Distance sampling: its relevance to wildlife management

R. M. Fewster

University of Auckland, Auckland, New Zealand

and S. T. Buckland

University of St Andrews, St Andrews, Scotland

Summary. Distance sampling is a widely used method of abundance estimation for wild animal populations. Recently, Barry and Welsh (2001) raised queries about the validity of the distance sampling estimators. We show that their conclusions were based on unrepresentative designs, and redo their evaluations in a representative framework. With respect to this framework, we show that the distance sampling estimators have the desired properties of design-unbiasedness and asymptotic consistency. We derive conditions on the spatial distribution of objects needed for successful estimation of the detection function, and examine analytic variance estimators. We also show that a completely model-based approach can improve precision in density estimation.

Keywords: Animal abundance estimation; Circular design; Design-based analysis; Line transect sampling; Model-based analysis; Object density.

Address for correspondence: Rachel Fewster, Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand.

E-mail: r.fewster@auckland.ac.nz

1 Introduction

Distance sampling (Buckland *et al.*, 1993) defines a suite of methods for estimating the abundance of objects, usually animals, in a study area. An observer stands at a point or travels along a line, counting the number of objects observed and recording their distances from the line or point. Detectability is assumed to diminish with increasing distance, so the observed object distances are modelled to estimate a detection function. Using the detection function, the true object abundance can be inferred from the detected abundance. The most popular approach to distance sampling is line transect sampling.

Distance sampling can be used for a wide range of wildlife applications, and can be based either on counts of the animals themselves, or on cues such as nests or dung piles. The methods are relatively inexpensive to apply, and the assumptions involved are often easily met by good design. For these reasons, distance sampling techniques are used extensively in practice. However, the validity of the methods and estimators was called into question in a recent paper by Barry and Welsh (2001). If their analysis is correct, distance sampling methods have limited value to practitioners. Because so many wildlife managers currently rely on the methods, it is important to evaluate the results of Barry and Welsh. Like them, we focus on line transect sampling, although the issues for point transects are similar.

1.1 Background and notation

Let A be a study area, containing an unknown number N of objects. For most of our analyses, we use a rectangular study area with a corner at the origin: $A = [0, \ell] \times [0, 1]$. The area of A is written $|A|$. The true density of objects in the area, which we wish to estimate, is $\Delta = N/|A|$.

We estimate Δ by surveying m line transects. Following Barry and Welsh, in most of our analyses transect i is the line $y = u_i$, and the length of the transect is $\ell_i = \ell$. This means that each transect covers the whole horizontal extent of A, so only the y-coordinates of the objects are important. The N objects have y–coordinates $y_1, \ldots, y_N$.

In surveying a transect, we establish a search strip of width w to either side of the transect. (w is sometimes called the half-width of the strip.) Objects within the search strip can be detected, but objects outside the strip are considered too far away and are ignored if seen. A total of n_i objects is detected from transect i, so the total number of detections is $n = \sum_{i=1}^{m} n_i$. The recorded data are $n_1, \ldots, n_m$, together with each distance $|y_j - u_i|$ of a detected object from the appropriate transect line.

The detection function, $g(r, \theta)$, is the probability that an object is detected, given that it is located at distance r from the transect line. Detectability is zero outside the search strip, so $g(r, \theta) = 0$ if $|r| > w$. A function is proposed for g, parametrized by θ, and θ is estimated from the observed distance data. One possibility is the half-normal detection function, $g(r, \theta) = \exp\{-r^2/(2\theta^2)\}\, I(r \le w)$, where $I(\cdot)$ is the usual indicator. A key assumption of basic distance sampling is that $g(0, \theta) = 1$, so all objects directly on the transect line are detected. It is also assumed that $g(-r, \theta) \equiv g(r, \theta)$.

Let

$$\bar{g}_\theta = \frac{1}{w} \int_0^w g(r, \theta)\, dr\, .$$

If the distribution of object distances from the transect line is uniform across the search strip, then $\bar{g}_\theta$ is the probability that an object is detected, given that it is in the strip. Suppose that the detection function parameter θ is known. The basic distance sampling estimator of object density from m transects is

$$\hat{\Delta} = \frac{n}{2\ell w m \bar{g}_\theta}\, .$$

This can be interpreted as estimating the total number of objects in the m search strips by $n/\bar{g}_\theta$, and scaling by the total area surveyed, $2\ell w m$, to estimate the overall object density.

If the detection parameter is not known, we estimate it by $\hat{\theta}$ and use density estimator $\hat{\Delta} = n/\left(2\ell w m \bar{g}_{\hat{\theta}}\right)$. Maximum likelihood is used to obtain $\hat{\theta}$. Let $f(r, \theta)$ be the p.d.f. of distances of detected objects: then $f(r, \theta)dr$ is the probability that the distance of a detected object from the transect line is in the interval $(r, r+dr)$. Using elementary probability arguments, we obtain $f(r, \theta) = g(r, \theta)/(w\bar{g}_\theta)$ if object distances are uniformly distributed across the search strip. We can therefore maximize the log-likelihood $\sum_{j=1}^{n} \log\{g(r_j, \theta)/(w\bar{g}_\theta)\}$ to obtain $\hat{\theta}$, where r_j is the observed distance of the jth detected object from the transect line.

1.2 Assumptions

In the derivation above, we made assumptions that $g(-r, \theta) \equiv g(r, \theta)$; that $g(0, \theta) = 1$; that the distribution of object distances across the search strip is uniform; and that objects are detected independently of each other. The first assumption is generally accepted.

The second is reasonable for many animal populations, but is problematic in some circumstances, for example marine surveys of diving animals or ground surveys of birds in high canopies. If we cannot assume that $g(0, \theta) = 1$, a more advanced distance sampling scheme is required, such as the double platform designs of Borchers *et al.* (1998).

The assumption that object distances are uniformly distributed across the search strip is stronger than necessary, as the same derivation holds if the p.d.f. is linear. We investigate this issue in more detail in Section 3. For the fourth assumption, violation of independence has limited effect on density estimation, but it can affect variance estimators.

A further assumption of distance sampling is that objects should be detected at their initial locations, before there is any movement in response to the observer. Careful attention to field methods is required to meet this assumption, and a more advanced sampling scheme may be necessary if responsive movement cannot be eliminated. Counting a moving object repeatedly from the same transect will also cause bias.

In addition to the theoretical assumptions, the performance of the method can be greatly enhanced by following some basic rules of survey design. Buckland *et al.* (2001) recommend that the number of transect lines m is at least 10 to 20, to ensure good coverage and improve estimates of variance. They suggest that at least $n = 60$ to 80 observations are made before the detection function is estimated. If a substantial gradient in object density is known to occur across A, transects should be placed parallel to the gradient to make the coverage more representative. This often happens in marine surveys, where animal density changes along transects perpendicular to the coastline. The placement of transect lines should be random with respect to the object distribution, and use of roads or tracks as transect lines should be avoided.

1.3 Inferential framework

Barry and Welsh introduced two frameworks for evaluating the distance sampling estimators. The first is the design-based framework, in which the positions $y_1, \ldots, y_N$ of the N objects are regarded as fixed. The randomness in the estimation of Δ and $\text{var}(\hat{\Delta})$ is due to the design, namely the positions $u_1, \ldots, u_m$ of the m transects. The second is the model-based framework, in which the transect position u is fixed, and randomness is due to the object positions $y_1, \ldots, y_N$. Objects are assumed to be distributed across A according to some probability density function $k(y)$.

In both the model-based and design-based frameworks, Barry and Welsh investigated the bias of $\hat{\Delta}$, $\hat{\theta}$, and various variance estimators. However, their investigations were flawed in two main respects. In the design-based framework, they used a design that did not provide representative coverage of the area A. Not surprisingly, this caused bias in the results. In the model-based framework, they fixed the transect positions in an unrepresentative region of A, and attempted to extrapolate their density estimates to the whole of A. Again, this was an unreasonable approach that understandably produced poor results.

Our approach to evaluating the estimators differs a little from that of Barry and Welsh, although we make use of both frameworks and build on many of their ideas. The philosophy behind distance sampling is twofold. Firstly, object density is modelled within an individual search strip, to obtain a density $\hat{\Delta}_i = n_i/(2w\ell\bar{g}_{\hat{\theta}})$ within that strip. Secondly, the m search strips are assumed to form a representative subset of A, which enables us to construct the overall density estimate of $\hat{\Delta} = n/(2w\ell m\bar{g}_{\hat{\theta}})$.

