Williams, C. B. 1944. Some applications of the logarithmic series and the index of diversity to ecological problems. *Journal of Ecology.* 32 (1), pp. 1-44.

Some Applications of the Logarithmic Series and the Index of Diversity to Ecological Problems

Author(s): C. B. Williams
Source: *Journal of Ecology*, Vol. 32, No. 1 (May, 1944), pp. 1-44
Published by: British Ecological Society
Stable URL: https://www.jstor.org/stable/2256485
Accessed: 24-10-2018 12:12 UTC

VOLUME 32, No. 1 MAY 1944

SOME APPLICATIONS OF THE LOGARITHMIC SERIES AND THE INDEX OF DIVERSITY TO ECOLOGICAL PROBLEMS

BY C. B. WILLIAMS, *Chief Entomologist, Rothamsted Experimental Station, Harpenden*

(*With twenty-five Figures in the Text*)

CONTENTS

1. INTRODUCTION

In a recent paper in the *Journal of Animal Ecology* (Fisher, Corbet & Williams, 1943) R. A. Fisher has discussed from the mathematical point of view the problem of the frequency of occurrence of species with different numbers of individuals in a random sample from a mixed population containing a large number of species; and in the same paper Corbet and myself have shown that Fisher's theories apply very well to observed random catches of large numbers of insects.

It had been previously noted on more than one occasion that if a sample of a number of individuals is taken from a mixed population, the number of species represented by only one individual will be comparatively high, the number of species represented by two individuals will be about half the number represented by one; the number represented by three individuals will be about one-third the number represented by one; and so on. This same distribution had also been stressed by Willis (1922) and Chamberlin (1924) in the frequency of the number of genera of plants and animals with one, two, three or more species.

It had, however, been generally assumed that the 'hollow curve' thus formed could be best represented by the harmonic series of the form

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

where n_1 is the number of groups with 1 unit.

Fisher *et al.* (1943), however, show that there is reason to believe that the facts are best represented by a logarithmic series

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

where x is a constant less than unity.

Several important biological and mathematical points follow from this. In the first place the new series, unlike the hyperbola, is convergent; and both the total number of groups (species) and the total number of units (individuals) are finite even for an infinite number of terms of the series. The total number of groups $= \frac{-n_1}{x} \log(1-x)$, and the total number of individuals $= \frac{n_1}{1-x}$. As there are two equations and only two 'unknowns' if we know for any sample the total number of species and the total number of individuals, we can rapidly find both n_1 and x, and hence can calculate the whole distribution.

Fisher has also shown a further development of great ecological interest. If a number of samples of different sizes be taken from the *same population*, it is found (both mathematically and by practical observation) that the ratio n_1/x is constant in all of them, and this ratio has been designated 'α' by Fisher.

It follows that if a sample is taken from a population, and the number of species and the number of individuals is counted, then n_1 and x can be calculated and from these α can be obtained. Once knowing α we can calculate the number of species that should be represented in any other sample of any size from the same population.

If α is large it indicates that the individuals are largely divided up into groups—there are more species for the same number of individuals. If α is small the individuals are divided up into fewer groups and, for the same number of individuals, the groups will be larger.

Since α expresses numerically the extent to which subdivision (or grouping) has taken place we have called it the 'Index of Diversity' of the population.

It is the purpose of the following paper to discuss the application of Fisher's logarithmic series and the index of diversity to a number of ecological problems connected with both botany and zoology and to show the wide application of the series to such problems, some of which have in the past been incompletely or only empirically explained.

It might be as well to begin by noting a few further points about the index of diversity:

(1) If several samples are taken from the same population, not only will they have the same index of diversity, but when two or more of them are added together the combined sample will still have the same index.

(2) Different populations may have, by chance, the same index of diversity, and samples from them will, of course, have the same α value. When these samples are added together, however, the α of the combined sample will be larger than that of either of the two or more original samples.

