A new method for the estimation of density by the spacing method applicable to non-randomly distributed populations

Masaaki Morisita
Department of Biology
Faculty of Science
Kyushu University

November 1957

Abstract

The document that follows is an English translation of Morista's article that appeared in *Physiology and Ecology*, 7:134144, 1957 in Japanese with an English summary. The translation was made by the United States Department of Agriculture, Division of Range Management in 1960.

This document has been cited numerous times in the literature. However, for more than a decade I was unable to locate a copy, even at the USDA. I could find no one who had actually seen a copy, including those who had cited it, myself included, until this copy was located in 2013. The document has been passed back to the USDA to post online. This cover page has been included to help search engines actually find the document.

Kevin Mitchell
Department of Mathematics and Computer Science
Hobart and William Smith Colleges
mitchell@hws.edu
2015.01.19

UNITED STATES DEPARTMENT OF AGRICULTURE

FOREST SERVICE

A NEW METHOD FOR THE ESTIMATION OF DENSITY BY THE SPACING METHOD APPLICABLE TO NON-RANDOMLY DISTRIBUTED POPULATIONS

(Translation)

Masaaki Morisita

Department of Biology, Faculty of Science, Kyushu University

Division of Range Management and Wildlife Habitat Research

Washington, D.C.

M—5123

U. S. GOVERNMENT PRINTING OFFICE

FOREWORD

The original of the following translation appeared in Japanese in PHYSIOLOGY AND ECOLOGY 7:134-144 (November 1957). The translation was made at our request by Mr. Hirotsuga Yasuda, a graduate student at State University College of Forestry, Syracuse University. A few minor changes in the translation were made by a statistician at Syracuse University and by us.

This paper briefly reviews the theory of the more simple distance measure methods and then presents the "angle-order" method, which is more complex. The "angle-order" method is now being tested in several vegetation types in cooperation with western Forest Service experiment stations, but the results are not yet conclusive. This translation should be useful to persons interested in methods studies.

Meredith J. Morris

James P. Blaisdell

Division of Range Management and Wildlife Habitat Research

Forest Service, U.S. Department of Agriculture

May 23, 1960

A new method for the estimation of density by the spacing method applicable to non-randomly distributed populations

Masaaki Morisita
Department of Biology, Faculty of Science, Kyushu University

SUMMARY

1. Although the spacing methods which utilize the spacing distance for the sampling of populations have been much developed in recent years, the theoretical considerations on the estimation of population density through the use of these methods are still concerned only with randomly distributed populations. The method of density estimation described in this paper seems to be available for ecological works since it is applicable to any populations, either randomly or non-randomly distributed.

2. The theoretical procedures of the method are as follows: Even if T individuals are distributed irregularly over an area (A), the area may possibly be divided into several small fractions (A_i, i=1, 2, 3, ..., δ) on which the individuals contained therein are distributed randomly or uniformly.

By placing N sample points randomly on the total area A, dividing the circle of infinite radius surrounding each sample point into k sectors, and measuring the distance (r) to the n-th nearest individual in each sector from the sample point, the following formulae can be derived when the individuals are distributed at random on A_i.

$$E\left(\frac{1}{r^2}\right) = \sum_{i=1}^{\delta} \frac{A_i}{A} \cdot \frac{2}{(n-1)!} \left(\frac{m_i}{k}\right) \int_0^{\infty} r^{2n-3} e^{-\frac{m_i}{k} r^2} dr = \frac{m}{k(n-1)} \qquad (1)$$

$$E\left(\frac{1}{\omega}\right) = \sum_{i=1}^{\delta} \frac{A_i}{A} \frac{m_i^{nk}}{(nk-1)!} \int_0^{\infty} \omega^{nk-2} e^{-m_i \omega} d\omega = \frac{m}{nk-1} \qquad (2)$$

where

$$m_i = \pi \frac{T_i}{A_i}, \quad T_i = \text{number of individuals on } A_i,$$

and

$$m = \pi \frac{T}{A}, \quad \omega = \frac{1}{k} \sum_{j=1}^{k} r_j^2.$$

By putting

$$\hat{m}_1 = \frac{n-1}{N} \sum_{i=1}^{N} \sum_{j=1}^{k} \frac{1}{r_{ij}^2} \qquad (3)$$

and

$$\hat{m}_2 = \frac{nk-1}{N} \sum_{i=1}^{N} \frac{k}{\sum_{j=1}^{k} r_{ij}^2}, \tag{4}$$

$\hat{m}_1$ and $\hat{m}_2$ become the unbiased estimates of m.

When the individuals are uniformly distributed on A_1, the following relation is expected;

$$\hat{m}_1 < m < \hat{m}_2.$$

Hence,

$$\hat{m}_0 = \frac{\hat{m}_1 + \hat{m}_2}{2} \tag{5}$$

may be used as an estimate of m when $\hat{m}_1 < \hat{m}_2$.

When $\hat{m}_1 > \hat{m}_2$, m should be estimated by $\hat{m}_1$, instead of $\hat{m}_0$.

Since we can get the values of $\hat{m}_1$ and $\hat{m}_2$ without using m_1 and A_1, the estimation of m by $\hat{m}_0$ or $\hat{m}_1$ seems to be applicable to any kinds of patterns of spatial distribution of individuals ($n \geq 3$).

