BIOMETRICS 48, 1195–1199
December 1992

An Adaptive Procedure for Sampling Animal Populations

Steven K. Thompson,[1] Fred L. Ramsey,[2] and George A. F. Seber[3]

[1] Department of Statistics, Pennsylvania State University, University Park, Pennsylvania 16802, U.S.A.

[2] Department of Statistics, Oregon State University, Corvallis, Oregon 97331, U.S.A.

[3] Department of Mathematics and Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand

SUMMARY

We consider an adaptive stratified sampling procedure for animal populations in which the sample size in a given stratum or primary unit depends on the observations obtained in previous strata or primary units. This method provides a more efficient way of sampling sparse but highly clustered populations. Two sampling approaches are considered. The first is a design unbiased strategy as the estimate of population density is unbiased through design-induced factors such as the random selection of sites in each stratum. The second is called a model unbiased strategy as the density estimate is unbiased under the model assumed for the spatial distribution of the population. The theory is demonstrated by reference to sampling from a shrimp population.

1. Introduction

Animal populations that are spread over an extensive area, and are highly clustered, present problems for the sampler. As the coefficient of variation of the number per unit area tends to be much greater than 1, large sample sizes are required to estimate the population density with a given precision. An example of such a population is that of krill in the ocean. Krill are an essential food for whales, and krill distribution will affect whale movements. Another population with which one of the authors (GAFS) was recently involved is the Chatham Rise orange roughy fishery in New Zealand. Orange roughy is a deep-sea fish that comes into shallower waters to spawn in large aggregations, usually around pinnacles rising from the ocean floor. An estimate of fish density is needed from survey methods so that fishing quotas can be determined. If a simple random sample of plots (net tows) is used for such a sparse but highly clustered population, then many of the tows will be empty and more tows will need to be taken to compensate for this. Clearly more efficient methods of sampling are required.

If the patterns of spatial variation in abundance are known a priori, conventional stratification can be used to increase precision for a given sample size by allocating higher sampling effort to strata with higher variability. With many natural populations, higher variability is associated with regions of higher abundance. When the patterns of abundance are not predictable before the survey, one approach is to use a two-phase sample in each stratum, letting the sample size at the second phase depend on the observed abundance or variability of the first-phase sample in that stratum. However, conventional estimators such as the stratified sample mean are subject to bias with such sampling procedures (Francis,

Key words: Adaptive sampling; Adaptive stratified sampling; Animal abundance estimation; Clustered populations; Design unbiased; Model unbiased; Rare species estimation.

1984; Kremers, Preprint No. 128, Institut für Mathematik, Universität Augsburg, 1987). A way of obtaining an unbiased estimate of the population mean with such a procedure was provided by Kremers using the Rao–Blackwell method, but the resulting estimator is computationally complex.

A different approach, described in this paper, bases the sample size in each primary unit or stratum other than the first on observations in previous primary units or strata. With this procedure, the sampling intensity may be increased in the vicinity of high observed abundance or decreased following observations of low abundance, yet conventional estimators remain unbiased. For example, in trawl surveys of shrimp conducted by the Alaska Department of Fish and Game, the net is towed for a measured distance along the ocean floor and shrimp density is estimated from the weight of shrimp caught and the area swept by the net, taking into account the width of the net and the distance towed. However, shrimp have distinct aggregation or schooling tendencies that can be represented statistically by exponential-shaped correlation functions of abundances by distance between tows. Although the aggregation tendency is well known from past data, it is not possible to predict in advance where within a survey area the shrimp will be concentrated. A common experience on surveys is for many consecutive tows to be made with zero or near-zero catches of shrimp. The catches then increase more or less steadily as an aggregation is approached. Two adaptive procedures that make use of this knowledge are now described.

The shrimp habitat in each survey area, such as a bay, is partitioned into a grid of primary units of approximately equal areas. One adaptive strategy is as follows. In the first primary unit, m_1 random half-mile tows are made. If an average of 50 pounds or less per mile are caught, then m_1 tows are carried out in the secondary primary unit. However, if more than 50 pounds per mile are caught, then m_2 tows ($m_2 > m_1$) are carried out in the second unit. If less than 50 pounds per mile are caught in the second unit, then m_1 tows are taken in the third unit; otherwise, m_2 are taken, and so on. A variation on this theme is to set $m_1 = 1$ and $m_2 = 2$ but with the m_2 tows merged into a single one-mile tow. These strategies are considered briefly in Section 4.

