Williams, C. B. 1950. The application of the logarithmic series to the frequency of occurrence of plant species in quadrats. *Journal of Ecology.* 38 (1), pp. 107-138.

THE APPLICATION OF THE LOGARITHMIC SERIES TO THE FREQUENCY OF OCCURRENCE OF PLANT SPECIES IN QUADRATS

By C. B. WILLIAMS, Sc.D., *Rothamsted Experimental Station, Harpenden*

(*With fourteen Figures in the Text*)

CONTENTS

I. Introduction

In recent years a number of studies have been made on the relative abundance of different species of animals, especially insects, in mixed wild populations (see Fisher, Corbet & Williams, 1943; Williams, 1944, 1947*b*).

It has been found that there is a comparatively simple mathematical relation between the number of species and the number of individuals in a series of random samples of different sizes taken from a single ecological community.

Further, it has been shown that in any random sample from a wild population the number of species each represented by only one individual is higher than the number represented by two, and that this in turn is higher than the number represented by three; and so on. The series of the numbers of species represented in the sample by one, two, three, etc., individuals forms a 'hollow curve', and it can be closely represented by the logarithmic series which has the form

$$n_1, \quad \frac{n_1 x}{2}, \quad \frac{n_1 x^2}{3}, \quad \frac{n_1 x^3}{4}, \text{ etc.,}$$

where n_1 is the number of species with one individual, and x is a constant (for the sample) less than unity.

It has already been shown (Williams, 1944; Bond, 1947) that some of the general principles of this work seem to apply to the number of species of plants found on areas of different sizes when selected within a single plant community.

One of the main differences between the treatment of animal and plant populations from this point of view arises from the difficulty of determining, in many species of plants, what is an 'individual'. In most animals this difficulty does not occur. On the other hand, while in plants it is quite easy to say that a particular plant came from a particular spot, even to within a few inches; with mobile and often winged insects it is not possible to tie them down to a particular area of origin. They may range in the course of their life over many miles, and may be found far away from the ecological community to which they belong. As a result of these differences the study of the detailed structure of an animal population has tended to be based on 'numbers of individuals', while in the botanical world the emphasis has been on 'area'.

There has thus grown up in botanical ecology a study of the distribution of species in 'quadrats' or small samples of fixed area (usually 1 sq.m. or less) in which only the presence or absence of particular species is recorded and not its numbers. Plant associations have been classified according to the 'dominant' species of the vegetation; that is, those which occur in a high proportion of quadrats. Much empirical work has been done in the study of these results (e.g. Raunkiaer, 1934), but relatively little with a strict mathematical background.

It is the object of the present paper to see how far the results obtained by the application of the logarithmic series in animal populations can be applied to the relation between plant species and area.

In the first place it must be realized that a 'quadrat' is—or should be—a small random sample of the ecological community under investigation. It contains a possibly unknown number of 'plant-units', but, if the quadrats are large enough in relation to the size of the plants themselves, the number of 'plant-units' in a series of quadrats should be proportional to the number of quadrats. Thus if there are N 'plant-units' on 1 quadrat, we may assume that there will be qN 'plant-units' on q quadrats.

The 'plant-unit' may be considered as that quantity of any species of plant present in the quadrat which is behaving, particularly from the point of view of distribution, as if it corresponded to one 'individual' in an animal population. In a plant species which produces a definite plant from each seed, and which has no tillers, stolons, or other vegetative means of lateral spreading, the 'unit' will correspond to the 'individual'. In species which produce compact masses by vegetative growth the unit may consist of a number of plants, originally, and perhaps still, in living contact, which behave in such close association as to resemble in some aspects an 'individual' in other plants.

It is important to note that to be a statistically reliable sample a quadrat must contain a 'reasonable' number of units. If in an association the plant units are large—e.g. trees or large vegetation aggregations—then the quadrats may have to be increased in size. This is also necessary for another reason, namely that the presence of one large plant in a quadrat of fixed area reduces the possible number of other 'units' on the rest of the area. Thus if large plants are present, even in small numbers, the small quadrat of 1 sq.m., or less, may be inadequate in size. The error of sampling in such a case may easily obscure the real biological differences which we wish to investigate.

II. The number of species in increasing numbers of quadrats

If we have a number of small random samples of fixed area—so-called quadrats—from a reasonably uniform plant association, then, on an average, each quadrat will contain the same number of species. Two quadrats together will have a chance of containing more species than one quadrat; three quadrats rather more than two; and so on. The chance of new species, however, gets steadily smaller as the number of quadrats increases. If p_1 be the number of species found on 1 quadrat, and p_2 the increase by adding a second quadrat, and p_3 the further increase by adding a third quadrat, and so on, then it follows that with 2 quadrats, there are

$2p_2$ species found on 1 quadrat only

and p_1-p_2 species found on both quadrats.

With 3 quadrats there are $3p_3$ species found on 1 quadrat only,

$3(p_2-p_3)$ species found on 2 quadrats,

and $(p_1-p_2)-(p_2-p_3)$ species found on all 3 quadrats;

for q quadrats the distribution of the species is

$$\text{Species in 1 quadrat only} = qp_q,$$

$$\text{Species in 2 quadrats only} = \frac{q(q-1)}{1.2}(p_{q-1}-p_q),$$

$$\text{Species in 3 quadrats only} = \frac{q(q-1)(q-2)}{1.2.3}(p_{q-2}-2p_{q-1}+p_q),$$

and the number of species in r quadrats out of the q

$$= {}^{q}C_r[p_{q-r+1} - {}^{r-1}C_1 p_{q-r+2} + {}^{r-2}C_2 p_{q-r+3} + \dots + (-1)^{r-1}p_q].$$

The actual estimation of these values can be simplified by putting the values of the number of species in 1, 2, 3, 4, etc., quadrats (S_1, S_2, S_3, etc.) in a vertical column as in Table 1; the differences between these, p_1, p_2, p_3, etc., in a second vertical column with p_1 opposite S_1; and then again the differences between them in a third column with p_1-p_2 opposite p_2; and so on: then any line read horizontally will give in the successive difference from p_n onwards the number of species in any *one* particular quadrat only; the number in any two particular quadrats; the number of species in any three particular quadrats; and so on.

To get the total number of species in all possible single quadrats; in all possible pairs; in all possible threes, etc., the column differences must be multiplied by ${}^{q}C_r$ as shown in heavy type numbers in the same table. This is independent of any particular theory of the relative abundance of the species.

Table 1. *Method of differences used in calculating the expected number of species to be found in* 1 *quadrat only, in* 2 *quadrats, in* 3 *quadrats, and so on*

Above general theory; below numerical example (based on the logarithmic series, where $\alpha = 10$ and $N = 100$). The figures in any one column are the differences between two figures in the previous column.

No. of quadrats	Total no. of species	First difference. Species in 1 quadrat only	Second difference. Species in 2 quadrats only	Third difference. Species in 3 quadrats only	Fourth difference. Species in 4 quadrats only
1	S_1	**1** $\times p_1$	—	—	—
2	S_2	**2** $\times p_2$	p_1-p_2	—	—
3	S_3	**3** $\times p_3$	**3** $\times (p_2-p_3)$	$(p_1-p_2)-(p_2-p_3)$	—
4	S_4	**4** $\times p_4$	**6** $\times (p_3-p_4)$	**4** $(p_2-p_3)-(p_3-p_4)$	difference
5	S_5	**5** $\times p_5$	**10** $\times (p_4-p_5)$	**10** $(p_3-p_4)-(p_4-p_5)$	**5** × difference
6	S_6	**6** $\times p_6$	**15** $\times (p_5-p_6)$	**20** $(p_4-p_5)-(p_5-p_6$	**15** × difference, etc.
1	23·98	23·98	—	—	—
2	30·44	**2** × 6·47	17·51	—	—
3	34·34	**3** × 3·89	**3** × 2·57	14·94	—
4	37·14	**4** × 2·80	**6** × 1·10	**4** × 1·47	13·47
5	39·32	**5** × 2·18	**10** × 0·61	**10** × 0·49	**5** × 0·99
6	41·11	**6** × 1·79	**15** × 0·39	**20** × 0·22	**15** × 0·26 etc.

III. The application of the logarithmic series

The logarithmic series can be written in two forms, either

$$n_1,\quad \frac{n_1x}{2},\quad \frac{n_1x^2}{3},\quad \frac{n_1x^3}{4},\quad \ldots,\ \text{etc.,}$$

or

$$\alpha x,\quad \frac{\alpha x^2}{2},\quad \frac{\alpha x^3}{3},\quad \frac{\alpha x^4}{4},\quad \ldots,\quad \text{etc.}$$

In each case the first term, n_1 or αx, is the number of species represented by one individual; the second term, $\frac{n_1x}{2}$ or $\frac{\alpha x^2}{2}$, is the number of species with two individuals, and so on.

'x' is a number less than unity which is a constant for each sample, and is dependent on the size of the sample when different sized samples are taken from the same population. The larger the sample the nearer 'x' approaches to unity.

'α' is another constant, which is however independent of the size of the sample, and is a property of the population sampled. It is high when the population is much diversified into species and low when the population is uniform with few different species. We have called this the 'Index of Diversity' and it appears to have considerable interest from an ecological point of view.

Evidence has been brought forward (Williams, 1944) to show that some of the properties of the logarithmic series are found to apply to samples taken from plant associations by the method of quadrats. For the moment, it is only necessary to point out that if we make the assumption that the logarithmic series can be applied to the problems under consideration, then we can use the mathematical properties of the series to suggest possible effects, and then test these against the field observations. If our deductions from the assumption fit our observed facts we have some justification for believing that our assumption might be correct.

It must be remembered that in what follows the 'individual' of the animal kingdom must be replaced by the somewhat elusive and perhaps fictitious 'plant-unit'.

One of the properties of the logarithmic series is that if a sample of N units is taken from a population with an Index of Diversity α, then the number of groups represented in the sample

$$S = \alpha \log_e \left(1 + \frac{N}{\alpha}\right).$$

For example, if a sample of 100 individuals is taken from a plant population with an Index of Specific Diversity of 10, then the number of species represented will be

$$10 \log_e (1 + \tfrac{100}{10}) = 10 \log_e 11 = 23{\cdot}98.$$

If we suppose this to represent an imaginary quadrat, then 2 quadrats of the same size will on an average, contain

$$10 \log_e (1 + \tfrac{200}{10}) = 10 \log_e 21 = 30{\cdot}45 \text{ species.}$$

By the same principle 3 quadrats will contain 34·34 species; 4 quadrats 37·14 species, and so on. This gives a series of values of S_1, S_2, S_3, and, by difference p_1, p_2, p_3, from which we can calculate the frequency of occurrence of species in quadrats by the method of differences discussed above.

Some values for this particular population are shown in the lower half of Table 1.

It follows from the above that the difference between the number of species in q quadrats and in $q-1$ quadrats

$$=\alpha \log_e\left(1+\frac{qN}{\alpha}\right)-\alpha \log_e\left(1+\left[\frac{q-1}{\alpha}\right]N\right).$$

If the number of units in the q quadrats is large compared with α, as it should be for a good sample, then we may neglect the '1', in comparison with qN/α. In this case the difference is

$$\alpha \log_e\left(\frac{qN}{\alpha}\right)-\alpha \log_e\left(\frac{q-1}{\alpha}\right)N=\alpha \log_e\frac{q}{q-1}.$$

But the value of $\log_e q/(q-1)$ rapidly approaches $1/q$ as q becomes large. Hence p_q, the difference between the number of species in q and $q-1$ quadrats, approaches α/q as the number of quadrats become large.

If we insert in the general formula for a number of quadrats (p. 109) the value of α/q for p_q, $\alpha/(q-1)$ for p_{q-1}, etc., we get the following results:

Number of species in 1 quadrat only $=qp_q=q\left(\frac{\alpha}{q}\right)=\alpha$,

Number of species in 2 quadrats $=\frac{q(q-1)}{1.2}\left(\frac{\alpha}{q-1}-\frac{\alpha}{q}\right)=\frac{\alpha}{2}$,

Number of species in 3 quadrats $=\frac{q(q-1)(q-2)}{1.2.3}\left(\frac{\alpha}{q-2}-\frac{2\alpha}{q-1}+\frac{\alpha}{q}\right)=\frac{\alpha}{3}$, and so on.

So that with a large number of quadrats selected at random from a population arranged in a logarithmic series with an Index of Diversity α, the number of species in 1, 2, 3, etc., quadrats approaches to a Hyperbolic (Harmonic) series with the first term (the number of species found in one quadrat only) equal to α.

This holds for the earlier part of the series only; in the later terms (species in a large proportion of the quadrats) the values are above the hyperbola, as will be seen in Figs. 1 and 2.

The value of any term derived from the successive differences of $\alpha \log_e\left(1+\frac{qN}{\alpha}\right)$ depends on α, and on the ratio between N and α. If we express the frequency of occurrence of species in quadrats in terms of α, the form of the distribution of the number of species in 1, 2, 3, etc., quadrats is determined only by ratio of N/α.

In the logarithmic series

$$N/\alpha=x/(1-x),\quad S/\alpha=-\log_e(1-x),$$

and

$$\frac{N}{S}=\frac{x}{(1-x)(-\log_e 1-x)}$$

So any particular value of N/α gives a particular value of x, and this in turn gives a particular value of S/α, and also of N/S.

Thus when one defines N/α for a population, or says that quadrats on two different associations should have the same N/α, the same must also be true for S/α.