3. It is considered that regular sampling, instead of random sampling, is also available for the density estimation by the method mentioned above, unless the individuals are distributed regularly on A_1.

4. The results of the estimation of m by $\hat{m}_0$ (k=4, n=3) applied to various artificial populations are shown in Table 1. It may be said from the results that the method is sufficient to be put to practical use at least when k=4 and n=3 are employed.

INTRODUCTION

The spacing, or distance method, which attempts to determine the distribution of individuals by the measurement of the distance between the sample point and the individual, or distance between the plants (or animals), has received much attention by ecologists in recent years because the method has the big advantage of ease of calculation and does not have the disadvantages of the quadrat method, e.g., the result depends on the size of quadrat. (Dice 1952, Cottam and Curtis 1949, 1953 etc.). Since the theoretical expression for a randomly distributed population was found by Moore, Morisita, Hopkins and Skellam, and Clark and Evans independently in 1954, many theoretical methods have been developed using the theoretical

expression. Examples are: the angle method by Morisita 1954, Clark and Evans 1954; the order method and n-th neighbour method by Morisita 1954, Thompson 1956; and the grouping method by Morisita 1955, Clark 1955 a, b, etc.

The theory of the spacing method, however, has been considered only for the spatial population in the normal hexagonal distribution and the square distribution, or for the randomly distributed population. It seems to be considered that it is very difficult and almost impossible to deal with nonrandomly distributed populations by a generalized procedure because there are so many different cases for nonrandomly distributed populations. Nothing has been done about this difficulty so far, and the theoretical equation of the spacing method for randomly distributed population has been applied for the estimation of density of biological species which usually has a nonrandomly distributed population. (e.g., Cottam and Curtis 1956).

The method described in this paper, however, can be applied for almost all cases, either randomly or nonrandomly distributed populations. Although the theoretical procedure for this method still lacks completeness, the estimation of density can be done practically and with good accuracy by using many sample points. Therefore, I am going to publish the general idea of this method, adapting the incomplete theory, and I should like to add some further considerations and developments in other papers.

FUNDAMENTAL THEORY

1. Randomly Distributed Populations

a) Shortest distance method[1]

If individuals are distributed randomly on a certain area, we can set the sample points randomly on the area. Then the shortest distance, r, between each sample point and the nearest individual to each of the sample points can be measured. The probability density of distance r can be expressed by the equation,

$$f(r) = 2mre^{-mr^2} \qquad (1)$$

where m is the average number of individuals in an area of π (that is, $m = \pi\rho$ where ρ is the average number of individuals per unit area) (Morisita 1954).

From equation (1),

$$E(r^2) = \frac{1}{m} . \qquad (2)$$

---

[1] Named by Morisita (1954).

If the total number of sample points is N, then $\hat{m}$ unbiased estimates of m can be expressed by the equation,

$$\hat{m} = \frac{N-1}{\sum_{i=1}^{N} r_i^2} \tag{3}$$

(Moore 1954). That is, the estimated value of m can be obtained by dividing (N-1) by the sum of the r^2. In this case, the variance of $\hat{m}$ is given by the equation,

$$\sigma^2_{\hat{m}} = \frac{m^2}{N-2} \tag{4}$$

b) Angle method

Dividing the area around the sample point into equiangle sectors, the distance r between the sample point and the nearest individual in each sector can be measured, and $\hat{m}$ can be expressed by the equation,

$$\hat{m} = k\,\frac{kN-1}{\sum_{i=1}^{kN} r_i^2} \tag{5}$$

where k is the number of sectors, and

$$\sigma^2_{\hat{m}} = \frac{m^2}{kN-2} \tag{6}$$

This corresponds to the case in the shortest distance method; here m, $\hat{m}$, and N are replaced by m/k, $\hat{m}$/k, and kN respectively. The reason for replacing m by m/k is that the average number of individuals within the circle of r=1 around the sample point in each sector is m/k. In like fashion N is replaced by kN because the number of measurements of r is kN for N sample points.

c) Order method

This method consists of measuring the distance between the sample point and the 2nd, 3rd, 4th, ..., n-th nearest individuals. The shortest distance method corresponds to the case where n=1 in this method. The probability density of r in this method is then expressed by the equation,

$$f(r) = \frac{2}{(n-1)!}\, m^n\, r^{2n-1}\, e^{-mr^2}\ ; \tag{7}$$

therefore

$$E(r^2) = \frac{n}{m} \qquad (8)$$

(Morisita 1954).

In this case, $\sum_{i=1}^{N} r_i^2$ follows the relationship of $\Gamma(nN, m)$. Replacing $r^2 = x$,

$$E\left(\frac{nN}{\sum_{i=1}^{N} r_i^2}\right) = nN\int_0^{\infty} \frac{m^{nN}}{(nN-1)!} \cdot \frac{1}{x} \cdot x^{nN-1} e^{-mx}\, dx = \frac{nN}{nN-1}\, m \; ; \qquad (9)$$

therefore $\frac{nN}{\sum_{i=1}^{N} r_i^2}$ is not the unbiased estimate of m.

If we substitute

$$\hat{m} = \frac{nN-1}{\sum_{i=1}^{N} r_i^2} \quad , \qquad (10)$$

then

$$E(\hat{m}) = E\left(\frac{nN-1}{\sum_{i=1}^{N} r_i^2}\right) = (nN-1)\, E\left(\frac{1}{\sum_{i=1}^{N} r_i^2}\right) \qquad (11)$$

$$= m$$

therefore $\hat{m}$ is the unbiased estimate of m.

