REVIEW / SYNTHÈSE

Systematic sampling of discrete and continuous populations: sample selection and the choice of estimator

Harry T. Valentine, David L.R. Affleck, and Timothy G. Gregoire

Abstract: Systematic sampling is easy, efficient, and widely used, though it is not generally recognized that a systematic sample may be drawn from the population of interest with or without restrictions on randomization. The restrictions or the lack of them determine which estimators are unbiased, when using the sampling design as the basis for inference. We describe the selection of a systematic sample, with and without restriction, from populations of discrete elements and from linear and areal continuums (continuous populations). We also provide unbiased estimators for both restricted and unrestricted selection. When the population size is known at the outset, systematic sampling with unrestricted selection is most likely the best choice. Restricted selection affords estimation of attribute totals for a population when the population size — for example, the area of an areal continuum — is unknown. Ratio estimation, however, is most likely a more precise option when the selection is restricted and the population size becomes known at the end of the sampling. There is no difference between restricted and unrestricted selection if the sampling interval or grid tessellates the frame in such a way that all samples contain an equal number of measurements. Moreover, all the estimators are unbiased and identical in this situation.

Résumé : L'échantillonnage systématique est simple, efficace et largement utilisé, bien qu'il ne soit pas communément accepté qu'on puisse prélever un échantillon systématique dans la population visée avec ou sans contraintes sur le caractère aléatoire de l'échantillonnage. La présence ou l'absence de contraintes détermine quels estimateurs sont non biaisés lorsqu'on utilise le plan d'échantillonnage comme base pour faire de l'inférence statistique. Nous décrivons la sélection d'un échantillonnage systématique avec et sans contraintes à partir de populations d'éléments discrets et à partir de continuums linéaire et aréal (populations continues). Nous fournissons également des estimateurs non biaisés à la fois pour une sélection avec ou sans contraintes. Lorsque la taille de la population est connue au départ, l'échantillonnage systématique par sélection sans contraintes est très probablement le meilleur choix. La sélection avec contraintes permet d'estimer les totaux des attributs pour une population lorsque la taille de la population est inconnue (par exemple l'aire d'un continuum aréal). Cependant, l'estimation par ratio est fort probablement une option plus précise lorsque le choix est soumis à des contraintes et que la taille de la population devient connue à la fin de l'échantillonnage. Il n'y a pas de différence entre la sélection avec contraintes et la sélection sans contraintes si l'intervalle d'échantillonnage et la grille du damier sont tels que tous les échantillons contiennent un nombre égal de mesures. En outre, tous les estimateurs sont identiques et sans biais dans cette situation.

[Traduit par la Rédaction]

Introduction

Systematic sampling designs are widely used in natural resource assessment and monitoring surveys. The regular dispersion of units within systematic samples can appreciably increase precision and almost always simplifies the collection of field data. Although these aspects of systematic sampling are well known, important statistical issues regarding sample selection and estimation bias appear to be less well understood, particularly when populations distributed over linear or spatial continuums are considered.

Standard treatments of systematic sampling (Cochran 1977; Schreuder et al. 1993) focus on discrete populations. Systematic sampling is generally presented as a 1-in-a routine, wherein a set of elements separated by an interval a are selected in tandem. In this context, randomization is

Received 8 September 2008. Accepted 10 February 2009. Published on the NRC Research Press Web site at cjfr.nrc.ca on 11 June 2009.

H.T. Valentine.[1] U.S. Forest Service, 271 Mast Road, Durham, NH 03824 USA.
D.L.R. Affleck. University of Montana, 32 Campus Drive, Missoula, MT 59812 USA.
T.G. Gregoire. Yale University, 360 Prospect Street, New Haven, CT 06511 USA.

[1]Corresponding author (e-mail: hvalentine@fs.fed.us).

commonly introduced by the uniform random selection of an integer between 1 and a, which indexes the initial selection and anchors the chain of elements forming the sample. The restriction imposed on random selection allows for systematic sampling in the absence of a priori knowledge of a population's size; it can also be appealing because it results in an equal probability design, every element of the population being observed with equal frequency. Nonetheless, an implication of this restriction is that when the population size is not evenly divisible by the sampling interval the sample average is a biased estimator of the population mean. Essentially, since some systematic samples contain more elements than others, the sample average weights too heavily elements selected in smaller samples and too lightly those found in larger samples.

Natural populations, being distributed over space or through time, are often most readily sampled from a continuous areal or temporal frame. Stands of trees, for example, are commonly sampled via the selection of spatial coordinates that serve to locate plots, lines, or points. Any bounded continuous frame can be fractioned into a finite number of disjoint units that can be treated collectively as a discrete population, but if the frame has irregular boundaries it may be impossible to divide it into cells of the same size and shape. Treating a frame as a continuous entity is appropriate also if point-sampling methods such Bitterlich or critical height sampling are used, if sampling designs that cannot tessellate the continuum are applied (e.g., circular fixed-area plots), or if the resource of interest is itself continuously distributed. Regardless of how the frame is specified, however, if its confines are known in advance then systematic samples can be chosen in two distinct ways: (*i*) with restricted randomization, where the first sample unit is drawn at random from a subset of the frame, or (*ii*) with unrestricted randomization, where the first sample unit is drawn at random from the full extent of the frame. Both approaches are used in practice, though it does not appear to be widely appreciated that they produce distinct sampling designs and recommend distinct unbiased estimators.

This article is partly review and partly new, motivated by our perceived need for a more thorough treatment of the connection between restrictions on systematic selection and unbiased estimation. Johnson (2000), Iles (2003), and Mandallaz (2008) describe restricted selection from discrete populations and continuums. Thompson (2002) briefly covers unrestricted and restricted selection in discrete populations, and Gregoire and Valentine (2008) cover restricted selection in discrete populations and unrestricted selection in continuous populations. No one, so far as we know, covers restricted and unrestricted selection of systematic samples from both discrete and continuous populations. Thus, we begin with 1-in-a sampling of discrete populations and extend this design to the systematic sampling of linear and areal continuums. Our prime objective is to demonstrate how restrictions on the selection of a systematic sample determine which estimators are unbiased, when using the sampling design as the basis for inference. Several practical issues are addressed in the discussion. Proofs of the unbiasedness of some estimators are provided in the appendix. An extensive bibliography on systematic sampling is provided by Gregoire (2009).

