Some Sampling Characteristics of a Population of Randomly Dispersed Individuals

Grant Cottam; J. T. Curtis; B. Wilde Hale


Stable URL:
http://links.jstor.org/sici?sici=0012-9658%28195310%2934%3A4%3C741%3ASSCOAP%3E2.0.CO%3B2-C

Ecology is currently published by The Ecological Society of America.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/esa.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact jstor-info@umich.edu.
SOME SAMPLING CHARACTERISTICS OF A POPULATION
OF RANDOMLY DISPERSED INDIVIDUALS

Grant Cottam, J. T. Curtis, and B. Wilde Hale

Department of Botany, University of Wisconsin, Madison, Wisconsin

INTRODUCTION

As our knowledge of plant communities increases, greater emphasis is being placed on the methods used to measure the characteristics of these communities. Succeeding decades have shown a trend toward the use of quantitative methods, with purely descriptive studies becoming less common. One reason for the use of quantitative techniques is that the resulting data are not tinged by the subjective bias of the investigator. The results are presumed to represent the vegetation as it actually exists; any other investigator should be able to employ the same methods in the same communities and secure approximately the same data. If such is not the case, the so-called quantitative methods are worse than valueless, since they lend an aura of false accuracy that may be very misleading. To make certain that his results are reproducible, the worker must be sure that his sample is adequate, that the location of his sample has been made without bias, and that the limitations of his method are understood.

The student of plant communities is concerned with the distribution of objects in space. His first problem is the determination of the kinds of plants present and their relative numbers. Since it is usually impossible to examine all of the plants present, he confines himself to some kind of a sample of the vegetation. This may be in the form of area samples, as quadrats or belt transects; line samples, as in the line intercept method; exclusion angle methods; or any of a variety of point samples. In each of these cases, the results are affected in their accuracy by the actual dispersion of the plants on the ground. The plants may be regularly distributed as on the corners of squares or hexagons or in other definite patterns, so that every individual is equally distant from each of its nearest neighbors; they may be grouped into distinct clumps, as in the aggregated or so-called contagious distributions (Short 1845, Goodall 1952); or they may be distributed in an intermediate or random fashion.

Many of the methods now in use presuppose a random distribution of individuals in the population and are satisfactorily accurate when such conditions prevail. However, since many herbaceous plants are not randomly distributed (Ashby 1948, Whitford 1949, Steiger 1930), it is essential that a knowledge of their actual distribution be determined before the usual measurements may be interpreted properly. To date, no completely suitable method for assessing the random or non-random nature of plant distributions has been proposed. Before such a method is developed, it will be necessary to understand in detail the characteristics of a random population. This paper is concerned with an analysis of some of the characteristics of such a random population.

Thanks are due Miss Margaret Gilbert and Mr. Harold Liebherr for their aid in several phases of this investigation. Mr. Fred Gruenberger, of the University Computing Service, kindly supplied the random numbers table used in the construction of the artificial population.

---

1 This work was supported in part by the Research Committee of the Graduate School of the University of Wisconsin from funds supplied by the Wisconsin Alumni Research Foundation and in part by a contract between the Office of Naval Research, Department of the Navy and the University of Wisconsin (NR 164-056).
Artificial versus natural populations

The ideal way to study random plant distributions would be to find such a distribution in nature and apply field techniques to it but for obvious reasons this cannot be done. A more practicable approach is to construct a map of a population from a set of coordinates derived from a random numbers table which has already been subjected to exhaustive tests for randomness. Fortunately, such random numbers tables exist. The use of an artificial population map has other advantages in that it can be sampled over and over again without the inevitable destruction that accompanies field sampling; sampling points on the map can be located quickly and with great accuracy; and the parameters of the population are known.

Random distributions

The concept of a random distribution is not easy to define. In general, it can be said that each of the objects making up the population is located independently of all others, so that any object has an equal and independent chance of occurring at any locus. There is an absence of fixed pattern—the objects are not located at fixed distances from one another nor is there a marked gradient from regions where the population is dense to regions where it is rather sparse. Any major portion of the population will have approximately the same number of objects as any other equal portion, but within each portion there will be areas of greater than average density and there will be distinct bare areas. No two random populations will be exactly alike, but all will have the above mentioned characteristics.