In other words, we can obtain the estimated value of m by replacing (N-1) terms in equation (3) of the shortest distance method by the term (nN-1). The distribution of $\hat{m}$ in this case is then given by the equation,

$$E\left(\frac{1}{\left(\sum_{i=1}^{N} r_i^2\right)^2}\right) = \int_0^{\infty} \frac{m^{nN}}{(nN-1)!} \cdot \frac{1}{x^2} \cdot x^{nN-1} \cdot e^{-mx}\, dx = \frac{m^2}{(nN-1)(nN-2)} \qquad (12)$$

and

$$E\left((\hat{m})^2\right) = E\left(\frac{(nN-1)^2}{\left(\sum_{i=1}^{N} r_i^2\right)^2}\right) = \frac{nN-1}{nN-2}\, m^2 \qquad (13)$$

then

$$\sigma^2_{\hat{m}} = E\left((\hat{m})^2\right) - E^2(\hat{m}) = \frac{m^2}{nN-2} \qquad (14)$$

d) Angle-order method[2]

This is a combination of the angle method and the order method. If we measure distances to the n-th individuals in each of the k sectors, the unbiased estimate of m, $\hat{m}$, can then be expressed by equation (15) which is similar to equation (5).

$$\hat{m} = \frac{nkN-1}{\sum_{i=1}^{kN} r_i^2} k \qquad (15)$$

$$\sigma^2_{\hat{m}} = \frac{m^2}{nkN-2} \qquad (16)$$

The above mentioned are methods of estimating the density of randomly distributed populations. The estimation with the greatest accuracy is the angle-order method.[3] This can easily be seen by comparing the $\sigma^2_{\hat{m}}$'s for each method. In this case if we take the larger k and n, we can get the smaller $\sigma^2_{\hat{m}}$, thus obtaining a greater accuracy of estimation. For the practical outdoor measurement, however, k=4, n=3~4 might be the most efficient method.

2. Aggregatedly Distributed Populations

If we take an area above a certain size, the distribution of biological individuals in a natural state usually tends to be aggregatedly distributed to a greater or less extent.

Aggregate distribution, in a general sense, can be defined as the state in which the distribution density of individuals is not homogeneous but depends on parts in a given space; in other words, the parts of high density and low density coexist beyond the extent where it can be of casual occurrence.

Assuming T individuals are distributed irregularly over an area A, the area can be divided theoretically into several small fractions in which no aggregated distribution can be observed. In this case, the density is not necessarily the same in the different fractions. If the number of small fractions thus obtained is δ, the number of individuals in each fraction is T_i ($i=1,2,3,\ldots,\delta$), the area of each fraction is A_i, and the density of individuals in each fraction is m_i/π (the number of individuals in a unit area), the following relation can be obtained:

$$m_i = \frac{\pi T_i}{A_i}, \qquad T = \sum_{i=1}^{\delta} T_i, \qquad A = \sum_{i=1}^{\delta} A_i .$$

---

[2] This name is given in this paper.

[3] It was mentioned by Morisita (1954).

The average density of individuals, (m/π), over the total area A is given by the equation,

$$\frac{m}{\pi} = \frac{T}{A} = \frac{1}{A} \sum_{i=1}^{\delta} \frac{m_i}{\pi} A_i \tag{17}$$

Here we are dealing with small fractions in which no more aggregation can be observed; therefore, the distribution of individuals in the area A_i should be either random or fairly uniform.[4] Now let us consider each of these possible two cases respectively.

a) Random case

Now placing sample points randomly in the total area A, the probability of finding one sample point in A_i is A_i/A. Whereas the distribution of individuals in area A_i is random, the probability density of the distance r between the sample point and the n-th nearest individual can be derived from equation (7),

$$f(r) = \frac{2}{(n-1)!} m_i^n \, r^{2n-1} \, e^{-m_i r^2} \tag{18}$$

Considering now the new variable $1/r^2$, the average value of $1/r^2$ over all A_i's (for total A) can be obtained from the equation,

$$E\left(\frac{1}{r^2}\right) = \int_0^{\infty} \frac{1}{r^2} \left[\sum_{i=1}^{\delta} \frac{A_i}{A} \frac{2}{(n-1)!} \, m_i^n \, r^{2n-1} \cdot e^{-m_i r^2} \right] dr$$

$$= \frac{1}{A(n-1)} \sum_{i=1}^{\delta} A_i m_i = \frac{m}{n-1} \tag{19}$$

The variance of $1/r^2$ is

$$\sigma^2_{1/r^2} = E\left(\frac{1}{r^4}\right) - E^2\left(\frac{1}{r^2}\right) = \frac{1}{A(n-1)(n-2)} \sum_{i=1}^{\delta} A_i m_i^2 - \frac{m^2}{(n-1)^2} \tag{20}$$

Assuming the total number of sample points in the total area A is N, and

$$\hat{m} = \frac{n-1}{N} \sum_{j=1}^{N} \frac{1}{r_j^2} \tag{21}$$

then

---

[4] Self-spacing arrangement by Torii (1952). In general it is called "under dispersion."