Table 1. Symbols for discrete populations.

Symbol	Definition
a	Number of sets in $\mathcal{P}_S$ (1-in-a systematic sample)
N	Number of elements in $\mathcal{P}_U$
n_j	Number of elements in set $\mathcal{S}_j$
p_j	Selection probability of $\mathcal{S}_j$
$\mathcal{P}_S$	Population of sets $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_a$
$\mathcal{P}_U$	Population of elements $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_N$
$\mathcal{S}_j$	The jth of a sets of elements in $\mathcal{P}_S$
t_j	Sum of attributes for elements in set $\mathcal{S}_j$
$\mathcal{U}_k$	The kth of N elements in $\mathcal{P}_U$
$\bar{y}_j$	Average attribute per element in set $\mathcal{S}_j$
y_k	Attribute of element $\mathcal{U}_k$
μ_y	Average attribute per element in $\mathcal{P}_U$
τ_y	Total attribute in $\mathcal{P}_U$ and in $\mathcal{P}_S$

Systematic sampling of a discrete population

Let $\mathcal{P}_U$ be a population comprising N discrete elements, $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_N$ (see Table 1). Each element, $\mathcal{U}_k$, manifests a measurable amount, y_k, of some trait or attribute of interest. Our objective is the estimation of the total amount of attribute, τ_y, or the average amount of attribute per element, μ_y, for $\mathcal{P}_U$, where

$$[1] \qquad \tau_y = \sum_{k=1}^{N} y_k$$

and

$$[2] \qquad \mu_y = \frac{\tau_y}{N}$$

Under the usual protocols of a 1-in-a systematic sampling design, the N discrete elements of $\mathcal{P}_U$ aggregate into a mutually exclusive sets of elements, set $\mathcal{S}_j$ comprising n_j elements, with different sets possibly containing different numbers of elements. Suppose, for example, that element $\mathcal{U}_1$ belongs to set $\mathcal{S}_1$. The other $n_1 - 1$ elements of set $\mathcal{S}_1$ are determined systematically by the design parameter a, i.e.,

$$[3] \qquad \mathcal{S}_1 = \{\mathcal{U}_1, \mathcal{U}_{1+a}, \mathcal{U}_{1+2a}, \mathcal{U}_{1+3a}, \ldots\}$$

More generally, if $\mathcal{U}_k \in \mathcal{S}_j$, then

$$[4] \qquad \mathcal{S}_j = \{\ldots, \mathcal{U}_{k-2a}, \mathcal{U}_{k-a}, \mathcal{U}_k, \mathcal{U}_{k+a}, \mathcal{U}_{k+2a}, \ldots\}$$

Thus, the design parameter, a, and the order of the elements in $\mathcal{P}_U$ determine a population, $\mathcal{P}_S$, comprising the a sets of elements; it is this latter population that is sampled when a 1-in-a design is applied. The expected number of elements per set is N/a, a real valued number. However, a set can comprise only whole elements, so the a different sets in $\mathcal{P}_S$ comprise the same number of elements only if N is an integer multiple of a. Regardless,

$$[5] \quad N = \sum_{j=1}^{a} n_j$$

The total amount of attribute in set $\mathcal{S}_j$ is t_j, i.e.,

$$[6] \quad t_j = \sum_{\mathcal{U}_k \in \mathcal{S}_j} y_k$$

whence

$$[7] \quad \tau_y = \sum_{j=1}^{a} t_j$$

In the systematic selection process, we ordinarily select an element, $\mathcal{U}_k$, at random from $\mathcal{P}_\mathcal{U}$ and then identify the other elements of the unique set to which $\mathcal{U}_k$ belongs. The systematic sample comprises this set of elements. How we go about selecting $\mathcal{U}_k$ from $\mathcal{P}_\mathcal{U}$ determines which estimators are unbiased.

Restricted selection

Systematic sampling with a random start is an equal probability design — each of the a sets of elements has the same chance of becoming the sample. The selection is restricted to the first a elements in $\mathcal{P}_\mathcal{U}$, that is, one of the elements $(\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_a)$ is selected uniformly at random with probability $1/a$. The set, $\mathcal{S}_j \in \mathcal{P}_\mathcal{S}$, to which the selected element $\mathcal{U}_j$ $(1 \le j \le a)$ belongs is our systematic sample. Alternatively, we could restrict the selection of an element to one of the first a elements after $\mathcal{U}_k$, i.e., $(\mathcal{U}_{k+1}, \mathcal{U}_{k+2}, \ldots, \mathcal{U}_{k+a})$, where $\mathcal{U}_{k+a} \le N$. In either case, $\mathcal{S}_j$ is selected with probability $p_j = 1/a$, so the target parameter, τ_y, is unbiasedly estimated with the Hansen–Hurwitz estimator:

$$[8] \quad \hat{\tau}_{y,\,\mathrm{rs}} = \frac{t_j}{p_j} = at_j$$

Because the single sample unit is actually the set, $\mathcal{S}_j$, the selection probability, p_j, and the inclusion probability, π_j, of $\mathcal{S}_j$ are identical, so τ_y is unbiasedly and identically estimated by the Horvitz–Thompson estimator:

$$[9] \quad \hat{\tau}_{y,\,\mathrm{rs}} = \frac{t_j}{\pi_j} = at_j$$

which can also be derived by noting that the inclusion probability of $\mathcal{U}_k$ is identically $1/a$ for all k. The average amount of attribute per element, μ_y, is unbiasedly estimated by $\hat{\tau}_{y,\,\mathrm{rs}}/N$. The ratio estimator is a third alternative (e.g., Thompson 2002):

$$[10] \quad \hat{\tau}_{y,\,\mathrm{rat}} = \frac{N}{n_j} t_j = N\bar{y}_j$$

where $y_j = t_j/n_j$ is the average amount of attribute per element in $\mathcal{S}_j$. The corresponding estimator of μ_y is $\hat{\mu}_{y,\,\mathrm{rat}} = \bar{y}_j$. The ratio estimator is unbiased for random-start systematic sampling only if N is an integer multiple of a, in which case $p_j = \pi_j = 1/a = n/N$ for all j.