As it is very much easier to count the number of species on a quadrat than the number of individuals, the use of S/α is of greater practical value. On the other hand, it must be remembered that changes in the size of the quadrat are directly proportional to N, but not to S.

Table 2. *Frequency distribution (in terms of 'α') of species in various numbers of quadrats up to 25 with different values of the ratio of N/α and S/α*

No. of quadrats	Total no. of species	No. of quadrats: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
$N/\alpha=2:S/\alpha=1{\cdot}10$																										
1	1·098	1·098	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
5	2·398	1·003	0·506	0·345	0·273	0·270	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
10	3·044	1·001	0·501	0·335	0·253	0·204	0·173	0·151	0·138	0·133	0·155	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
15	3·434	1·000	0·501	0·334	0·251	0·202	0·169	0·145	0·128	0·115	0·105	0·097	0·092	0·089	0·091	0·114	—	—	—	—	—	—	—	—	—	—
20	3·714	1·000	0·500	0·334	0·251	0·201	0·168	0·144	0·126	0·113	0·102	0·093	0·086	0·080	0·076	0·072	0·069	0·068	0·068	0·071	0·092	—	—	—	—	—
25	3·932	1·000	0·500	0·334	0·250	0·200	0·167	0·144	0·126	0·112	0·101	0·092	0·085	0·079	0·073	0·069	0·065	0·062	0·059	0·057	0·055	0·054	0·054	0·055	0·059	0·078
$N/\alpha=10:S/\alpha=2{\cdot}40$																										
1	2·398	2·398	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
5	3·932	1·091	0·613	0·485	0·494	1·248	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
10	4·615	1·043	0·547	0·385	0·308	0·267	0·246	0·241	0·255	0·318	1·005	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
15	5·172	1·028	0·529	0·365	0·284	0·236	0·206	0·186	0·173	0·165	0·161	0·163	0·172	0·195	0·259	0·896	—	—	—	—	—	—	—	—	—	—
20	5·303	1·021	0·522	0·356	0·274	0·225	0·193	0·171	0·155	0·143	0·134	0·128	0·124	0·122	0·121	0·124	0·130	0·142	0·166	0·227	0·829	—	—	—	—	—
25	5·525	1·016	0·517	0·351	0·268	0·219	0·186	0·163	0·147	0·134	0·124	0·116	0·110	0·105	0·102	0·099	0·098	0·097	0·098	0·100	0·104	0·111	0·124	0·148	0·206	0·782
$N/\alpha=50:S/\alpha=3{\cdot}93$																										
1	3·932	3·932	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
5	5·525	1·111	0·639	0·523	0·564	2·690	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
10	6·217	1·051	0·557	0·396	0·321	0·282	0·265	0·265	0·291	0·386	2·403	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
15	6·621	1·033	0·536	0·371	0·291	0·244	0·215	0·195	0·183	0·177	0·176	0·181	0·196	0·230	0·325	2·267	—	—	—	—	—	—	—	—	—	—
20	6·909	1·025	0·526	0·361	0·279	0·230	0·198	0·177	0·161	0·150	0·142	0·136	0·133	0·132	0·134	0·139	0·148	0·165	0·200	0·292	2·182	—	—	—	—	—
25	7·132	1·020	0·520	0·355	0·272	0·223	0·191	0·168	0·151	0·139	0·129	0·121	0·116	0·111	0·108	0·107	0·106	0·106	0·108	0·112	0·119	0·129	0·147	0·181	0·270	2·121

The relation between the value of these different expressions for different values of N/α is shown below, and the general relation between N/α (on a logarithmic scale) and S/α is shown in Fig. 1:

N/α	x	S/α	N/S
2	0·6666	1·10	1·82
10	0·9091	2·40	4·17
50	0·9804	3·93	12·72
100	0·9901	4·62	21·67
400	0·9975	5·99	66·73
1000	0·9990	6·91	144·7

With these relations in view, Table 2 and Figs. 2 and 3 have been prepared showing the relative frequency of species in up to 25 quadrats, when $N/\alpha=2$, 10 and 50, or $S/\alpha=1{\cdot}1$, 2·4 and 3·93.

In the figures the vertical scale is in terms of α, and in the tables the numbers given must be multiplied by 'α' to give the actual number of species.

Fig. 4 shows the distribution of the species (in terms of α) in any number of quadrats out of 25, with calculated values for $S/\alpha=1{\cdot}1$, 2·4, 3·93 and 5·99, and other values by interpolation.

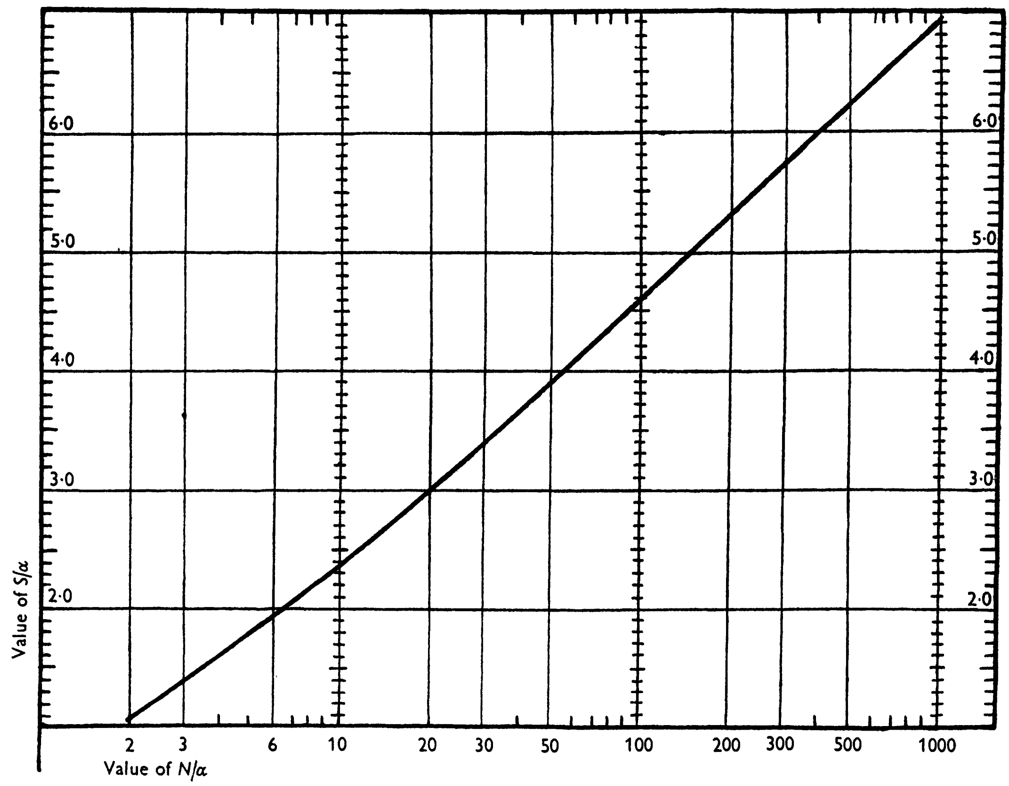

Fig. 1. Relation between values of N/α and S/α.

It will be seen from Fig. 4 that once the samples are large (i.e. $S/\alpha=3$ or more), the ratio makes little difference to the numbers of species in quadrats up to nearly all of the total; but very great differences to the number of species found in *all* the quadrats, whatever the total number may be.

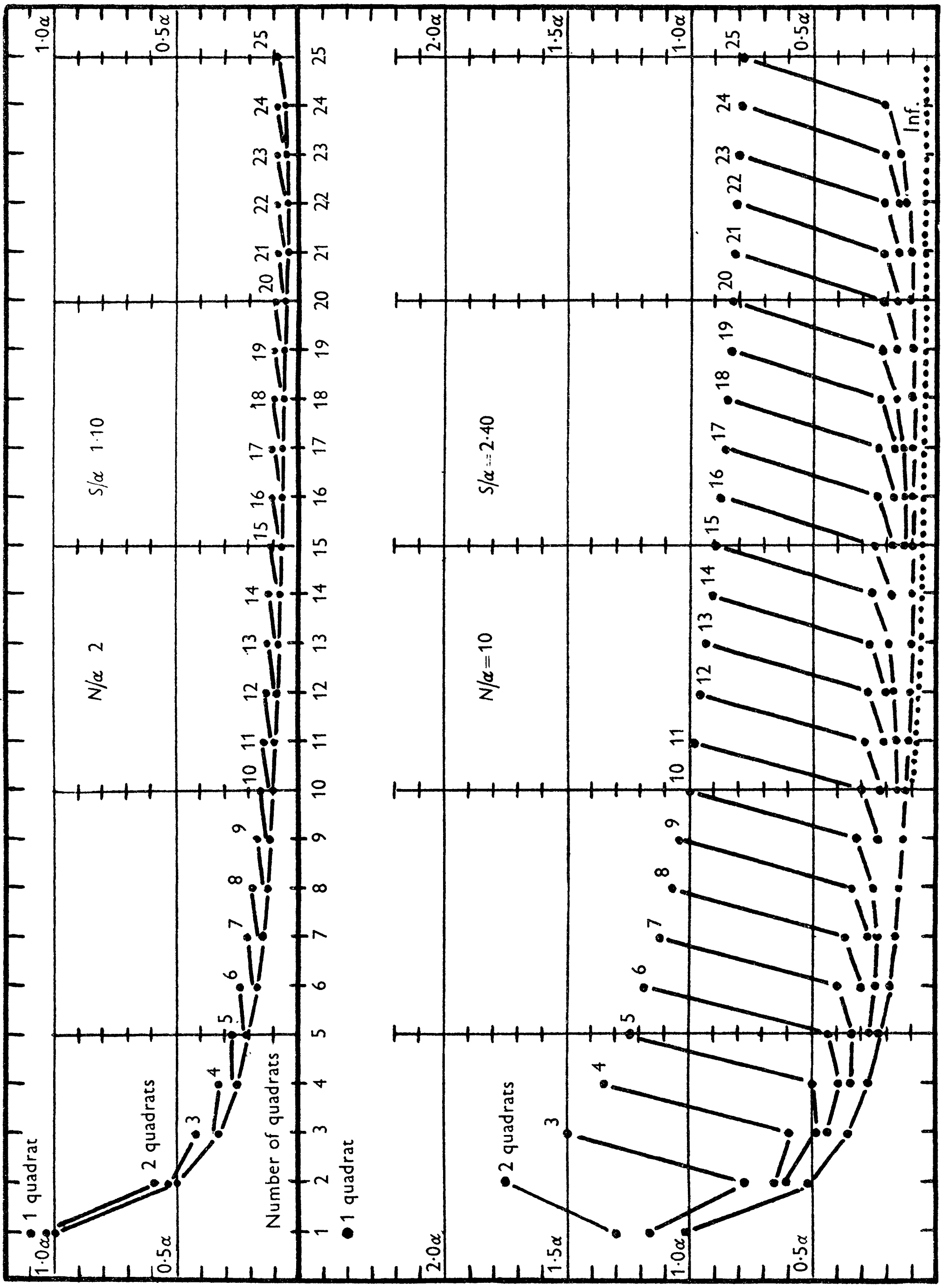

Fig. 2. Theoretical frequency distribution of species (in terms of α) in 1–25 quadrats for values of $N/\alpha=2$ and $N/\alpha=10$, or $S/\alpha=1{\cdot}1$ and $2{\cdot}4$.

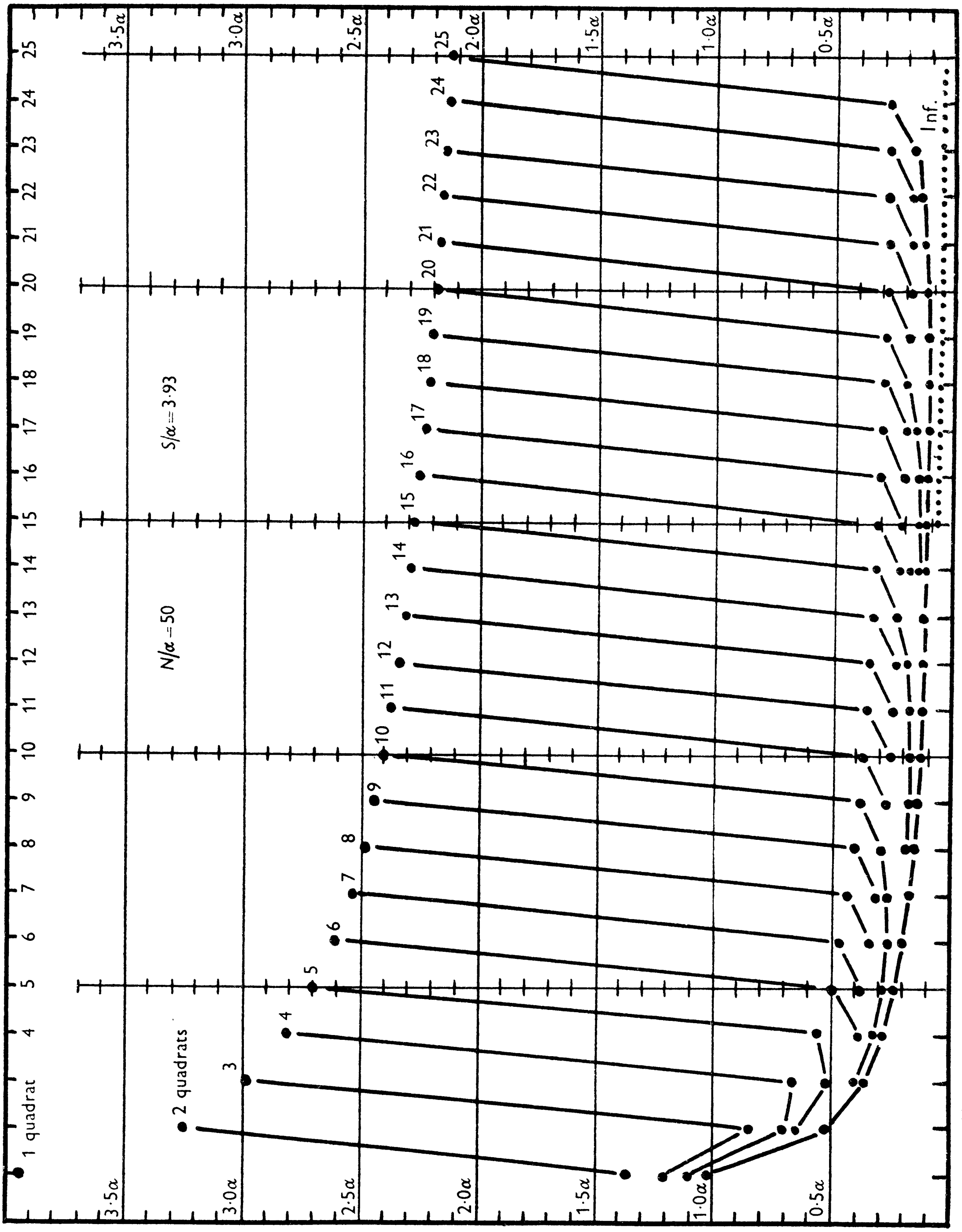

Fig. 3. Theoretical frequency distribution of species (in terms of α), in 1–25 quadrats for value of $N/\alpha = 50$, or $S/\alpha = 3{\cdot}93$.