$$E(\hat{m}) = E\left(\frac{n-1}{N}\sum_{j=1}^{N}\frac{1}{r_j^2}\right) = \frac{n-1}{N}\sum_{j=1}^{N} E\left(\frac{1}{r_j^2}\right) = \frac{n-1}{N} N \frac{m}{n-1} = m \qquad (22)$$

that is, $\hat{m}$ is the unbiased estimate of m. The variance of $\hat{m}$ is

$$\sigma^2_{\hat{m}} = \frac{(n-1)^2}{N^2}\sum_{j=1}^{N}\sigma^2_{1/r_j^2} = \frac{1}{N}\left\{\frac{1}{A}\left(1+\frac{1}{n-2}\right)\sum_{i=1}^{\delta} A_i m_i^2 - m^2\right\} \qquad (23)$$

From equation (21) it is seen that if the distribution of individuals in small fractions is random, the estimate of m can be obtained by placing sample points randomly over the total area, ignoring the difference of density between fractions, and measuring the distance to the n-th individual and then simply summing up the $1/r^2$ terms and multiplying by n-1/N. The individuals can be distributed in any fashion depending on the combination of A_i's and m_i's. Equation (21), however, does not contain either A_i or m_i; therefore the estimation using equation (21) can be applicable not only to any aggregated population but also to the population in which A_i and m_i are unknown (providing that the distribution of the individuals in A_i is random). It can be seen that equation (21) is very useful and effective to estimate the density of individuals in the population where the type of distribution is unknown. The estimate of total number of individuals over A, $\hat{T}$, can be obtained by the use of the equation

$$\hat{T} = \frac{\hat{m}}{\pi} A \qquad (24)$$

Assuming $\hat{M}$ as the estimate of the average area for one individual, then

$$\hat{M} = \frac{\pi}{\hat{m}} \qquad (25)$$

The more accurate method of estimation can be obtained by combining the angle method just as we have done for randomly distributed populations. In this case we have two possible procedures.

(i) Calculate $1/r^2$ for each sector.

By substituting $\hat{m}_1$ as the estimate of m for this method and k for the number of sectors, we can calculate $\hat{m}_1$ by replacing $\hat{m}$ in equation (21) by $\hat{m}_1/k$ and replacing N by kN. That is,

$$\hat{m}_1 = k\,\frac{n-1}{kN}\sum_{j=1}^{kN}\frac{1}{r^2_j} = \frac{n-1}{N}\sum_{j=1}^{kN}\frac{1}{r^2_j} \qquad (26)$$

$$\sigma^2_{\hat{m}_1} = \frac{1}{kN}\left\{\frac{1}{A}\left(1+\frac{1}{n-2}\right)\sum_{i=1}^{\delta} A_i m_i^2 - m^2\right\} \qquad (27)$$

(ii) Calculate $\sum_{j=1}^{k} r_j^2$ for each sample point.

By substituting the average value of r^2 for each of the sample points as ω, we obtain

$$\omega = \frac{1}{k} \sum_{j=1}^{k} r^2_j \tag{28}$$

If the distribution of individuals in A_i is random, ω should satisfy the relation $\Gamma(kn, m_i)$; therefore

$$E\left(\frac{1}{\omega}\right) = \sum_{i=1}^{\delta} \frac{A_i}{A} \frac{1}{(nk-1)!} m_i^{nk} \int_0^{\infty} \omega^{nk-2} e^{-m_i \omega} d\omega = \frac{m}{nk-1} \tag{29}$$

$$\sigma^2_{1/\omega} = \frac{1}{A(nk-1)(nk-2)} \sum_{i=1}^{\delta} A_i m_i^2 - \frac{m^2}{(nk-1)^2} \tag{30}$$

By substituting $\hat{m}_2$ as the estimate of m in this method, we obtain

$$\hat{m}_2 = \frac{nk-1}{N} \sum_{i=1}^{N} \frac{1}{\omega_i} = \frac{nk-1}{N} \sum_{i=1}^{N} \frac{k}{\sum_{j=1}^{k} r^2_{ij}} \tag{31}$$

$$\sigma^2_{\hat{m}_2} = \frac{1}{N} \left\{ \frac{1}{A} \left(1 + \frac{1}{nk-2} \sum_{i=1}^{\delta} A_i m_i^2 - m^2 \right) \right\} \tag{32}$$

The main difference between method (i) and (ii) can be summarized as follows: In method (i), the density in each sector is calculated first, and then the density over the total area is estimated from the average of the densities of the sectors for all the sample points. In method (ii) the average density around the sample point is calculated first, and then the density over the total area is estimated from the average unit of the average densities for all the sample points. If a sample point is placed far away from the center of an aggregate, or if there is a small aggregate which can be in a sector, that is, if there is a significant difference in density between sectors of a sample point, method (i) would be more suitable. In general, the method which has the better accuracy can be judged by whether the value of $(\sigma^2_{\hat{m}_1} - \sigma^2_{\hat{m}_2})$ is less than or greater than zero.

From equation (27) and (32), if

$$\frac{\frac{1}{A} \sum_{i=1}^{\delta} A_i m_i^2}{m^2} < \frac{(n-2)(nk-2)}{n^2k - 2n(k+1) + 2}$$

then $\sigma^2_{\hat{m}_1} > \sigma^2_{\hat{m}_2}$; therefore method (ii) has better accuracy (providing that $n \geq 3$, $k \geq 2$). If the distribution of individuals over the total area is random, method (ii) has greater accuracy in the case where $n \geq 3$, and $k \geq 2$.