Despite its bias, $\hat{\tau}_{y,\,\mathrm{rat}}$ often is more precise than an unbiased alternative. For example, Table 2 contains the bias and standard errors of estimators of total leaf area based on 1-in-6 and 1-in-8 systematic samplings from a population of $N = 64$ leaves, where $\tau_y = 1582$ cm^2 (from Barrett and Nutt 1979). The ratio estimator is biased but more precise than the Horvitz–Thompson estimator for the 1-in-6 sampling. The two estimators are identical and unbiased for the 1-in-8 sampling because the population size, $N = 64$, is an integer multiple of the sampling interval, $a = 8$.

Table 2. Bias and standard errors (SE) of estimators of total leaf area ($\tau_y = 1582$ cm^2) from 1-in-a samples of 64 leaves.

Restriction	a	Estimator	Bias (cm^2)	SE (cm^2)*
Yes	6	$\hat{\tau}_{y,\,\mathrm{rs}} = at_j$	0.0	112.4
		$\hat{\tau}_{y,\,\mathrm{rat}} = N\bar{y}_j$	−1.5	56.8
Yes	8	$\hat{\tau}_{y,\,\mathrm{rs}} = at_j$	0.0	97.5
		$\hat{\tau}_{y,\,\mathrm{rat}} = N\bar{y}_j$	0.0	97.5
No	6	$\hat{\tau}_{y,\,\mathrm{rs}} = at_j$	4.6	56.4
		$\hat{\tau}_{y,\,\mathrm{us}} = N\bar{y}_j$	0.0	58.4
No	8	$\hat{\tau}_{y,\,\mathrm{rs}} = at_j$	0.0	97.5
		$\hat{\tau}_{y,\,\mathrm{us}} = N\bar{y}_j$	0.0	97.5

*Square root of the estimator's variance.

Unrestricted selection

Systematic sampling with unrestricted selection is an unequal probability design because a set, $\mathcal{S}_j$, is selected from $\mathcal{P}_\mathcal{S}$ as the systematic sample unit with probability proportional to the number of elements in the set. An element, $\mathcal{U}_k$, is selected uniformly at random with probability $1/N$ from $\mathcal{P}_\mathcal{U}$, which selects set $\mathcal{S}_j \ni \mathcal{U}_k$ from $\mathcal{P}_\mathcal{S}$ as the systematic sample with probability $p_j = n_j/N$. Accordingly, the unbiased Hansen–Hurwitz estimator of τ_y is

$$[11] \quad \hat{\tau}_{y,\,\mathrm{us}} = \frac{t_j}{p_j} = \frac{N}{n_j} t_j = N\bar{y}_j$$

As with restricted selection, $p_j = \pi_j$, so the Horvitz–Thompson estimator of τ_y is identical:

$$[12] \quad \hat{\tau}_{y,\,\mathrm{us}} = \frac{t_j}{\pi_j} = N\bar{y}_j$$

At the elemental level, this estimator can also be motivated by the fact that the inclusion probability of $\mathcal{U}_k \in \mathcal{S}_j$ is n_j/N under unrestricted selection.

Note that both the unbiased Hansen–Hurwitz and Horvitz–Thompson estimators for unrestricted selection are identical to the ratio estimator for restricted selection. Also, if N is an integer multiple of a, then all the estimators of τ_y are identical and unbiased with either method of selection. Moreover, μ_y is unbiasedly estimated by $\bar{y}_j$ in this situation.

On the other hand, if N is not an integer multiple of a, then the unbiased estimators for restricted selection are biased for unrestricted selection, and the unbiased estimators for unrestricted selection are biased for restricted selection (as shown in Table 2). The bias of the unrestricted estimators under restricted selection is equivalent to the bias of the ratio estimator. However, the bias is small relative to the standard error. Estimators with "Hansen–Hurwitz" and "Horvitz–Thompson" labels are, by definition, unbiased, so

[8] and [11] are not Hansen–Hurwitz estimators and [9] and [12] are not Horvitz–Thompson estimators when applied in situations for which they are biased.

Overall, the message is clear. If one knows N at the outset of sampling, then unrestricted selection is most likely the best choice (compare rows 1 and 6 in Table 2). If N becomes known only after sampling with restricted selection, then precision is gained through the use of the ratio estimator, though this gain is attended by a small bias.

A circular systematic sampling design also prescribes unrestricted selection of one of the N elements from $\mathcal{P}_U$ (e.g., Gregoire and Valentine 2008, Chap. 3). Selection of the other elements proceeds in a circular fashion, so when the end of the frame ($\mathcal{U}_N$) is reached, the remainder of the sampling interval to the next element in the sample continues from the start of the frame ($\mathcal{U}_1$). This "circular protocol" provides for a systematic sample with a fixed number elements, n, but it does not define a mutually exclusive sets, unless $a = N/n$. Accordingly, none of our estimators are unbiased for circular selection when $a \neq N/n$; however, all of our estimators are unbiased when $a = N/n$.

Systematic sampling of a linear continuum

A linear continuum, $\mathcal{L}$, of any length, L, comprises a set of infinitely many points. For convenience, we specify that $\mathcal{L}$ is contained in [0, L]. Of prime interest is the total amount of some attribute, τ_ρ, that is distributed along the continuum. Let x denote the location of a point in [0, L], and let $\rho(x)$ be the attribute density (the amount of attribute per unit length) at x. Then,

$$[13] \qquad \tau_\rho = \int_{\mathcal{L}} \rho(x)\, dx$$

The mean attribute density (average amount of attribute per unit length) in $\mathcal{L}$ is

$$[14] \qquad \mu_\rho = \frac{\tau_\rho}{L}$$

A list of symbols for this section is provided in Table 3.