Before developing a theory for adaptive techniques like the above, we note that other adaptive strategies have been proposed by Francis (1984) and Thompson (1990, 1991a, 1991b). The importance of adaptive sampling has been stressed by Seber (1986) and Cormack (1988). The theory in this paper is based on Thompson and Ramsey (Technical Report 82, Department of Statistics, Oregon State University, 1983).

2. Design Unbiased Adaptive Sampling

The notation used is similar to that of Cochran (1977). A study region is partitioned into L large primary units (strata) of area A_h ($h = 1, 2, \ldots, L$), and primary unit h is further partitioned into N_h secondary units each of area a_h. The total number of secondary units is $N = \Sigma_h N_h$. Our objective is to estimate the density $D = Y/A$, where Y is the total number of animals in the region and A [$= \Sigma_h A_h$] is the area of the study region. A random sample of n_h secondary units is taken from primary unit h and y_{hj} is the number of animals counted on the jth unit of the sample ($j = 1, 2, \ldots, n_h$). Let $y_{h.} = \Sigma_j y_{hj}$ be the total number observed in primary unit h and let $\bar{y}_h = y_{h.}/n_h$. The observed density in primary unit h will be $\hat{D}_h = \bar{y}_h/a_h$, an estimate of $D_h = Y_h/A_h$.

The sampling design is as follows: Select at random a fixed number n_1 of secondary units within the first primary unit. Next choose a random sample of n_2 secondary units from the second primary unit, but with n_2 depending on $\hat{D}_1$. We continue in this fashion, selecting n_h secondary units at random from the primary unit h with n_h depending on the observed density $\hat{D}_{h-1}$ in the preceding primary unit. A natural estimate of D is

$$\hat{D} = \sum_{h=1}^{L} W_h \hat{D}_h, \tag{1}$$

where $W_h = A_h/A$. If c_h denotes "the set of all $\hat{D}_k$ for $k < h$," then, since c_h determines n_h, it follows from the theory of simple random sampling with fixed n_h that

$$\mathrm{E}[\hat{D}_h \mid c_h] = \mathrm{E}[\bar{y}_h \mid n_h]/a_h = D_h. \tag{2}$$

Taking expectations with respect to the condition, we see that $\hat{D}_h$, and therefore $\hat{D}$, is unconditionally unbiased. Also

$$\mathrm{var}[\hat{D}_h] = \mathrm{E}_{c_h}\{\mathrm{var}[\hat{D}_h \mid c_h]\} + \mathrm{var}_{c_h}\{\mathrm{E}[\hat{D}_h \mid c_h]\} \tag{3}$$

with the second term being zero. Now, for $h < i$,

$$\mathrm{E}[(\hat{D}_h - D_h)(\hat{D}_i - D_i) \mid c_i] = 0, \tag{4}$$

and this result holds unconditionally. Hence, by (4) and (3),

$$\begin{aligned}
\mathrm{var}[\hat{D}] &= \mathrm{E}[\{\Sigma_h W_h(\hat{D}_h - D_h)\}^2] \\
&= \Sigma_h W_h^2 \mathrm{E}_{c_h}\{\mathrm{var}[\hat{D}_h \mid c_h]\} && (5) \\
&= \Sigma_h W_h^2 \mathrm{E}_{n_h}\left[\frac{1}{a_h^2}\{\mathrm{var}[\bar{y}_h \mid n_h]\}\right]. && (6)
\end{aligned}$$

Again, by the theory of stratified random sampling with fixed sample sizes, and defining $s_h^2 = \Sigma_j(y_{hj} - \bar{y}_h)^2/(n_h - 1)$, we have

$$v_D = \Sigma_h W_h^2 \frac{s_h^2}{a_h^2 n_h}\left(1 - \frac{n_h}{N_h}\right) = \Sigma_h W_h^2 v_h, \quad \text{say}, \tag{7}$$

is an unbiased estimate of $\mathrm{var}[\hat{D}]$. Hence the standard formula for nonadaptive stratified sampling still applies to the adaptive situation. Furthermore, unbiasedness is a function of the design and does not depend on assumptions about the population, so that $\hat{D}$ is design unbiased.