I have previously mentioned that the methods mentioned above are applicable to any aggregated population, but it must be emphasized that they are applicable providing the distribution of individuals in A_i is random. In many cases, however, the distribution of biological species are relatively uniform within their aggregate, even if they have an aggregated form as a whole. Therefore, as a general method of estimating density which is applicable to such a case, the method mentioned above containing only $\hat{m}_1$ or $\hat{m}_2$ is not sufficient. Now let us consider this point in the next chapter.

b) Fairly uniform case

In this case it is very difficult to obtain the unbiased estimate of m by the method previously mentioned, and we have to be satisfied with an appropriate estimate with relatively less variance.

If the distribution of individuals on A_i is not random but is more or less uniform, the variance of r taken k times would be smaller than the variance in the random case,[5] because if the distribution is uniform, the range that the value of r can take is much smaller than the range for the random case. Therefore, if the m's are the same for both cases, the value of $\sum_{j=1}^{k} r^2_j$ for a uniform distribution is smaller than the value of $\sum_{j=1}^{k} r^2_j$ for a random distribution. Similarly the value of $\sum_{j=1}^{k} \frac{1}{r^2_j}$ is smaller for a uniform distribution than for a random distribution. Then, if the distribution of individuals on A_i is uniform, the value of $\hat{m}_1$ obtained by

$$\hat{m}_1 = \frac{n-1}{N} \sum_{j=1}^{kN} \frac{1}{r^2_j} = \frac{n-1}{N} \sum_{i=1}^{N} \sum_{j=1}^{k} \frac{1}{r^2_{ij}}$$

is smaller than m, and the value of $\hat{m}_2$ obtained from equation (31),

$$\hat{m}_2 = \frac{nk-1}{N} \sum_{i=1}^{N} \frac{k}{\sum_{j=1}^{k} r^2_{ij}}$$

---

[5] It has been made clear also from the results of the normal hexagonal distribution and artificial uniform distribution which will be mentioned in a later part of this paper.

is expected to be greater than m. That is, if the distribution of individuals on A_1 is regular, the following relation is rationalized

$$\hat{m}_1 < m < \hat{m}_2 ,$$

and it is possible to estimate m by the value of $\hat{m}_0$ which is defined by the equation

$$\hat{m}_0 = \frac{a\hat{m}_1 + b\hat{m}_2}{a+b} \tag{33}$$

In this case, however, it is impossible to obtain the values of a and b by theoretical considerations. For the observed value in a hexagonal distribution which will be mentioned in a later part of this paper, we can obtain sufficiently good estimates by letting $a = b = 1$. Therefore, it seems reasonable to use $a = b = 1$ for the general distribution with less uniformity (closer to a random distribution), if we take $n \geqq 3$ and $k \geqq 4$. That is, practically we can use equation (34) to estimate the value of m.

$$\hat{m}_0 = \frac{\hat{m}_1 + \hat{m}_2}{2} \tag{34}$$

We considered $\hat{m}_0$ of equation (34) as an estimate of m when the individuals are uniformly distributed on A_1. However, we can use $\hat{m}_0$ in either case; when the total distribution is random, or when the total distribution is uniform. In other words, we can apply this method to any type of distribution without considering the nature of aggregation or type of distribution. However, when $\hat{m}_1 > \hat{m}_2$, that is, the density varies depending on the sectors even if the sample point is identical, $\hat{m}_1$, which is based on $1/r^2$ for each sector, should be used as the estimate of m.

3. Restriction on the value of n

It is to be seen from equation (21) and (26) that the above mentioned methods of estimating the density from $\hat{m}_0$ or $\hat{m}_1$ cannot be applied if n=1. It is also clear from equation (23) and (27) that they are not applicable when n=2 because the variances of $\hat{m}$ or $\hat{m}_1$ become infinite. Therefore, the estimation from $\hat{m}_0$ or $\hat{m}_1$ should take at least n=3, that is, we have to estimate the density by the measurement of distances to at least the third nearest individual.

4. The method which uses regularly spaced sample points

The methods I have mentioned are based on the theoretical consideration of using randomly selected sample points. In the practical outdoor measurement, however, it needs additional procedure to select the sample

points randomly because we have to decide the randomized position beforehand by some means, e.g., using the randomized table. Many practical outdoor measurements have been carried out without deciding the randomized point beforehand just by walking a certain distance from one position to another arbitrarily. These arbitrarily chosen sample points tend to be more regularly placed in position than randomly chosen points. From this view, it seems to be worthwhile to consider the method of using regularly spaced sample points.

When we place sample points regularly on the area A, if the distribution of individuals is random, the estimation of m by use of equation (26) and (31) would be close enough because the relation between any one of the sample points on A_i and the individuals around it is quite accidental, although there is a certain spatial relation between the sample points. In this point it is in the same situation as in randomly selected sample points. Furthermore, we can keep the following relation for any small positive number ϵ by taking a larger N, e.g.,

$$\Pr.\left\{\left|\frac{N_i}{N}-\frac{A_i}{A}\right|>\epsilon\right\}\to 0$$

If $N_i/N = A_i/A$, it can be easily proven that $\hat{m}_1$ and $\hat{m}_2$, as obtained from equation (26) and (31), are the unbiased estimates of m, even if the sample points are regularly placed.