To put the linear continuum in a context familiar to foresters, let $\mathcal{L}$ be the straight central axis of a tree bole of length L. Let $\rho(x)$ be the volume per unit length (cross-sectional area) at x on the central axis; then τ_ρ is the volume of the bole. Another forestry example is represented by the traditional strip cruise, where cruise lines (transects) are equally spaced and perpendicular to a continuous baseline, $\mathcal{L}$, with length L that spans a tract of interest. Each transect yields a measurement of the attribute of interest, but dividing this measurement by the width of the search interval (i.e., the strip width) provides the attribute per unit length, $\rho(x)$, for the point $x \in \mathcal{L}$ from which the cruise line emanates. Thus, the total amount of attribute, τ_ρ, distributed along $\mathcal{L}$ is equivalent to the total amount of attribute distributed over the tract of interest. In an ecophysiological context, $\mathcal{L}$ may be a continuous stream of time, and $\rho(x)$, the net exchange of carbon per unit time between an ecosystem and the atmosphere, in which case τ_ρ is the amount of carbon sequestered by the ecosystem in the time interval [0, L].

Our objective is the estimation of τ_ρ by a continuous version of systematic sampling. The continuous version of simple random sampling is called crude Monte Carlo. In unreplicated crude Monte Carlo, a single sample point, $s \in \mathcal{L}$, is selected uniformly at random at x_s with probability density $f(x_s) = 1/L$. Heuristically, a probability of 1 is stretched uniformly across $\mathcal{L}$, giving each point $x \in \mathcal{L}$ a probability density (probability per unit length) of $1/L$. Consequently, $x_s = uL$, where u obtains from Uniform[0,1]. The total attribute in $\mathcal{L}$ is unbiasedly estimated by

$$[15] \qquad \hat{\tau}_\rho = \frac{\rho(x_s)}{f(x_s)} = L\rho(x_s)$$

which we shall call the MC estimator. The mean attribute density, μ_ρ, is unbiasedly estimated by $\hat{\mu}_\rho = \rho(x_s)$.

Table 3. Symbols for linear continuums.

Symbol	Definition
$f(x)$	Probability density at $x \in \mathcal{L}$ for unrestricted selection
$f_\phi(x)$	Probability density at $x \in \mathcal{L}_\phi$ for restricted selection
ℓ	Interval between measurement points
L	Length of the linear continuum $\mathcal{L}$
$\mathcal{L}$	Linear continuum in [0, L]
$\mathcal{L}_\phi$	Subdomain of $\mathcal{L}$ in $[x_\phi, x_\phi + \ell]$
$n(x_s)$	Number measurement points in $\mathcal{T}_s$
s	Sample point at x_s
$\mathcal{T}_s$	Systematic sample comprising a set of measurement points anchored by s
u	Uniform [0,1] random number
μ_ρ	Mean attribute density along $\mathcal{L}$
$\rho(x)$	Attribute density at $x \in \mathcal{L}$
$\bar{\rho}(x_s)$	Average attribute density in $\mathcal{T}_s$
τ_ρ	Total attribute distributed over $\mathcal{L}$
$\phi(x_s)$	Sum of attribute densities in $\mathcal{T}_s$

Restricted selection

In the continuous version of systematic sampling, the location of a sample point, x_s, and a measurement interval, ℓ, determine a unique set of measurement points, $\mathcal{T}_s$, evenly spaced across $\mathcal{L}$. The measurement interval is a design parameter. If L is not an integer multiple of ℓ, the number of measurement points in $\mathcal{T}_s$ depends on x_s, so let $n(x_s)$ be the number of points in the set $\mathcal{T}_s$ determined by x_s and ℓ. Hence, the systematic sample comprises the $n(x_s)$ measurement points in $\mathcal{T}_s$.

To perform systematic sampling under restricted selection, we do not have to know L. We restrict the location of the sample point, s, to a subdomain $\mathcal{L}_\phi \subset \mathcal{L}$ that spans an interval $[x_\phi, x_\phi + \ell]$. Ordinarily, we use $x_\phi = 0$, so the sample point is selected in $[0, \ell]$. In either case, the probability density function, $f_\phi(x)$, is defined for the subdomain, such that $f_\phi(x) = 1/\ell$ for all $x \in \mathcal{L}_\phi$. Consequently, the sample point is selected at $x_s = x_\phi + u\ell$ with probability density $f_\phi(x_s) = 1/\ell$. Measurement points in $\mathcal{L}$ at $x_s + j\ell$ and $x_s - j\ell$, $j = 1, 2, \ldots$, complete the systematic sample, $n(x_s)$ points in total.

For estimation, we define a function $\phi(x)$ for the subdomain $\mathcal{L}_\phi$. Each $x \in \mathcal{L}_\phi$ belongs to the unique set of $n(x)$ points in $\mathcal{L}$ determined by x and ℓ. Let $\phi(x)$ be the sum of the attribute densities for the set of points in $\mathcal{L}$ to which $x \in \mathcal{L}_\phi$ belongs, i.e., for all $x \in \mathcal{L}_\phi$,

$$[16] \quad \phi(x) = \rho(x) + \sum_{\substack{x+j\ell \in \mathcal{L}'_\phi \\ j=1,2,\ldots}} \rho(x+j\ell) + \sum_{\substack{x-j\ell \in \mathcal{L}'_\phi \\ j=1,2,\ldots}} \rho(x-j\ell)$$

where $\mathcal{L}'_\phi$ is the complement of $\mathcal{L}_\phi \subset \mathcal{L}$. Hence,

$$[17] \quad \begin{aligned} \int_{\mathcal{L}_\phi} \phi(x)\,dx &= \int_{\mathcal{L}_\phi} \rho(x)\,dx + \int_{\mathcal{L}'_\phi} \rho(x)\,dx \\ &= \int_{\mathcal{L}} \rho(x)\,dx \end{aligned}$$

and, therefore, $\int_{\mathcal{L}_\phi} \phi(x)\,dx = \tau_\rho$. Consequently, under systematic sampling with restricted selection, the target parameter, τ_ρ, is unbiasedly estimated with a MC estimator, i.e.,