To evaluate the estimators, we need to make rigorous the definition of a representative design that validates the second of these steps. We do this in Section 2, where we conduct the same design-based analyses of the estimators as Barry and Welsh. We show that their design was unrepresentative, and that with respect to a representative design, the estimators have the properties desired.

For the estimation of object density within a strip, we move to the model-based framework. This enables us to investigate conditions on the object distribution within a strip that will yield unbiased strip estimates of Δ and θ. If all strips satisfy the conditions for unbiased within-strip density estimation, we can rely on a representative design with large m for successful estimation of Δ in the whole area A. Our model-based analysis is performed in Section 3. We also present results from a completely model-based framework in which density is modelled across the whole of A, rather than relying on the representative nature of the design.

Finally we examine the variance estimators. For these, it is inappropriate to confine ourselves either to the model-based or the design-based framework, because the variability in the estimator $\hat{\Delta}$ has components due to both object distribution and transect location. We investigate the properties of the traditional variance estimators in Section 4, and also derive a new estimator.

2 Design-based analysis of $\hat{\Delta}$

2.1 Design-based analysis when the detection parameter is known

Suppose that the detection function $g(r, \theta)$ and the parameter θ are known. The study area is $A = [0, \ell] \times [0, 1]$, and we consider a single horizontal transect with length ℓ and y-coordinate u. The density estimator from this transect is $\hat{\Delta} = n/(2\ell w \bar{g}_\theta)$, where n is the number of detections and $\bar{g}_\theta = w^{-1} \int_0^w g(r, \theta)\, dr$.

$\hat{\Delta}$ is design-unbiased for Δ if $E_u(\hat{\Delta} \,|\, \boldsymbol{y}) = \Delta$, where the expectation is taken over the random transect position u, and conditions on the object positions $\boldsymbol{y} = (y_1, \ldots, y_N)$. Whether or not $\hat{\Delta}$ is design-unbiased depends upon the design. In the current framework, with fixed length and orientation of transects, the design governs the distribution of u and also specifies how to handle search strips that extend outside the area A. We are only interested in the design-based properties of $\hat{\Delta}$ with respect to designs that provide representative coverage of the objects in A. We define this to mean that every object in A has equal *a priori* probability of being sampled, before the transect position is known.

Barry and Welsh consider a design with $u \sim \mathrm{U}[0, 1]$, under which any parts of the search strip $[u - w, u + w]$ that lie outside of A are ignored. We shall call this a linear design on $[0, 1]$. The linear design is not representative, because when $w < 0.5$ the sampling probability is lower for points y in $[0, w]$ or in $[1 - w, 1]$ than for points in $[w, 1 - w]$. Similarly, for $w > 0.5$, points in $[0, 1 - w]$ or $[w, 1]$ have lower sampling probability than points in $[1 - w, w]$. Barry and Welsh use a number of different object configurations to evaluate the properties of $\hat{\Delta}$ under the linear design. Many of these place the majority of objects into $[1 - w, 1]$, so the effect of under-sampling of objects at the edges is extreme. In part, this is the cause of the severe design bias that they report in $\hat{\Delta}$ under the linear design.

To obtain a representative design with $u \sim \mathrm{U}[0, 1]$, we must eliminate the problem of under-sampling at the edges. This can be achieved by a circular design that conceptually glues A onto the surface of a cylinder with the horizontal edges $y = 0$ and $y = 1$ joined. Any part of the search strip lying beyond one edge of A reappears at the other edge. As long as $w \leq 0.5$, this means that a transect position at u effectively establishes three non-overlapping search strips, with centrelines at $u - 1$, u, and $u + 1$. We survey all parts of the search strips that overlap with A.

We can readily show that the circular design is representative and $\hat{\Delta}$ is design-unbiased when $w \leq 0.5$. The *a priori* probability of detecting a point at position y is

$$\int_{\max(y-w,0)}^{\min(y+w,1)} g\Big(|y-u|,\theta\Big)\,du \;+\; \int_{\min(y-w+1,1)}^{1} g\Big(|y-(u-1)|,\theta\Big)\,du \;+\; \int_{0}^{\max(0,y+w-1)} g\Big(|y-(u+1)|,\theta\Big)\,du\,,$$

where the three terms correspond to the parts of A overlapping the three search strips. This reduces after appropriate changes of variable to $\int_{y-w}^{y+w} g(|y-r|,\theta)\,dr = 2\int_0^w g(r,\theta)\,dr$, showing that every point y has equal probability $2\int_0^w g(r,\theta)\,dr$ of being sampled. We then have

$$E_u(\hat{\Delta}) = \frac{1}{2\ell w\bar{g}_\theta}\sum_{j=1}^{N} E_u\left\{I(\text{object at } y_j \text{ detected})\right\} = \frac{2N\int_0^w g(r,\theta)\,dr}{2\ell\int_0^w g(r,\theta)\,dr} = \frac{N}{\ell\times 1} = \Delta\,,$$

so the estimator $\hat{\Delta}$ is design-unbiased.

If $w > 0.5$ we cannot use the circular design on $[0,1]$, because the edges of the search strip overlap when wrapped around the notional cylinder. For a representative design when $w > 0.5$, we can extend the sampled area to $A' = [0,\ell]\times[\frac{1}{2}-w,\frac{1}{2}+w]$, and use $u \sim \mathrm{U}[\frac{1}{2}-w,\frac{1}{2}+w]$. The density in A' is $\Delta' = N/(2w\ell) = \Delta/(2w)$. The search strip now fits into the area A' exactly, so by the working above we have $E_u(\hat{\Delta}') = \Delta'$. Thus $E_u(\hat{\Delta}) = E_u(2w\hat{\Delta}') = 2w\Delta' = \Delta$, so the extended circular design is again unbiased for Δ.

The poor results for $E_u(\hat{\Delta})/\Delta$ in Table 1 of Barry and Welsh can be explained by two effects that are corrected by the circular design. The first is the reduced sampling probability for objects y in $[0,w]$ or $[1-w,1]$. This effect causes underestimation of density for $w < 0.5$ when the objects are concentrated in the edge regions: for example when $y \sim \mathrm{Beta}(5,1)$, Beta(5, 0.5), or Beta(5, 0.25). The good results for $y \sim \mathrm{Beta}(3,3)$ occur because this distribution focuses objects away from the edges.

The second problem is lack of attention to the area sampled. When w is large the search strip $[u-w,u+w]$ can extend substantially outside of the area A. As the density outside A is zero, the object density in the sampled area is lower than Δ. This effect can be seen from the entry in Table 1 of Barry and Welsh for $\theta = 2$, $w = 1$, and $y \sim \mathrm{Beta}(1,1)$. The entry is $E_u(\hat{\Delta})/\Delta = 0.510$. With these parameters, all objects lie in the search strip, and most are detected, but the area sampled is twice the size of A and therefore the density in the sampled region is half the intended density Δ. In this situation, either the area or the density estimate should be amended before reporting results.

The design-unbiasedness of $\hat{\Delta}$ with the circular design holds for any detection function and any configuration of objects on $[0, 1]$. In practice, a circular design may be inconvenient to implement, because a search strip at one edge of A must be wrapped around to the opposite edge. For an easy implementation, we can define a buffer zone of width w around A, and survey the region $A'' = [0, \ell] \times [-w, 1 + w]$. Any objects in the buffer zone are ignored, so overall density is $\Delta'' = N/[(1 + 2w)\ell]$. The sampling protocol for a circular design on A'' is identical to that for the more convenient linear design, because no objects are recorded in the wrapped part of the search strip so we can simply ignore it. The estimated density in A'' is $\hat{\Delta}''$, which is adjusted to $\hat{\Delta} = (1 + 2w)\hat{\Delta}''$.

In almost all real surveys, the design-based bias of $\hat{\Delta}$ under the linear design is negligible. Only when w is large compared with A, or most objects are concentrated at the edges of A, does it become important to use a circular design or buffer zone. Even a strip width of $w = 0.05$ (5% of the width of A) is large compared with likely values in the field. Note also that, although $\hat{\Delta}$ is design-unbiased for any set of objects in $[0, 1]$, the density estimate from an individual transect will be most reliable if the transect is placed parallel to the density gradient. For the extreme Beta distributions used by Barry and Welsh, the transects occur at right angles to the density gradient, and the variance of $\hat{\Delta}$ will be high.

