & \frac{y_{0}+y_{1}}{6}  \tag{21}\\
& +\frac{1}{2}\left[\mathrm{E}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)\right\}+\mathrm{E}\left\{\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}\right]
\end{align*}
$$

Next evaluate the two expectations in the brackets by using (9).

$$
\begin{aligned}
\mathrm{E}\{ & \left\{\left(1-x_{h}\right) y\left(x_{l}\right)\right\} \\
& =120 \iint_{\substack{0<x<1 \\
0<z<1-x}}(1-z) x z(1-x-z) y(x) \mathrm{d} x \mathrm{~d} z \\
& =120 \int_{0}^{1} x y(x) \int_{0}^{1-x} z(1-z)(1-x-z) \mathrm{d} z \mathrm{~d} x
\end{aligned}
$$

By integrating with respect to $x$, the following expression is obtained.

$$
\begin{equation*}
\mathrm{E}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)\right\}=10 \int_{0}^{1}\left(x-2 x^{2}+2 x^{4}-x^{5}\right) y(x) \mathrm{d} x \tag{22}
\end{equation*}
$$

Likewise,

$$
\begin{aligned}
\mathrm{E}\left\{\left(1-x_{l}\right)\right. & \left.y\left(1-x_{h}\right)\right\} \\
= & 120 \iint_{\substack{0<x<1-z \\
0<z<1}}^{1}(1-x) x z(1-x-z) y(1-z) \mathrm{d} x \mathrm{~d} y \\
= & 120 \int_{0}^{1} z y(1-z) \int_{0}^{1-z} x(1-x)(1-x-z) \mathrm{d} x \mathrm{~d} z
\end{aligned}
$$

By integrating with respect to $x$, and then substitute $1-z$ with $x, \mathrm{~d} x=-\mathrm{d} z$, we get.

$$
\begin{equation*}
\mathrm{E}\left\{\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}=10 \int_{0}^{1}\left(2 x^{3}-3 x^{4}+x^{5}\right) y(x) \mathrm{d} x \tag{23}
\end{equation*}
$$

By combining (21), (22) and (23), an expression for $\mathrm{E}(T)$ is obtained.

$$
\begin{align*}
\mathrm{E}(T) & =\frac{1}{6}\left(y_{0}+y_{1}\right)+5 \int_{0}^{1} x(1-x)[1-x(1-x)] y(x) \mathrm{d} x  \tag{24}\\
& =\frac{1}{6}\left(y_{0}+y_{1}\right)+5 \int_{0}^{1}\left(x-2 x^{2}+2 x^{3}-x^{4}\right) y(x) \mathrm{d} x
\end{align*}
$$

Formula (24) has been checked with the functions $a+b x, x^{2}$ and $(1-x)^{2}$ for $y(x)$. These results are identical with the results obtained by deriving $\mathrm{E}(T)$ directly from (8) and using (10).

Deriving formula (12) is a bit tedious. It follows from (8) that

$$
\begin{align*}
4 T^{2}= & \left\{y_{0} x_{l}+y_{1} x_{h}+\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}^{2} \\
= & \left(y_{0} x_{l}+y_{1} x_{h}\right)^{2}+\left(1-x_{h}\right)^{2} y^{2}\left(x_{l}\right)+\left(1-x_{l}\right)^{2} y^{2}\left(1-x_{h}\right) \\
& +2\left(y_{0} x_{l}\left(1-x_{h}\right)+y_{1} x_{h}\left(1-x_{h}\right)\right) y\left(x_{l}\right)  \tag{25}\\
& +2\left(y_{0} x_{l}\left(1-x_{l}\right)+y_{1} x_{h}\left(1-x_{l}\right)\right) y\left(1-x_{h}\right) \\
& +2\left(1-x_{h}\right)\left(1-x_{l}\right) y\left(x_{l}\right) y\left(1-x_{h}\right)
\end{align*}
$$

Then, the expectation of each term is derived by using (9) and (10) and trying to split double integrals.