$$[18] \quad \hat{\tau}_{\rho,\,\mathrm{rs}} = \frac{\phi(x_s)}{f_\phi(x_s)} = \ell\phi(x_s)$$

where $\phi(x_s)$ is the sum of the attribute densities in the systematic sample, $\mathcal{T}_s$. The mean attribute density in $\mathcal{L}$ is unbiasedly estimated by $\hat{\mu}_{\rho,\,\mathrm{rs}} = \hat{\tau}_{\rho,\,\mathrm{rs}}/L$. If L is known at the outset, we can fix n and calculate $\ell = L/n$, in which case, $f_\phi(x_s) = 1/\ell = n/L$ for all $x \in \mathcal{L}_\phi$. Then τ_ρ is unbiasedly estimated by

$$[19] \quad \hat{\tau}_{\rho,\,\mathrm{rs}} = \frac{L}{n}\,\phi(x_s) = L\bar{\rho}(x_s)$$

where $\bar{\rho}(x_s)$ is the average of the n attribute densities in $\mathcal{T}_s$. Moreover, μ_ρ is unbiasedly estimated by $\bar{\rho}(x_s)$. If $\ell \neq L/n$, we can use a ratio estimator,

$$[20] \quad \hat{\tau}_{\rho,\,\mathrm{rat}} = \frac{L}{n(x_s)}\,\phi(x_s) = L\bar{\rho}(x_s)$$

The ratio estimator is biased, however, because $f(x) = 1/\ell$ for all $x \in \mathcal{L}_\phi$, but there exist $x \in \mathcal{L}_\phi$ for which $n(x)/L \neq 1/\ell$.

Unrestricted selection

Unrestricted systematic selection is easiest if we know L at the outset. A sample point at x_s is selected in $\mathcal{L}$ with probability density $f(x_s) = 1/L$, so $x_s = uL$. Additional points in $\mathcal{L}$, at $x_s + j\ell$ and $x_s - j\ell$, $j = 1, 2,\ldots$, fill out the systematic sample, $\mathcal{T}_s$. Because the selection of any point in a set selects the set as the systematic sample, the unrestricted uniform selection of a sample point from $\mathcal{L}$ is, in effect, a continuous analog of selecting a set with probability proportional to set size. Unrestricted selection also can be accomplished, even though L is unknown, by employing von Neumann's acceptance–rejection method: Imagine a continuum $\mathcal{L}^*$ of length L^* that is sure to include all of $\mathcal{L}$. We select $x^* = uL^*$, which becomes the sample point, x_s, if x^* occurs in $\mathcal{L}$. Otherwise, we reject x^* as the sample point and repeat the procedure with a new random value of u.

For estimation, we define a function, $\bar{\rho}(x)$, for all $x \in \mathcal{L}$, where $\bar{\rho}(x)$ is the average attribute density for the unique set of $n(x)$ points to which each x belongs. Hence, $\tau_\rho = \int_{\mathcal{L}} \bar{\rho}(x)\,dx$, and therefore, an unbiased estimate of τ_ρ, under unrestricted selection, is provided by the MC estimator,

Table 4. Symbols for areal continuums.

Symbol	Definition
A	Horizontal area of areal continuum $\mathcal{A}$
$\mathcal{A}$	Areal continuum with horizontal area A
$\mathcal{A}_\phi$	Subdomain of $\mathcal{A}$ with shape and horizontal area of a grid cell
c	Horizontal area of a grid cell
$f(x, z)$	Probability density at $(x, z) \in \mathcal{A}$ for unrestricted selection
$f_\phi(x, z)$	Probability density at $(x, z) \in \mathcal{A}_\phi$ for restricted selection
$\mathcal{G}_s$	Systematic sample comprising a set of grid points anchored by s
h, ℓ	Design parameters for grid point spacing
$n(x_s, z_s)$	Number of grid points in $\mathcal{G}_s$
r, R	Apothem and circumradius of a regular hexagon
s	Sample point at (x_s, z_s)
μ_ρ	Mean attribute density in $\mathcal{A}$
$\rho(x, z)$	Attribute density at (x, z)
$\bar{\rho}(x_s, z_s)$	Average attribute density in $\mathcal{G}_s$
τ_ρ	Total amount of attribute in $\mathcal{A}$
$\phi(x_s, z_s)$	Sum of attribute densities in $\mathcal{G}_s$

$$[21] \quad \hat{\tau}_{\rho,\,\mathrm{us}} = \frac{\bar{\rho}(x_s)}{f(x_s)} = L\bar{\rho}(x_s)$$

where $\bar{\rho}(x_s)$ is the average of the $n(x_s)$ attribute densities in $\mathcal{T}_s$. The mean attribute density, μ_ρ, is unbiasedly estimated by $\hat{\mu}_{\rho,\,\mathrm{us}} = \bar{\rho}(x_s)$.

If L is an integer multiple of ℓ, then $n(x) = n$ for all $x \in \mathcal{L}$, in which case [18], [19], and [21] are equivalent unbiased estimators, and μ_ρ is unbiasedly estimated by $\bar{\rho}(x_s)$.

Systematic sampling of an areal continuum

We consider an areal continuum, $\mathcal{A}$, with a closed boundary and horizontal area A (see Table 4). We allow $\mathcal{A}$ to comprise a region of interest surrounded by a buffer region with a closed boundary, in which case A is the area of the region of interest plus the area of the buffer. The areal continuum comprises infinitely many location points, with each location point identified by its coordinates (x, z). Of interest is the amount of some attribute, τ_ρ, that is distributed across $\mathcal{A}$. Let $\rho(x, z)$ be the attribute density (the amount of attribute per unit horizontal area) at (x, z), then

$$[22] \quad \tau_\rho = \iint_{\mathcal{A}} \rho(x, z)\,dx\,dz$$

The mean attribute density across $\mathcal{A}$ is $\mu_\rho = \tau_\rho/A$.

The systematic sampling of $\mathcal{A}$ involves the selection of a sample point, s, at (x_s, z_s). The sample point anchors a systematic grid of measurement points that span $\mathcal{A}$. The interval(s) between the grid points and their systematic spatial pattern depend on design parameters. Square, rectangular, and equilateral triangular grid patterns are popular choices. The number of grid points in $\mathcal{A}$ may change with the location of the sample point, though the set of points in any grid is fixed by the location of the sample point and the design parameters. Each point in $\mathcal{A}$ belongs to one and only one set of grid points.