Numerous tests have been used to determine whether or not a population is random. Perhaps the most well-known is comparison with a Poisson distribution, which is used with area samples (Blackman 1935). Briefly stated, if a number of area sampling units of equal size are laid down on the population, it is possible to determine from the Poisson distribution how many of them should include no individuals, how many should include one individual, two individuals, etc. A comparison of the actual results obtained from an area sample with the expected results from a Poisson distribution calculated from the sample mean, using a suitable test of significance, makes it possible to determine whether or not the results differ significantly from those that would be expected from a population with the individuals distributed at random. With artificial populations, the dispersion of the individuals is checked by dividing the map of the population into many equal areas by means of grid lines; each area is then considered to be a sample of the population equivalent to a quadrat of the same size. It should be pointed out that populations other than random populations can be constructed which will fit a Poisson distribution when sampled in this manner. An arrangement with the individuals distributed in a regular manner, but with the densities of different areas varying along a gradient from dense to sparse will fit a Poisson. Other populations could be constructed with the individuals grouped into various sizes of clumps that also might fit a Poisson distribution. In view of these facts, it is apparent that additional tests must be employed. One such test concerns geographical homogeneity. The population map may be divided into quarters, fifths, or into rows and columns, and the numbers of individuals present in each portion compared with the expected numbers by means of the chi-square test of homogeneity (Snedecor 1946). In this case it is assumed that each major area of the map will have an equal number of individuals, and the expected numbers are the total number of individuals divided by the number of equal areas into which the map has been divided. Greig-Smith (1952) has suggested grouping the individual grid units into blocks and using analysis of variance to determine variation between blocks as a method for determining geographical homogeneity.
An arrangement of the individuals showing absolute regularity, as exemplified by the trees in an orchard, will possess absolute geographical homogeneity, so this test also is insufficient by itself to determine randomness. Both of these tests are in a sense negative since failure of a population to conform to either of them definitely indicates that the population is not random, but the fact that a population does fit a Poisson distribution or does possess geographic homogeneity is no real assurance that the population is random. Use of both of these tests gives considerably more assurance than the use of either one of them alone, but the results can never be absolutely positive. However, the possibility of factors other than chance entering into the distribution of individuals is much less in an artificial population constructed from an adequately tested random numbers table than in a natural population in the field.

Methods

Construction of the population map

Several criteria were used in the planning of the population map employed in this paper. It was desired that the population contain a fairly large number of individuals and that the individuals be so designated that they could be grouped into species containing varying percentages of the total population. It was also desired that the ultimate grid units be at least as small as the finest unit used to measure distances. It was decided to have 1,000 individuals placed on a population map one square meter in area. The finest unit of the grid was one millimeter, so that there were one million points available for occupancy, only 0.1 per cent of which were actually used. The use of such a low percentage of the total possible points is an insurance against regularity. A coarse grid with most of the points occupied will have tendencies toward a regular distribution with the distance between individuals equal to the distance between points on the grid. The individuals were placed on the grid by means of a set of random coordinates supplied by the University of Wisconsin Computing Service. These coordinates were taken from a set of 400,000 random digits punched on IBM cards. Each IBM card had 40 columns of digits, and 6 of these columns were used. The columns to be used were chosen at random, and the deck of IBM cards was entered at a random point, so that the choice of numbers was entirely a matter of chance. The digits were tabulated directly onto a paper tape, and each set of 6 digits which made up a single pair of coordinates was numbered serially. The entire operation was performed by the machine so that there was no chance of personal bias in the choice or printing of the numbers.

The coordinates consisted of 6 numbers, such as 036 874. The first 3 represent the number of millimeters from the point of origin along the x axis, the last 3, the number of millimeters from the point of origin along the y axis. The base map was constructed on a piece of millimeter-ruled graph paper one meter square. Each decimeter line was heavily ruled and each centimeter line had a lighter ruling, but this ruling was heavier than the ultimate millimeter rulings. The decimeter lines represented the first digit of each coordinate and the centimeter lines the second digit. Through the use of these accented lines, and with the aid of an L-shaped ruler, ruled in centimeters, it was possible to locate each point on the map with fair speed. The intersections of millimeter lines on the graph paper were the exact locations of the points designated by the coordinates. Each individual was numbered serially as it was placed on the population, so that it was possible to establish species having any number of individuals from one to 1,000 by using blocks of the numbered individuals to represent a species.

After the base map had been completed, a tracing paper overlay was made, and this was reproduced by means of the
Ozalid process. The finished map contained none of the grid lines used to establish the locations of the individuals. It was deemed necessary to sample the population by means of a grid that was independent of the grid used to establish the locations of the individuals, since use of the same grid might result in a number of sampling points being placed directly on top of the individuals in the population.

The population was tested for randomness by means of the two tests previously described. For the geographical tests, the map was divided into quarters, fifths, 10 vertical columns, 10 horizontal rows, rectangular areas equal to 1/20 of the map area each, and square areas equal to 1/25 of the map area each. In all of these cases, the probability values were .50 or greater by the chi-square test for homogeneity, indicating that in a similar sample of another random population, the chances of getting a better fit were in all cases no better than even.