(3) If two populations are entirely apart—for example, a population of Geometrid and one of Noctuid moths; a population of Primulaceae and one of Violaceae; or plants in two widely separated areas with no species in common—then the α value of one sample from each combined will be the sum of the two α values for each sample separately.

If there is some overlap between the populations, as for example the flora of Spain and of England, the combined two samples will have an α value larger than either sample but smaller than the combined total. Thus the index of diversity of two combined samples, when compared with the values for the two samples separately, becomes a measure of the resemblance of the two populations from which the samples were obtained.

(4) If two samples, not of very small size, are taken from the same population the larger sample being p times the smaller, then the number of species in the larger sample will be $\alpha \log_e p$ more than in the smaller. It follows that if one sample is double the size of the other it will contain $0{\cdot}69\alpha$ more species (as $\log_e 2 = 0{\cdot}69$); and that α is the number of species added to a sample by multiplying its size by $e = 2{\cdot}718$.

This gives a method of finding the index of diversity of a population from two samples when their relative size, but not their absolute size in numbers of individuals, is known; as, for example, in the case of numbers of species on quadrats of different sizes. If two quadrats were used, one 1 m. sq. and the other 165 cm. sq. ($= \sqrt{2{\cdot}718}$), the difference in the average number of species in each would give α direct.

(5) For samples of different sizes for the same population x becomes gradually closer to unity as the size of the sample increases. For example, in the discussion below on British birds (p. 22), where the size of the sample is forty million pairs, $x = 0{\cdot}9999997182$. But we have seen that $n_1 = x\alpha$; so that as the size of a sample increases the number of species with one individual (or groups with one unit) gradually approaches α, but can never exceed it. The index of diversity of any population is therefore theoretically the limiting value of n_1, or the maximum number of species with one individual that can be obtained from one sample however large.

2. Problems of species, area and numbers of individuals

Botanical

Number of species of plants on a grassland area in Britain

G. E. Blackman (1935) gives the number of plants found on quadrats of different sizes in a grassland formation at Jeallots Hill, Bracknell, Berkshire, England, in August 1932. The data are shown in Table 1. It will be seen that the increase in number of species found by each successive doubling of the area is varying round 2·05; the first three values are rather below this figure.

Table 1

Size of quadrat in sq. in.	No. of species of plants	Additional no. obtained by doubling the area
2	5·9	—
4	7·8	1·9
8	9·5	1·7
16	11·1	1·6
32	13·6	2·5
64	16·1	2·5
128	18·2	2·1
	Av. increase	2·05

I have already shown (Fisher *et al.* 1943, p. 51) that, if the logarithmic series applies to a sample of a population, then the number of additional species obtained by doubling the size of the sample from the same population is approximately constant—provided the sample is not too small—and approximates to $\alpha \log_e 2$, which equals $0{\cdot}6931\alpha$ or alternatively $\alpha = 1{\cdot}443n$. If the sample is small the increase is less than this, but the smaller the value of α the smaller can be the size of the sample without large alteration in the size of the increase.

The figures in Table 1 give a definite indication that such conditions prevail, so that we can suggest that they represent samples from a population with an index of diversity rather above $1{\cdot}443 \times 2{\cdot}05$, or approximately 3·0.

If there is an average number of p plants per square inch in the quadrat then the actual number of individuals in each quadrat is p times the number of square inches; or, if we