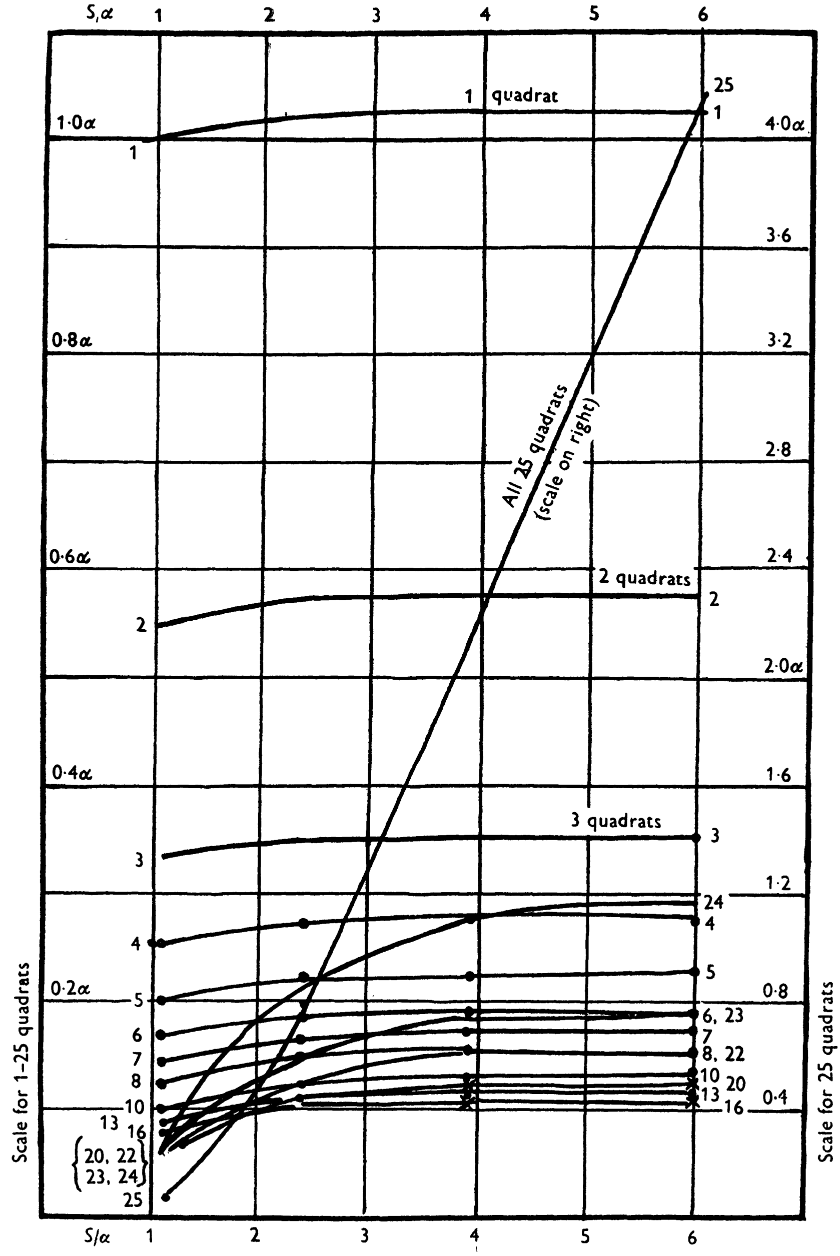

Fig. 4. Diagram showing the number of species (in terms of α) found on any number of quadrats out of 25, for any value of S/α from 1 to 6. Calculated values at 1·1, 2·4, 3·93 and 5·99: rest by interpolation. Values for 9, 11, 12, 14, 15, 17, 18, 19 and 21 quadrats can be obtained by interpolation. Scale for 1–24 quadrats on left; for 25 quadrats on right.

It follows from the above that if we have two plant associations, one richer than the other, they will give different relative frequency distributions of species in 1, 2, 3, etc., quadrats, unless the size of the quadrat, in plant units or in number of species, has been chosen in proportion to the Index of Diversity.

It will be seen from the tables, that if, for example, in a population with an N/α ratio of 10, ($S/\alpha=2{\cdot}4$) 5 quadrats are studied, there will be $1{\cdot}09\alpha$ species in 1 quadrat only; $0{\cdot}61\alpha$ species in 2 quadrats; $0{\cdot}49\alpha$ in three; $0{\cdot}49\alpha$ in four; and $1{\cdot}25\alpha$ in all 5 quadrats. If, however, the size of the quadrats had been increased five times ($N=50\alpha$) the number of species would be $1{\cdot}11\alpha$, $0{\cdot}64\alpha$, $0{\cdot}52\alpha$, 0·56 and $2{\cdot}70\alpha$ respectively. An increase in the size of the quadrats greatly increases the relative number of species common to all quadrats.

It is important to note that the numbers of species which occur in x quadrats out of a total of y, is not the same as the number which occur in $2x$ quadrats out of $2y$. This mistake has been made by several botanists, when they were unable to get the full numbers of quadrats that they wished in one particular association.

IV. The number of species found in different percentages of quadrats

Certain botanists, and particularly Raunkiaer (1934), have developed a convention of dividing up the species found on a number of quadrats into five groups as follows:

I. Species found on 1–20% of the quadrats; the rarest species.
II. Species found on 21–40% of the quadrats; not so rare.
III. Species found on 41–60% of the quadrats.
IV. Species found on 61–80% of the quadrats.
V. Species found on 81–100% of the quadrats; these are the common or dominant species which are considered as characteristic of the association which is being sampled.

Table 3. *Number of species in terms of α, and as percentages of total species, in Raunkiaer's five groups of quadrats* (*i.e. species in* 1–20, 21–40, 41–60, 61–80 and 81–100% *of the total observed quadrats*)

No. of quadrats	Total no. of species	Group I		Group II		Group III		Group IV		Group V	
		No.	%	No.	%	No.	%	No.	%	No.	%
				$N/\alpha=2:S/\alpha=1{\cdot}1$							
5	2·40α	1·00α	42	0·51α	21	0·35α	14	0·27α	11	0·27α	11
10	3·04	1·50	49	0·59	19	0·38	12	0·29	9	0·29	9
15	3·43	1·84	53	0·62	18	0·39	11	0·29	9	0·29	9
20	3·71	2·09	56	0·64	17	0·39	11	0·30	8	0·30	8
25	3·93	2·28	58	0·65	17	0·40	10	0·30	8	0·30	8
				$N/\alpha=10:S/\alpha=2{\cdot}4$							
5	3·93α	1·09	28	0·61	16	0·49	12	0·49	13	1·25	32
10	4·62	1·59	34	0·69	15	0·51	11	0·50	11	1·32	29
15	5·02	1·92	38	0·73	14·5	0·52	10	0·50	10	1·35	27
20	5·30	2·17	41	0·74	14	0·53	10	0·50	9	1·36	26
25	5·53	2·37	43	0·75	14	0·53	10	0·50	9	1·37	25
				$N/\alpha=50:S/\alpha=3{\cdot}93$							
5	5·53α	1·11	20	0·64	12	0·52	9	0·56	10	2·69	49
10	6·22	1·61	26	0·72	12	0·55	9	0·56	9	2·79	45
15	6·62	1·94	29	0·75	11	0·56	8	0·55	8	2·82	43
20	6·91	2·19	32	0·77	11	0·56	8	0·55	8	2·84	41
25	7·13	2·39	34	0·78	11	0·56	8	0·55	8	2·84	40
				$N/\alpha=400:S/\alpha=5{\cdot}99$							
5	7·60α	1·12	15	0·64	8	0·53	7	0·58	8	4·73	62
10	8·29	1·61	19	0·72	9	0·56	7	0·57	7	4·83	58
15	8·70	1·95	22	0·75	9	0·56	6	0·57	7	4·87	56
20	8·99	2·20	24	0·77	9	0·57	6	0·57	6	4·89	54
25	9·21	2·39	26	0·78	9	0·57	6	0·57	6	4·90	53

It will be interesting to apply the theories developed above to see how far such divisions are reliable criteria of the population; and how far, on the contrary, they depend on (*a*) the size of the quadrat, (*b*) the number of quadrats, and (*c*) the total area sampled.

From the tables already given, the information in Table 3 and Fig. 5 has been extracted to show the theoretical number of species expected in the five groups in the case of 5, 10, 15, 20 and 25 quadrats with different values of N/α or S/α. All the numbers are in terms of α.

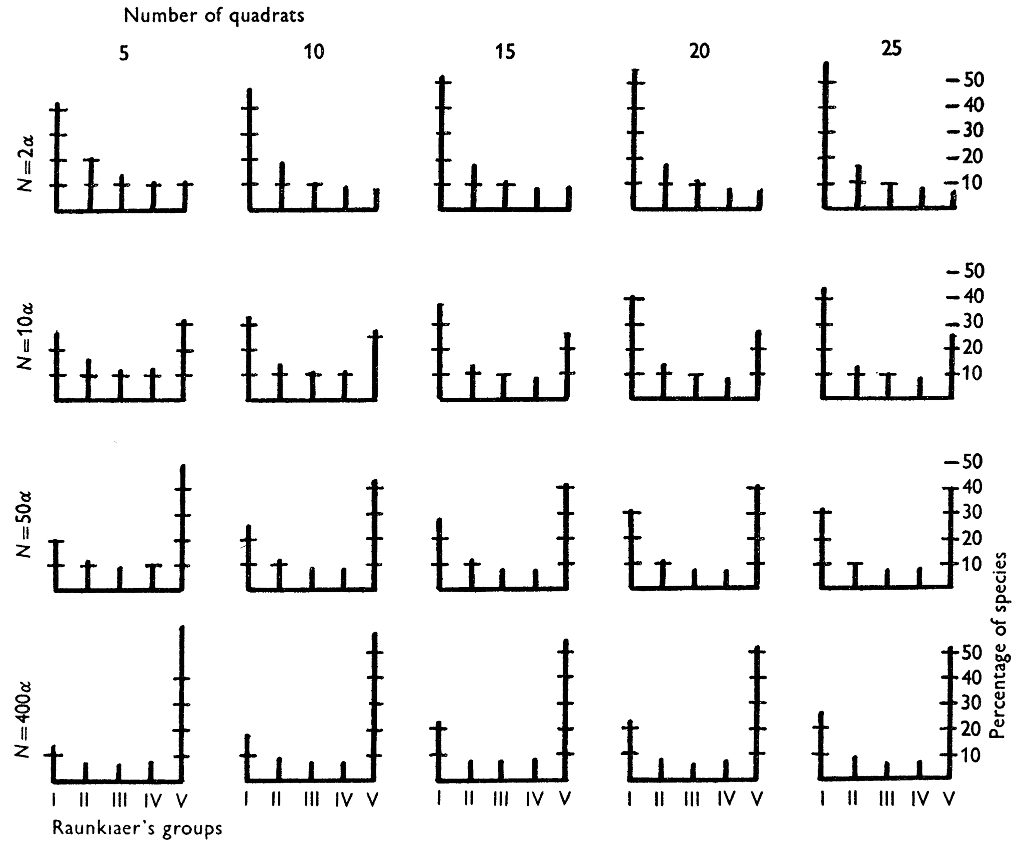

Fig. 5. Theoretical percentage of total species which occur in Raunkiaer's five groups of quadrats, according to variations in the size of the quadrat, and to the number of quadrats.