$$
\begin{equation*}
\mathrm{E}\left\{\left(y_{0} x_{l}+y_{1} x_{h}\right)^{2}\right\}=\frac{1}{21}\left(3\left(y_{0}^{2}+y_{1}^{2}\right)+4 y_{0} y_{1}\right) \tag{26}
\end{equation*}
$$

$$
\begin{aligned}
\mathrm{E}\{ & \left.\left(1-x_{h}\right)^{2} y^{2}\left(x_{l}\right)\right\} \\
& =120 \iint_{\substack{x>0, z>0 \\
x+2<1}}(1-z)^{2} x z(1-x-z) y^{2}(x) \mathrm{d} x \mathrm{~d} z \\
& =120 \int_{0}^{1} x y^{2}(x) \int_{0}^{1-x}\left\{(1-x) z(1-z)^{2}-z^{2}(1-z)^{2}\right\} \mathrm{d} z \mathrm{~d} x
\end{aligned}
$$

By evaluating the integral with respect to $z$ and contracting, the following result are obtained:

$$
\begin{align*}
& \mathrm{E}\left\{\left(1-x_{h}\right)^{2} y^{2}\left(x_{l}\right)\right\}  \tag{27}\\
= & 2 \int_{0}^{1}\left(3 x-5 x^{2}+5 x^{5}-3 x^{6}\right) y^{2}(x) \mathrm{d} x
\end{align*}
$$

Next
$\mathrm{E}\left\{\left(1-x_{l}\right)^{2} y^{2}\left(1-x_{h}\right)\right\}$
$=120 \iint_{\substack{x>0, z>0 \\ x+z<1}}(1-x)^{2} x z(1-x-z) y^{2}(1-z) \mathrm{d} x \mathrm{~d} z$
$=120 \int_{0}^{1} z y^{2}(1-z) \int_{0}^{1-z}\left\{(1-z) x(1-x)^{2}-(1-x)^{2} x^{2}\right\} \mathrm{d} x \mathrm{~d} z$
Integrating with respect to $x$, contracting and changing the integrator variable $z$ with $1-x$, gives

$$
\begin{align*}
& \mathrm{E}\left\{\left(1-x_{l}\right)^{2} y^{2}\left(1-x_{h}\right)\right\}  \tag{28}\\
= & 2 \int_{0}^{1}\left(10 x^{3}-20 x^{4}+13 x^{5}-3 x^{6}\right) y^{2}(x) \mathrm{d} x
\end{align*}
$$

## Continuing

$$
\begin{aligned}
& \mathrm{E}\left[x_{l}\left(1-x_{h}\right) y\left(x_{l}\right)\right] \\
& \quad=120 \iint_{\substack{x>0, \gg 0 \\
x+z<1}} x(1-z) x z(1-x-z) y(x) \mathrm{d} x \mathrm{~d} z \\
& \quad=120 \int_{0}^{1} x^{2} y(x) \int_{0}^{1-x}\left[(1-x) z(1-z)-z^{2}(1-z)\right] \mathrm{d} z \mathrm{~d} x
\end{aligned}
$$

Evaluating and contracting gives

$$
\begin{align*}
& \mathrm{E}\left[x_{l}\left(1-x_{h}\right) y\left(x_{l}\right)\right] \\
= & 10 \int_{0}^{1}\left(x^{2}-2 x^{3}+2 x^{5}-x^{6}\right) y(x) \mathrm{d} x \tag{29}
\end{align*}
$$

The next term follows.

$$
\begin{aligned}
\mathrm{E} & \left\{x_{h}\left(1-x_{h}\right) y\left(x_{l}\right)\right\} \\
& =120 \iint_{\substack{x>0, z>0 \\
x+z<1}} z(1-z) x z(1-x-z) y(x) \mathrm{d} x \mathrm{~d} z \\
& =120 \int_{0}^{1} x y(x) \int_{0}^{1-x}\left\{(1-x)\left(z^{2}(1-z)-z^{3}(1-z)\right\} \mathrm{d} z \mathrm{~d} x\right.
\end{aligned}
$$