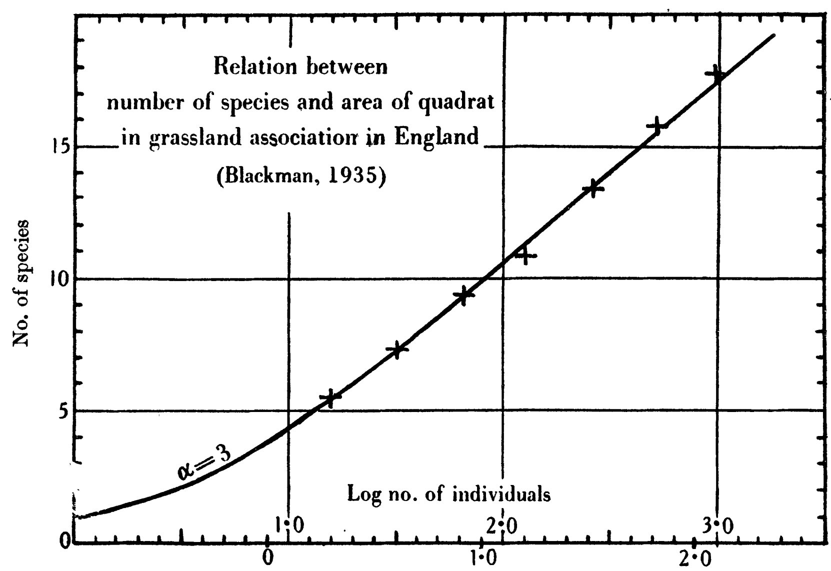


Fig. 1.

use the log of the number of individuals in the quadrat, it must equal $\log p + \log$ number of square inches. Hence for samples from any one population the log number of individuals = the log of the area + a constant.

If we plot theoretically the log number of individuals against the number of species in various sized samples from a population with $\alpha = 3$ we get the line shown in Fig. 1.

If we take the log number of square inches in Blackman's quadrats and plot them against the number of species we get seven points almost in a straight line.

If to the log number of square inches in Blackman's quadrats we add 0·9 (= log of 8), these seven points fall almost exactly on the line for $\alpha = 3$ (Fig. 1).

Thus Blackman's results can be interpreted by the logarithmic series, and the index of diversity suggested by R. A. Fisher, on the assumption that we are dealing with a population with an index of diversity of 3 and with about 8 'individual plants' per square inch. This is about 288 per $6 \times 6n$ quadrat, 1152 per sq. ft. or 50 million per acre; or just over 1 plant per sq. cm.

It is not possible to say if these figures are anywhere near correct, as it is difficult in most plants to decide where an 'individual' begins or ends. It is interesting to note, however, that on p. 757 of Blackman's paper he indicates that the five dominant species on the area includes about 250 tillers per 6×6 quadrat. If we consider therefore a tiller as an individual, the suggestion that there should be 288 units of all plants on the quadrat is not inconsistent with reality.

Since we have already reached the portion of the curve where $\alpha = 3$ is practically a straight line, it is possible to make some interesting suggestions by extrapolation. Thus 1 acre = approximately $2^{22 \cdot 7}$ sq. in. The largest quadrat taken was 2^7 sq. in. and contained 18 species. By doubling this area 15·7 successive times we get 1 acre, and therefore if each doubling adds approximately 2·1 species, the acre should contain $18 + (2 \cdot 1 \times 15 \cdot 7)$ or approximately 52 species. Of these 3 should be represented by only one individual each (since α is the limiting value of n_1); about 10 species should be represented by 10 individuals or less, and about 15 species by 100 individuals or less.

Unfortunately, no record was taken at the time of the total number of species per acre, but Dr W. E. Brenchley informs me that it is not an unlikely number for an acre of fairly uniform natural grassland.

Number of species of plants on areas of different sizes in an aspen association in North Michigan

H. A. Gleason (1922) has discussed the number of species of plants which he found on 240 quadrats of 1 sq. m. each in an aspen association in North Michigan, U.S.A.