For the comparison with a Poisson distribution, 5 centimeter grid lines were used to divide the map into 400 equal areas, and the number of individuals was counted in each of these areas. The results are shown in Table I. The probability in this case is about .40, indicating that 60 times out of 100, a random population will give a better fit. It can be seen from the table that the deviations from the expected values follow no pattern. Regular distributions have too few samples with no individuals, too many samples with the mean number of individuals, and no samples with many individuals. Aggregated distributions, on the other hand, have too many samples with a large number of individuals and too many samples with no individuals. No marked trend toward either regularity or aggregation can be found in Table I. Rather, the deviations appear to be random fluctuations, with the greatest deviation being too few samples with 3 individuals, which is counterbalanced by too many samples with 4 and 5 individuals. Another Poisson distribution was run with the map divided into 100 equal areas. The probability was about the same, and again no trends were discernible.

The sampling of an artificial population presents one problem which is not usually serious when sampling in the field. This is the problem of the border. The individuals of the population have as much chance of being placed near the border as at any other location on the map. The points at which the sampling units fall can similarly be at any locus. Hence, it is possible for quadrats to be placed with parts of them extending over the edge of the map. The same difficulty arises when sampling with the exclusion angle methods used in this paper.

This problem can be avoided by discarding all sampling units, parts of which extend over the edge of the map, but this might result in an inadequate sample of the areas immediately adjacent to the edge. A more satisfactory alternative is to delimit an area immediately adjacent to the edge of the map as a border area. This area is considered to be an extension of the population within which no samples are laid down, but which is equal in density and species composition to the population in the interior. The actual population for which parameters are de-

**Table I. Comparison of the distribution of the individuals on the population with a Poisson distribution**

<table>
<thead>
<tr>
<th>Number of individuals per quadrat</th>
<th>Observed number of quadrats</th>
<th>Expected number of quadrats</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>31</td>
<td>32.9</td>
</tr>
<tr>
<td>1</td>
<td>87</td>
<td>82.2</td>
</tr>
<tr>
<td>2</td>
<td>104</td>
<td>102.6</td>
</tr>
<tr>
<td>3</td>
<td>71</td>
<td>85.5</td>
</tr>
<tr>
<td>4</td>
<td>61</td>
<td>53.4</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>26.6</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>11.1</td>
</tr>
<tr>
<td>7 and over</td>
<td>4</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>400</td>
</tr>
</tbody>
</table>

$\bar{x}=2.205$, $\chi^2=6.35$, $P=.40$
terminated and within which all the sampling points fall is thus reduced by the size of the border, but any samples which extend over the population into the border are counted just as though they were entirely within the population proper.

The width of the border was determined by balancing the desire to have the usable population as large as possible against the need to minimize border effects. It was decided that the border should be sufficiently wide that the distance from an individual in the interior of the population to its closest neighbor would be less than the distance from that individual to the outer edge of the border. A border equal in width to the square root of the mean area satisfies this requirement.

The parameters of the population with the border area excluded are as follows: area—8835.96 sq. cm.; number of individuals—882; mean area—10.018 sq. cm.; square root of mean area—31.65 mm. Mean area is defined as the average area occupied by a single individual and is obtained by dividing the area of the map by the number of individuals on it (Curtis and McIntosh 1950). The mean area of a species is the area of the map divided by the number of individuals of that particular species. The term, then, is an expression of the spacing of the individuals. In this paper, when the mean area of the entire population is being considered, mean area will be called population mean area.

When the individuals in the population are divided into 10 species by grouping serial numbers 1-100, 101-200, etc., relative densities vary from 9.41% to 10.54% as compared to the 10.00% value for each in the total population and mean areas of the species vary from 95.01 to 106.46 sq. cm. as compared to the 100.00 sq. cm. value for each in the total population. It can be seen that the exclusion of the border area has not greatly changed the parameters of the population.

**Sampling methods**

Since the individuals on the population were considered to be points without area, no measure involving cover could be used. Instead determinations of frequency and density were made. The methods used can be considered under two categories, the area methods and the exclusion angle methods involving distance measurements. No attempt was made to include a great variety of methods since the primary aim was not a study of methods, but rather a determination of the characteristics of the population as measured by various procedures. It was known on the basis of previous work with area samples (Cottam 1951) and with exclusion angle methods (Lieberr 1952) that related methods behave similarly on random populations.

The area samples were square quadrats of five different sizes. The sizes used were equal to 1/4, 1/2, 1, 2, and 4 times the population mean area. Since the quadrat method is well-known to all ecologists, its mechanics will not be discussed.

The exclusion angle methods rely on the measurement of distance to arrive at density per unit area. Simplest of these is the use of a sampling point to choose an individual on the population—the individual closest to the point being taken—and the measurement of the distance from this individual to its nearest neighbor. This method will be called the closest individual method.

The other exclusion angle methods use other criteria for the determination of the individuals to which distances are measured. They are of two types, those in which the distance from the sampling point to various objectively chosen individuals is measured, and those in which the distance from the individual closest to the point to one or more objectively chosen individuals is measured. There are a wide variety of these methods—11 are discussed by Lieberr (1952)—and one of each type is used in the present
study. They are the point centered quadrant and the random pairs methods.