If we consider first the data for 25 quadrats it will be seen that as the quadrat size increases from $N=2\alpha$ to $N=400\alpha$ there is very little alteration in the number of species in group I (1–5 quadrats), from $2{\cdot}28\alpha$ to $2{\cdot}39\alpha$ only; small difference in groups II and III (from $0{\cdot}65\alpha$ to $0{\cdot}78\alpha$, and $0{\cdot}40\alpha$ to $0{\cdot}57\alpha$); proportionally a larger difference, but actually a small difference in group IV (from $0{\cdot}30\alpha$ to $0{\cdot}57\alpha$); but a very large increase in the number of species in group V (from $0{\cdot}30\alpha$ to $4{\cdot}90\alpha$). In fact, of the total increase of $5{\cdot}28\alpha$ species, due to the increase of quadrat size, $4{\cdot}60\alpha$ are added to the species in group V.

The actual species in group I have probably all changed, as the sample size has been increased 200 times, and many if not most of the species originally in the rarer groups will have moved up into the dominant group. It is, of course, obvious that if the size of the quadrat was very large, say one acre, nearly all the species would be in nearly all the quadrats. So if we increase the size of the quadrats we increase the number and proportion of apparently dominant species.

The effect of altering the number of quadrats without altering their size is also quite distinctly seen in both Table 3 and Fig. 5. This effect is the reverse of the above. With increase in number of quadrats the number of species in group I, the rare species, is increased without any noticeable increase in the other four groups. Thus if $N=2\alpha$, of the $1{\cdot}53\alpha$ species added by increasing the number of quadrats from 5 to 25, $1{\cdot}28\alpha$ are added to group I and only $0{\cdot}3\alpha$ to group V; if $N=400\alpha$, of the $1{\cdot}61\alpha$ species added $1{\cdot}27\alpha$ are added to group I and only $0{\cdot}17\alpha$ to group V.

It is extremely difficult to calculate, on the assumption of the logarithmic series, the distribution of species in a greater number of quadrats than 25. Fig. 6, which has been

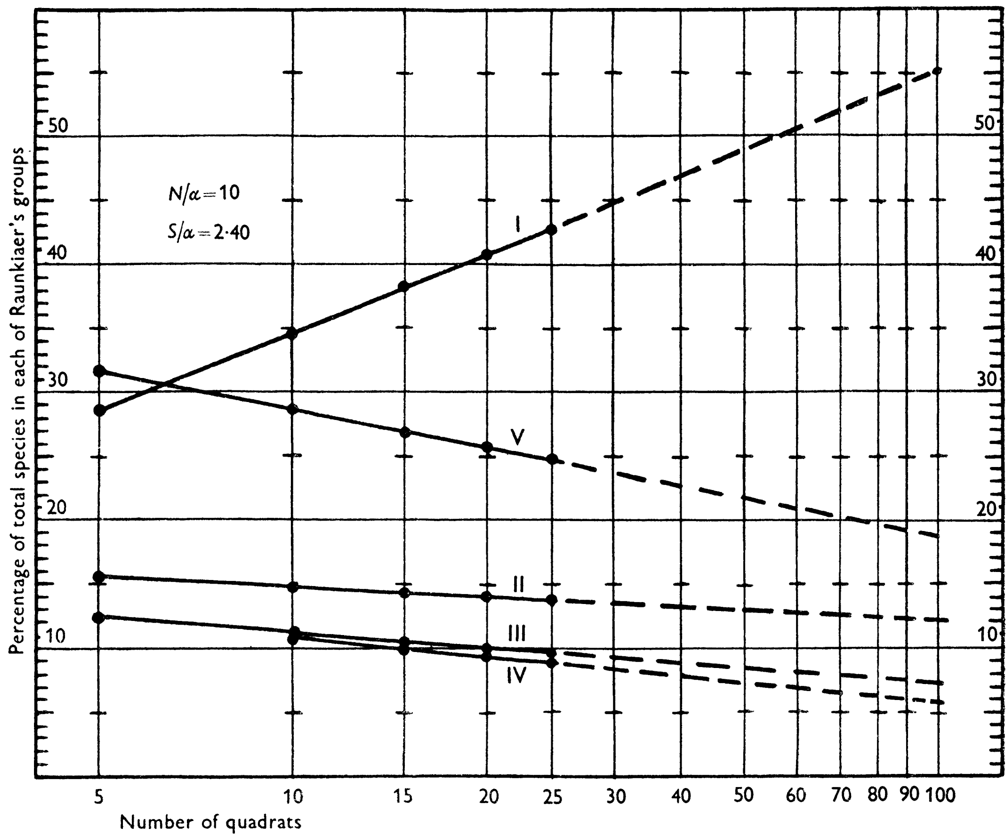

Fig. 6. The change in the percentage of the total species found in each of Raunkiaer's groups as the number of quadrats is increased, for the same association in which $N/\alpha=10$. Up to 25 quadrats by calculation, above 25 by extrapolation.

prepared from the data for $N/\alpha=10$ in Table 3, shows, however, that there is an almost straight line relation between the percentage of the total species in each of Raunkiaer's group and the logarithm of the number of quadrats up to 25. The figure has therefore been extended to show by extrapolation the percentage of species in the groups up to 100 quadrats.

For example, for 100 quadrats, when $N/\alpha=10$ (and $S/\alpha=2{\cdot}40$) the percentage of total species in the five groups is approximately 55, 12, 7, 6 and 9. On these 100 quadrats there should be $6{\cdot}90\alpha$ species.

Thus, according to the number of quadrats and their size (in relation to richness), the number and proportion of rare or dominant species can be altered within very wide limits.

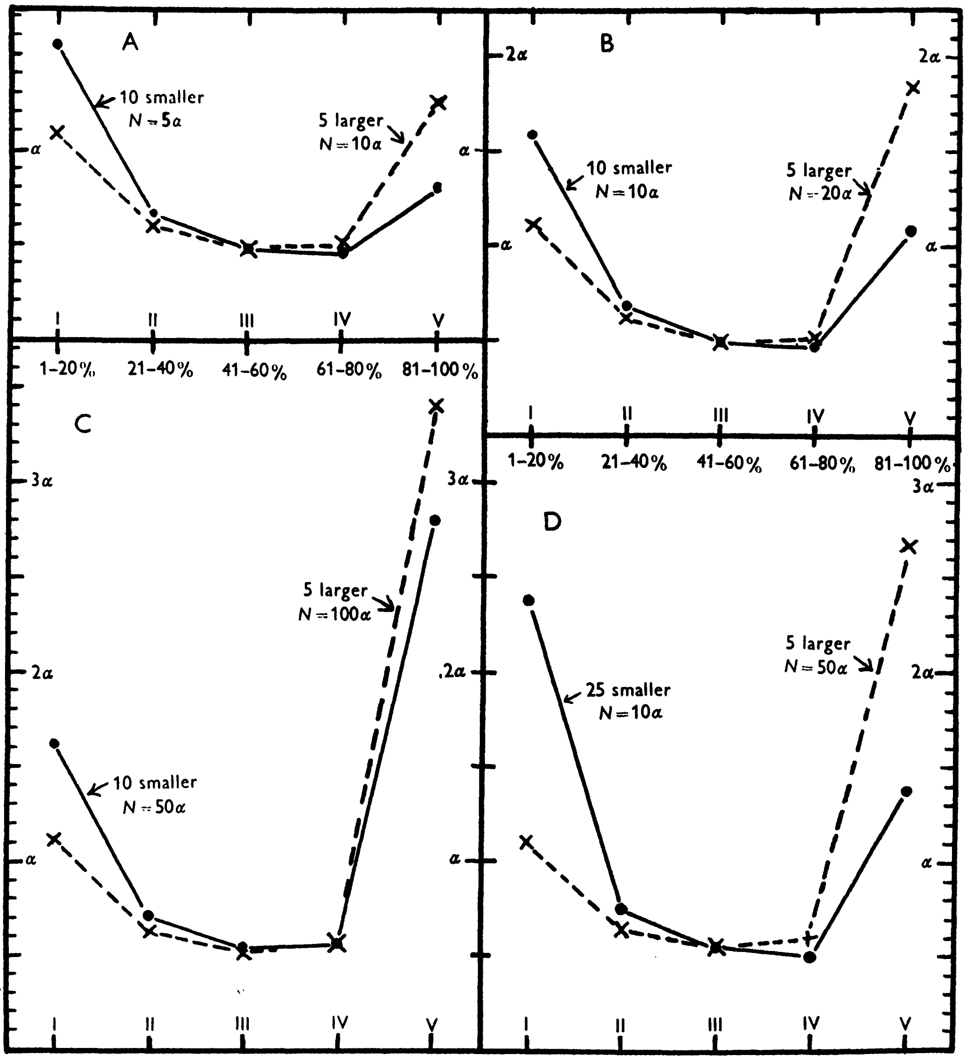

Fig. 7. Variations in the distribution of species in Raunkiaer's five groups of quadrats when a certain area is divided up into more smaller quadrats, or fewer larger ones.

It has already been shown (p. 111) that with a very large number of quadrats, the number of species in the relatively smaller number of quadrats 1, 2, 3, 4, etc., approaches very closely to a hyperbola starting with α species in 1 quadrat and $\alpha/2$ in 2, etc. The number of species in 1–5 quadrats would therefore be $\alpha(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5})=2{\cdot}28\alpha$ species. In our calculations for 25 quadrats and $N/\alpha=10$ it will be seen that there should be $2{\cdot}37\alpha$ species in the first group of 1–5 quadrats. Thus with only 25 quadrats the number of species in 1–5 quadrats has fallen to within 0·09α of the theoretical number for an infinite number of quadrats.

If we had 100 quadrats we would expect the first group (1–20) to contain just over 3·60α species; and if 200 quadrats the first group (1–40) would contain 4·28α species. If, however, we had doubled the size of the 100 quadrats the number of species in group I would scarcely have altered.

If we alter both the number and size of the quadrats, so that one compensates for the other and the total area sampled remains the same, we still get inconsistent results as shown in Table 4 and Fig. 7.

Table 4. *Same total area divided up into different numbers of quadrats*

	Quadrats	Group ...	I	II	III	IV	V
A	10	$N/\alpha=5$	1·57	0·67	0·47	0·43	0·80
	5	$N/\alpha=10$	1·09	0·61	0·49	0·49	1·25
B	10	$N/\alpha=10$	1·59	0·69	0·51	0·50	1·32
	5	$N/\alpha=20$	1·10	0·63	0·51	0·53	1·84
C	10	$N/\alpha=50$	1·61	0·72	0·55	0·56	2·79
	5	$N/\alpha=100$	1·11	0·64	0·53	0·57	3·36
D	25	$N/\alpha=10$	2·37	0·75	0·53	0·50	1·37
	5	$N/\alpha=50$	1·11	0·64	0·52	0·56	2·69

It will be seen again that the two results are quite different in each example. The larger number of smaller quadrats gives more rare species in group I, and the smaller number of larger quadrats gives more dominant species in group V.

So we see that the number and proportion of species in the five groups of quadrats as defined by botanists depend entirely on the number and size of the observed quadrats. To get comparable results in a number of studies the number of quadrats must be kept constant, and within a single association the size of the quadrat must also be defined. If different ecological associations are to be compared one with another then the size of the quadrat as measured either by the number of individuals or the number of spaces must bear a fixed relation to the Index of Diversity of the flora. The quadrats must be larger in 'rich' areas and smaller in 'poor' areas, or, in the terms of the present theory and notation, N/α and S/α must be constant.

V. POSSIBLE SOURCES OF ERROR IN FIELD OBSERVATIONS

The above calculations, as has been mentioned, are based on a theoretically perfect distribution of units and on a perfect randomization of the quadrat samples.

Before testing the results against field observations, it is necessary to consider in what direction errors are likely to occur in practice and then we can see if the differences between observed and calculated results are in the direction expected, and if they are large enough to obscure any possibility of seeing the mathematical laws behind the variability.

We have assumed a perfectly random distribution of the units over the area, and that each quadrat is identical ecologically in every way with all the others, all differences between quadrats being due to mathematical chance of distribution of individuals of the plant species.

With regard to the distribution of the units this is liable to be interfered with by aggregation, which causes a particular number of plants to be found on fewer quadrats. It must be recollected, however, that in the conception of 'plant-unit' as opposed to 'individual' a certain allowance has been made for aggregation, particularly when a close aggregation of plants is behaving from the point of view of distribution as an indivisible block.

With regard to the sampling by quadrats, it is obvious that in practically all natural botanical populations there is a greater difference between any 2 quadrats than would be expected by chance distribution of the units. No natural area, however small, is really uniform. One quadrat may be a little lower than another; a little more shaded, a little better drained, etc. Even the fact that some animal excreta, a cow-pat for example, has fallen within a quadrat perhaps a year before, will alter its flora and make it not strictly comparable with the others.

The effect of all such differences will be to reduce the number of species common to a large number of quadrats, and particularly those found on all the quadrats; because among 10 quadrats or more one or two may differ sufficiently to be unsuited for an otherwise generally distributed plant. This effect will be greater with relatively small quadrats.

We have therefore to expect in natural field samples errors due mainly to lack of uniformity in the distribution of the flora.

A second error which may occur in field observations is the occasional missing of a rare and inconspicuous species which occurs in perhaps only one of a number of quadrats. This is likely to reduce slightly the total number of observed species, and also by a greater proportion the number observed on a single quadrat.

A more serious, but avoidable, source of error in field work, from the point of view of statistics, is that botanists do not always attempt to 'randomize' their quadrats properly over the area to be studied. Small numbers of quadrats are often taken in contact with one another, and while these may be excellent representations of the actual area covered, they are not fair samples of the whole area of the association.