We note that the above theory can be extended. If we now assume that the size a_h of the secondary unit in primary unit h also depends on c_h, then when a_h is conditionally fixed, (2) still holds and $\hat{D}$ is still unbiased. Also $\mathrm{var}[\hat{D}]$ is again given by (5). Equation (6) holds if the expectation is with respect to both a_h and n_h. With this change it then follows from (6) that

$$\mathrm{E}[v_D] = \Sigma_h W_h^2 \mathrm{E}_{n_h, a_h} \mathrm{E}[v_h \mid n_h, a_h] = \mathrm{var}[\hat{D}],$$

and v_D is still unbiased.

3. Model Unbiased Adaptive Sampling

In this case we initially assume the same setup of primary and secondary units as before. However, we now develop a model for the spatial distribution of the animals. Even though the population is clustered, we assume that the clustering is due to variations in density from primary unit to primary unit and not due to clustering within a primary unit. Such an assumption could hold when the animal patches are much larger than a primary unit. We therefore treat this model as an "apparent" rather than a "true" contagious model (Seber, 1982, p. 48).

Within primary unit h we assume that we have a spatial Poisson process with intensity (mean per unit area) λ_h: the λ_h are assumed to represent observations from some spatial process with mean λ. For two different primary units h and i, λ_h and λ_i are not necessarily independent. However, conditional on λ_h and λ_i, the y_{hj} and the primary unit total Y_h are independent of the y_{ik} and Y_i. Under this model Y and D are now random variables.

The sample size n_h of secondary units within primary unit h will again depend on the observed density in the preceding primary unit. However, we need not assume now that the secondary units are randomly selected. They may be systematically or purposively located as randomness is now imposed by the model. We use the same density estimate $\hat{D}$ given by (1).

Assuming the above model, $y_{h.}$ is conditionally Poisson with mean $n_h a_h \lambda_h$ so that $\mathrm{E}[\hat{D}_h \mid \lambda_h, n_h] = \lambda_h$, and $\mathrm{E}[\hat{D}] = \lambda$. Depending on whether we are estimating the parameter λ or the random variable D, we need to determine $\mathrm{var}[\hat{D}]$ or the mean squared error $\mathrm{E}[(\hat{D} - D)^2]$, respectively (Thompson and Ramsey, 1987). The latter would seem more appropriate here. We note that for $h \neq i$,

$$\mathrm{E}[(\hat{D}_h - D_h)(\hat{D}_i - D_i) \mid \lambda_h, \lambda_i] = 0, \tag{8}$$

so that the unconditional expectation is zero. Furthermore, it can be shown that, conditional on λ_h, the distribution of y_{hj} is independent of the condition c_h except for the dependence of its mean on n_h and a_h. Hence writing

$$\hat{D}_h - D_h = \frac{1}{a_h}\left\{y_{h.}\left(\frac{1}{n_h} - \frac{1}{N_h}\right) - \frac{Y_h - y_{h.}}{N_h}\right\},$$

where $y_{h.}$ and $Y_h - y_{h.}$ are conditionally independent Poisson variables with respective means (and variances) $n_h a_h \lambda_h$ and $(N_h - n_h)a_h\lambda_h$, we find that

$$\mathrm{E}[(\hat{D}_h - D_h)^2 \mid \lambda_h, c_h] = \mathrm{var}[\hat{D}_h - D_h \mid \lambda_h, c_h] = \frac{\lambda_h}{a_h n_h}\left(1 - \frac{n_h}{N_h}\right).$$

Therefore using (8),

$$\mathrm{E}[(\hat{D} - D)^2] = \Sigma_h W_h^2 \mathrm{E}[(\hat{D}_h - D_h)^2] = \Sigma_h W_h^2 \mathrm{E}_{c_h}\left\{\frac{1}{a_h n_h}\left(1 - \frac{n_h}{N_h}\right)\mathrm{E}[\lambda_h \mid c_h]\right\}. \tag{9}$$

Since for the Poisson distribution $\mathrm{E}[s_h^2 \mid \lambda_h, c_h] = \lambda_h a_h$, we find that v_D of (7) is once again a (model) unbiased estimate of $\mathrm{E}[(\hat{D} - D)^2]$. For the Poisson model we can use v_D', in which each s_h^2 is replaced by $\bar{y}_h$.