Evaluating and compressing gives

$$
\begin{align*}
& \mathrm{E}\left\{x_{h}\left(1-x_{h}\right) y\left(x_{l}\right)\right\} \\
& \quad=2 \int_{0}^{1}\left(2 x-5 x^{2}+10 x^{4}-10 x^{5}+3 x^{6}\right) y(x) \mathrm{d} x \tag{30}
\end{align*}
$$

Next term
$\mathrm{E}\left\{x_{l}\left(1-x_{l}\right) y\left(x_{h}\right)\right\}$
$=120 \iint_{\substack{x>0, z>o \\ x+z<1}} x^{2}(1-x) z(1-x-z) y(1-z) \mathrm{d} x \mathrm{~d} z$
$=120 \int_{0}^{1} z y(1-z) \int_{0}^{1-x}\left\{(1-z)\left(x^{2}(1-x)-x^{3}(1-x)\right\} \mathrm{d} x \mathrm{~d} z\right.$
Evaluating and compressing

$$
\begin{align*}
& \mathrm{E}\left\{x_{l}\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}  \tag{31}\\
= & 2 \int_{0}^{1}\left(5 x^{4}-8 x^{5}+3 x^{6}\right) y(x) \mathrm{d} x
\end{align*}
$$

The next last term

$$
\begin{aligned}
& \mathrm{E}\left[x_{h}\left(1-x_{l}\right) y\left(x_{h}\right)\right] \\
& \quad=120 \iint_{\substack{x>0, z>0 \\
x+z<1}} z(1-x) x z(1-x-z) y(1-z) \mathrm{d} x \mathrm{~d} z \\
& \quad=120 \int_{0}^{1} z^{2} y(1-z) \int_{0}^{1-x}\left[(1-z) x(1-x)-x^{2}(1-x)\right] \mathrm{d} z \mathrm{~d} x
\end{aligned}
$$

Integrating and changing $z$ with $1-x$ in the integral with respect to $z$ gives

$$
\begin{align*}
& \mathrm{E}\left\{x_{h}\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\} \\
& \quad=10 \int_{0}^{1}\left(2 x^{3}-5 x^{4}+4 x^{5}-x^{6}\right) y(x) \mathrm{d} x \tag{32}
\end{align*}
$$

And finally the last term

$$
\begin{align*}
& \mathrm{E}\left\{\left(1-x_{l}\right)\left(1-x_{h}\right) y\left(x_{l}\right) y\left(1-x_{h}\right)\right\}  \tag{33}\\
& =120 \iint_{\substack{x>0, \gg 0 \\
x+z<1}}(1-x)(1-z) x z(1-x-z) y(x) y(1-z) \mathrm{d} x \mathrm{~d} z
\end{align*}
$$

This formula could be used for the last term in (12). However, it is possible to derive alternative versions. One is given here.

$$
\begin{aligned}
& \mathrm{E}\left\{\left(1-x_{l}\right)\left(1-x_{h}\right) y\left(x_{l}\right) y\left(1-x_{h}\right)\right\} \\
= & 120 \int_{0}^{1} z(1-z) y(1-z) \int_{0}^{1-z} x(1-x)(1-x-z) y(x) \mathrm{d} x \mathrm{~d} z \\
= & 120 \int_{0}^{1} z(1-z) y(1-z) \\
& \quad \int_{0}^{1-z}\left\{(1-z) x(1-x)-x^{2}(1-x)\right\} y(x) \mathrm{d} x \mathrm{~d} z \\
= & 120 \int_{0}^{1}(1-z) z y(z) \int_{0}^{z}\left\{z\left(x-x^{2}\right)-x^{2}+x^{3}\right\} y(x) \mathrm{d} x \mathrm{~d} z
\end{aligned}
$$

Hence

$$
\begin{align*}
& \mathrm{E}\left\{\left(1-x_{l}\right)\left(1-x_{h}\right) y\left(x_{l}\right) y\left(1-x_{h}\right)\right\} \\
& =120 \int_{0}^{1}(1-z) z y(z)  \tag{34}\\
& \times \int_{0}^{z}\left\{\left(x-x^{2}\right) z-x^{2}+x^{3}\right\} y(x) \mathrm{d} x \mathrm{~d} z
\end{align*}
$$

Inserting (26) - (33) or (34) into (25) gives (12).
Formula (12) has been controlled with a linear function for $y(x) . \operatorname{Var}(T)=0$ for linear functions, and
this were obtained by the control. The derivation of (11) and (12) do not involve complicated mathematics, but rather an extensive algebraic task with risk of making errors.