Now the variance of a number of sample points, N_i, which are in a particular area A_i, should be considered. If the sample points are placed randomly, the probability distribution of N_i on repeated sampling follows the binomial distribution of $p=A_i/A$, and its variance is $Npq(q=1-p)$; therefore if A_i/A is small, the variance comes close to the average value. While the variance of the probability distribution of N_i in regularly spaced sample points would be much smaller than the average value,[9] the probability distribution of N_i is the same as the probability distribution of individuals for the area when we apply random sampling to the sample points within the area of A_i. In a general aggregated population the deviation within A_i would be smaller for regularly placed sample points than for the randomly placed case, because A_i/A would be expected to be the smaller value in many cases. Therefore, if $E(N_i/N) \approx A_i/A$, the

---

[9] An example; placing many points on the square distribution, taking 2.5 cm. as the distance between two points, and place the circles of r=6 cm. randomly on this figure. The result of counting the points which are in the circles;

The calculated average value 18.096

The observed average value taken 120 times 18.025

The variance = 1.758 (variance)/(average) = less than 0.1%

accuracy of density estimation over the total area A can be possibly considered better for regularly spaced sample points than for randomly selected sample points. When $N_1/N = A_1/A$, it is easily proven that the regular spacing has better accuracy than the random selection.

The most difficult case in the method with regularly spaced sample points is when the distribution of the individuals on A_1 is uniform. As an extreme case, consider when the distribution of the individual is the same uniform distribution as the sample points, then the above mentioned density estimation from $\hat{m}_0$ cannot be applied because there is the same spatial relation between the sample point and the individual. As another case, consider when the space between the individuals is constant, and if the ratio of space between individuals and space between sample points is an integral number, and the orientation of both are also parallel, then the spatial relation between the sample point and the individual become uniform, and the estimation from $\hat{m}_0$ is not applicable. In the case of the distribution of plants, however, such special cases are rare except in the case of forest plantation even if we deal with such regularly distributed plants, it hardly happens to be the above mentioned special cases, unless we arrange the distance, directions, etc. of sample points in such a way on purpose. Thus, if we take $n \geqq 3, k \geqq 4$, the variance of estimate caused by regularity of sample points can be considered practically small enough. This can be seen by the examples shown in the latter part of this paper.

To eliminate the troubles mentioned above the following method can be considered, that is, by dividing the total area into many small plots of equal area, and placing one or more sample points randomly on each plot.[7] By this method, we can avoid the possibility of placing many sample points in one small area which may happen when we place sample points randomly over the total area. Furthermore, we can make the variance of estimate smaller (although the distribution of sample points over the total area becomes rather regular) because we can avoid the uniform spatial relation between the sample points and the individuals on A_1, even if the distribution of individuals is uniform. The method with regularly spaced sample points, however, seems to be more practical, considering the ease of spacing procedure of sample points, if the individuals are distributed less regularly in each aggregate.

PRACTICAL PROCEDURE OF DENSITY ESTIMATION

The method of estimating density from $\hat{m}_0$ (when $\hat{m}_1 < \hat{m}_2$) or $\hat{m}_1$ (when $\hat{m}_1 > \hat{m}_2$) is applicable to any population where the biological species are distributed randomly, aggregatedly, or uniformly. Therefore, even if the distribution of individuals is unknown, the estimate of the average density can be obtained by the following mechanical procedure.

---

[7] This method is also recommended for the square method (e.g., Torii 1952).

1. Determine N positions as the sample points randomly or regularly over the total area, or divide the total area into many equal sized plots, and place an equal number of sample points randomly in each plot.

2. Divide the circle surrounding each sample point into k equiangle sectors. Practically k=4; placing two pieces of rope or rods at right-angles on the sample point would be the easiest way.

3. Measure the distance r between a sample point and the n-th nearest individual in each sector. $n \geqq 3$ should be satisfied. If k=4, then n=3~4 would be the most practical. Of course we can obtain greater accuracy by taking a larger n.

4. Calculate $1/r^2$ for the each value of r which is taken k times for each sample point; sum up the $1/r^2$'s for each sample point, and then sum up the $1/r^2$'s for all the sample points.

5. Multiple the sum of $1/r^2$ for all the sample points obtained in step 4 by n-1/N. This is the value of $\hat{m}_1$.

6. Calculate $\sum_{i=1}^{k} r^2_i$ for each sample point, and find $\frac{1}{\sum_{i=1}^{k} r^2_i}$; then sum up these values for all the sample points.

7. Multiple the value obtained in step 6 by k(nk-1)/N. This is the value of $\hat{m}_2$.

8. Compare $\hat{m}_1$ and $\hat{m}_2$;

If $\hat{m}_1 \leqq \hat{m}_2$, find $\hat{m}_0 = \hat{m}_1 + \hat{m}_2/2$, and then calculate $\hat{m}_0/\pi$. This is the estimate of the average number of individuals per unit area. The total number of individuals in the total area can be obtained by multiplying this value by the total area.

If $\hat{m}_1 > \hat{m}_2$, $\hat{m}_1/\pi$ instead of $\hat{m}_0/\pi$, is the estimate of the average number of individuals per unit area.