The quadrant method divides the area around the sampling point into four equal parts. On artificial population maps this is easily accomplished by the use of a piece of clear sheet plastic with two perpendicular lines engraved on it. The sheet is placed with the intersection of the two lines directly over the sampling point; distances are measured from this point to the closest individual in each of the four quadrants. At each point four distances and the species of the four individuals are obtained. In terms of numbers of individuals sampled, this method is comparable to a quadrate of a size equal to four times the population mean area. Variations of this method include using the closest individual rather than the point as the center, and dividing the circle into 3, 6, or any other number of equal parts (Dice 1952).

The random pairs method (Cottam and Curtis 1949) involves the use of an angle of exclusion to determine the individuals between which distances are measured. The individual closest to the point is located and an angle delimited with its vertex at the point and its bisector on this closest individual. Any other individual falling within this angle is excluded, and the distance is measured from the individual closest to the point to the individual nearest to it outside the angle of exclusion. This method has been used in the field with angles of exclusion of 160° and 180°. In this study, angles varying from 0° to 340° have been used.

The population was sampled through the use of a type of stratified random sample called restricted randomization. After the border had been delimitated, the remainder of the population map was gridded into 100 equal squares. Two points were placed at random by means of a random numbers table in each of these squares, making a total of 200 such points. These points were further subdivided into four groups of 50. Measurements with each of the methods were taken at the same set of points. This was done to make the measurements taken for each of the methods more strictly comparable. The advantages of this procedure are discussed by Goodall (1952) in connection with measuring the changes that take place in the vegetation of an area over a period of years.

Results

Density

As used in this paper, area density refers to number of individuals per unit area, i.e. per quadrat, per acre, per 100 square meters, etc., while relative density is a percentage of the total number of individuals present. Area density may be determined directly from quadrat data; the mechanics of converting distance measurements obtained with the exclusion angle methods to area density will be discussed later. Relative density is obtained directly from the original data for all methods. The total number of individuals of a species is simply divided by the total number of individuals of all species, and the result multiplied by 100 to convert it to per cent.

The differences between the various methods in determining density may be illustrated through the use of an analogy. Let us assume that the plants in the population are presented by colored jelly beans, each color representing a different species. These colored beans are interspersed at random among a much larger number of white jelly beans, which represent unoccupied space. The quadrat method is then analogous to taking a scoop and counting the number of colored beans, by color, which are picked up in each scoopful. From the volume of the scoop, the number of colored beans per unit volume may be determined. The relative number of each color may also be determined by dividing the number of each color of beans by the total number of colored beans. The exclusion angle methods arrive at relative density, in the terms of the analogy, by taking a handful
of the beans and recording by color—2 colored beans for the random pairs method or 4 colored beans for the quadrant method. Density per unit volume in this analogy would be achieved by obtaining a measure of the amount of empty space, most simply by counting the number of white beans encountered between successive colored beans—a rather rough analogy to the measurement of the distance between individuals.

**Relative density**

Determinations of the accuracy and variability of relative density measurements in a random population were made using the quadrat (two sizes), quadrant, and random pairs (180° angle) methods.

**Table II. Comparison of variation with changes in the relative density of the species and in the number of sampling units employed.** Figures in the body of the table are standard errors expressed as a per cent of the mean, of the values of ten replicate samples for each combination of actual relative density and number of units.

<table>
<thead>
<tr>
<th>Actual relative density in population</th>
<th>Number of units</th>
<th>Random pairs</th>
<th>2M Quadrats</th>
<th>Quadrants</th>
<th>4M Quadrats</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>65.3</td>
<td>71.4</td>
<td>51.4</td>
<td>49.6</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>45.7</td>
<td>54.2</td>
<td>31.2</td>
<td>34.5</td>
</tr>
<tr>
<td>40</td>
<td>10</td>
<td>30.4</td>
<td>35.1</td>
<td>21.3</td>
<td>23.6</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>45.8</td>
<td>49.7</td>
<td>35.6</td>
<td>35.8</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>33.8</td>
<td>42.8</td>
<td>21.1</td>
<td>26.1</td>
</tr>
<tr>
<td>40</td>
<td>20</td>
<td>19.4</td>
<td>23.2</td>
<td>14.4</td>
<td>18.2</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>28.3</td>
<td>32.5</td>
<td>22.5</td>
<td>25.3</td>
</tr>
<tr>
<td>20</td>
<td>50</td>
<td>22.7</td>
<td>25.2</td>
<td>15.4</td>
<td>17.0</td>
</tr>
<tr>
<td>40</td>
<td>50</td>
<td>12.2</td>
<td>18.8</td>
<td>10.6</td>
<td>12.6</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>19.2</td>
<td>19.9</td>
<td>19.5</td>
<td>18.6</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>13.1</td>
<td>16.2</td>
<td>14.1</td>
<td>11.2</td>
</tr>
<tr>
<td>40</td>
<td>100</td>
<td>8.9</td>
<td>13.3</td>
<td>7.9</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Data were taken with each method at each of the 200 sampling points described earlier. Results are shown in Table II. The two quadrat sizes chosen were those most directly comparable to the two angle methods used. The angle methods differ from the area sample methods in that the number of individuals in each subsample, rather than the size of the sampling area, is constant. The quadrant method always measures 4 individuals at each point, the random pairs method 2 individuals. A quadrant of a size equal to 2 times the population mean area includes an average of 2 individuals every time it is laid down, and is thus comparable to the random pairs method. A similar relationship exists between the quadrant and the 4 times population mean area quadrat. Table II was constructed by subdividing the population into 10 species and recording the number of individuals of each species obtained in the first 10 quadrats of a given size, and in each succeeding set of 10 quadrats. Data for the random pairs and quadrant methods were obtained in a similar manner. Data for the 20 and 40 per cent species were obtained by combining 2 or 4—10 per cent species, and data for the larger number of sampling units were obtained by combining successive groups of 10 sampling units. The figures in the body of the table are coefficients of variation of the means (per cent standard error) of 10 sets of the number of sampling units and actual relative density of species indicated.