2.2 Design-based analysis when the detection parameter is estimated

Barry and Welsh present an asymptotic analysis of the design-based properties of $\hat{\Delta}$ when the detection function is known only up to an unknown parameter θ. The sightings data from which θ is estimated are the distances r_j of observed objects from the randomly selected transect lines. The function $\text{crit}(t)$ is defined as $\text{crit}(t) = \sum_{j=1}^{n} \log\{g(r_j, t)/(w\bar{g}_t)\}$. Under certain object distributions and independence of detections, $\text{crit}(t)$ is the exact log-likelihood function; otherwise it is an approximation. As the number of transects $m \to \infty$, the strong law of large numbers states that $\text{crit}(t) \to E_u(\text{crit}(t) \mid \boldsymbol{y})$. The quantity $\theta_0(\boldsymbol{y}) = \arg\max_t \{E_u(\text{crit}(t) \mid \boldsymbol{y})\}$ is defined to be the maximum of the asymptotic log-likelihood.

On the whole, the asymptotic design-based properties of $\hat{\Delta}$ are of limited interest, as the number of transects m is not usually large enough for asymptotic results to apply. Nonetheless, under the circular design it can be shown that $\theta_0(\boldsymbol{y}) = \theta$ for any detection function g, as long as $g(r, \theta) \equiv g(-r, \theta)$. This contrasts with the poor results in Table 2 of Barry and Welsh, which give $\theta_0(\boldsymbol{y})/\theta$ for the linear design.

When θ is estimated, we can write $\hat{\Delta} = \hat{\Delta}(\hat{\theta}) = \hat{\Delta}(\theta_0(\boldsymbol{y})) + (\hat{\theta} - \theta_0(\boldsymbol{y}))\hat{\Delta}'(\theta_0(\boldsymbol{y})) + \ldots$. Using the circular design, we have $\theta_0(\boldsymbol{y}) = \theta$, from which the leading term approximation to $E_u(\hat{\Delta}(\hat{\theta}))$ is $E_u(\hat{\Delta}(\theta)) = \Delta$. Again, this differs from the results in Table 3 of Barry and Welsh for the linear design.

We return to the estimation of θ in greater detail in the model-based framework, where an asymptotic treatment is more appropriate and the issues are more interesting. There, we can establish conditions on the spatial distribution of objects that will ensure that $\hat{\theta}$ and $\hat{\Delta}$ are asymptotically unbiased. Note that when θ is estimated we do not expect in either framework that $\hat{\Delta} = n/(2\ell w \bar{g}_{\hat{\theta}})$ will be unbiased in a non-asymptotic sense, because maximum likelihood estimation of θ does not produce an unbiased estimate of $1/\bar{g}_{\theta}$.

3 Model-based analysis of $\hat{\Delta}$

For a model-based analysis of the distance sampling estimators, we consider the effect of object distribution when the transect position is fixed. For convenience, we place the transect line at $u = 0$ and sample objects in the strip $[-w, w]$. This means we no longer consider the area A, but restrict attention to the search strip only. We define $k(y)$ to be the p.d.f. of object positions in the strip, so $\int_{-w}^{w} k(y)\, dy = 1$, and we aim to study conditions on $k(y)$ necessary for successful estimation of Δ in the strip.

Barry and Welsh present results of a similar model-based analysis in Tables 4 and 6 of their paper, but their results are badly distorted by extrapolation of the strip density estimates into the whole area A. With a fixed search strip, it is impossible to obtain an unbiased estimate of the density in A unless the search strip is representative of A. The usual philosophy of distance sampling is that each search strip should yield a model-unbiased estimate of density within the strip, while the design ensures that the total region sampled is representative of the overall area. Accordingly, in the model-based framework we focus on density estimation within the strip alone. The properties of a representative design were established in Section 2.

In Figure 1 we plot some of the configurations used by Barry and Welsh in their model-based analysis, from which it is clear that the sampled areas are highly unrepresentative of A. Such experiments do not tell us anything useful about desirable properties of $k(y)$.

Note that in each of their configurations, the search strip could be made representative of A by changing the orientation of the transect, as in Figure 1(d). Aligning the transect parallel to the density gradient in this way is recommended by Buckland *et al.* (1993), page 5, Fig 1.4, and page 298.

3.1 Model-based analysis when the detection parameter is known

For simplicity, we take $w = 1$ and select a range of values for the detection parameter θ, from $\theta = w/4$ (limited detection), to $\theta \gg w$ (perfect detection). We study four different shapes for $k(y)$ on the strip $y \in [-1, 1]$: 1. Uniform, $k(y) = 0.5$; 2. Linear, $k(y) = 0.5 + 0.25y$; 3. $\cap$-Quadratic, $k(y) = 0.6 - 0.3y^2$; 4. $\cup$-Quadratic, $k(y) = \frac{1}{3} + 0.5y^2$. We use the notation $\cap$-Quadratic to indicate a peaked quadratic shape, which might occur in the field if there is responsive movement of animals towards the transect. Likewise, $\cup$-Quadratic indicates a U-shaped quadratic that can represent responsive movement away from the transect. We selected the functions for their shapes, so the exact specifications are arbitrary. Figure 2 shows the four functions $k(y)$ and the range of detection functions considered.

Let Δ be the density within the search strip: then $\Delta = N/(2w\ell)$, where N is the number of objects in the strip. The estimator is $\hat{\Delta} = n/(2w\ell\bar{g}_\theta)$ as usual. If θ is known, then

$$\frac{E_y(\hat{\Delta})}{\Delta} = \frac{\int_{-1}^{1} g(|y|, \theta)k(y)\, dy}{\bar{g}_\theta}\,.$$

This ratio is tabulated in Table 1 for the various choices of $k(y)$ and θ. When detection is high, namely $\theta > w$, the ratio is approximately unity for each function k. This is because $\bar{g}_\theta \simeq 1$, so the estimated number of objects is roughly equal to the observed number.

As detection decreases, bias appears for the quadratic object distributions but not for the linear and uniform distributions. The bias arises for the quadratic distributions because the object density changes beyond the part of the strip where the change can be properly detected. To some extent this can be counteracted by truncating the observations to a smaller value of w, but not entirely. The estimator still performs moderately well but the bias is not negligible.

For linear and uniform object distributions, $\hat{\Delta}$ will always be model-unbiased. In fact a sufficient condition for model-unbiasedness is that $k(-y)+k(y)$ is constant, or equivalently that $k(y) - k(0)$ is an odd function of y. The proof is straightforward. Firstly we note that if $k(-y)+k(y)=c$, then the condition $\int_{-w}^{w} k(y)\,dy = 1$ forces $c = w^{-1}$. Then

$$\frac{E_y(\hat{\Delta})}{\Delta} = \frac{\int_{-w}^{w} g(|y|,\theta)k(y)\,dy}{\bar{g}_\theta} = \frac{\int_0^w g(y,\theta)\,\{k(-y)+k(y)\}\,dy}{w^{-1}\int_0^w g(y,\theta)\,dy} = 1\,.$$

Intuitively, when $k(-y)+k(y)$ is constant, an overabundance of objects on one side of the line is exactly compensated for by an underabundance on the other side, so the overall density estimate is correct.

The differing shapes for $k(y)$ correspond to different situations in the field. Because the width of the search strip is usually very small compared with that of the survey area (often as little as 0.1%), any large-scale trends in object density over the survey area are likely to appear approximately linear over the search strip. As long as there are no other factors influencing density gradient across the strip, we have high confidence in the estimator $\hat{\Delta}$. The large-scale changes in Δ across the survey area are allowed for by surveying several strips, following a representative design.

When objects move in response to the presence of an observer, the object distribution across the strip can become non-linear in such a way that $k(-y)+k(y)$ is not constant. This situation can be problematic. Issues of responsive movement are covered at length in Buckland *et al.* (1993), for example pages 18–19 and pages 31–34.

3.2 Model-based analysis when the detection parameter is estimated

We now examine the asymptotic properties of $\hat{\Delta}$ when the detection parameter θ is estimated. Once again we restrict attention to a single strip of width w centred at $u=0$. There are N objects in the strip, at positions $y_1,\ldots,y_N$. The sightings data are the object positions y of detected objects. We use θ for the true value of the detection parameter.

We define the function $\operatorname{crit}(t)$ as

$$\operatorname{crit}(t) = \sum_{j=1}^{N} \log\left(\frac{g(|y_j|,t)}{\int_0^w g(r,t)\,dr}\right) I(\text{object } j \text{ detected})\,,$$

where $I(\cdot)$ represents the usual indicator. If $k(-y)+k(y) = \text{constant} = w^{-1}$, and sightings are independent, then $\operatorname{crit}(t)$ is the log-likelihood of the sightings data. The argument t denotes the detection parameter. As $k(\cdot)$ is unknown, we use $\operatorname{crit}(t)$ as the log-likelihood for all object distributions.