Systematic grids are used widely to sample tracts of land,

ranging in scale from small fields or wood lots to landscapes to entire countries. Grid points, for example, may serve as points from which Bitterlich or perpendicular distance sampling is conducted. They may serve as center or corner points for plots or plot clusters, or as center points or end points of transects for line intersect or line intersect distance sampling. Most of the specialized sampling methods that are applied on tracts of land, including those just mentioned and many others, provide attribute densities (amount of attribute per unit land area) for discrete elements of interest at any grid point (see Gregoire and Valentine (2008), Chap. 10, or Mandallaz (2008), Chap. 4). Consequently, the simple estimators that we provide below for areal continuums have wide applicability, regardless of whether τ_ρ is an attribute of a continuous entity or the sum of the attributes for a population of discrete elements that occurs within the continuum. How we select the sample point that anchors the grid determines which estimators are unbiased.

To sample $\mathcal{A}$ by unreplicated crude Monte Carlo, we select a single sample point at (x_s, z_s) uniformly at random with probability density (probability per unit area) $f(x_s, z_s) = 1/A$. This selection is most easily accomplished by the acceptance–rejection method. Imagine a rectangle, with dimensions $X \times Z$, which is large enough to include all of $\mathcal{A}$. Draw u_x and u_z from Uniform [0, 1] and test whether $(x_s = u_x X, z_s = u_z Z)$ occurs in $\mathcal{A}$. If so, accept (x_s, z_s); if not, draw new random numbers and repeat. The target parameter, τ_ρ, is unbiasedly estimated by

$$[23] \qquad \widehat{\tau}_\rho = \frac{\rho(x_s, z_s)}{f(x_s, z_s)} = A\rho(x_s, z_s)$$

and μ_ρ is unbiasedly estimated by $\widehat{\mu}_\rho = \rho(x_s, z_s)$.

Restricted selection

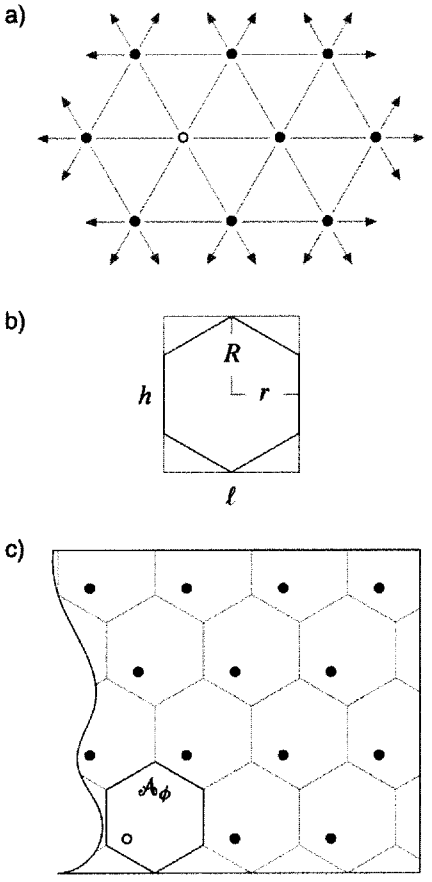
For systematic sampling with restricted selection, we need not know A, but we use the design parameters that define the shape and spacing of the systematic grid.

For illustrative purposes, we let ℓ and h, respectively, be the intervals between the x coordinates and z coordinates of points in a rectangular $(\ell \neq h)$ or square $(\ell = h)$ grid. The corresponding grid cell is a $\ell \times h$ rectangle, and the location of the sample point is restricted to a subdomain, $\mathcal{A}_\phi \subset \mathcal{A}$, of this shape and area. For an equilateral triangular grid (Fig. 1*a*), $\mathcal{A}_\phi$ is a regular hexagon, so ℓ may be the distance between grid points (i.e., $\ell = 2r$, where r is the apothem of a hexagon), and h may be twice the circumradius of the hexagon (i.e., $h = 2R$, where R is the circumradius), in which case the hexagon can be circumscribed by a $\ell \times h$ rectangle (Fig. 1*b*).

The subdomain $\mathcal{A}_\phi$ may occur anywhere in $\mathcal{A}$. For example, we may arbitrarily select any convenient location point (x_ϕ, z_ϕ) in $\mathcal{A}$ that can serve as a vertex of a $\ell \times h$ rectangle. The subdomain $\mathcal{A}_\phi$ is either coincident with, or circumscribed by, this rectangle. In the former case, the sample point for a rectangular grid is selected at $(x_s = x_\phi + u_x\ell, z_s = z_\phi + u_z h)$ with probability density $f_\phi(x_s, z_s) = 1/c$, where $c = \ell h$.

For an equilateral triangular grid, the acceptance–rejection method selects the sample point at (x_s, z_s) within the hexagonal subdomain, $\mathcal{A}_\phi$, circumscribed by the $\ell \times h$ rectangle.

Fig. 1. (*a*) Triangular grid anchored by a sample point (○); (*b*) hexagon, with circumradius R and apothem r, circumscribed with a $\ell \times h$ rectangle; (*c*) the design parameters and the location of the sample point (○) within the hexagonal subdomain, $\mathcal{A}_\phi$, determine a unique set of triangular grid points in $\mathcal{A}$.



The sample point is selected with probability density $f_\phi(x_s, z_s) = 1/c$, where $c = 3\ell h/4 = 3rR$ is the area of the hexagon (Fig. 1*c*). The resultant systematic grid, $\mathcal{G}_s$, anchored by the sample point, s, at (x_s, z_s), contains $n(x_s, z_s)$ measurement points in $\mathcal{A}$.