This section is concluded with deriving (14). There may be several ways to obtain an estimator for $\operatorname{Var}(T)$. The version derived here is rather straightforward.

Since $\operatorname{Var}(T)=\mathrm{E}\left(T^{2}\right)-\mathrm{E}(T)^{2}$, we try to estimate both terms. As $T^{2}$ is an unbiased estimator of $\mathrm{E}\left(T^{2}\right)$, we can use $T^{2}$ and subtract an estimator for $\mathrm{E}^{2}(T) \equiv[\mathrm{E}(T)]^{2}$.
$T^{2}$ is given by (25), but the expectation of the first term $\left(y_{0} x_{l}+y_{1} x_{h}\right)^{2}$ is computable from the data. Therefore $\left(y_{0} x_{l}+y_{1} x_{h}\right)^{2}$ is replaced by its expected value $\frac{1}{21}\left(3\left(y_{0}^{2}+y_{1}^{2}\right)+4 y_{0} y_{1}\right)$ in the expression for $T^{2}$. This will reduce the variability of the estimator. This is confirmed by trying both versions during sampling artificial signal functions, and the version with the first terms equal to $\left(3\left(y_{0}^{2}+y_{1}^{2}\right)+4 y_{0} y_{1}\right) / 21$ is smallest on all occasions.

An expression for $\mathrm{E}^{2}(T)$ follows from (21).

$$
\begin{aligned}
\mathrm{E}^{2}(T) & =\frac{1}{36}\left(y_{0}+y_{1}\right)^{2} \\
& +\frac{1}{6}\left(y_{0}+y_{l}\right) \mathrm{E}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\} \\
& +\frac{1}{4} \mathrm{E}^{2}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}
\end{aligned}
$$

Using the arguments in the expectation operators as estimators for the expectations, and performing the subtraction $T^{2}-\hat{\mathrm{E}}^{2}(T)$ the following estimator is obtained.

$$
\begin{aligned}
& \operatorname{Var}(T)=\frac{1}{126}\left(y_{0}^{2}+y_{1}^{2}-13 y_{0} y_{1}\right)+\frac{1}{4}\left\{\left(1-x_{h}\right)^{2} y^{2}\left(x_{l}\right)+\left(1-x_{l}\right)^{2} y^{2}\left(x_{h}\right)\right\} \\
& +\left\{\frac{y_{0}}{2}\left(x_{l}-\frac{1}{3}\right)+\frac{y_{1}}{2}\left(x_{h}-\frac{1}{3}\right)\right\}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\} \\
& +\frac{1}{2}\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)-\frac{1}{4}\left[\mathrm{E}^{2}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}\right]
\end{aligned}
$$

This expression is identical with (14). The last line cannot be estimated unbiased by using $\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}^{2}$, as this would yield an over estimate of $\mathrm{E}^{2}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}$, This follows from the general result,
$\operatorname{Var} X=\mathrm{E}\left(X^{2}\right)-\mathrm{E}^{2} X$, that holds for any random variable $X$. However,

$$
\begin{aligned}
& \mathrm{E}^{2}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\} \\
&= \mathrm{E}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}^{2} \\
&-\operatorname{Var}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}
\end{aligned}
$$

and it is believed that the variance term is small relative to the expectation terms in general. Therefore, (14) with $-\frac{1}{4}\left\{\left(1-x_{h}\right) y\left(x_{l}\right)+\left(1-x_{l}\right) y\left(1-x_{h}\right)\right\}^{2}$ as the last line was used in Appendix II.