EXAMPLES OF APPLICATION OF THE METHOD FOR ARTIFICIAL POPULATIONS

1. Completely Regular Distribution

As an example of regular distribution, let us consider the distribution where individuals are placed at each apex and the center of a normal hexagon. Placing 96 sample points almost uniformly in a diamond which is formed by the nearest four points of normal hexagonally distributed individuals (the distance between two individuals is 4 cm.) and taking k=4, measurement of the distance of n=3 was repeated in six directions for each

sample point, rotating the two axes which divide 4 plots on each sample point 0°, 15°, 30°, 45°, 60°, 75° respectively. The results are summarized in Table 1, A. According to this result, the difference between $\hat{m}_0$ (the estimate of m) and m is less than 0.1%.

Table 1. Results of the density estimation through the use of the spacing method. (n=3, k=4). m = average number of individuals in the area π. $\hat{m}_0$ = estimate of m.

Distribution		Number of sample points	$\hat{m}_1$	$\hat{m}_2$	Estimated value $\hat{m}_0$	Estimated value Total number of individuals	m	Actual value Total number of individuals
A Regular hexagonal distribution		96*	0.19924	0.25408	0.22666	—	0.22672	—
B Uniform distribution		188	0.01271	0.01553	0.01412	1,573	0.01438	1,602
Aggregated distribution (Fig. 1)	I	36	0.14695	0.16923	0.15909	1,701	0.16406	1,755
	II	36	0.16031	0.17493	0.16762	1,793		
	III	56	0.15065	0.17122	0.16094	1,721		
	C_1** I + II + III	128	0.15214	0.17153	0.16184	1,731		
	I'	21	0.15361	0.17634	0.16499	1,765	0.16406	1,755
	II'	21	0.16477	0.18055	0.17266	1,847		
	C_2*** I'+ II'+ III	98	0.15432	0.17418	0.16425	1,757		

* The measurement was repeated in 6 directions for each sample point.

** Sample points I, II, III and I + II + III (Fig. 2) were used respectively.

*** Sample points on the border line of the examined area (Fig. 2) were excluded.

2. Fairly Uniform Distribution

1602 points were marked fairly uniformly on one sheet of cross-section paper of 50 x 70 $cm.^2$ In order to know the degree of uniformity, dividing the total area into 1 $cm.^2$ and 4 $cm.^2$ sections, the frequency of occurrence in each section was compared with the Poisson distribution. The results are summarized in Table 2.

The ratios of (variance)/(average value) are 0.594 for the 1 $cm.^2$ section and 0.401 for the 4 $cm.^2$ section; this means the distribution is fairly uniform.

Placing the sample points randomly by the use of the randomized table on this figure, distances were measured taking n=3 and k=4. Table 1, B shows the result of measurements for 188 sample points excepting the sample points which are nearer to the edge of the cross-section paper than to the 3rd nearest individuals. According to this result, the difference between actual density and estimate is only 1.8% of actual density.

Table 2. Comparison of the distribution of individuals of an artificial population distributed uniformly over an area of 50 x 70 $cm.^2$, (Table 1,B) with Poisson distribution.

Number of individuals per quadrat	1 $cm.^2$ quadrat		4 $cm.^2$ quadrat	
	Observed number of quadrats	Expected number of quadrats	Observed number of quadrats	Expected number of quadrats
0	1,939	2,214	33	140
1	1,520	1,014	281	257
2	41	232	389	235
3	0	35	146	143
4	0	4	25	66
5	0	0	1	24
6 and over	0	0	0	10
Total	3,500	3,499	875	875
$P\{X^2\}$	< 0.01		< 0.01	
$\frac{\text{Variance}}{\text{Mean}}$	0.594		0.401	

3. Aggregated Distribution

As an example of the aggregated distribution, an artificial figure shown in Fig. 1[8/] was made. When we made this figure, we paid attention only to making aggregates of varying density and size and did not pay attention to the distribution of points in an aggregate. As a result, however, it is a typical aggregated distribution as a whole, but the distribution of points within the aggregates can be considered as a fairly uniform distribution (Table 3). Placing three types of sample point systems regularly on this figure, Fig. 2,[8/] the density was measured in the area shown by the broken line on Fig. 2 taking k=4 and n=3. Some sample points in system I and system II are placed on the edge of the rectangle;

Table 3. Comparison of the distribution of individuals of the artificial population shown in Fig. 1 with Poisson distribution.

Number of individuals per quadrat(1 $cm.^2$)	(a) Total population*		(b) A clump**	
	Observed number of quadrats	Expected number of quadrats	Observed number of quadrats	Expected number of quadrats
0	29	1.8	0	
1	45	9.5	0	
2	37	24.8	0	
3	26	43.1	0	
4	40	56.2	3	14.8
5	22	58.7	0	
6	19	51.0	2	
7	21	38.1	4	
8	24	24.8	9	9.6
9	23	14.4	9	
10	13	7.5	6	
11	9	3.6	2	11.6
12	7	1.5	1	
13 and over	21	0.9	0	
Total	336	335.9	36	36.0
$P\{X^2\}$	< 0.01		< 0.01	
Variance / Mean	3.28		0.41	

* The area enclosed by the broken line in Fig. 2 was examined.
** The area enclosed by the dotted line in Fig. 2 was examined.