Now as $N \to \infty$, $\mathrm{crit}(t) \to E_y(\mathrm{crit}(t))$ by the strong law of large numbers. The detection of object j involves the true detection parameter θ, so we obtain

$$E_y(\mathrm{crit}(t)) = N \int_{-w}^{w} g(|y|,\theta)\, \log\{g(|y|,t)\}\, k(y)\, dy - N \left(\int_{-w}^{w} g(|y|,\theta) k(y)\, dy \right) \log \left(\int_0^w g(r,t)\, dr \right).$$

If $k(-y) + k(y) = \text{constant} = w^{-1}$, this becomes

$$E_y(\mathrm{crit}(t)) = \frac{N}{w} \int_0^w g(y,\theta)\, \log\{g(y,t)\}\, dy - \frac{N}{w} \left(\int_0^w g(y,\theta) dy \right) \log \left(\int_0^w g(r,t)\, dr \right).$$

Differentiating with respect to t, we find that $\frac{\partial}{\partial t} E_y(\mathrm{crit}(t)) \Big|_{t=\theta} = 0$. Thus θ_0, the maximum of the asymptotic likelihood function, is equal to the true parameter θ. That is, for any detection function g such that $g(-r,t) \equiv g(r,t)$, and for any object distribution k such that $k(-y) + k(y) = \text{constant}$, we have

$$\theta_0 = \arg \max_t E_y(\mathrm{crit}(t)) = \theta\,.$$

Expanding the estimator $\hat{\Delta}(\hat{\theta})$ about $\theta_0 = \theta$, the leading term in $E_y(\hat{\Delta}(\hat{\theta}))$ is $E_y(\hat{\Delta}(\theta)) = \Delta$. It follows that the estimator $\hat{\Delta}$ is asymptotically unbiased when $k(-y) + k(y)$ is constant.

If $k(-y) + k(y)$ is not constant, notably when there is responsive movement, the estimator $\hat{\Delta}$ is subject to the same problems as it is when θ is known. When responsive movement is severe, practitioners should avoid using the uncorrected estimator $\hat{\Delta}$ and consider alternative approaches such as those suggested by Buckland and Turnock (1992).

3.3 Completely model-based approach

Suppose we again consider a large area A of which only a small region is sampled. We have so far considered a design-based approach to density estimation in A, by randomly selecting transects u_i ($i = 1, \ldots, m$) of length ℓ. The density estimator for A is $\hat{\Delta} = \sum_{i=1}^{m} n_i/(2\ell w m \bar{g}_{\hat{\theta}})$, where n_i is the number of objects detected from transect i. Here we investigate an alternative approach, suitable when there are large-scale trends in density across A. We fix the transect positions, and explicitly model the density gradient. This can be thought of as a completely model-based approach to distance sampling. We present a simple simulation indicating that the approach can improve precision when the density gradient across the region is substantial.

For the simulation we use $A = [0,1] \times [0,1]$. We sample from $m = 10$ transects at positions $u = 0.05,\ 0.15, \ldots, 0.95$, each with length $\ell = 1$. The strip width is $w = 0.005$, so 10% of the region is sampled, and for the true values of θ we use $\theta = 0.001,\ 0.002,\ 0.004$, and 0.006. The total number of objects in the region is $N = 10000$ to ensure reasonable sample sizes for each transect, at least in the more regular object configurations. Object configurations considered are $y \sim \text{Beta}(1,\ 1)$, $\text{Beta}(3,\ 3)$, $\text{Beta}(5,\ 2)$, and $\text{Beta}(5,\ 1)$; we omit the configurations $\text{Beta}(5,\ 0.5)$ and $\text{Beta}(5,\ 0.25)$ because these are so extreme that the upper limit of the search region ($u + w = 0.955$) still misses half or more of the objects.

We estimate the detection parameter θ by pooling sightings data from all ten search strips. Density is estimated separately for each search strip, giving point estimates $\hat{\Delta}(u_i) = n_i/(2w\ell\bar{g}_{\hat{\theta}})$ where $u_i = 0.05, 0.15, \ldots, 0.95$. The full density gradient $\hat{\Delta}(u)$ for $u \in [0,1]$ is obtained by fitting an interpolating cubic spline to the ten point estimates, extrapolating the output from $u = 0.05$ to $u = 0$ and from $u = 0.95$ to $u = 1$ by straight lines. The estimate of overall density $\hat{\Delta}$ is given by $\hat{\Delta} = \int_{u=0}^{1} \hat{\Delta}(u)\, du$. The true overall density is $\Delta = 10000$. The method requires w to be small enough for the density in the search strip $[u - w, u + w]$ to be a good estimate of the instantaneous density $\Delta(u)$.

Typical outputs for the four object configurations and $\theta = 0.002$ are shown in Figure 3. In each case, positions of 10000 objects are generated from $y \sim \text{Beta}(\beta_1, \beta_2)$, and sightings are generated according to $g(|y - u_i|, \theta) = \exp\{-(y - u_i)^2/(2\theta^2)\}\, I(|y - u_i| \leq w)$ for $i = 1, \ldots, 10$. The interpolating spline method is successful when there is a clear density gradient, as in parts (b), (c), and (d). In part (a), the true density is uniform across the strip and the patterns in the spline curve are spurious. Linear extrapolation to $u = 0$ and $u = 1$ can also be poor when the true density is uniform.

Table 2 summarizes results from 1000 simulations for each combination of θ and (β_1, β_2). A separate y vector is generated for each simulation. The table gives the mean and standard deviation of the 1000 estimates of $\hat{\theta}$ and $\hat{\Delta}/\Delta$ after the interpolating spline has been fitted. For comparison, the corresponding results from the design-based approach are included, under a linear design on $[0,1]$. The design-based density estimate is $\hat{\Delta} = \sum_{i=1}^{10} n_i/(20w\ell\bar{g}_{\hat{\theta}})$. The choice of linear design is intentional because the circular correction is generally unnecessary for realistic search strip dimensions.

The results suggest that both the completely model-based approach and the design-based approach give unbiased estimates of $\hat{\Delta}$, but the model-based results have much lower variance for the non-uniform object configurations. Both methods allow θ to be estimated with similar precision and no bias, except for $\theta = 0.006$ when high detectability can lead to non-identifiability of θ. The bias in $\hat{\theta}$ does not cause bias in $\hat{\Delta}$ when this occurs.

We also investigated a more sophisticated model-based approach, in which a linear density gradient was fitted within each search strip. The log-likelihood is:

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} \log\left\{g(|y_{ij}-u_i|,\theta)\left(\frac{1}{2w}+\gamma_i(y_{ij}-u_i)\right)\right\} - \sum_{i=1}^{m} n_i \log\left\{\int_{u_i-w}^{u_i+w} g(|r-u_i|,\theta)\left(\frac{1}{2w}+\gamma_i(r-u_i)\right)dr\right\},$$

where parameters $\gamma_1, \ldots, \gamma_m$ govern the slope of the linear density gradient in the search strips $1, \ldots, m$, and y_{ij} is the jth sighting from transect i. No improvement was found in either $\hat{\Delta}$ or $\hat{\theta}$ due to including the γ parameters, even after increasing the strip width to $w = 0.05$ with $m = 5$.

We conclude that a completely model-based approach with fixed transects, even with this simplistic implementation, can lead to better precision than design-based placement of transects. The use of distance sampling estimates to construct spatial models of object distribution is considered by Hedley (2000).

4 Variance estimators

Returning to the usual distance sampling framework, we now look at estimating the variance of the estimator $\hat{\Delta}$. The customary analytic estimator for $\mathrm{var}(\hat{\Delta})$ is based on the delta method (Seber, 1982, pages 7–9). Following Barry and Welsh, we summarize this as

$$\widehat{\mathrm{var}}(\hat{\Delta}) = \hat{\Delta}^2\left(\frac{\widehat{\mathrm{var}}(n)}{n^2} + \frac{\widehat{\mathrm{var}}(\hat{f}_0)}{\hat{f}_0^{\,2}}\right),$$

where n is the total number of sightings and $\hat{f}_0 = 1/(w\bar{g}_{\hat{\theta}})$. Use of the delta method is justified by Buckland *et al.* (1993), pages 53–54, and is valid if $E(\hat{f}_0 \,|\, n) = f_0 = 1/(w\bar{g}_\theta)$. Alternatively, variance can be estimated by bootstrap methods (Buckland *et al.*, 1993, page 94), which render this requirement unnecessary. If the delta method is used, however, we need viable estimators for $\mathrm{var}(n)$ and $\mathrm{var}(\hat{f}_0)$. Here we restrict attention to these estimators.