He first obtained the number of species on each of the 240 quadrats and then combined them in two different ways to get the number of species in larger areas. First, he found the average number of species in areas consisting of a number of contiguous quadrats; and secondly, he found the number of species in the same number of quadrats scattered as widely as possible over the whole area. His data is shown in Table 2. It will be seen

Table 2. *Number of species of plants found on areas of different sizes in an aspen association in North Michigan. Data from H. A. Gleason*

Area in sq. m.	No. of species: Contiguous quadrats	No. of species: Scattered quadrats	Area in sq. m.	No. of species: Contiguous quadrats	No. of species: Scattered quadrats
1	4·38	4·38	16	12·00	16·67
2	5·82	6·67	20	12·92	17·67
3	6·90	8·75	24	13·50	19·40
4	7·60	10·37	30	15·13	20·00
5	8·21	11·29	40	16·17	20·33
6	8·95	12·65	60	19·75	22·75
8	9·67	13·80	80	20·00	22·67
10	10·33	14·88	120	23·5	25·00
12	11·25	16·25	240	27·0	27·00
15	12·25	16·94			

that while the average number of species per quadrat (4·38) and the total number on the whole 240 quadrats (27·0) are the same for both series, the series showing contiguous areas rises more slowly at first than the scattered areas, but more rapidly later. This, according to Gleason, is due to the fact that the individuals of the different species are not truly randomized over the whole area, but numbers of one species tend to occur together. If the sample quadrats had been grouped by randomizing, the figures would be intermediate between the two results obtained.

It should be noted that while the number of species for 1 sq. m. is the average of 240 estimations, the number of species in the 240 sq. m. is obtained only from a single observation and so is subject to a greater error.

Gleason's data are shown graphically in Fig. 2, the contiguous series by dots and the scattered series by crosses. The number of species is plotted against the log of the number of square metres covered by the sample. Gleason himself has pointed out that there is an approximately straight line relationship and he has extrapolated his results to suggest that there should be 56 species on the 25 sq. km. over which the formation extended.

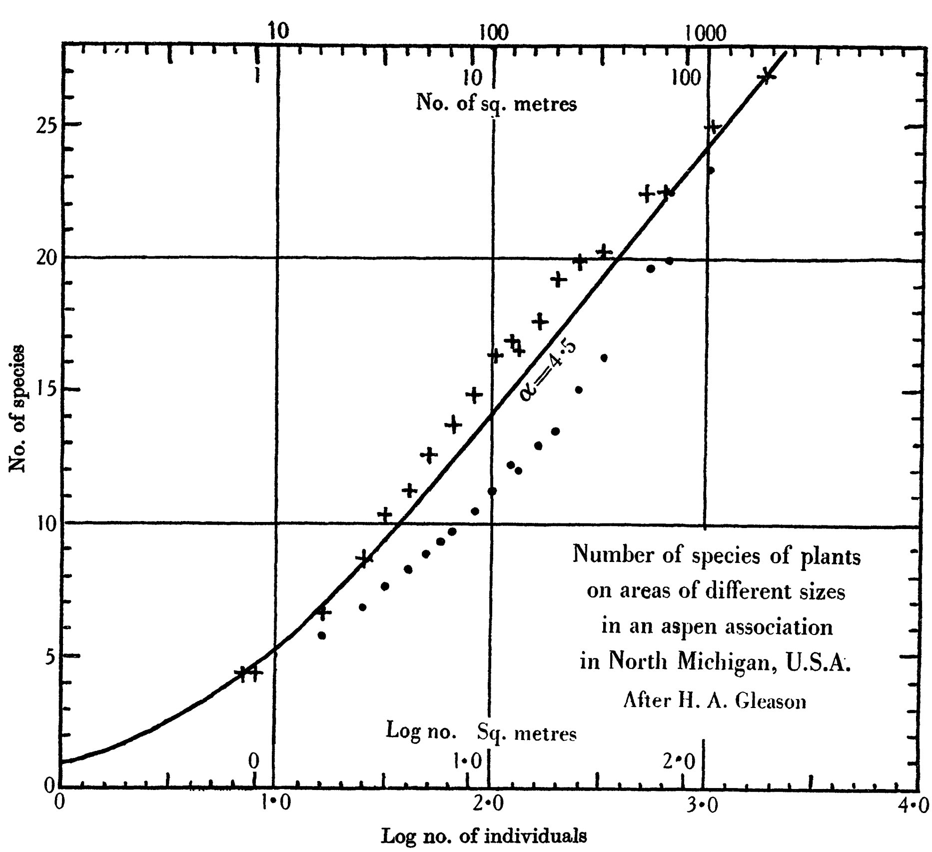


Fig. 2.