A slightly different approach is to regard the primary totals Y_h (and hence the D_h) as fixed. Then, under the above Poisson assumptions, the conditional distribution of $y_{h.}$ is binomial(Y_h, n_h/N_h) and it can be shown that

$$\mathrm{var}[\hat{D}] = \Sigma_h W_h^2 D_h \, \mathrm{E}_{c_h}\left[\frac{1}{a_h n_h}\left(1 - \frac{n_h}{N_h}\right)\right],$$

which is estimated unbiasedly by v_D'. Since the y_{hj} ($j = 1, 2, \ldots, n_h$) are conditionally multinomial, we find that v_D is still unbiased, as might be expected.

4. Shrimp Example

Without data and some experience with adaptive procedures, it is difficult at this stage to provide practical guidelines for the design of such experiments. The size of the primary unit, and the choice of m_1 and the ratio $r = m_2/m_1$ will depend on the length of time available for fishing and the degree of aggregation in the population. Intuitively, the greater the aggregation, the greater r needs to be to make up for the increase in the number of low catches. To find the best adaptive design for the shrimp example we need to minimize var$[\hat{D}]$ subject to a given expected total tow length. This problem is under investigation at present. Once a design has been chosen, then for the first adaptive strategy in Section 1, we can use $\hat{D}$ of equation (1) with $a_h = a$, the area swept by the net over a half-mile tow. A variance estimate is given by (7) with n_h either m_1 or m_2.

The other adaptive strategy suggested in Section 1, where the tows are either half-mile or one-mile, is much easier to carry out at sea because of the random allocation of tows required, but we can no longer use (7) as we have no replications for computing s_h^2. However, if we can make the strong assumption that the population is "locally" Poisson, as described in Section 3, then we can use a model unbiased design with v_D' as a variance estimate where each n_h is 1 or 2. If the Poisson model holds, then it does not matter that the two half-mile tows are added together to form a one-mile.

Acknowledgement

We are grateful to a referee for some helpful comments. This work was partially supported by National Science Foundation Grant DMS-9016708.

Résumé

Nous considérons une procédure adaptative d'échantillonnage stratifié pour des populations animales dans laquelle la taille de l'échantillon d'une strate donnée ou d'une unité primaire dépend des observations obtenues dans les strates précédentes ou des autres unités primaires. Cette méthode fournit un moyen plus efficace d'échantillonner des populations peu denses mais fortement agrégées. Deux procédures d'échantillonnage sont examinées. La première est une stratégie non biaisée basée sur un plan comme l'estimation de la densité de population est sans biais lorsque le plan inclus des éléments tels que la sélection aléatoire des sites dans chacune des strates. La seconde est appelée stratégie non biaisée basée sur un modèle comme l'estimation de la densité est sans biais sous l'hypothèse d'un modèle de répartition spatiale de la population. La théorie est démontrée à partir d'exemples d'échantillonnage de populations de crevettes.

References

Cochran, W. G. (1977). *Sampling Techniques*, 3rd edition. New York: Wiley.

Cormack, R. M. (1988). Statistical challenges in the environmental sciences: A personal view. *Journal of the Royal Statistical Society, Series A* **151**, 201–210.

Francis, R. I. C. C. (1984). An adaptive strategy for stratified random trawl surveys. *New Zealand Journal of Marine and Freshwater Research* **18**, 59–71.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters*, 2nd edition. London: Griffin.

Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics* **42**, 267–292.

Thompson, S. K. (1990). Adaptive cluster sampling. *Journal of the American Statistical Association* **85**, 1050–1059.

Thompson, S. K. (1991a). Adaptive cluster sampling: Designs with primary and secondary units. *Biometrics* **47**, 1103–1115.

Thompson, S. K. (1991b). Stratified adaptive sampling. *Biometrika* **78**, 389–397.

Thompson, S. K. and Ramsey, F. L. (1987). Detectability functions in observing spatial point processes. *Biometrics* **43**, 355–362.

Received September 1990; *revised May* 1991; *accepted July* 1991.