I have further been informed that some botanists, if they are taking quadrat samples on an '*Erica*-Heath' association would not include any quadrat which did not contain an *Erica* plant. Such a selection, or rather elimination, would fail to give any idea of the frequency in the area of this dominant *Erica* species; in fact it is forced to be in 100 % of the samples.

Samples should be taken absolutely at random over the area, with no conscious selection or intelligent reasoning used to determine the exact spot. It is sometimes useful to take two samples close together at each of a series of randomized spots. It is then possible to compare the differences in distribution between quadrats close together and quadrats further apart.

Consciously selected quadrats, will usually increase the number of species in many quadrats, and reduce the number in a few, because the tendency will be to select either quadrats close together, or 'typical' areas, which means those which conform with a fixed conception of the sampler.

VI. Jaccard's observations in Switzerland

To be suitable for comparison with the above theoretical results, information is required on the presence or absence of a large number of species of plants on a large number of quadrats. I am indebted to Dr W. B. Turrill for drawing my attention to the work of Paul Jaccard, which, although not perfect for this purpose, does give a considerable amount of suitable data.

In 1908 Jaccard published a paper entitled 'Nouvelles Récherches sur la Distribution Florale', in which he discussed the distribution of 92 species of plants on 52 quadrats of 1 sq.m. each, in the Alpine valley of the River Grand Eau near Ormont Dessus, in the department of Vaud, Switzerland; and at a height of about 4000 ft. (1200 m.). The 52 quadrats were in nine groups of from 2 to 8 quadrats each, the quadrats in each group being in contact with one another on at least one side, and the different groups being 'about a kilometre apart up the valley'; except for groups IV and V which were 'in the same meadow'. In group IV the quadrats were in a line up and down, and in group V across the slope; the last group IX was 'about 30 m.' from group VIII.

He gives a table showing the occurrence of all the species in the different quadrats, and uses the information as the basis of a statistical discussion. Unfortunately, he chose for much of his discussion characteristics which were dependent on the size of the quadrat and not true properties of the ecological plant community (see Williams, 1949), but the data he gives are suitable for testing in relation to the logarithmic series.

Jaccard's table of data, rearranged in order of frequency of occurrence in quadrats, is reproduced in Fig. 8, and the list of the 92 species, as named by him, is given separately in Table 5.

It is first necessary to see what evidence exists that the samples obtained by Jaccard conform to the properties of the logarithmic series.

It has already been pointed out that if N is the number of plant-units in a quadrat taken in a log series population then the number of species on q quadrats is given by

$$S_q = \alpha \log_e \left(1 + q\frac{N}{\alpha}\right),$$

where α is the Index of Diversity of the population.

If this is the case there should be an approximately straight-line relation between S and $\log q$ (provided that qN is large compared with α); that is to say between the number of species and the logarithm of the number of quadrats.

Table 6 shows the average number of species on 1 quadrat, on 2, on 4, and on all the quadrats for Jaccard's nine groups, and Fig. 9 shows the same data in graphical form plotted as number of species, and log number of quadrats.

It will be seen that (excluding group IX which only consists of 2 quadrats), five groups (II, III, IV, VI and VII) show almost a straight-line relation between log area and number of species, but the other three groups (I, V and VIII) have a tendency to a curved-line relation with the 2 and 4 quadrats having a little too many species, or the 1, 7, and 8 quadrats a little too few. The most definitely curved are groups I and VIII, but even in these the error from a straight line is usually less than two species.

This curvature would however, be expected, if species which would normally have occurred in most or all of the quadrats, are prevented from occurring on the expected number by lack of uniformity in the area sampled.

So we see that there is some evidence that Jaccard's data conform with this property of the log series. As the slope of the line is a measure of the Index of Diversity, it appears also that group I has a higher Index, i.e. a 'richer' flora, than group III, and group III is richer than group VIII, etc.

It follows from the log series that, provided the samples are not small, the increase in

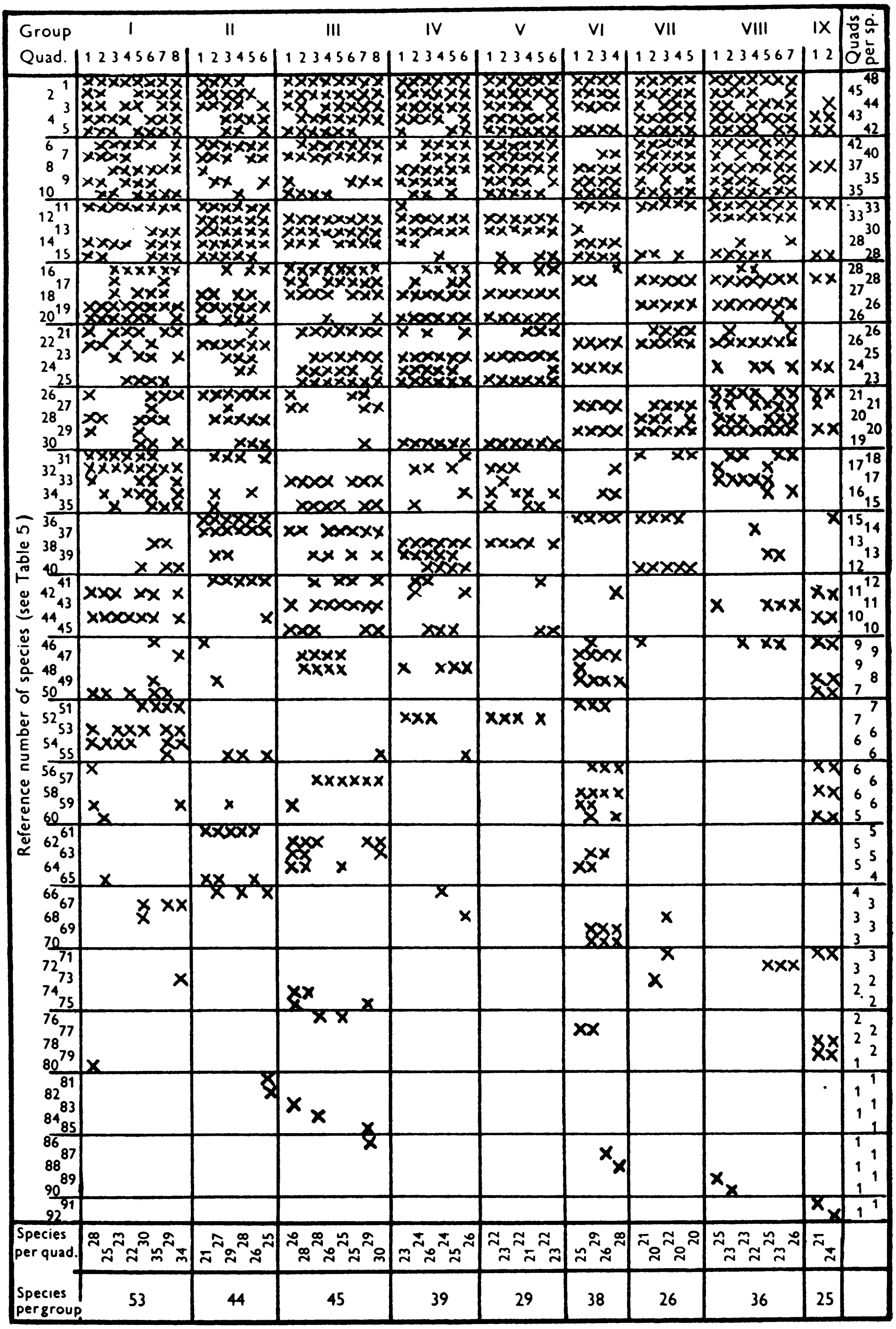

Fig. 8. The distribution of 92 species of plants on 52 quadrats of 1 sq.m. as observed by Jaccard in an Alpine valley in Switzerland (redrawn from Jaccard, 1908).

Table 5. *List of species in Jaccard's 52 quadrats, rearranged in order of frequency of occurrence in quadrats with number of occurrences in brackets*

1 Trifolium pratense (48)
2 Alchemilla pratensis (45)
3 Chrysanthemum leucanthemum (44)
4 Festuca pratensis (43)
5 Leontodon hispidus (42)
6 Dactylis glomerata (42)
7 Campanula romboidalis (40)
8 Taraxacum officinale (37)
9 Lathyrus pratensis (35)
10 Ranunculus acer (35)
11 Colichicum autumnalis (33)
12 Trisetum flavescens (33)
13 Geranium silvaticum (30)
14 Brunella vulgaris (28)
15 Avena pubescens (28)
16 Anthoxanthum odoratum (28)
17 Vicia cracca (28)
18 Anthriscus sylvestris (27)
19 Alectorolophus hirsutus (26)
20 Cynosurus cristatus (26)
21 Poa pratensis (26)
22 Carum carvi (26)
23 Rumex acetosa (25)
24 Trifolium repens (24)
25 Polygonium bistorta (23)
26 Sanguisorba minor (21)
27 Lotus corniculatus (21)
28 Medicago lupulina (20)
29 Plantago media (20)
30 Crepis taraxifolia (19)
31 Briza media (18)
32 Plantago lanceolata (17)
33 Cerastium caespitosum (17)
34 Tragapogon pratense (16)
35 Phyteuma spicatum (15)
36 Silene inflatà (15)
37 Knautia arvensis (14)
38 Melandrum silvestre (13)
39 Bromus erectus (13)
40 Agrostis canina (12)
41 Myosotis silvatica (12)
42 Primula elatior (11)
43 Deschampsia caespitosa (11)
44 Carex sempervivens (10)
45 Vicia sepium (10)
46 Centaurea jacea (9)
47 Festuca ovina (9)
48 Achillea millefolium (9)
49 Alectorolophus minor (8)
50 Linum catharticum (7)
51 Campanula rotundifolia (7)
52 Veronica chamaedrys (7)
53 Bellis perennis (6)
54 Equisetum avense (6)
55 Galium asperum (6)
56 Anthyllis vulneria (6)
57 Ranunculus montanus (6)
58 Gentiana campestris (6)
59 Thymus serphyllum (6)
60 Polygala vulgaris (5)
61 Ranunculus bulbosus (5)
62 Trifolium badium (5)
63 Carex pallescens (5)
64 Potentilla silvestris (5)
65 Viola tricolor (4)
66 Centaurea montana (4)
67 Geum urbanum (3)
68 Ranunculus aconitifolius (3)
69 Ajuga reptans (3)
70 Picia excelsa (3)
71 Phyteuma orbiculaire (3)
72 Polygonium aviculare (3)
73 Cirsium oleraceum (2)
74 Crepis paludosa (2)
75 Luzula silvatica (2)
76 Hieraceum pilosella (2)
77 Gentiana verna (2)
78 Carlina acaulis (2)
79 Hippocrepis comosa (2)
80 Viola hirta (1)
81 Agropyrum repens (1)
82 Hypericum quadrangulum (1)
83 Aposeris foetida (1)
84 Trollius europaeus (1)
85 Astransia major (1)
86 Juncus lamprocarpus (1)
87 Veronica teucrium (1)
88 Crepis aurea (1)
89 Listera ovata (1)
90 Veratum album (1)
91 Brachypodium pinnatum (1)
92 Biscutella laevigata (1)

Table 6. *Number of species in different numbers of quadrats in Jaccard's seven groups, together with calculation of constants on the basis of the logarithmic series*

Average no. of species	Jaccard's group nos. I	II	III	IV	V	VI	VII	VIII	IX
In 1 quadrat	28·3	26·0	27·1	24·7	22·2	27·0	20·6	23·9	22·5
In 2 quadrats	38·4	33·5	34·6	30·8	25·6	32·5	23·3	29·1	25·0
In 4 quadrats	46·2		40·3			38·0		33·4	
In all quadrats	53	44	45	39	29	—	26	36	—
No. of quadrats	8	6	8	6	6	4	5	7	2
Index of Diversity (α)	12·1	10·1	8·7	8·0	3·8	8·0	3·4	6·3	3·6