8/ [Figures 1 and 2 are not reproduced. Fig. 1 is a contagious or aggregated distribution of points having several clumps of varying density. Fig. 2 shows the arrangement of sample points for the density estimation of the artificial population which was shown in Fig. 1. This arrangement consisted of three overlapping grids (2 diamond, 1 rectangular)].

on these sample points distances were measured only in the sectors which are inside of the rectangle. By this treatment some sample points were used only in one or two sectors out of four; therefore, the following treatment was applied for calculation.

Substituting the total number of sectors used for measurement as K, and the number of sectors used for measurement in each sample point as k', $\hat{m}_1$ can be obtained by the equation,

$$\hat{m}_1 = \frac{k(n-1)}{K} \sum_{i=1}^{N} \sum_{j=1}^{k'} \frac{1}{r_{ij}^2} \tag{35}$$

Substituting the value of $\hat{m}_2$ at each sample point as $\hat{m}_{2\cdot i}$ (that is, $\hat{m}_2$ when N=1), $\hat{m}_{2\cdot i}$ is given by the equation,

$$\hat{m}_{2\cdot i} = k(k'n-1) \frac{1}{\sum_{j=1}^{k'} r_{ij}^2} \tag{36}$$

By taking the average of $\hat{m}_{2\cdot i}$ with the appropriate weighting for the number of sectors used, $\hat{m}_2$ can be obtained by the equation,

$$\hat{m}_2 = \frac{\sum_{i=1}^{N} k' \hat{m}_{2\cdot i}}{K} \tag{37}$$

When we use the sample points which are used only in one or two sectors, however, the value of $\hat{m}_0$ tends to be too small if the distribution is uniform. In the example of the aggregated population mentioned above, the $\hat{m}_0$ has the same tendency because the distribution of individuals in an aggregate is fairly uniform as shown in Table 3(b) (It is also seen by the fact $\hat{m}_1 < \hat{m}_2$). In order to avoid this tendency, it is better not to use sample points on the edge of the area. The result of the calculation without sample points on the edge is shown in Table 1, C_2. The results of (I′ + II′ + III′) are closer to the observed value than the value of (I + II + III) in C_1, in spite of decreased number of sample points. It is worthwhile to note that the error of estimation is about 5% in II′ and less than 1% in I′ in spite of the fact that the estimates were calculated from only 21 sample points. Measuring the density on the area, shown by the dotted line in Fig. 2, Table 3(b), we obtained the following result from the 8 sample points in this area.

$$\hat{m}_1 = 0.23132$$

$$\hat{m}_2 = 0.29366$$

$$\hat{m}_0 = 0.26249$$

and the estimate of the number of individuals calculated from $\hat{m}_0$ is 301, while there are 299 individuals; therefore, it can be said that they coincide very well.

Thus, we obtained good estimation of the density of Fig. 1, despite the uniformity of the distribution within the aggregate, by the regular spacing of the sample points. That is, this method is sufficient to be put to practical use, unless the distribution within the aggregate is completely uniform.

CONCLUSION

The method of estimating the density mentioned above has a great advantage in being applicable to any kind of distribution of individuals. Furthermore, it would be possible to show the density around each sample point with good accuracy by calculating $\hat{m}_0$ (or $\hat{m}_1$) for each sample point assuming N=1, if we take k and n large enough. It is also possible to show the distribution of the individuals graphically by drawing contour lines for the same density levels, using the estimates of the density of each sample point. Another method to obtain the general idea of the distribution is from $\hat{m}_3$:

$$\hat{m}_3 = k \frac{knN-1}{\sum_{i=1}^{N} \sum_{j=1}^{k} r_{ij}^2} .$$

If $\hat{m}_3$ is greater than, less than, or equal to $\hat{m}_0$, the type of total distribution is aggregate, random, or uniform.

For more detailed analysis of distribution, the method using $\hat{m}_1 - \hat{m}_2$ and $\hat{m}_2 - \hat{m}_3$ can be used to obtain the distribution within and between the aggregates. The appropriate expression to obtain a parameter which indicates the character of distribution has been found, and it is possible to analyse the various types of distributions by the combination of these parameters to some extent. The details of these methods will be published in other papers.

LITERATURE CITED

Clark, P. J., and F. C. Evans
1954. Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology 35 :445-453.

_____ and _____
1955. On some aspects of spatial pattern in biological populations. Science 121 :397-398.

Clark, P. J.
1956. Grouping in spatial distributions. Science 123 :373-374.

Cottam, G. and J. T. Curtis
1949. A method for making rapid surveys of woodlands by means of pairs randomly selected trees. Ecology 30 :101-104.

_____, _____ and B. W. Hale
1953. Some sampling characteristics of a population of randomly dispersed individuals. Ecology 34 :741-757.

_____ and _____
1956. The use of distance measures in phytosociological sampling. Ecology 37 :451-460.

Dice, L. R.
1952. Measure of the spacing between individuals within a population. Contrib. Lab. Vert. Biol. Univ. Mich. No. 55.

Hopkins, B.
1954. A new method for determining the type of distribution of plant individual. Ann. Bot. N.S. 18 :213-227.

Moore, P. G.
1954. Spacing in plant populations. Ecology 35 :222-227.

Morisita, M.
1954. Estimation of population density by spacing method. Mem. Fac. Sci. Kyushu Univ. Ser. E 1 :187-197.

[in Japanese, not translated]

Thompson, H. R.
1956. Distribution of distance to n-th neighbour in a population of randomly distributed individuals. Ecology 37 :391-394.

[in Japanese, not translated]