Estimator (14) depends on $x_{l}$ and $x_{h}$ even if the signal function is linear. This means that it takes values that in general are different from zero even when the variance is zero, as in the linear case. To try estimator (14), a program that calculates it has been made. Results from using (14) in the main paper are given in Appendix II.

Estimator (14) depends on $x_{l}$ and $x_{h}$ even if the signal function is linear. This means that it takes values that in general are different from zero even when the variance is zero, as in the linear case. To try estimator (14), a program that calculates it has been made. Results from using (14) in the main paper are given in Appendix II.

In some difficult cases the variance estimator takes negative values. This happens on some sampling occasions with 5 strata, but also more seldom with 10 strata on the signal functions a and b in Appendix II. The standard deviation is set to zero in such cases.

## Appendix II

## Sampling man made functions

The two variance formulas based on (13) and given by (14) have been tried by sampling several artificial signal functions. An immediate impression from these investigations is that the two variance estimates show different properties with the variance estimator based on (13) being the smallest.

The advantage with sampling known artificial signal functions is that the parameter to be estimated can be calculated as well as the expected value (11) and variance (13) of the estimator (8) used. From this, the bias and deviation between estimated and true value follows. In the present case, however, the variance (13) was not computed. From repeated independent samples the variability of the estimate is shown, and this indicates something about the theoretical variability of the estimator.

Results from 30 repeated independent samples of four different signal functions are given. Each signal function is sampled with 5 and 10 strata for comparison.

The signal functions were constructed by means of the 10 first terms in a Fourier series and added some higher frequencies. To ensure that the signal functions are non-negative, they were put equal to zero if they were negative. There are many parameters in such functions, and it is possible to make any function shape. However, since real signal functions are integrated echo intensities over depth and along a fixed direction in the area, they will seldom show very sharp variations.

The strata widths are constant and equal to one for 5 strata, and half width for 10 strata. Then the integral over the signal function is the same for 5 and 10 strata.

However, when fish concentrate to schools, real signal functions may be difficult to sample, in particular when the area of distribution is small. The signal functions a and b shown in Figure 4 and 6 are of this kind.
The results of 30 repeated independent samplings are shown in Figure 5. Note that the y-axis is similar in these plots.

It is seen from Figure 5 that the error interval based on (14) seems not to be wider than the error interval


Figure 4. The first signal function (a) sampled with 5 and 10 strata, respectively. On the first sampling occasion, the function was observed at the locations where there are vertical lines, including the grid.


Figure 5. Estimate, true value, expected value and estimates of two error bounds for 30 independent samples of signal function a based on sampling with 5 (upper) and 10 strata (lower graph).
based on (13). It is also seen that the error interval based on (14) has zero width several places, especially for 5 strata.
For each signal function that was sampled 30 times, it was counted how many times the error intervals contained the true value, and how many times they contained the expected value of the estimator. For signal function a, these numbers are given in Table 1.
The next signal function (b) is shown in Figure 6. This is also a difficult function to estimate.

The result of 30 independent samples of signal function b is shown in Figure 7.

Here is seen that the two error intervals seems to be of the same order for 5 strata, but for 10 strata the error interval based on (14), main paper, is the biggest.

The numbers of times the error intervals contain the true value and the expected value of the estimator are given in Table 2. See Table 1 for more information.

The next signal function is shown in Figure 8. This
Table 1: Numbers of times the error interval contained the true value and expected value, respectively out of 30 independent samples of signal function a.

|  | 5 strata |  | 10 strata |  |
| :--- | ---: | ---: | ---: | ---: |
| Estimate $\pm 2$ 2SD based on (13) | 13 | 13 | 24 | 21 |
| Estimate $\pm$ SD based on (14) | 11 | 7 | 17 | 18 |



Figure 6. Signal function b sampled with 5 and 10 strata, respectively. The first sampling locations within each stratum are shown as red vertical lines.


Figure 7. Estimate, true value, expected value and two estimates of error bound based on 30 independent samples of signal function b based on sampling with 5 (upper) and 10 strata (lower graph).