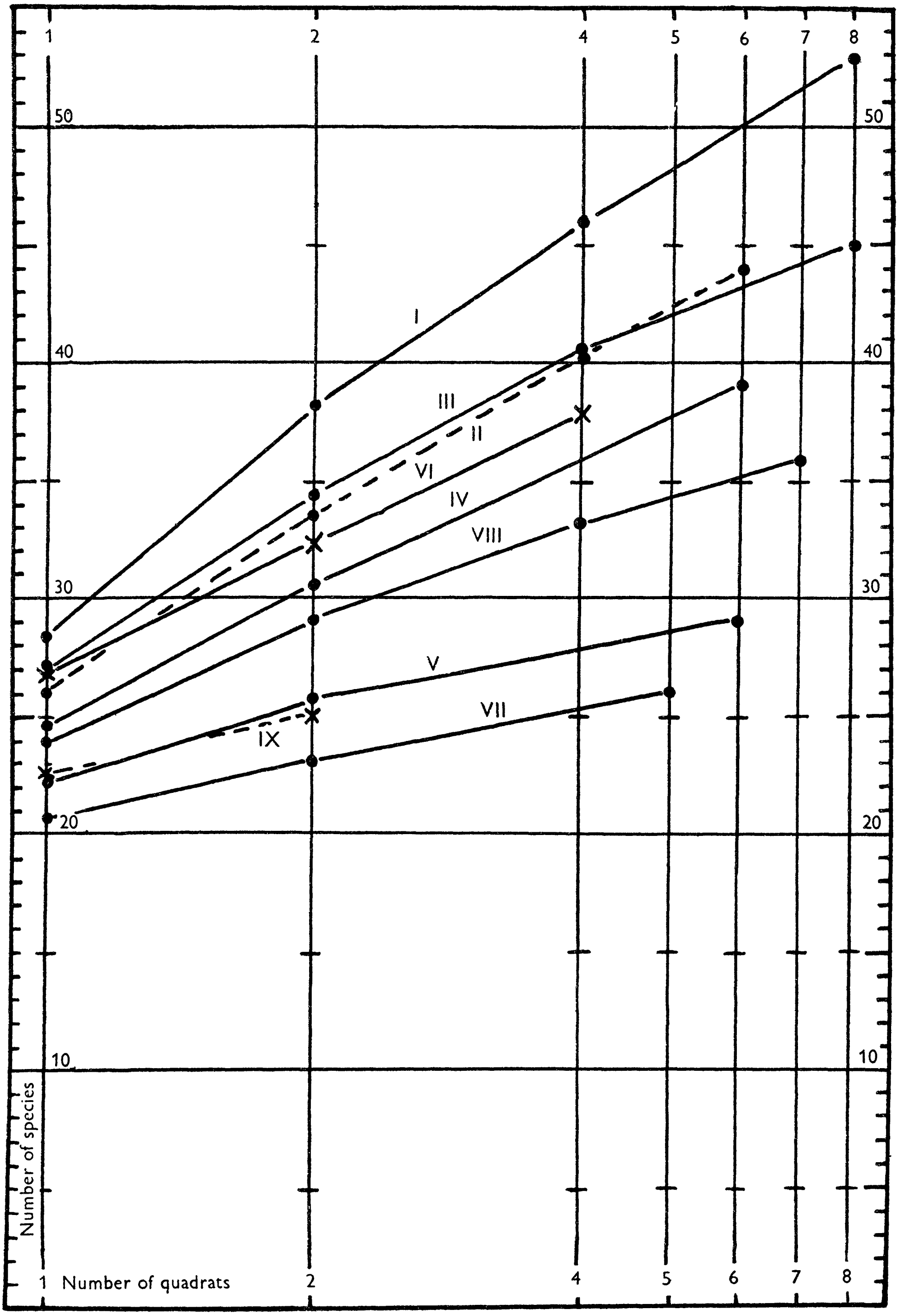

Fig. 9. The relation between the number of species and log area (number of quadrats) in the nine groups into which Jaccard's 52 quadrats were divided.

the number of species by multiplying the size of the sample by K is approximately $\alpha \log_e K$, or

$$\alpha = \frac{\text{increase in number of species}}{\log_e K}$$

Applying this to Jaccard's data we find for group I that there is an average of 28·25 species in 1 quadrat, and 53 for all 8 quadrats, hence

$$\alpha = \frac{53 - 28{\cdot}25}{\log_e 8} = \frac{24{\cdot}75}{2{\cdot}079} = 12 \text{ approx.}$$

Hence the rate of increase in the number of species in Jaccard's group I can be explained by the log series as samples from a population with an Index of Diversity of about 12. The calculated values of α for all nine groups are shown in Table 6.

Table 7 shows the average number of species in each of Jaccard's nine groups which occur in 1, 2, 3, etc., quadrats, together with the numbers calculated on the theories outlined above. The two sets of data are shown graphically in Fig. 10, the observed data by vertical columns of dots and the calculated by a light broken line.

Table 7. *The number of species on each of Jaccard's nine groups, which occur in* 1, 2, 3, etc., *quadrats, together with the numbers calculated on the assumption that the population is arranged in a logarithmic series*

Jaccard's group no.		No. of species which occur only in quadrats								Total no. of species in all quadrats
		1	2	3	4	5	6	7	8	
I	Obs.	11	4	4	8	6	11	5	4	53
	Cal.	12·6	6·8	4·9	4·0	3·7	3·7	4·4	13·0	
II	Obs.	11	3	8	5	7	10	—	—	44
	Cal.	10·9	6·0	4·5	4·1	4·5	13·2	—	—	
III	Obs.	7	4	3	7	2	5	11	6	45
	Cal.	9·2	5·0	3·6	3·0	2·8	2·9	3·5	15·1	
IV	Obs.	8	3	7	4	5	12	—	—	39
	Cal.	8·7	4·8	3·7	3·4	3·8	14·4	—	—	
V	Obs.	3	1	4	3	4	14	—	—	29
	Cal.	4·1	2·1	1·8	1·7	2·0	17·1	—	—	
VI	Obs.	8	8	4	18	—	—	—	—	38
	Cal.	9·1	5·5	5·1	18·0	—	—	—	—	
VII	Obs.	4	0	3	5	14	—	—	—	26
	Cal.	3·8	2·2	1·8	2·0	16·1	—	—	—	
VIII	Obs.	4	6	2	3	4	6	11	—	36
	Cal.	6·8	3·7	3·0	2·3	2·4	2·9	15·1	—	
IX	Obs.	5	20	—	—	—	—	—	—	25
	Cal.	5·6	19·6	—	—	—	—	—	—	

It will be seen, particularly from the figure, that in general there is a resemblance between the observed and calculated results. The fit is particularly good in groups V–IX, when the observed numbers both in one and in all quadrats, are fairly close to the calculated. Group IV is fairly close at both ends, but somewhat irregular in the middle. The poorest fits are found in groups I and III, and this is particularly so with the species on all the quadrats.

In group I we expect about 12 species in all 8 quadrats and find only 4. In group III we expect about 18 and find only 6. But these two groups contain the largest number of quadrats and we have already pointed out that with a larger number of quadrats it becomes increasingly unlikely that they will represent samples of a really uniform area.

The low numbers of species found in all quadrats would be expected from a lack of uniformity in the area sampled.

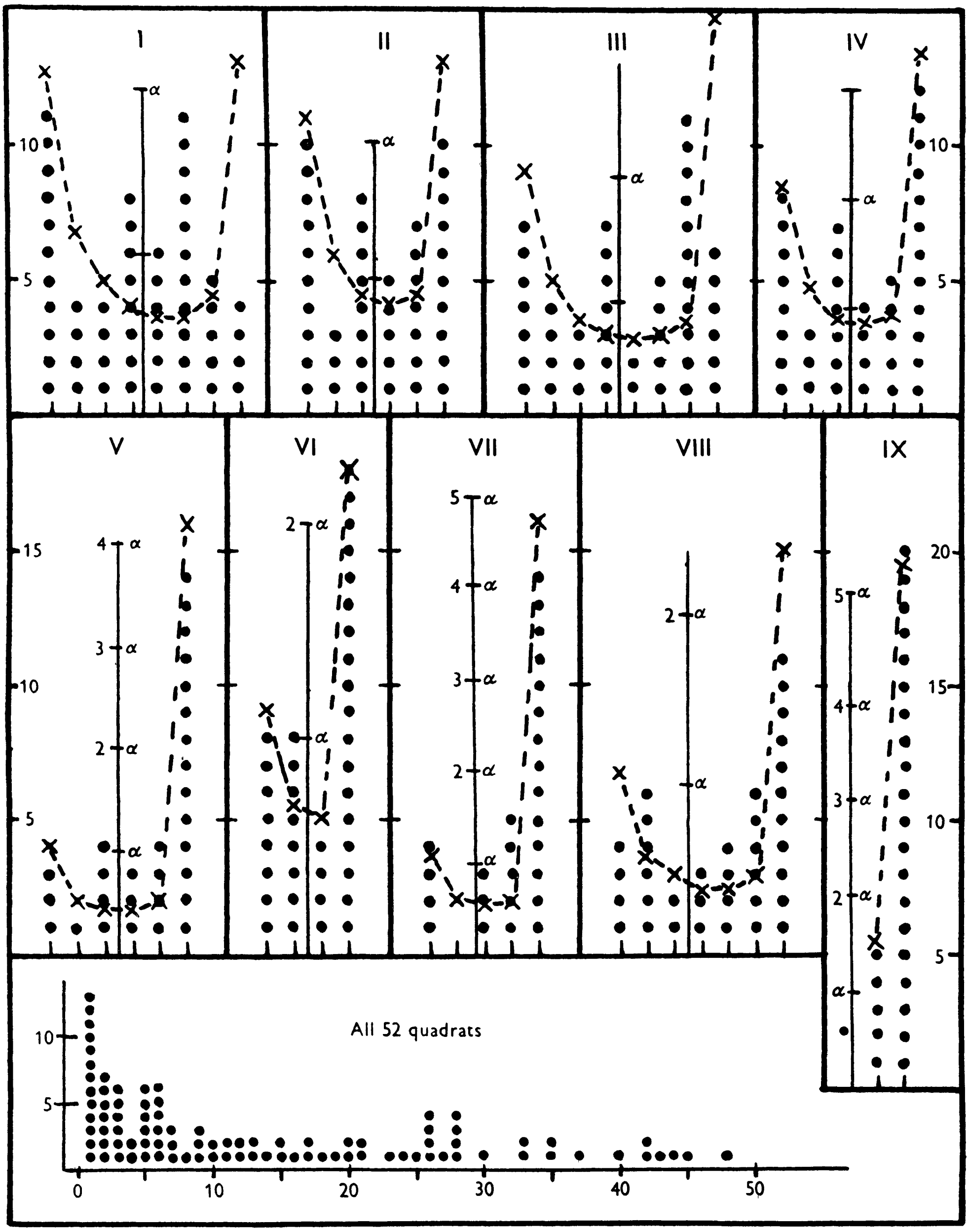

Fig. 10. The frequency distribution of species among the quadrats in the nine groups of quadrats observed by Jaccard. Dots represent Jaccard's observations; broken lines represent calculations on the basis of the logarithmic series. Also, below, the distribution of species in all 52 quadrats.

It should also be noticed that the observed numbers of species in one quadrat only is a little below the calculated in seven out of the nine of Jaccard's groups, in one case is almost exactly correct, and in only one case is slightly above. Thus the direction of difference again corresponds to the possible error of missing in the field one or two of the rarest species.

The theoretical distribution of the species in Raunkiaer's five groups can be partly tested by Jaccard's observations over the whole area. Of the 92 species that he found on the 52 quadrats, 49 occurred in 10 quadrats or less, which corresponds to Raunkiaer's 1–20% group. The Index of Diversity for the whole 52 quadrats calculated from his data is 17. Therefore there were $2{\cdot}88\alpha$ species in the first group. On the theory, assuming that the first 10 of 50 quadrats will be very close to a hyperbola, we should have found just over $2{\cdot}93\alpha$ species, which is remarkably close to the observed figure. In the second group, 11–20 quadrats, Jaccard found 16 species, which is approximately $0{\cdot}95\alpha$, and the calculated figure from the hyperbola is something over $0{\cdot}67\alpha$. The higher observed figure is to be expected because in group II the values would be distinctly above the hyperbola, and also because the total area sampled by Jaccard was far from uniform. Close agreement in the higher groups would not be expected for this same reason.

VII. Pidgeon and Ashby's observations in New South Wales

For reasons already discussed, it is usually difficult to obtain field records which show the numbers of individual plants for each species in a given area. However, Pidgeon & Ashby (1940), studying arid land in New South Wales where the vegetation was sparse, have obtained some figures which are considered below.

The survey was carried out to see the effect of fencing to exclude rabbits and stock from natural vegetation near Broken Hill, where there was trouble from drifting sand. An experimental enclosure of 22 acres had shown a rapid and beneficial increase in grasses and shrubs.

Four areas were studied: (1) West Reserve fenced; (2) West Reserve unfenced; (3) South Reserve fenced; and (4) South Reserve unfenced. In addition, there was (5) a permanent quadrat of 50 metres square in one of the fenced reserves. The quadrats were of elongate shape: 10 m. × 15 cm. with an area of 1·50 sq.m. Fifty contiguous quadrats were studied on areas (1) and (2); forty on (3); and thirty-nine on (4).

Pidgeon & Ashby give three sets of data which are of interest in the present discussion:

(1) The number of species observed in 1, 5, 10, 50, 100 and 200 sq.m. for areas 1, 2, 3 and 5 (Table 8).

Table 8. *Number of species of plants on 1–200 sq.m. in different areas at Broken Hill, New South Wales*

Area (sq.m.)	West fenced	West unfenced	South fenced	Permanent quadrat
1	12	8	5	5
5	18	16	12	9
10	19	17	13	10
50	23	24	17	17
100	28	26	21	21
200	31	29	27	24
12,000 (3 acres)	—	37	—	—
32,400 (7 acres)	50	—	47	—
α cal. from 1 and 200 m.^{2}	4·0	3·6	4·2	3·6

(2) The number of individuals in each species on 39–50 quadrats in areas 1–4. They unfortunately multiplied these figures to give the number of individuals on 100 sq.m.,

but as this is not desirable, I have recalculated back to the original areas sampled (Table 9).

(3) The distribution of the species in different percentages of quadrats according to Raunkiaer's five groups (Table 10).

Unfortunately the data in Table 8 do not seem to have been taken from the same sets of quadrats as in Tables 9 and 10. In the three areas mentioned in Tables 8 and 9 the numbers of species on 60–75 sq.m. in Table 9 are all greater than the numbers on 200 sq.m. in Table 8. This makes it very difficult to compare the results.

Pidgeon & Ashby used an empirical formula to relate the number of species to the size of the area sampled. In the notation of the present paper it can be written.

$$S_q - S_1 = m \log_{10} q.$$

They considered that this gives a very close representation of their results, but they do not give any calculated values of their constant 'm'.