The recommended estimator for $\mathrm{var}(\hat{f}_0)$ is obtained by the usual Hessian method, and will be valid as long as the sightings are independent and the p.d.f. of object distribution within each search strip satisfies $k(-y) + k(y) = \text{constant}$. This second condition ensures that $\text{crit}(t)$ is the correct form for the log-likelihood function, and we consider it to be a reasonable approximation to reality for most real surveys in the absence of responsive movement.

The estimator recommended for $\text{var}(n)$ by Buckland *et al.* (1993), pages 88 and 90, is

$$\hat{V}_n = \begin{cases} qn & \text{when } m = 1, \\ \dfrac{L}{m-1} \displaystyle\sum_{i=1}^{m} \ell_i \left(\frac{n_i}{\ell_i} - \frac{n}{L} \right)^2 & \text{when } m > 1. \end{cases}$$

Here, m is the number of transects surveyed, ℓ_i is the length of transect i, and $L = \sum_{i=1}^{m} \ell_i$ is the total transect length. When $m = 1$, the multiplier q is either $q = 1$ or $q = 2$ according to the user's choice, although in practice we recommend that surveys should be designed with no fewer than ten transects. Before examining the properties of $\hat{V}_n$, we establish the framework and context of its derivation.

Suppose we have a large area A to be surveyed by m transects. The number of sightings, n, depends upon the placement of transect lines, the random detection of objects within the search strips, and the object density within the strips. An estimator for $\text{var}(n)$ should take into account all these sources of variability. We should not condition on either the positions of the transect lines or the positions of the objects, as this would remove some of the variability we aim to quantify. Focusing on either the model-based or the design-based framework alone would therefore be inappropriate for $\widehat{\text{var}}(n)$, so instead we consider an unconditional framework.

The derivations that follow apply to any object distribution and any orientation of transects. Let $\boldsymbol{x} = (x, y)$ be a point in A, and suppose that search strip i has fixed length ℓ_i, width w, and is centred on point $\boldsymbol{x}$. Let $\Delta(\boldsymbol{x}, \ell_i)$ be the average object density within the search strip, and n_i be the number of objects detected in the strip. Now $n_i \sim \text{Binomial}\left(N, \frac{2w\ell_i\bar{g}_\theta\Delta(\boldsymbol{x},\ell_i)}{N}\right)$. For large N and a small search strip area $2w\ell_i$, this becomes $n_i \mathrel{\dot{\sim}} \text{Poisson}(2w\ell_i\bar{g}_\theta\Delta(\boldsymbol{x}, \ell_i))$. If the value of $2w\ell_i\bar{g}_\theta\Delta(\boldsymbol{x}, \ell_i)$ is not sufficiently small, this Poisson approximation inflates the variance of n_i.

Under the Poisson approximation, $E(n_i \mid \boldsymbol{x}) \simeq \text{var}(n_i \mid \boldsymbol{x}) \simeq 2w\ell_i\bar{g}_\theta\Delta(\boldsymbol{x}, \ell_i)$. This gives

$$E(n_i) = E_{\boldsymbol{x}} \left\{ E(n_i \mid \boldsymbol{x}) \right\} \simeq 2w\ell_i\bar{g}_\theta\Delta \,, \tag{1}$$

$$\text{var}(n_i) = E_{\boldsymbol{x}} \left\{ \text{var}(n_i \mid \boldsymbol{x}) \right\} + \text{var}_{\boldsymbol{x}} \left\{ E(n_i \mid \boldsymbol{x}) \right\} \simeq 2w\ell_i\bar{g}_\theta\Delta + 4w^2\ell_i^2\bar{g}_\theta^2\text{var}_{\boldsymbol{x}}\left(\Delta(\boldsymbol{x}, \ell_i)\right) . \tag{2}$$

The rationale behind the estimator $\hat{V}_n$ now becomes clear. If objects are uniformly distributed over the region A, then $\text{var}_{\boldsymbol{x}}(\Delta(\boldsymbol{x}, \ell_i)) = 0$. This means that $\text{var}(n_i) \simeq E(n_i)$. When $m = 1$ and we have no better way of estimating variance, we can therefore put $\widehat{\text{var}}(n) = n$. This explains the case $q = 1$ for $\hat{V}_n$. The case $q = 2$ was suggested by Burnham *et al.* (1980), page 55, to accommodate super-Poisson variation. It is unlikely that objects will be exactly uniformly distributed, so using $q = 2$ provides some leeway.

When $m > 1$, we can obtain better variance estimates using the full sample of transects. For uniformly distributed objects, we can write $E(n_i) = \beta\ell_i$ and $\text{var}(n_i) = \sigma^2\ell_i$, where β and σ^2 are unknown, as from equations (1) and (2). In fact, $\beta = \sigma^2$ when objects are exactly uniform, but by estimating β and σ^2 separately we enhance the performance of $\hat{V}_n$ when objects are not uniform.

Because $n = \sum_{i=1}^m n_i$ and $L = \sum_{i=1}^m \ell_i$, we have $E(n) = \beta L$, so we estimate β by n/L. Furthermore, in the uniform case, $\sigma^2 = \text{var}(n_i/\sqrt{\ell_i})$ for all i, so the random variables $\frac{n_i}{\sqrt{\ell_i}} - \beta\sqrt{\ell_i}$ $(i = 1, \ldots, m)$ each have mean 0 and variance σ^2. This suggests the estimator

$$\hat{\sigma}^2 \quad = \quad \frac{1}{m-1}\sum_{i=1}^m \left(\frac{n_i}{\sqrt{\ell_i}} - \frac{n\sqrt{\ell_i}}{L}\right)^2 \quad = \quad \frac{1}{m-1}\sum_{i=1}^m \ell_i\left(\frac{n_i}{\ell_i} - \frac{n}{L}\right)^2 .$$

Hence $\hat{V}_n$ is derived as

$$\hat{V}_n \quad = \quad \hat{\sigma}^2 L \quad = \quad \frac{L}{m-1}\sum_{i=1}^m \ell_i\left(\frac{n_i}{\ell_i} - \frac{n}{L}\right)^2 .$$

We expect $\hat{V}_n$ to perform well when objects are uniformly distributed across A, but we need to investigate its performance for other object configurations. For any object distribution, we can rewrite equations (1) and (2) as $E(n_i) = \beta\ell_i$ and $\text{var}(n_i) = \beta\ell_i + \gamma\ell_i^2$. Let $S = \sum_{i=1}^m \ell_i^2$. Under the Poisson approximation, the true variance of n is $\text{var}(n) = \sum_{i=1}^m \text{var}(n_i) = \beta L + \gamma S$.

Expanding the expression for $\hat{V}_n$, and using $E(n_i) = \beta\ell_i$, $\text{var}(n_i) = \beta\ell_i + \gamma\ell_i^2$, we obtain

$$E(\hat{V}_n) = \beta L + \gamma\left(\frac{L^2 - S}{m-1}\right) .$$

This shows that $\hat{V}_n$ is not unbiased for $\text{var}(n)$, although the bias is rarely severe. The expression $(L^2 - S)/(m-1)$ can be rewritten as $\sum_{i=1}^m \ell_i(\overline{\ell_{-i}})$, where $(\overline{\ell_{-i}})$ denotes the mean of all line lengths excluding ℓ_i. Comparing $E(\hat{V}_n) = \beta L + \gamma\sum_{i=1}^m \ell_i(\overline{\ell_{-i}})$ with

$\mathrm{var}(n) = \beta L + \gamma \sum_{i=1}^{m} \ell_i^2$, we see that the bias in $\hat{V}_n$ is small if line lengths are not too disparate, and zero when line lengths are equal.

The expression above suggests a new, unbiased estimator for $\mathrm{var}(n)$ under the Poisson approximation. Let

$$\hat{W}_n = n + (\hat{V}_n - n) \left(\frac{m-1}{L^2 - S} \right) S\,.$$

Then $E(\hat{W}_n) = \beta L + \gamma S$. Under any object configuration, $\hat{W}_n$ is unbiased for $\mathrm{var}(n)$ if the Poisson approximation holds. In practice, the Poisson approximation is not exact so we expect that $\hat{W}_n$ will overestimate $\mathrm{var}(n)$ by some amount.