The unbiased MC estimator of τ_ρ is the two-dimensional analog of [18]. Each location point $(x, z) \in \mathcal{A}_\phi$ belongs to a unique set of $n(x, z)$ grid points in $\mathcal{A}$, so we define $\phi(x, z)$ in $\mathcal{A}_\phi$ to be the sum of the attribute densities for the set of grid points in $\mathcal{A}$ to which $(x, z) \in \mathcal{A}_\phi$ belongs. Consequently, $\iint_{\mathcal{A}} \rho(x, z)\, dx\, dz = \iint_{\mathcal{A}_\phi} \phi(x, z)\, dx\, dz = \tau_\rho$, which is unbiasedly estimated by

$$[24] \qquad \widehat{\tau}_{\rho,\,\mathrm{rs}} = \frac{\phi(x_s, z_s)}{f_\phi(x_s, z_s)} = c\phi(x_s, z_s)$$

where $\phi(x_s, z_s)$ is the sum of the attribute densities across the $n(x_s, z_s)$ grid points in the set $\mathcal{G}_s$. An unbiased estimator of μ_ρ is $\widehat{\mu}_{\rho,\,\mathrm{rs}} = c\phi(x_s, z_s)/A$.

If A is the attribute of interest, i.e., $\tau_\rho = A$, then $\rho(x, z) = 1$ (unit of area per unit area) for all $(x, z) \in \mathcal{A}$, so A is unbiasedly estimated by $\widehat{A} = an(x_s, z_s)$. If A is known, we can estimate τ_ρ with a ratio estimator

Table 5. Summary table of estimators for systematic sampling.

Discrete elements	Linear continuum	Areal continuum
Restricted selection		
$\widehat{\tau}_{y,\,\mathrm{rs}} = at_j$	$\widehat{\tau}_{\rho,\,\mathrm{rs}} = \ell\phi(x_s)$	$\widehat{\tau}_{\rho,\,\mathrm{rs}} = c\phi(x_s, z_s)$
$\widehat{\tau}_{y,\,\mathrm{rat}} = N\bar{y}_j$	$\widehat{\tau}_{\rho,\,\mathrm{rat}} = L\bar{\rho}(x_s)$	$\widehat{\tau}_{\rho,\,\mathrm{rat}} = A\bar{\rho}(x_s, z_s)$
Unrestricted selection		
$\widehat{\tau}_{y,\,\mathrm{us}} = N\bar{y}_j$	$\widehat{\tau}_{\rho,\,\mathrm{us}} = L\bar{\rho}(x_s)$	$\widehat{\tau}_{\rho,\,\mathrm{us}} = A\bar{\rho}(x_s, z_s)$
Components		
$t_j = \sum_{U_k \in \mathcal{S}_j} y_k$	$\phi(x_s) = \sum_{x \in \mathcal{T}_s} \rho(x)$	$\phi(x_s, z_s) = \sum_{(x,z) \in \mathcal{G}_s} \rho(x, z)$
$\bar{y}_j = \frac{t_j}{n_j}$	$\bar{\rho}(x_s) = \frac{\phi(x_s)}{n(x_s)}$	$\bar{\rho}(x_s, z_s) = \frac{\phi(x_s, z_s)}{n(x_s, z_s)}$

[25] $\widehat{\tau}_{\rho,\,\mathrm{rat}} = A\bar{\rho}(x_s, z_s)$

where $\bar{\rho}(x_s, z_s)$ is the average of the $n(x_s, z_s)$ attribute densities. The ratio estimator is unbiased if $n(x, z)/A = 1/c$ for all $(x, z) \in \mathcal{A}_\phi$. Or, to put it another way, the ratio estimator is unbiased if $\mathcal{A}$ is conterminous with the union of the grid cells.

Unrestricted selection

For unrestricted selection, we have no need of $\mathcal{A}_\phi$. The acceptance–rejection method selects the sample point, s, anywhere in $\mathcal{A}$ with probability density $f(x_s, z_s) = 1/A$. The sample point anchors the grid $\mathcal{G}_s$ with $n(x_s, z_s)$ measurement points in $\mathcal{A}$. By analogy to [21], the target parameter, τ_ρ, is unbiasedly estimated with

[26] $$\widehat{\tau}_{\rho,\,\mathrm{us}} = \frac{\bar{\rho}(x_s, z_s)}{f(x_s, z_s)} = A\bar{\rho}(x_s, z_s)$$

Moreover, μ_ρ is unbiasedly estimated by $\widehat{\mu}_{\rho,\mathrm{us}} = \bar{\rho}(x_s, z_s)$.

The three estimators, [24], [25], and [26], coincide if $\mathcal{A}$ tessellates completely into n grid cells, each with area c.

By this point, the extension of our results to sampling a three-dimensional container should be obvious. If not, Baddeley and Jensen (2005) cover systematic sampling in three dimensions with restricted selection.

Discussion

A systematic sample ordinarily comprises a single set of elements or measurement points drawn from a population of sets. In discrete populations and linear continuums, some sets may contain one more element or point than the other sets. In an areal continuum, the variation in the number of points among potential grids may be much greater, depending on the shape of the areal continuum and the shape and size of grid cells. Whether this variation in set size leads to bias in an estimator depends on the restrictions on randomization for the sample selection. To wit, the estimators that are unbiased for restricted selection are biased for unrestricted selection, and the estimators that are unbiased for unrestricted selection are biased for restricted selection. The magnitude of the estimation bias, however, will be small in relation to the variance for most natural populations. Unbiasedness, nonetheless, retains a certain appeal in natural resource surveys, particularly those conducted and defended by public agencies. That the bias in systematic designs can be eliminated by pairing either sample selection method with a suitable estimator therefore remains noteworthy.

It is evident from Table 5 that the collection of estimators derived above resolve into two basic forms. Horvitz–Thompson theory and ratio estimation for discrete populations produce distinct estimators when randomization is restricted, but these two strategies lead to the same basic rule when no restrictions are imposed. There are also obvious discrete-population analogs for each of the MC estimators, the latter being generalizations recognizing the continuity of the population and sampling interval. For example, in a discrete population, the sampling intensity is one element per sequential set of a elements, and in linear and areal continuums, the intensities, respectively, are one point per interval of length ℓ and one point per grid cell with area c. Moreover, the MC estimators for both randomization strategies can be derived with a continuous analog of Horvitz–Thompson theory (Cordy 1993).