It will be seen that this relation is exactly the same as is found in the logarithmic series when the sample size is large enough to neglect the 1 in relation to N/α in the formula

$$S_q = \alpha \log_e = \left(1 + \frac{N}{\alpha}\right),$$

i.e. when the line showing graphically the relation between the number of species and the log of the area sampled has become approximately straight.

Their 'm' is therefore equal to $\alpha \log_e 10$ or $2{\cdot}30\alpha$, and their formula is identical with that of the log series for large samples, but is not correct for small quadrats.

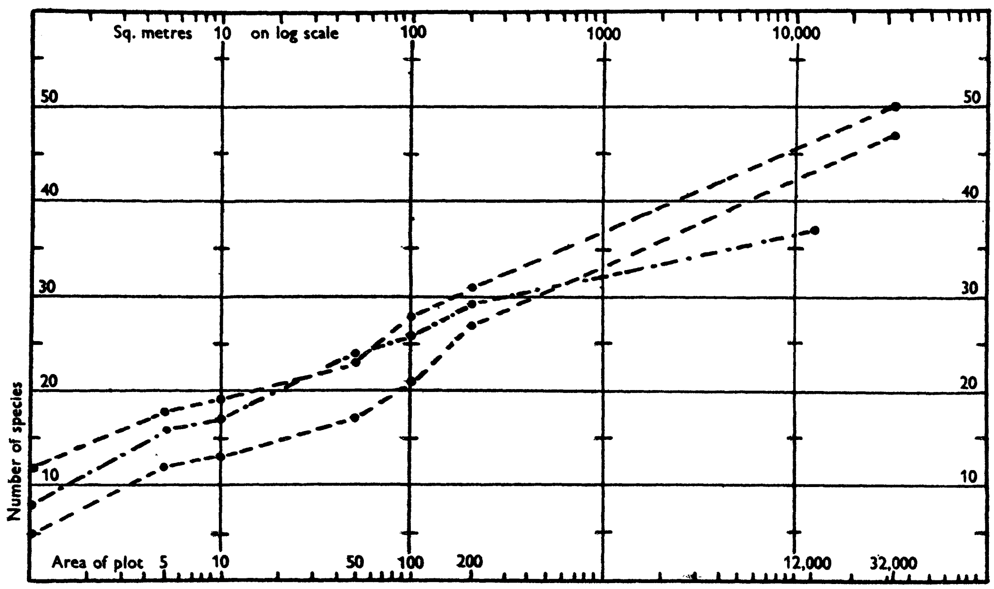

Fig. 11. Numbers of species on different plots observed by Pidgeon & Ashby, plotted against log scale of area.

Fig. 11 shows the data from Table 8 with a logarithmic scale for area of sample. The results are somewhat irregular, owing, I think, to there being only a single observation at each level. They are however, with the exception of the 3-acre number in West Unfenced, seldom more than one species out from the straight line which is expected from the log series, and so lend support to the idea that the population may be arranged in some distribution of this nature.

The information given in Table 9 and graphically in Fig. 12 is more detailed, and shows in each case more species with one individual than with two or any larger number, and a general resemblance to the hollow curve of the logarithmic series.

Table 9. *Number of species with different numbers of individuals in the four areas sampled*

No. of individuals per species	West fenced 50 quadrats, 75 sq.m.	West unfenced, 50 quadrats, 75 sq.m.	South fenced, 40 quadrats, 60 sq.m.	South unfenced, 39 quadrats, 58 sq.m.
1	7	6	9	4
2	0	3	5	3
3	6	3	2	2
4	0	0	0	0
5	3 (16)	4 (16)	0 (16)	0 (9)
6	1	0	0	3
7	0	1	1	1
8	0	1	1	2
9	0	1	3	1
10	1 (2)	0 (3)	0 (5)	1 (8)
	And also at 12, 17 (2), 23, 35, 38, 44 (2), 48, 98, 99, 118, 279, 316, 335, 719 and 1849	And also at 16, 17, 45, 48, 73, 98, 131, 154, 224, 758 and 1155	And also at 12, 20, 22, 24, 27 (2), 44, 66, 110, 131, 234, 346 and 667	And also at 11, 12, 13 (3), 15, 23, 34, 51, 107, 274, 337, 366, 460 and 698
Total plants (qN)	4147	2784	1797	2503
Total species S_q	35	30	34	32
Average plants:				
Per species	118	93	51	78
Per quadrat	83	56	45	64
Per sq.m.	55	37	30	43
Index of Diversity	5·3	4·8	6·2	5·2
α from qN and S_q from Table 8	3·6	4·0	4·2	—
n_1 calculated	5·3	4·8	6·2	5·2
n_1 observed	7	6	9	4
n_1–n_{10} cal. approx.	15·5	14·1	18·1	15·3
n_1–n_{10} observed	18	19	21	17

It will be seen that the calculated values of the Index of Diversity from these data are a little above those from Table 8. This would be expected as, as already mentioned, the numbers of species in 60–75 sq.m. in Table 9 are all above those for 200 sq.m. in Table 8 for the same areas.

The observed numbers of species with a single plant each are in three cases above the calculated, and in the fourth case below.

Pidgeon & Ashby separate in their original data the perennial from the annual plants, and I find that the perennial plants above give a better fit to the log series than either the annual plants or the whole lot together. The calculated n_1 for the perennials for all plots is 2·7 and the observed 2, and the calculated total for all species up to 10 individuals (i.e. $n_1 - n_{10}$) is 7·9 and the observed is 8.

This may be a coincidence, but it raises the question whether competition between and within the annuals, and between and within the perennials in a plant population may not be of a different order from the competition between the two groups.

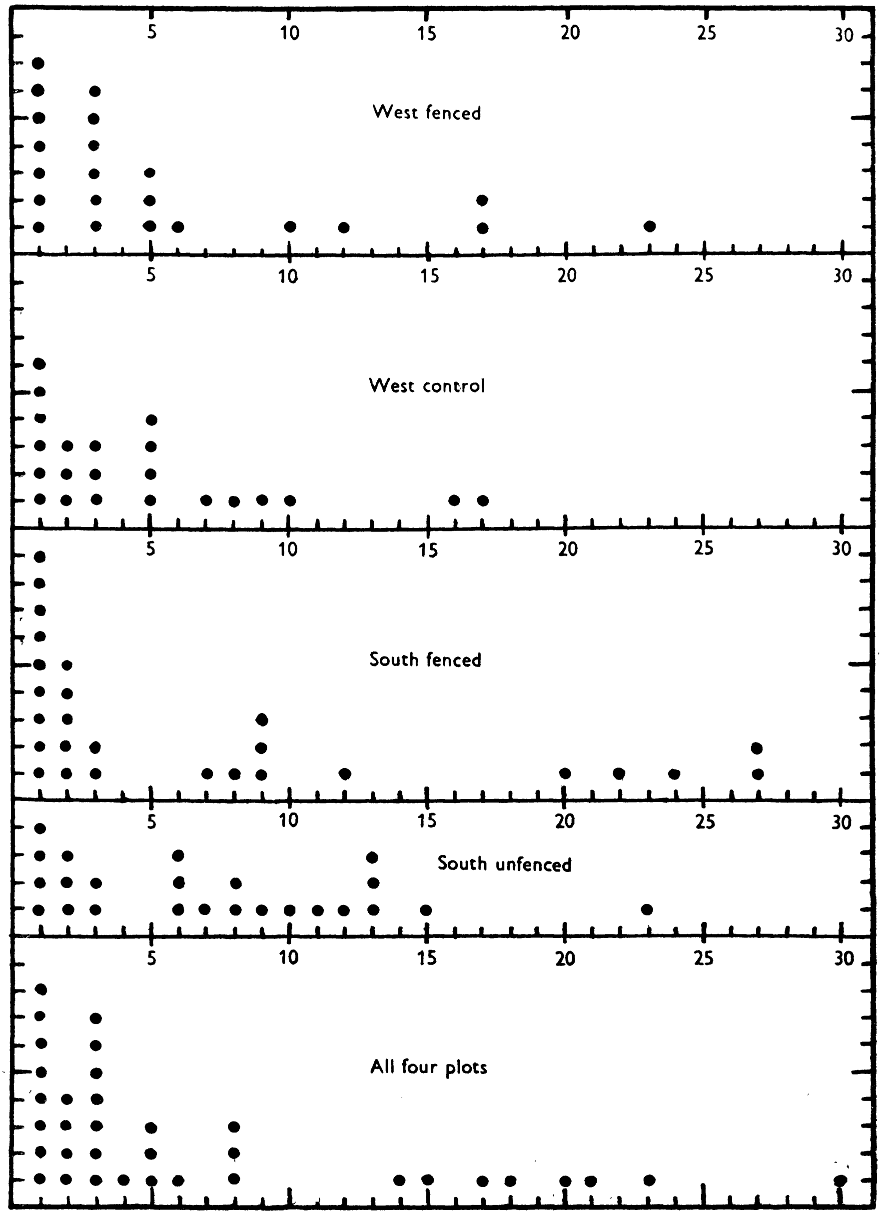

Fig. 12. Frequency distribution of species which occur in 1, 2, 3, etc., quadrats; in Pidgeon & Ashby's observations in New South Wales.

Table 10 gives Pidgeon & Ashby's results as classified in Raunkiaer's five groups.

Table 10. *Pidgeon & Ashby's results classified in Raunkiaer's groups*

	No. of quadrats	No. of species	Raunkiaer's groups I	II	III	IV	V
West fenced	50	35	19	6	3	2	5
West unfenced	50	30	19	2	4	2	3
South fenced	40	35	23	6	3	0	3
South unfenced	39	32	17	5	4	1	5

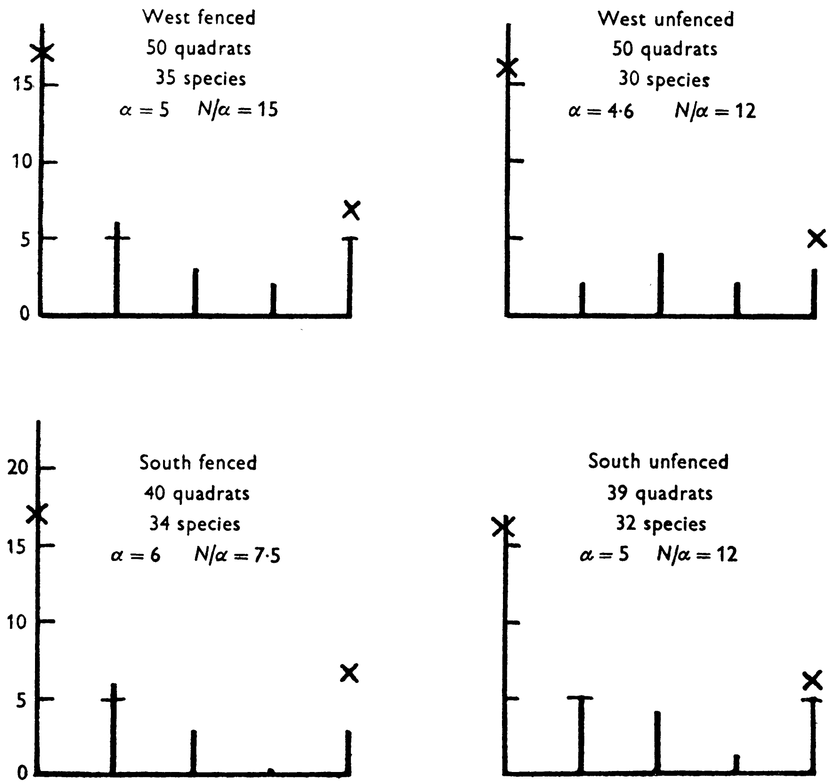

Fig. 13. Pidgeon & Ashby's New South Wales observations classified in Raunkiaer's five groups.

These results are shown graphically in Fig. 13, on which are marked approximate values for groups I and V, calculated by extrapolation from Figs. 4 and 5.

It will be seen that the observed figures for group I are too high and for group V are too low, otherwise there are more rare species than are required by the theory.

VIII. Discussion

The botanist, by the use of quadrat samples, is trying to get information about the structure of plant associations, and particularly under the following headings:

(1) The relative frequency of different species in a population.

(2) The comparison of two or more populations.

(3) The possibility of separating off certain more abundant species as dominant in, and therefore characteristic of, the association.

We have reason to believe that the relation between the distribution of individuals and of species in the original population, and in a series of random samples or quadrats, depends on three factors or variables:

(1) The size of the quadrat, which is measured either by the number of plant-units (N) which it includes; or indirectly by the number of species.

(2) The number of the quadrats (q).

(3) The Index of Diversity of the population (α).

The first two of these are under the control of the observer and can be altered to suit the particular problem, or his convenience. The richness or poorness of the flora, as measured by the Index of Diversity, is on the contrary a property of the association studied, and the value is fixed by natural causes. Two different associations may have different Indices, and this makes their comparison by quadrats complex.

The practical problem therefore is to see whether, by altering the size and number of the quadrats, it is possible to get comparable results from associations with different Diversity.

N can be kept constant or altered in proportion to α, q can be varied either independently of N, thereby altering the total area sampled, or in inverse ratio to N, so that the total number of plant-units sampled (Nq) remains constant.

In the earlier portion of this paper we have shown that the proportional distribution of species in the quadrats depends on q, and the ratio N/α or S/α, so that if the number of quadrats is fixed, and the size of the quadrat (in units) is varied in proportion to α, comparable results can be obtained from different associations, i.e. the same proportion of species will occur in the same proportion of groups.