The table shows that there is an increase in accuracy with an increase in size of sample and actual relative density of the species in question. This is true regardless of the method used. It appears that the chief factor determining accuracy is the total number of individuals of the species encountered. A high total number of individuals may be achieved by using many sampling units in the case of the angle methods, with the number of sampling units necessary increasing as the relative density of the species in question decreases. An additional possibility with the quadrats is an increase in the size of the quadrat.

Figure 1 shows the values in Table II plotted against N, the average number of individuals encountered in the sample. The solid line is \( \sqrt[3]{N} \). The data appear to follow the trend indicated by this curve, which assumes a straight-line relationship on log-log paper. This relationship is
Fig. 1. The values in the body of Table II plotted against the average number of individuals encountered in each sample. The solid line is the square root of the number of individuals per sample divided by the number of individuals per sample.

commonly encountered in similar sampling problems (Hoppe and Holbert 1936).

When the individuals in a randomly dispersed population represent several species, the relative proportions of each species in the population may be determined with suitable accuracy by either area or angle methods. The main factor affecting such accuracy is the number of individuals that appear in the total sample. If a given species predominates in the population, the proportions may be determined accurately by a small sample. For species representing lesser and lesser fractions of the total, more or larger sampling units are needed to assess their correct proportions. As a general statement, it appears that about 30 individuals of a particular species should be encountered in the total sample before confidence may be placed in statements about its relative density.

*Arca density by quadrats*

The determination of density per unit area by means of quadrat samples needs little discussion here. Results using five sizes of quadrats on the artificial population are shown in Table III. As was the case with the relative density figures, large quadrats give results more nearly approaching the correct values than do small quadrats, when the number of quadrats is held constant. The relative variability of subsamples of 50 quadrats is also less with the large quadrats than with the small ones.
If the population is truly random, the number of individuals per quadrat should follow a Poisson distribution. Poisson distributions were calculated using mean number of individuals per quadrat of \( \frac{1}{4} \), \( \frac{1}{2} \), 1, 2, 4, and 8. These distributions are shown in Figure 2. When the mean number of individuals per quadrat is low, the distribution is markedly skewed to the right, but as the mean gets larger, the distribution becomes broader, flatter, and more symmetrical. When curves for the actual distributions obtained from the quadrat data for the artificial population were compared with the theoretical Poisson distributions by the chi-square test of goodness of fit, the results did not differ significantly from those expected from a random distribution. Perhaps the most important feature of these are-density characteristics of a random population is the very great variation to be expected in number of individuals per quadrat when large quadrats (4 \( \times \) mean area or larger) are used. When the mean density is 8 individuals per quadrat, 2 per cent of the samples will have 2 individuals or less, and 2 per cent will have 15 or more, but only about 40 per cent will have from 7 to 9 individuals.

![Frequency polygons of Poisson distributions with different mean numbers of individuals per quadrat.](image-url)
Area density by angle methods

With the angle methods, distance measurements are used to provide an estimate of density per unit area. The theory behind their use is that distances will vary with the density of the population, becoming smaller as density per unit area increases, and that the distances obtained by any objective method will bear a fixed relationship to the density of a random population. The conversion of these distance measurements to density per unit area is based on the assumption that the spacing of the individuals fluctuates at random from the condition in which they are an equal distance apart. The individuals may be considered to be located in the centers of squares, in which case the distance between individuals will be equal to the square root of the mean area. This is the distance moyenne of Braun-Blanquet (1932). The individuals may also be considered to be located in the centers of regular hexagons. Huesing (1932) has found tendencies toward such a distribution in some shrub communities. It has the theoretical advantage that each individual then occupies an area which is more nearly isodiametric. The distances in the case of a hexagonal distribution will be equal to 1.075 times the square root of the mean area (Cottam and Curtis 1949).