He writes: 'Since the counted quadrats were located only in the treeless portions of the association, and since several additional species are confined to the aspen thickets, this theoretical result corresponds closely to the known total of approximately 80 species.'

It is interesting to note that if the relationship were a true straight line it could be extended in the opposite direction, with the curious result that in an area of about 0·4 m. there would be no species! This alone would make one suspect the validity of the straight-line association at very small values.

We notice from the table that an increase in the log number of square metres from 0 to 2·38 (an increase of 2·38) has been associated with an increase in the number of species from 4·38 to 27 (an increase of 23·63). If the linear relation holds an increase in the number of species corresponding to a log size of plot increase of 0·4342 ($=\log_{10} e$) would be 4·31. If the logarithmic series can apply to this data, then this is the value of α, or the index of diversity of the population (see p. 2). Since, however, we are dealing with rather small samples at first, α is probably rather larger than this.

If the curve of relation between numbers of species and log number of individuals for $\alpha = 4{\cdot}5$ is superimposed on the previous sets of points for species and log number of square metres in the plot; and if the ordinates are shifted so that log 1 sq. m. corresponds to log 8 individuals (i.e. $=0{\cdot}9$) we get the fit shown in Fig. 2. The 'scattered quadrats' lie to the left of and above the line of $\alpha = 4{\cdot}5$; the contiguous quadrats lie to the right and below. There is little doubt that truly randomized quadrats would give very close fit.

In other words, the data fit reasonably well to the theory of the index of diversity and the logarithmic series on the assumption that $\alpha =$ approximately 4·5 and that there are about 8 individual plants per sq. m. No information is available on this latter point, but there must be a comparatively small number, as on an average only 4·38 species are

Table 3

Area in sq. m.	No. of repetitions	No. of species: Tract 1	Tract 2	Tract 3
1	256	4·82	4·88	3·06
4	64	10·69	10·67	7·31
16	16	17·38	19·00	14·81
64	4	25·00	28·00	22·75
256	1 only	32·00	33·00	31·0
Index of diversity	...	5·3	5·9	6·1

Table 4

Tract no. ...	4	5	6	7	8	9*	10†
Av. no. of species per sq. m.	4·19	4·62	4·75	4·15	4·15	2·93	3·73
Total species per 100 sq. m.	44	34	44	39	41	19	30
Total area of tract in sq. m.	8492	4350	8895	6142	44,331	6400	2500
Observed total flora	84	59	77	70	91	32	46
Total flora calculated by Gleason	82·4	58·1	82·3	70·2	89·8	33·5	48·3
Index of diversity	8·8	6·4	—	7·5	8·0	3·4	5·6

* Tract 9; in beech, sugar-maple forest of about 40 acres near shore of Lake Michigan in Emmet County.
† Tract 10; part of hardwood forest in Antrim County. Forest originally extended over 100 sq. miles.

found per sq. m. This problem is of course complicated by the association of plants of the same species and the difficulty of deciding where an 'individual' plant begins and ends.

If we extrapolate on this basis we get the number of species of plants on 25 sq. km. as 79, which is much closer to the observed figure than the calculation of Gleason; perhaps it is too high. Of these 79 species about four or five should be represented by only a single individual.

In a later paper Gleason (1925) gives further information about the aspen association in North Michigan, and other areas, to which the logarithmic series can be applied. His data is shown in a condensed form in Tables 3 and 4.

The data for the three aspen association tracts in Table 3 is shown diagrammatically in Fig. 3, fitted to curves for $\alpha = 5$, 5·3 and 6. It will be seen that there is a very good fit for tract 1 to 5·3 and that tracts 2 and 3 are both close to $\alpha = 6$.