We investigate the performance of $\hat{V}_n$ and $\hat{W}_n$ by a simulation exercise. As before, we use $A = [0,1] \times [0,1]$. We select a set of $m = 10$ transects, with start-points uniformly distributed on A. For convenience, all transects are parallel with the x-axis. Each transect is assigned a length between 0.1 and 1.0, and the width of each search strip is $w = 0.02$. On average, this means that about 18% of A is surveyed.

We use a circular design, both horizontally and vertically, to ensure that the full length ℓ_i of each search strip is surveyed. We generate $N = 1000$ objects at positions governed by Beta distributions, so the x-coordinates satisfy $x \sim \mathrm{Beta}(\beta_1, \beta_2)$ and the y-coordinates satisfy $y \sim \mathrm{Beta}(\beta_3, \beta_4)$. Sightings are generated for each transect using the half-normal detection function with parameter $\theta = 0.01$. With $w = 0.02$, this means that detectability is moderate ($\bar{g}_\theta = 0.6$). The 10 transects are enumerated and the total number n of detections is found. The values of $\hat{W}_n$ and $\hat{V}_n$ are calculated. Keeping the same line lengths and object positions, but with different start-points for the transects, we repeat the experiment a total of 10000 times. We can estimate the true value of $\mathrm{var}(n)$ by the sample variance of the 10000 observations. We can also evaluate the bias and variability of the estimators $\hat{W}_n$ and $\hat{V}_n$ by taking the mean and standard deviation of their 10000 estimates.

The results are presented in Table 3. As expected, $\hat{W}_n$ appears to overestimate variance, while $\hat{V}_n$ underestimates. On the whole it is preferable to overestimate variance than to underestimate, but both estimators are reasonable and $\hat{V}_n$ has slightly better precision than $\hat{W}_n$. Note that the uniform object configuration has $\mathrm{mean}(n) \simeq \mathrm{var}(n)$, suggesting that the Poisson approximation is good in this case.

Barry and Welsh, page 41, recommend a further estimator for $\mathrm{var}(n)$, namely ms_n^2 which is correctly written as $m(m-1)^{-1} \sum_{i=1}^{m} \{(n_i - n/m)^2\}$. Their assertion that this estimator is design-unbiased for $\mathrm{var}(n \,|\, \boldsymbol{y})$ is mistaken, because it relies upon their equation (13) which overlooks the possibility of unequal line lengths.

5 Concluding remarks

Our results have shown that the concerns of Barry and Welsh about the validity of distance sampling estimators are largely unfounded. As long as the survey meets certain requirements, the methods perform well. The first requirement is a representative design, which we define as one under which all objects have equal probability of detection. This can be achieved either through a circular design, in which search strips outlying the study area on one side are wrapped round to the other side, or by a buffer zone design, which enables an easy implementation of the circular design by sampling a region outside of A. In practical distance sampling, the search strips are generally so narrow compared with the width of A that there is negligible bias if the design is not exactly representative at the edges of A.

The second requirement is that $k(r)$, the p.d.f. of object distances from the transect line, should have $k(-r) + k(r)$ constant for $-w \leq r \leq w$. When this is the case, $\hat{\theta} \to \theta$ as $N \to \infty$, and $\hat{\Delta}$ is asymptotically unbiased. This condition is likely to be approximately fulfilled for search strips that are narrow compared with A, as long as external effects such as responsive movement are absent. More advanced methods may be required in the event of responsive movement.

Our results indicate that a completely model-based approach to density estimation can give better precision in $\hat{\Delta}$ than a design-based approach. This is an active research field. For variance estimation, we favour bootstrap techniques, because these free us from assumptions of independence between the estimated components of $\hat{\Delta}$. Nonetheless, we have shown that the estimator $\hat{V}_n$ for $\mathrm{var}(n)$ performs with reasonably low bias even for highly non-uniform object distributions. We have also derived an alternative estimator, $\hat{W}_n$. Neither estimator is exactly unbiased. $\hat{W}_n$ gives higher estimates and has slightly poorer precision than $\hat{V}_n$.

Barry and Welsh, page 52, assert without justification that the estimator

$$\hat{\Delta}_R = \frac{1}{2\ell w m} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{1}{g(|u_i - y_j|, \theta)}$$

is preferable to the standard distance sampling estimator $\hat{\Delta} = n/(2\ell w m \bar{g}_\theta)$. In fact, if detectability is universally high, the performance of $\hat{\Delta}_R$ is very similar to that of $\hat{\Delta}$. However, if any detection probabilities $g(|u_i - y_j|, \theta)$ are close to zero, $\hat{\Delta}_R$ becomes unstable and has much lower precision than $\hat{\Delta}$. For this reason we recommend that $\hat{\Delta}_R$ is avoided except in special circumstances.

Finally, we note that distance sampling is a general framework that is often implemented in difficult field situations. Compromises to the theoretical design requirements are inevitable, yet we have seen that the methods are robust. The design parameters investigated by Barry and Welsh are quite exceptional in practice. Most real surveys involve at least 20 transect lines, and have strip widths w of far less than 5% of the width of the survey region. The problems that Barry and Welsh highlight under the linear design only become consequential under these somewhat unrealistic search strip dimensions. For these reasons, we consider that their findings are not relevant in the majority of distance sampling applications.

References

Barry, S. C. and Welsh, A. H. (2001) Distance Sampling Methodology. *J. R. Statist. Soc.* B, **63**, 31–53.

Borchers, D. L., Zucchini, W. and Fewster, R. M. (1998) Mark-recapture models for line transect surveys. *Biometrics*, **54**, 1207–1220.

Buckland, S. T. and Turnock, B. J. (1992) A robust line transect method. *Biometrics*, **48**, 901–909.

Buckland, S. T., Anderson, D. R., Burnham, K. P. and Laake, J. L. (1993) *Distance Sampling.* London: Chapman and Hall.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L. and Thomas, L. (2001) *Introduction to Distance Sampling.* Oxford: Oxford University Press.

Burnham, K. P., Anderson, D. R. and Laake, J. L. (1980) Estimation of density from line transect sampling of biological populations. *Wildlife Monograph Number 72.*

Hedley, S. L. (2000) Modelling heterogeneity in cetacean surveys. Ph.D. thesis, University of St Andrews, UK.

Seber, G. A. F. (1982) *The estimation of animal abundance and related parameters.* London : Charles Griffin.

θ	$E_y(\hat{\Delta})/\Delta$ *for the following data configurations:*			
	1. Uniform	2. Linear	3. $\cap$–Quadratic	4. $\cup$–Quadratic
0.25	1	1	1.16	0.73
0.50	1	1	1.08	0.86
1	1	1	1.03	0.96
2	1	1	1.01	0.99
100	1	1	1.00	1.00

Table 1: $E_y(\hat{\Delta})/\Delta$ when the half-normal detection function with parameter θ is known. Results are given for strip width $w = 1$ and the data configurations shown in Figure 2.

		model-based				*design-based*			
		$\hat{\Delta}/\Delta$		$\hat{\theta}$		$\hat{\Delta}/\Delta$		$\hat{\theta}$	
Data	θ	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
(1, 1)	0.001	1.00	0.078	0.0010	4.5×10^{-5}	1.01	0.078	0.0010	4.3×10^{-5}
	0.002	1.00	0.055	0.0020	7.6×10^{-5}	1.01	0.058	0.0020	8.1×10^{-5}
	0.004	1.00	0.049	0.0040	3.6×10^{-4}	1.00	0.051	0.0040	3.4×10^{-4}
	0.006	1.00	0.055	0.0143	2.0×10^{-1}	1.00	0.054	0.0155	2.0×10^{-1}
(3, 3)	0.001	1.01	0.076	0.0010	4.5×10^{-5}	1.00	0.217	0.0010	4.4×10^{-5}
	0.002	1.01	0.056	0.0020	7.7×10^{-5}	1.00	0.212	0.0020	7.5×10^{-5}
	0.004	1.00	0.049	0.0040	3.6×10^{-4}	1.00	0.211	0.0040	3.6×10^{-4}
	0.006	1.01	0.059	0.0279	3.5×10^{-1}	1.00	0.212	0.0356	4.0×10^{-1}
(5, 2)	0.001	1.01	0.075	0.0010	4.5×10^{-5}	1.01	0.296	0.0010	4.4×10^{-5}
	0.002	1.01	0.055	0.0020	8.0×10^{-5}	1.01	0.289	0.0020	8.0×10^{-5}
	0.004	1.01	0.050	0.0040	3.6×10^{-4}	1.01	0.289	0.0040	4.0×10^{-4}
	0.006	1.01	0.052	0.0239	2.9×10^{-1}	1.01	0.291	0.0311	3.5×10^{-1}
(5, 1)	0.001	1.00	0.077	0.0010	4.5×10^{-5}	1.00	0.416	0.0010	5.1×10^{-5}
	0.002	1.00	0.054	0.0020	7.8×10^{-5}	1.00	0.415	0.0020	8.8×10^{-5}
	0.004	1.00	0.050	0.0041	3.6×10^{-4}	1.00	0.413	0.0040	4.3×10^{-4}
	0.006	1.00	0.060	0.0273	3.5×10^{-1}	1.00	0.415	0.0149	6.9×10^{-2}