The distances derived through the use of either of these assumptions about the theoretical distribution of plants can be seen to bear a fixed mathematical relationship to one another. It is, therefore, not necessary, for the purpose of the paper, to consider which of these two hypotheses is the more valid. It is only necessary to have a standard by which the various distances may be compared. The simplest standard is the square root of the mean area, and this will be used throughout the paper.

Closest individual measurements

Because of the simplicity of this method, it was possible to take 100% samples rather than the samples from 200 points on which the data from the other methods are based. The distance was measured from each of the 882 individuals in the interior of the population to its closest neighbor. When this was done, it was found that many of the values were duplicates, individual A being closest to individual B, and individual B having individual A as its closest neighbor. The other alternative is for individual A to have individual B as its closest neighbor, but for individual B to be closest to some other individual. When these “paired individuals” were encountered, the distance between them was necessarily recorded twice, so that the data included 882 distances.

The most interesting results arising from the use of this method involve the consistency with which the measurements reflect the density per unit area of the population. By considering separate blocks of 100, 200, and 500 individuals to represent complete populations, it was possible to check the values obtained against population mean areas varying from 10 sq. cm. to 100 sq. cm. The results are shown in Table IV. It can be seen from the table that, in spite of a great variation in number of individuals per

| Table IV. Average distance between an individual and its closest neighbor for a series of populations differing in density per unit area |
|-------------------|-------------------|-------------------|-------------------|
| No. of individuals in population | Square root of mean area | Av. dist. between individuals | Column 3 as per cent of column 2  |
| 34 | 137.1 mm | 66.2 mm | 48.28 |
| 41 | 125.1 | 67.0 | 53.55 |
| 54 | 108.9 | 54.3 | 49.86 |
| 68 | 103.1 | 53.0 | 51.40 |
| 75 | 92.3 | 46.8 | 50.70 |
| 152 | 69.6 | 34.6 | 49.71 |
| 403 | 45.4 | 22.3 | 49.11 |
| 425 | 44.2 | 20.8 | 47.05 |
| 882 | 31.6 | 15.3 | 48.41 |

unit area, and a corresponding variation in the square root of the mean area of these populations, the results in terms of
per cent of the square root of the mean area are remarkably constant.

A comparison of the shape of the frequency polygons of the distances may be seen in Figure 3. Here the distances have been reduced to a common denominator by using varying class intervals to give an approximately equal number of distance classes for each population. The greater irregularity of the less dense populations is largely due to the fact that the size of the sample of these populations is smaller than is the case with the populations of higher density. Aside from this irregularity, the shape of the frequency polygons is very similar for each population density. They all possess a pronounced positive skewness. This type of distribution is often normalized by a square root transformation, and in this case square root transformations give values which, when tested by the $g_1$ statistic (Snedecor 1946) show no significant skewness.

Distances between closest individuals in a random population, then, bear a constant relation to the square root of the mean area. They are arranged about their mean in a Gaussian distribution when suitably transformed and hence may be subjected to further statistical treatment. Perhaps the most important attribute of these distances is their high coefficient of variation—averaging 54%.

**Quadrants**

The quadrant distance measurements were segregated according to the distance from the sampling point. For each sample, the quadrant in which the shortest distance was measured was designated $Q_1$, the quadrant in which the next shortest distance was measured was designated $Q_2$, etc. This provided four separate sets of data, the $Q_1$ data representing the distance between the individual closest to the sampling point and the point itself. The other quadrant figures do not necessarily measure the distance between the second, third and fourth closest individuals and the point, since any two, or even all four, of the closest individuals may be found in the same quadrant. In all cases, only the measurement between the point and the individual closest to the point in each quadrant was taken. Orientation of the sample was fixed, with the perpendicular lines separating the quadrants parallel to the grid lines on the population.

Table V gives a summary of these data. That the values for the means must increase from $Q_1$ to $Q_4$ is inherent in the method of collecting the data. In general,

**Table V. Summary of quadrant data. The last row in the table presents the results when the 4 distances obtained at each point are averaged**

<table>
<thead>
<tr>
<th>Quadrant</th>
<th>Mean (mm)</th>
<th>Mean as % of sq. root of mean area</th>
<th>Standard deviation</th>
<th>Standard error</th>
<th>Coefficient of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_1$</td>
<td>16.19</td>
<td>51.19%</td>
<td>8.02 mm</td>
<td>0.57 mm</td>
<td>49.54%</td>
</tr>
<tr>
<td>$Q_2$</td>
<td>35.82</td>
<td>115.75%</td>
<td>10.34</td>
<td>0.73</td>
<td>28.79%</td>
</tr>
<tr>
<td>$Q_3$</td>
<td>50.10</td>
<td>158.32%</td>
<td>15.16</td>
<td>1.07</td>
<td>30.25%</td>
</tr>
<tr>
<td>Combined quadrants</td>
<td>32.05</td>
<td>101.30%</td>
<td>7.60</td>
<td>0.53</td>
<td>23.71</td>
</tr>
</tbody>
</table>