If the units are grouped among species in a logarithmic series, then there must in any population be a perfectly smoothly falling series of values from the species represented by a single unit, and those represented by many. There are more species with one unit than with two, more with two than with three, more with three than with four, and so on. At no point in the series is there any mathematical break. Each species has a position in the series and so can be considered as rarer or commoner than another—but even this position may alter from time to time and from place to place in the association.

There is no mathematical justification for this separation of the species into those above or below any particular division as in Raunkiaer's five groups.

The first step in any analysis should be a rough calculation of the Index of Diversity. This can be obtained rapidly by finding the number of species on several areas of different sizes, preferably differing geometrically, e.g. each twice the size of the previous one. With this knowledge the size and number of quadrats can be chosen to suit the particular circumstances. The species of which the sample is made up can then be ranged in order of relative frequency of occurrence and quadrats; the more quadrats the finer the grading and the more species at the 'rare' end; the larger the quadrats the less the separation at the 'common' end, and the larger the number of 'dominant' species. It is then for the botanist to select any particular point in the series for division, realizing however, that all such divisions are arbitrary.

Further field work is obviously the next step in this investigation. A theoretical mathematical basis has been proposed, where before there were only empirical rules; so it should be possible to take more critical field observations, with more attention to randomization than has been done in the past.

It seems that the conception of the 'richness' of the flora, as here measured by the Index of Diversity, is the essential part of the interpretation of the structure of Plant Associations. In fact, if the logarithmic series is the true interpretation of the distribution of units into species, when we know the size of the quadrat, either in species or units, and the Index of Diversity, we know the whole pattern of the structure of the population: it is for the systematic botanist then to determine the species from which the pattern is made.

Two difficulties, however, remain. First, the logarithmic series, although giving approximate agreement, may not be the true interpretation; other mathematical series may be found to give even closer agreement. Secondly, the allowance made by the conception of 'plant-unit', instead of individual, for aggregation and dispersal of individuals in an association may not be sufficient to allow for the disturbing effects on distribution brought about by these factors. Only further research, both in the field and in mathematical analysis, can throw light on these points.

IX. SUMMARY

The object of the present paper is to show the possible application of certain statistical methods, which have already been found useful in the study of the structure of animal populations, to the structure of plant populations.

The main difference between the two problems is that with animals it is easy to define an individual, but often difficult to say to which particular area it belongs; whereas with the plants it is often difficult to define an individual but easy to say where it is growing. To overcome difficulties arising from this a 'plant-unit' is considered which, when not an easily defined individual plant, is a quantity of a species which in distribution behaves as an individual. It takes into consideration a certain amount of aggregation. The number of units in a series of samples from the same association is considered to be proportional to the area sampled.

Section II (p. 108) discusses the increase of the number of species represented as we increase the number of quadrats, and a formula is presented, based on the difference between the number of species in different sized samples (i.e. different numbers of quadrats), and the differences between these differences, from which it is possible to get values for the numbers of species found on any 1 particular quadrat only, or on any 2, or 3, etc., particular quadrats. From this it is possible to calculate the frequency distribution of species found only in 1 quadrat, in 2 quadrats or in 3 quadrats, and so on, out of any number of quadrats, if the total number of species is known for each additional quadrat added.

In section III it is shown that if the plant units in the population are distributed among the species in a logarithmic series, and if the quadrats are true random samples of a population distributed by chance; and further if we know: (1) the size of the quadrat as measured either by the number of species or the number of units, and (2) the Index of Diversity of the population, we can calculate the theoretical frequency distribution of species in any number of quadrats out of any total number. The mathematical formulae are given and Table 2 shows actual results for up to 25 quadrats, for certain values of the ratio between N and α or S and α.

Section IV is a discussion of the distribution of species in certain percentages of the total number of quadrats. Botanists have divided the species in an association into five groups: (1) those which occur in 1–20% of the quadrats, (2) in 21–40%, (3) on 41–60%, (4) on 61–80%, and (5) the common species which occur in 81–100% of the quadrats. Theoretical considerations, based on the assumption of distribution in a logarithmic series, indicate that the result of any such classification depends on the number of quadrats, on the size of the quadrats, and on the Index of Diversity of the population. Increase in the size of the quadrat increases the proportion of species in group V, and particularly those found in all of the quadrats. Increase in the number of quadrats increases the proportion of species found in a few only (group I) and reduces those in group V. No deductions can be made from the distribution of species in these groups unless all three factors are taken into consideration.

In section V there is a discussion of the sources of error in the field that might be expected to make observed values differ from the theoretical. The most important of these is lack of uniformity in the environment, which will reduce the numbers of species found on a very high proportion of quadrats.

In section VI observations by Jaccard on the distribution of 92 species of plants in 52 quadrats on an Alpine Valley is analysed, and the results are shown to fit moderately well to the calculated figures. The differences found between the two are in the direction that would be expected from the known sources of error. The values obtained for the Index of Diversity are quite consistent.

In section VII observations in New South Wales by Pidgeon & Ashby, which include the 'number of individuals' of each species, are shown to fit moderately well to the distribution calculated from the logarithmic series. A point of great interest is that Pidgeon & Ashby had developed empirically a formula relating the number of quadrats and the number of species, and this formula is identical with that deduced from the logarithmic series provided that the number of plant units is large, i.e. the quadrats are large in size or in numbers or both.

Section VIII is a discussion of various problems arising, particularly pointing out how it is impossible to avoid the conception of some Index of Diversity as a measure of the richness of the Flora, and how many of the difficulties of botanists in the past in the analysis of quadrat results have been due to neglecting this property of a population.

APPENDIX

A POSSIBLE TREATMENT OF THE ABOVE PROBLEMS WITHOUT ASSUMING THE EXISTENCE OF A LOG SERIES DISTRIBUTION

Some of the deductions made above are dependent on the assumption of the existence of a log series distribution in the individual-species frequency of the population. Section II, however, is independent of this assumption.

If one is not prepared to accept the log series as a basis of discussion; let us make only the single assumption that over a considerable range there is a straight-line relation between log area and number of species.

This actually follows from the log series, but it could also follow from other possible series. It has been accepted as true by Gleason (1922), and by Pidgeon & Ashby (1940), and I have given several applications of its occurrence in other botanical quadrat surveys (Williams, 1944).

The law cannot possibly be true for very small samples as its strict application to these implies that a very few individuals may represent no species or even a negative number. Good quadrats, however, should never be as small as this.

At the upper limit the curve may flatten out, but this will probably be above the area that is usually considered in a series of quadrats.

If we make the assumption that

$$\text{the number of species} = d \times \log \text{total area sampled}$$

then if the area is increased by ten times, the number of species is increased by 'd'.

Some populations will give a higher value of d, some a lower value. Fig. 14, AB shows a possible relation with 5 species on 1 quadrat and 10 on 10, so that $d=5$. The line CD is $d=10$, and EF is $d=20$. The number of species on 1 quadrat can of course be varied at will by altering the size of the quadrat; but the slope of the line—i.e. the value of d—is a property of the population sampled.

It is obvious that a population in which there is a rapid increase of species with increase of size of sample must (other things being equal) be richer in species than one in which there is a slower increase. Thus our factor 'd' is a measure of the diversity of the population.

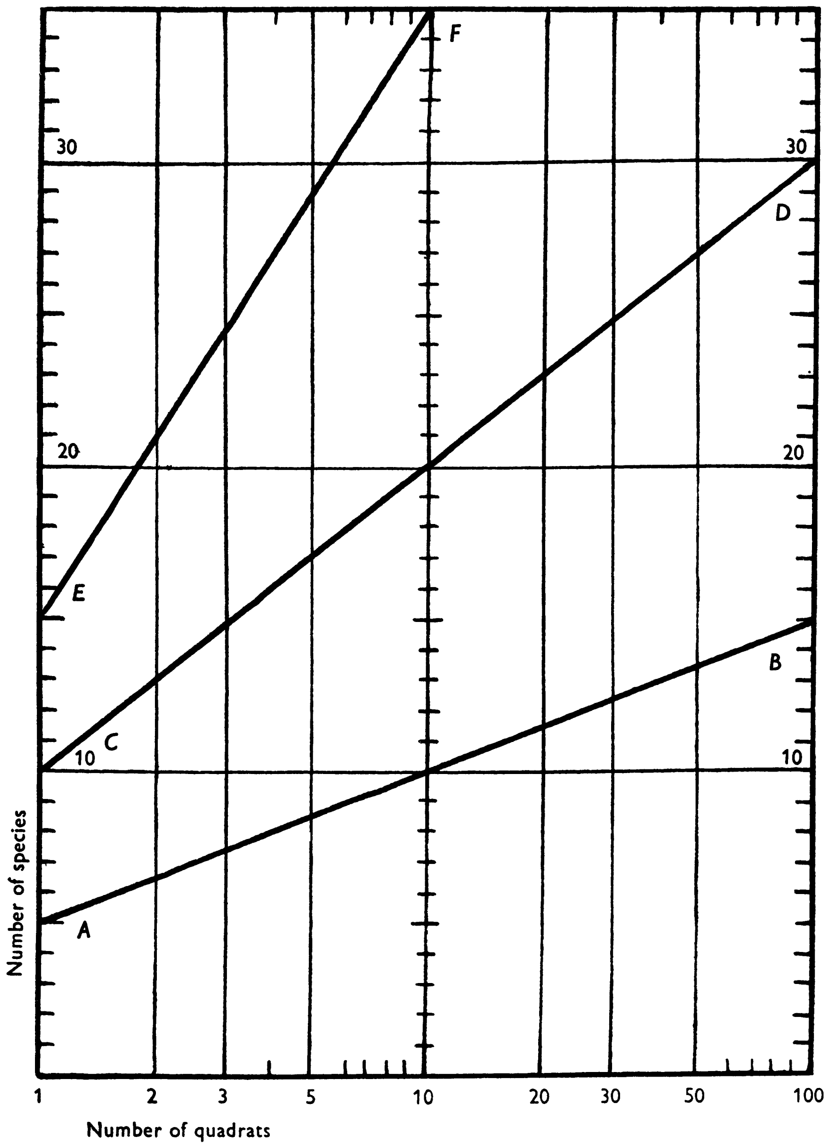

Fig. 14. Relation of log area to number of species in three theoretical populations.

If S_1 be the number of species on 1 quadrat, then the number on 2, 3, 4, etc., quadrats is given by

$$S_1 = S_1,$$
$$S_2 = S_1 + d \log 2,$$
$$S_3 = S_1 + d \log 2 + d \log 3/2,$$
$$S_4 = S_1 + d \log 2 + d \log 3/2 + d \log 4/3,$$

and $$S_n = S_1 + d(\log 2 + \log 3/2 + \log 4/3 \ldots + \log n/n-1).$$

which of course $= S_1 + d \log n$. In other words, the rate of increase of species, and hence (from section II above) the frequency distribution of species in quadrats, and so in Raunkiaer's groups of quadrats, is dependent only on d and S_1.

If, when comparing sets of quadrats from two different associations with different values of d, we choose the size of the quadrat so that in each association the number of species on 1 quadrat is proportional to the value of d; then S_1 would be Kd_1 in the first association and Kd_2 in the second.

Thus the frequency distribution of species in other quadrats would now depend in each case on $$d(K + \log 2 + 3/2 \times \log 4/3 \ldots),$$

i.e. only on d, since all the others are mathematical constants. This is exactly the same result that was reached by assuming the truth of the logarithmic series; namely that the frequency distribution of species occurring in different numbers of quadrats, and hence in the five Raunkiaer's groups, depends on the number of quadrats, the size of quadrats, and a measure of the diversity of the population.

To get rid of the effect of quadrat size (which is an accidental effect and not a property of the population) quadrat size in different associations should be chosen so that the number of species on a single quadrat is proportional to the diversity of the population.

Undoubtedly the main ecological property of the population that is being studied in all these distributions is the diversity.

REFERENCES

Bond, T. E. T. (1947). Some Ceylon examples of the logarithmic series and the Index of Diversity of plant and animal populations. *Ceylon J. Sci.* **12** (4), 195–202.

Fisher, R. A., Corbet, A. S. & Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12**, 42–58.

Gleason, H. A. (1922). On the relation between species and area. *Ecology*, **3**, 156–62.

Jaccard, P. (1908). Nouvelles récherches sur la distribution florale. *Bull. Soc. vaud. Sci. nat.* **44**, 223–70.

Pidgeon, I. M. & Ashby, E. (1940). Studies in applied Ecology. I. A statistical analysis of regeneration following protection from grazing. *Proc. Linn. Soc. N.S.W.* **65**, 123–43.

Raunkiaer, C. (1934). *Life Forms and Statistical Plant Geography.* Oxford (English Translation of many of his publications.)

Williams, C. B. (1944). Some applications of the logarithmic series and the Index of Diversity to ecological problems. *J. Ecol.* **32** (1), 1–44.

Williams, C. B. (1947a). The logarithmic series and the comparison of island floras. *Proc. Linn. Soc. Lond.* **158** (2), 104–8.

Williams, C. B. (1947b). The logarithmic series and its application to biological problems. *J. Ecol.* **34** (2), 253–72.

Williams, C. B. (1949). Jaccard's generic coefficient and coefficient of floral community, in relation to the logarithmic series and the Index of Diversity. *Ann. Bot.* N.S. **13** (49), 53–8.