Table 2: Results from model-based and design-based approaches to density estimation when $\Delta = 10000$, $m = 10$, and $w = 0.005$. For the model-based approach, $\hat{\Delta} = \int_{u=0}^{1} \hat{\Delta}(u)\, du$ and $\hat{\Delta}(u)$ is given by an interpolating cubic spline fitted to the 10 point-wise density estimates. In the design-based approach, $\hat{\Delta} = \sum_{i=1}^{m} n_i/(2mw\ell\bar{g}_{\hat{\theta}})$. Each entry summarizes the results from 1000 different $\boldsymbol{y}$ vectors.

Data configuration x	y	mean(n)	var(n)	mean($\hat{W}_n$)	mean($\hat{V}_n$)	sd($\hat{W}_n$)	sd($\hat{V}_n$)
(1, 1)	(1, 1)	122	122	120	120	65	55
(3, 3)	(3, 3)	117	973	1120	869	516	389
(5, 5)	(5, 5)	125	2040	2340	1880	1290	1030
(1, 3)	(5, 2)	190	3520	3560	3350	1370	1290
(1, 3)	(5, 0.5)	109	8190	8770	6620	14200	10700
(5, 0.5)	(5, 0.5)	145	14300	16400	13900	30100	25500

Table 3: Evaluation of the estimators $\hat{W}_n$ and $\hat{V}_n$. Objects are generated with x and y coordinates following the Beta distributions in columns 1 and 2 respectively. The values of mean(n) and var(n) are the sample mean and variance from 10000 simulations. The mean and standard deviation of the estimators $\hat{W}_n$ and $\hat{V}_n$ are obtained from the corresponding 10000 point estimates. Results are reported to three significant figures.

Captions to figures

Figure 1

Selection of object configurations used in Tables 4 and 6 of Barry and Welsh. In each case, the area A is shown with overall density 100, and the sampled areas are shaded. In (a), $y \sim \text{Beta}(3,3)$, and the density in the sampled region is greater than 100: e.g. $\Delta = 186$ within the shaded strip ($w = 0.05$). In (b), (c), and (d), $y \sim \text{Beta}(5, 0.25)$, and the sampled region for both $m = 1$ (part (b)) and $m = 5$ (part (c)) has object density much less than 100. Part (d) shows a transect parallel to the density gradient, following recommended practice, and giving $E_y(\hat{\Delta} \,|\, u)/\Delta = 1$.

Figure 2

Object p.d.f.s, $k(y)$, and detection functions, $g(y, \theta)$, for the experiments in Table 1. The functions $k(y)$ are 1. Uniform; 2. Linear; 3. $\cap$–Quadratic; 4. $\cup$–Quadratic. The detection functions are half-normal with parameter θ as shown.

Figure 3

Density functions from a completely model-based approach. The solid lines represent the true function $\Delta(u)$. The dashed lines give the estimated function $\hat{\Delta}(u)$, obtained by fitting an interpolating cubic spline to the values of $\hat{\Delta}(u)$ obtained at ten different transect positions (marked with dots). In each case, $N = 10000$, $w = 0.005$, and $\theta = 0.002$. The results are (a) $\beta_1 = \beta_2 = 1$, $\hat{\Delta}/\Delta = 1.00$, $\hat{\theta} = 0.0020$; (b) $\beta_1 = \beta_2 = 3$, $\hat{\Delta}/\Delta = 1.02$, $\hat{\theta} = 0.0020$; (c) $\beta_1 = 5, \beta_2 = 2$, $\hat{\Delta}/\Delta = 1.01$, $\hat{\theta} = 0.0020$; (d) $\beta_1 = 5, \beta_2 = 1$, $\hat{\Delta}/\Delta = 1.00$, $\hat{\theta} = 0.0021$.