The availability of an array of selection methods and estimators raises the question of which should be adopted in any given application. Unrestricted randomization selects systematic samples with probability proportional to the number of elements or measurement points in the sample; in principle this suggests improved precision. This selection strategy also allows one to use the sample mean without incurring a design bias. Where restricted randomization is used, perhaps because population size is not known in advance, the ratio estimator, though biased, is likely to have lower variance, as it corrects for realized number of measured elements or points. However, the area of an areal continuum may remain unknown even after the completion of a systematic sampling, precluding ratio estimation. In this case, restricted randomization provides for unbiased estimation of the total amount of attribute that is distributed over the continuum of unknown extent.

Acknowledgments

We thank Jim Flewelling, Jeff Gove, Daniel Mandallaz, and the Associate Editor for their reviews of the manuscript and their helpful suggestions.

References

Baddeley, A., and Jensen, E.B.V. 2005. Stereology for statisticians. Chapman & Hall/CRC, Boca Raton, Fla.

Barrett, J.P., and Nutt, M.E. 1979. Survey sampling in the environ-

mental sciences: a computer approach. COMPress, Inc., Wentworth, N.H.
Cochran, W.G. 1977. Sampling techniques. Wiley, New York.
Cordy, C.B. 1993. An extension to the Horvitz–Thompson theorem to point sampling from a continuous universe. Stat. Probab. Lett. **18**: 353–362. doi:10.1016/0167-7152(93)90028-H.
Gregoire, T.G. 2009. Bibliography, systematic sampling [online]. Available from http://forestry.yale.edu/biometrics/Bibliographies.html [accessed 5 January 2009].
Gregoire, T.G., and Valentine, H.T. 2008. Sampling strategies for natural resources and the environment. Chapman & Hall/CRC, Boca Raton, Fla.
Iles, K. 2003. A sampler of inventory topics. Kim Iles & Associates, Nanaimo, B.C.
Johnson, E.W. 2000. Forest sampling desk reference. CRC Press, Boca Raton, Fla.
Mandallaz, D. 2008. Sampling techniques for forest inventories. Chapman & Hall/CRC, Boca Raton, Fla.
Schreuder, H.T., Gregoire, T.G., and Wood, G.B. 1993. Sampling methods for multiresource forest inventory. Wiley, New York.
Thompson, S.K. 2002. Sampling. 2nd ed. Wiley, New York.

Appendix A An estimator of a parameter is unbiased if the expectation of the estimator equals the parameter. The proofs of unbiasedness of the estimators for restricted or unrestricted selection are straightforward. For example, for the restricted estimator $\hat{\tau}_{y,\,rs} = at_j$,

$$[A1] \quad \begin{aligned} E[\hat{\tau}_{y,\,rs}] &= \sum_{j=1}^{a} \pi_j a t_j = \sum_{j=1}^{a} \frac{1}{a}\, a t_j \\ &= \sum_{j=1}^{a} t_j = \sum_{k=1}^{N} y_k = \tau_y \end{aligned}$$

For $\hat{\tau}_{\rho,\,rs} = c\phi(x_s, z_s)$,

$$[A2] \quad \begin{aligned} E[\hat{\tau}_{\rho,\,rs}] &= \iint_{\mathcal{A}_\phi} f_\phi(x, z)\, c\phi(x, z)\, dx\, dz \\ &= \iint_{\mathcal{A}_\phi} \frac{1}{c}\, c\phi(x, z)\, dx\, dz \\ &= \iint_{\mathcal{A}_\phi} \phi(x, z)\, dx\, dz \\ &= \iint_{\mathcal{A}} \rho(x, z)\, dx\, dz = \tau_\rho \end{aligned}$$

Similarly, for the unrestricted estimator $\hat{\tau}_{y,\,us} = N\bar{y}_j$,

$$[A3] \quad \begin{aligned} E[\hat{\tau}_{y,us}] &= \sum_{j=1}^{a} \pi_j N \bar{y}_j = \sum_{j=1}^{a} \frac{n_j}{N} N \bar{y}_j \\ &= \sum_{j=1}^{a} n_j \bar{y}_j = \sum_{k=1}^{N} y_k = \tau_y \end{aligned}$$

And, for $\hat{\tau}_{\rho,\,us} = L\bar{\rho}(x)$,

$$\begin{aligned} E[\hat{\tau}_{\rho,\,us}] &= \int_{\mathcal{L}} f(x) L\bar{\rho}(x)\, dx = \int_{\mathcal{L}} \frac{1}{L}\, L\bar{\rho}(x)\, dx \\ &= \int_{\mathcal{L}} \bar{\rho}(x)\, dx \end{aligned}$$

Since $\bar{\rho}(x)$ is the average attribute density for the unique set of points to which x belongs

$$[A4] \quad E[\hat{\tau}_{\rho,\,us}] = \int_{\mathcal{L}} \bar{\rho}(x)\, dx = \int_{\mathcal{L}} \rho(x)\, dx = \tau_\rho$$

The ratio estimators are biased if the population size is not an integer multiple of the the sampling interval. For example, unless $N = an_j$ for all j, then $\hat{\tau}_{y,\,rat}$ is biased, i.e.,

$$[A5] \quad \begin{aligned} E[\hat{\tau}_{y,\,rat}] &= \sum_{j=1}^{a} \pi_j N \bar{y}_j = \sum_{j=1}^{a} \frac{1}{a} N \bar{y}_j \\ &= \sum_{j=1}^{a} \frac{N}{an_j}\, t_j \neq \tau_y \end{aligned}$$

On the other hand, the ratio estimator is unbiased if $N = an_j$ for all j, i.e.,

$$[A6] \quad \begin{aligned} E[\hat{\tau}_{y,\,rat}] &= \sum_{j=1}^{a} \frac{N}{an_j}\, t_j = \sum_{j=1}^{a} \frac{N}{N}\, t_j \\ &= \sum_{j=1}^{a} t_j = \tau_y \end{aligned}$$

Analogous arguments hold in the continuous case.