the standard deviation increases as the means of the distance increase. To quote Snedecor (1946), “It seems rather characteristic that large things vary much and small things vary little.” To remove the inherent variability due to differences in the length of the distances, the coefficients of variation are shown in the last column of the table. The coefficient of variation is much greater for the $Q_1$ data than for the other quadrants. This is probably due to the fact that the distances obtained in $Q_1$ are more greatly influenced by the location of the sampling point than are the distances obtained for the other quadrants. The sampling point can theoretically land on an individual, which means that zero distances for $Q_1$ are possible. Values for the other quadrants can never be this small. The only possibility for small values for $Q_2$ is to have the sampling point land between two individuals that are close together.
This is an exceedingly rare occurrence. Usually, two individuals that are close together will be in the same quadrant, and distance will be measured to only one of them. The principal factor governing distance measurements for $Q_1$, $Q_3$, and $Q_4$ is the dispersion of the individuals on the population, and these values should vary as the density of the individuals varies from place to place on the population. $Q_1$ is also influenced by the density of the individuals on the population, but superimposed on this is the chance variation of the location of the sampling points with respect to the location of the individuals.

Values for the individual quadrants are of interest from the theoretical standpoint, but of more practical importance are the combined values. When the four quadrant distances from each sampling point are averaged, and these average values plotted, it is found that the frequency distribution is approximately symmetrical, with a $g_1$ value of 0.0311. The last line in Table V shows that these combined values are not only within 2 per cent of the value for the square root of the population mean area, but in addition possess a comparatively small standard error and coefficient of variation.

**Random pairs**

In many ways the random pairs method is similar to the quadrant method, and it would be expected that the results would show strong resemblances. The random pairs results are summarized in Table VI. That the mean should increase with increasing angles is inherent in the method. The mechanics of exclusion which bring this about are different from those used with the quadrants, but the end result is similar. Standard deviations increase with an increase in the mean, as before. The coefficients of

![Frequency polygons of the distances obtained through the use of the closest individual method. Variable class intervals have been used to show the similarities in the shape of the polygons.](image-url)

Fig. 3. Frequency polygons of the distances obtained through the use of the closest individual method. Variable class intervals have been used to show the similarities in the shape of the polygons.
variation also show similar trends, and for about the same reasons.

Column 3 of the table is plotted in Figure 4. The increase in the mean with increasing angles of exclusion is approximately linear until the two largest angles are reached, when the mean increases rapidly. Had larger angles of exclusion been used, the rate of increase would have been even more rapid. The curve will theoretically become asymptotic with the 360° ordinate, since angles of exclusion approaching 360 degrees will leave only an infinitely small area of the population in which individuals to which the distance is measured can be located, and these individuals will be, on the average, very far from the individual closest to the point.

Frequency distributions for the four quadrants and four selected angles of exclusion are shown in Figure 5. The most obvious changes that take place as the excluded area is increased are that the entire frequency distribution moves to the right and the distribution becomes broad and flat. There is an indication that the shape of the quadrant curves change. "Q" is skewed to the right, not because of a long right hand tail but because of

![Graph](image)

**Fig. 4.** Distances, expressed as per cent of the square root of the population mean area, obtained through the use of several angles of exclusion with the random pairs method. The increase of distance with increase of angle is approximately linear, except for the 2 largest angles used.
Fig. 5. Frequency polygons of the distances obtained for the 4 quadrants, and for the random pairs method using 4 selected angles of exclusion.

the absence of a left hand tail. The curve is truncated—cut off on the left hand side. What there is of it appears to be symmetrical. \( Q_2 \) looks much like \( Q_1 \) except that it is moved to the right and the missing tail has been added. It is approximately symmetrical. \( Q_3 \) and \( Q_4 \) are increasingly skewed to the right. Except for the truncated nature of the curve for the 0 degree angle, there is little indication that the random pairs distributions show a similar increase in skewness with increasing angle of exclusion.

**Discussion and Conclusions**

The investigation reported in this paper was conducted for the purpose of ascertaining some of the characteristics of a random population. An artificial popu-
lation was used since such a population is free from most of the uncontrollable, and often unknown factors that cause peculiarities in plant distributions in the field. In a sense, this population represents an idealized and simplified map of the plants in a community. It was felt that a knowledge of such an idealized community would be of value as a standard with which existing plant communities could be compared. The results have been discussed as they have been presented. There remains a discussion of the applications of the information herein contained to results obtained by plant ecologists in the field.