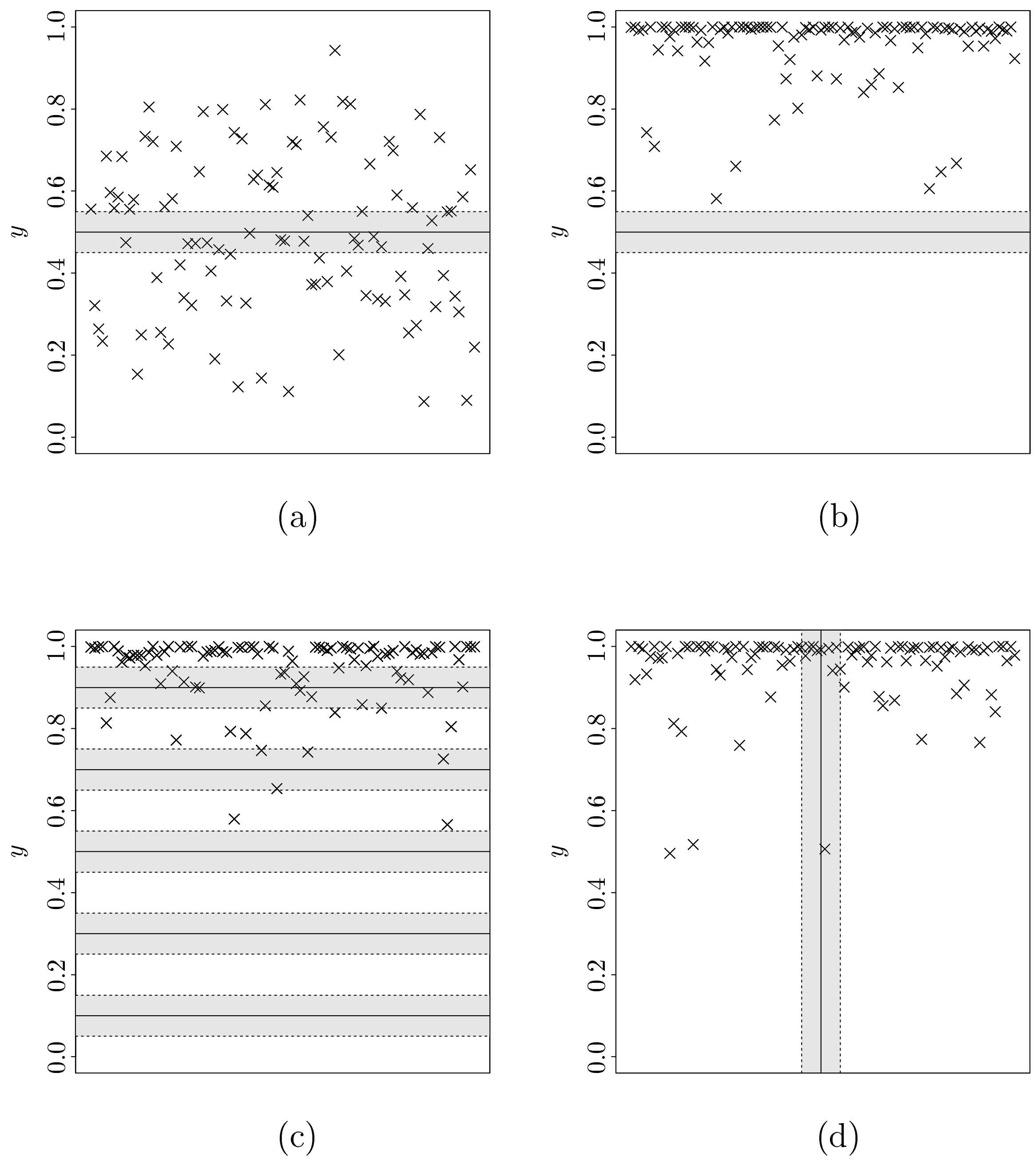

Figure 1

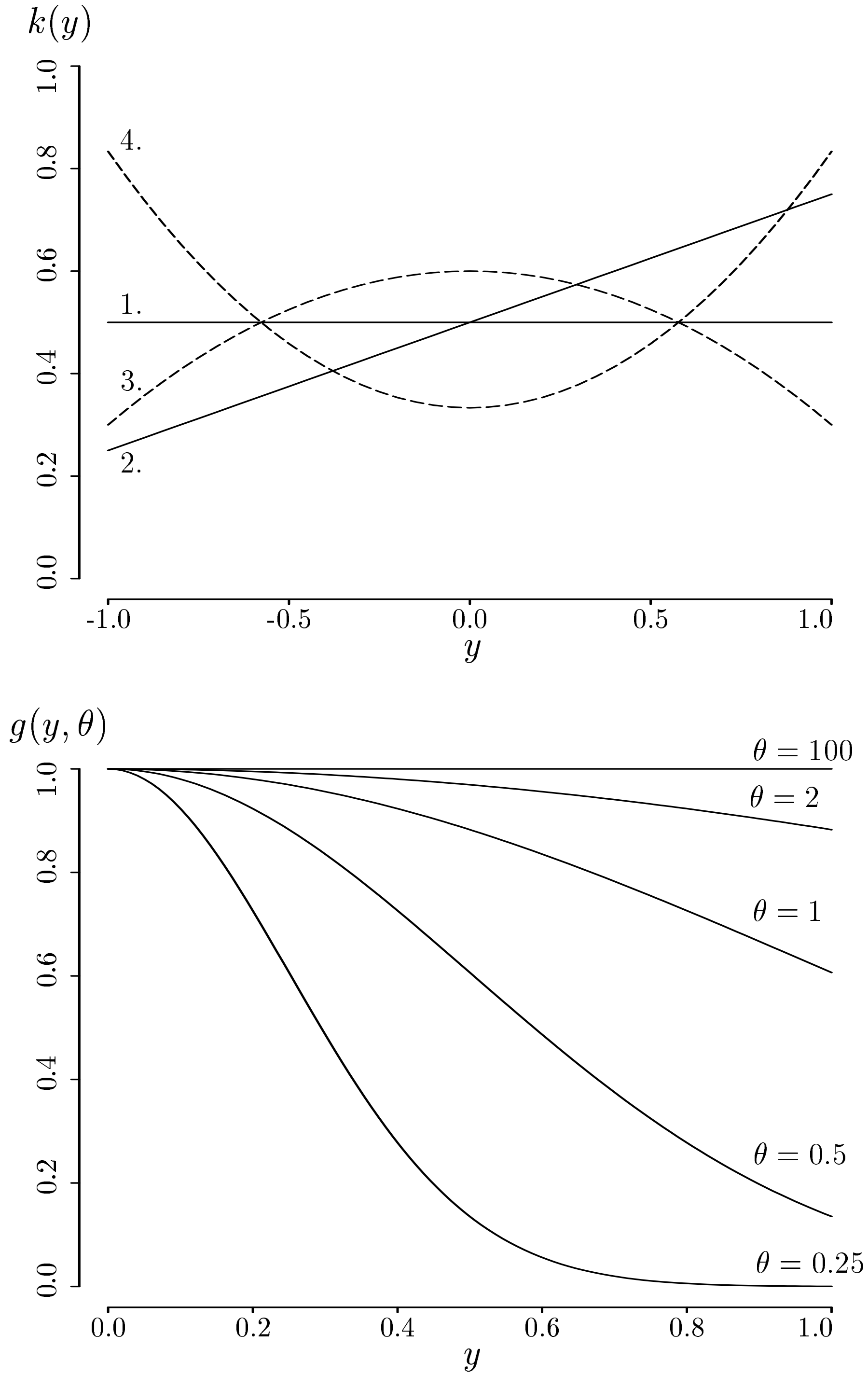
$k(y)$
1.0
0.8
0.6
0.4
0.2
0.0
4.
1.
3.
2.
-1.0
-0.5
0.0
0.5
1.0
y
$g(y,\theta)$
1.0
0.8
0.6
0.4
0.2
0.0
$\theta = 100$
$\theta = 2$
$\theta = 1$
$\theta = 0.5$
$\theta = 0.25$
0.0
0.2
0.4
0.6
0.8
1.0
y

Figure 2

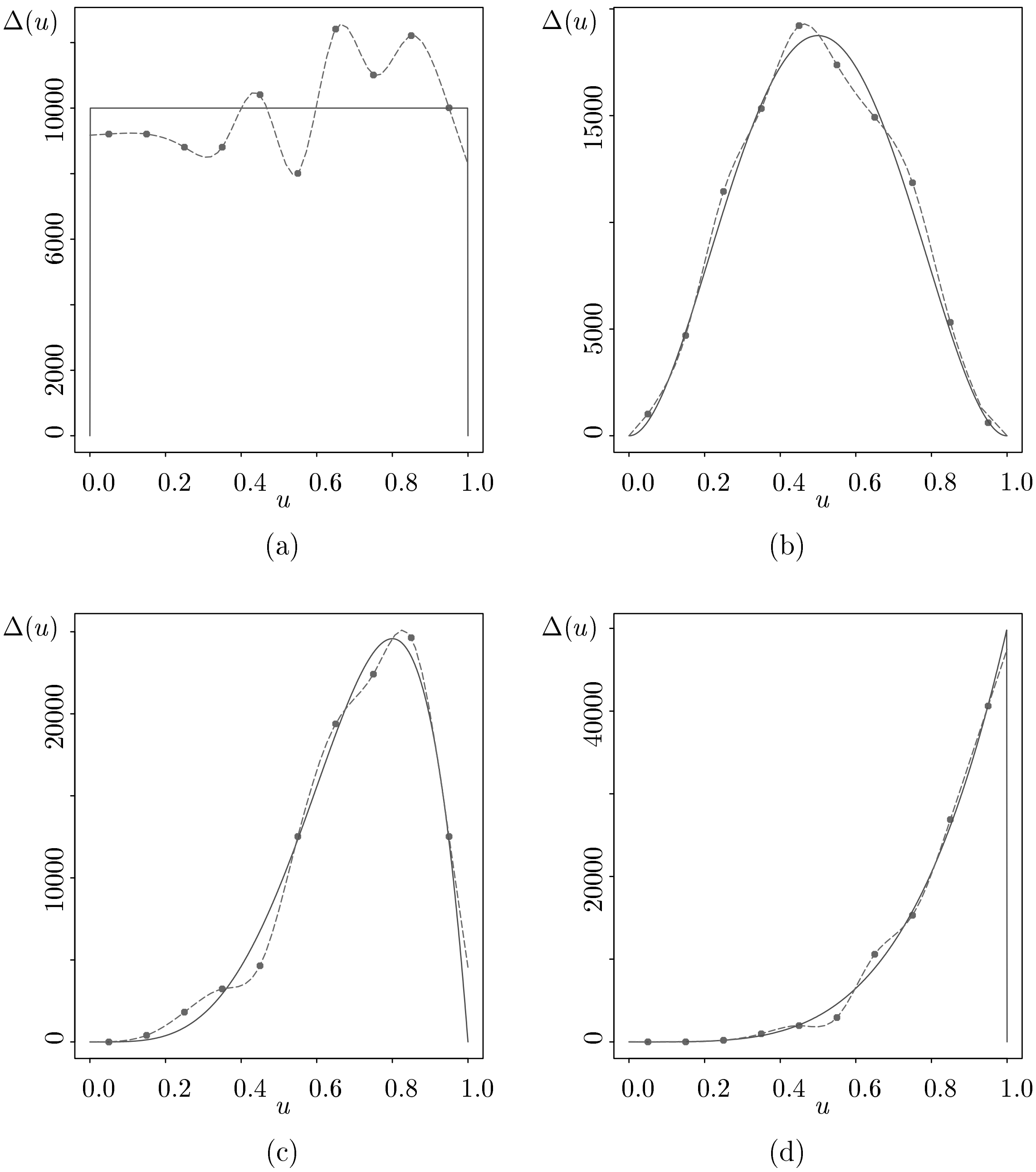
Δ(u)
0
2000
6000
10000
0.0
0.2
0.4
0.6
0.8
1.0
u
(a)
Δ(u)
0
5000
15000
0.0
0.2
0.4
0.6
0.8
1.0
u
(b)
Δ(u)
0
10000
20000
0.0
0.2
0.4
0.6
0.8
1.0
u
(c)
Δ(u)
0
20000
40000
0.0
0.2
0.4
0.6
0.8
1.0
u
(d)

Figure 3