Table 2: Numbers of times the error interval contained the true value and expected value, respectively out of 30 independent samples of signal function $\mathbf{b}$.

|  | 5 strata |  | 10 strata |  |
| :--- | :--- | :--- | :--- | :--- | :--- |
| Estimate $\pm 2$ SD based on (13) | 13 | 13 | 24 | 25 |
| Estimate $\pm$ SD based on (14) | 15 | 17 | 30 | 30 |



Figure 8. Signal function c sampled with 5 and 10 strata, respectively. The sampling positions in each stratum are shown as vertical lines.
function start with low values reaches a maximum before it declines and goes to zero.

The results of estimating the area under signal function c is shown in Figure 9.

Here the error interval based on (14) is the biggest both for 5 and 10 strata. This case is seldom in that the bias for 10 strata is bigger that the bias for 5 strata.

The numbers of times the error intervals contain one of two parameters are given in Table 3. See Table 1 for explanation.

The last signal function (d) is shown in Figure 10. This is a function with small sharp variations, and is an easy function to estimate.

The results of estimating the area under signal function d are shown in Figure 11.

Figure 11 and others indicate that the error intervals stabilize, gets narrower as well as less variable when the functions are sampled with 10 strata compared with 5 strata. But it also seems that the error interval based on (13) decreases faster than that based on (14) when the sampling density increases.

The numbers of times the error intervals for signal function d contain one of two parameters are given in Table 4. See Table 1 for more explanation.


Figure 9. Estimate, true value, Expected value and two estimates of the error bounds for 30 independent samples of signal function $c$ based on sampling with 5 (upper) and 10 strata (lower graph).

Table 3. Numbers of times the error interval contained the true value and expected value, respectively out of 30 independent samples of signal function $c$.

|  | 5 strata |  | 10 strata |  |
| :--- | :---: | :---: | :--- | :---: |
| Estimate $\pm 2$ SD based on (13) | 11 | 14 | 11 | 18 |
| Estimate $\pm$ SD based on (14) | 30 | 30 | 30 | 30 |



Figure 10. Signal function d sampled with 5 and 10 strata, respectively. The function is observed at the $x$-positions of the vertical lines, including the grid at the first sampling occasion.


Figure 11. Estimate, true value, Expected value and two estimates of the error bounds for 30 independent samples of signal function d based on sampling with 5 (upper) and 10 strata (lower graph).

Table 4. Numbers of times the error interval contained the true value and expected value, respectively out of 30 independent samples of signal function $d$.

|  | 5 strata |  | 10 |  |
| :--- | :---: | :---: | :---: | :---: |
| strata |  |  |  |  |
| Estimate $\pm 2$ SD based on (13) | 22 | 25 | 8 | 26 |
| Estimate $\pm$ SD based on (14) | 29 | 30 | 30 | 30 |

The results from sampling signal function d shows that the estimator based on (13) shrinks considerable from sampling with 5 strata to sampling with 10 strata. That the unknown value is within the error interval in only 8 cases for 10 strata is not good, but the expected value is contained in the error interval 26 of 30 times. Since the expected value is rather close to the true value here, it may be concluded that not many estimates are far from the true value. But this may, nevertheless, indicate an unfortunate property with this estimator in that it may be considerably to small if the sampling density is high while the signal function is smooth.

The Figures in Appendix II indicate that the variance estimator based on (13) is negatively biased. It is not easy, although not impossible, to derive the expectation of this estimator. Therefore, the estimator is tried on arti-
ficial signal functions.
The estimator given by (14) has not good properties, as the estimates within strata may be negative. When all within strata estimates are summed over several strata, the resulting variance has better properties, but based on the present figures it is hard to believe that this estimator is nearly unbiased. Note that the error interval based on (14) is calculated as Estimate $\pm$ estimated standard deviation, while that based on (13) is calculated as Estimate $\pm$ two times the estimated standard deviation.

Some of the chosen signal functions are difficult to estimate, in particular signal function $a$ and $b$ as shown in the Figures 4 and 6. This may occur in fisheries acoustics when fish concentrate to schools. To improve precision in such cases, the sampling effort has to be increased.