**Relative density**

The accuracy of relative density measurements of random populations appears to be dependent on the number of individuals sampled regardless of the method used to sample the individuals. A minimum of 30 individuals of any species is necessary before any statement can be made regarding its relative density in the population. In terms of usual field practice in forest sociology with the common 10 meter by 10 meter quadrat, only those species with a frequency of 26 per cent or more in a sample of 100 quadrats can be used for accurate statements of relative composition. With only 50 quadrats, this lower limit should be 45 per cent, while with the low number of 25 quadrats the frequency limit should be 70 per cent. The only recourse for species below these limits is to increase the number of quadrats, since an increase in size is not usually practicable in the field. Use of the exclusion angle methods discussed in this paper are subject to the same limitations, and with these methods the only possible way to increase the number of individuals is an increase in number of sampling points, since the number of individuals measured at each point is fixed.

In stands with one or two dominants, as in boreal conifer forests and certain deciduous forests of north temperate regions, the usual sample of 25 to 100 quadrats should give an accurate picture of species composition of the common trees and, since these make up the great preponderance of the individuals present, the lesser species will be rather narrowly defined as well. It does not make much difference in the characterization of a community whether some rare species is listed as having 2 per cent relative density or 4 per cent, although the error, so far as that particular species is concerned, may be very great. When the number of dominants increases, as in the southern Appalachians, then the size of the sample must increase proportionately. No fixed statements as to the size of the sample are now possible but it would appear that quantities approaching the upper feasible limit might be needed. In the tropical rain forests, with 60 to 90 species per hectare (Black, *et al.* 1950) the number of quadrats or points needed for even reasonable accuracy becomes impossibly great. The commonly employed method of using a 100% sample on an area as large as 2 or 3 acres is exceedingly time-consuming, but in spite of this, is still not very precise. This failure of the usual sampling methods to give accurate determinations of such a basic, simple characteristic as relative composition is no doubt responsible for the emphasis placed by the tropical ecologists on physiognomic classification (Beard 1944, Richards 1952).

**Area density**

Since quadrats are area samples, their use to measure density per unit area is relatively simple. In spite of the simplicity of the quadrat method, its use in the field is both tedious and time-consuming. The number of trees encountered in one 10 by 10 meter quadrat averages four or less in most upland deciduous forests of temperate regions. The amount of work required to lay out such a quadrat with any degree of ac-
curacy is much greater than that required to measure four trees by either of the exclusion angle methods discussed in this paper. In terms of sample size, the angle methods are capable of yielding greater returns per man hour than is the quadrat method. One of the disadvantages of the angle methods is that they have not been the subject of much study and their potentialities and characteristics are not well-known. The results presented in this paper show that distance measurements obtained through the use of these methods follow predictable patterns. The average of the four distances obtained at each point with the quadrant method is very close to the square root of the population mean area. To convert these distances to density per unit area, it is necessary only to square the average distance and divide the result into the number of square feet being used as the unit area. In the case of the random pairs method, average distances have been shown to increase with an increase in the angle of exclusion. This increase is linear over a wide range of exclusion angles, varying from 58 per cent of the square root of the population mean area with no angle of exclusion to 155 per cent with an angle of exclusion of 260 degrees. Any angle throughout this range may be used for field work, although angles of exclusion less than 100 degrees yield data with a higher coefficient of variation than do the larger angles. The distances so obtained must be multiplied by the appropriate correction factor to convert them to the square root of the mean area.

One characteristic of random populations which is often overlooked or at least not completely understood by practicing ecologists is the extreme variability that may be encountered. Regardless of the method used, this variation may be so large as to render any conclusions based on small samples greatly in error. Of equal importance is the fact that aggregated populations display even greater variability. Since this variability is inherent in the populations, the only recourse for the student of plant communities is to increase the size of his sample.

Summary

Some of the characteristics of a randomly distributed population were determined by the use of a map containing 1000 individuals located by means of rectangular coordinates taken from a random numbers table. The randomness of the population was checked by comparison with a Poisson distribution and by areal homogeneity tests. The artificial population was sampled by restricted randomization of sampling points and by both area sample methods and exclusion angle methods. In the former there were varying numbers of individuals in constant areas, while in the latter, a constant number of individuals per sample varied in their spacing distances.

Variability of relative density measurements was very great with small samples. Reasonable accuracy was attained when the number of individuals of the species in question exceeded 30 in the total sample. This was true regardless of sampling method and could be achieved by increasing either the number or the size of the sampling units.

Distances between individuals in the random population measured by any of the exclusion angle methods showed a Gaussian frequency distribution, either directly or after simple transformation. The average of all four distances obtained with the quadrant method was within 2 per cent of the parameter distance obtained from the square root of the population mean area. There was a linear increase in distances obtained with increasing exclusion angles in the random pairs method over the range from zero degrees to 260 degrees.

Perhaps the most important characteristic of a random population demonstrated in this paper is the extreme variability
of the results for any of the measures of plant communities studied. This variability is inherent in random populations and its presence necessitates careful planning of sampling procedure to insure that the results obtained by any sampling method approximate the parameters of the population.

The relation of these findings to actual field practice are discussed, with the general recommendation that larger samples than those generally employed are necessary for satisfactory results.

References