In both 1 and 2 the number of species on 1 sq. m. corresponds to 0·9 on the log number of individuals, indicating a population density of about 8 individual plants per sq. m.

Tract 3, however, with an average of only just over 3 species per sq. m., fits best at a point indicating only about 4 plants per sq. m. Without fuller information about the plots it is difficult to say if this has any special meaning or any confirmation in the tract conditions.

Tracts 4–10, dealt with in Table 4, are shown diagrammatically in Fig. 4. In this case the horizontal ordinate is the log number of square metres in the sample. It will be seen that, as Gleason pointed out, an almost straight-line relationship exists between the different-sized plots on the same tract.

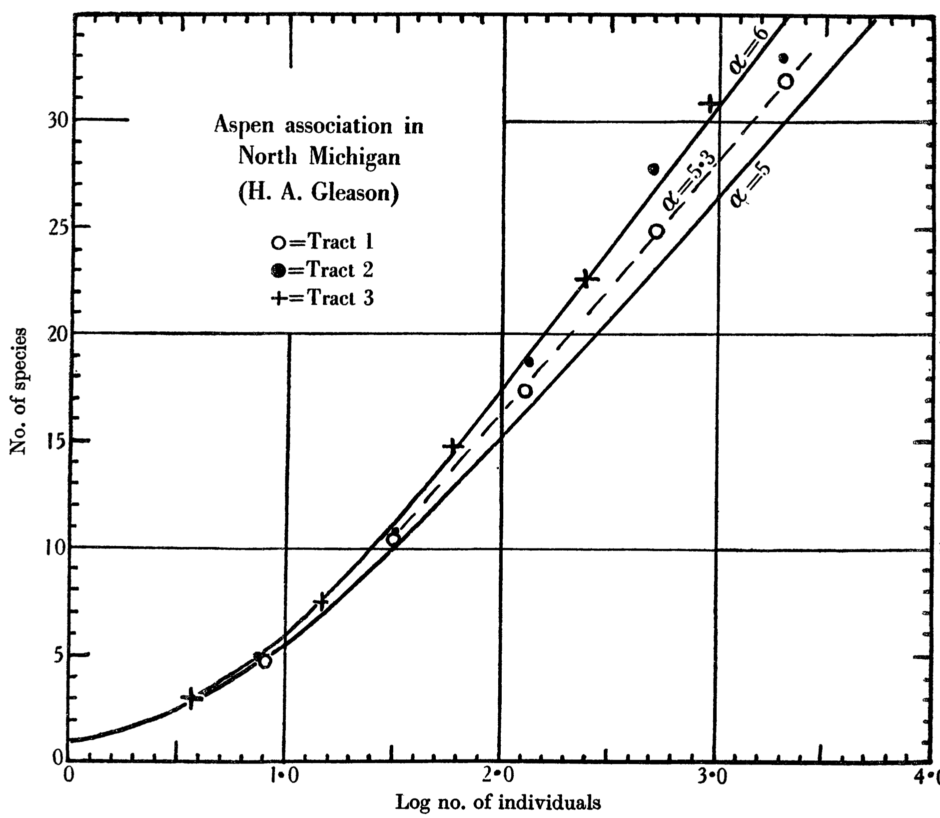


Fig. 3.

On the diagrams are marked by one or two short vertical lines the position on each line corresponding to the theoretical positions of 1000 and 10,000 individuals for the corresponding α values. It will be seen that in all cases except tract 4, the 1000 individuals line corresponds very closely to log 1·9 or 1·95, indicating that 1000 individuals is equivalent to between 80 and 90 sq. m., or 11–12 individuals per sq. m. Plot 4 indicates only about 8–9 plants per sq. m.

Plot 6 does not give a good straight line and has not been calculated.

Tracts 9 and 10, which give the two lowest figures for the index of diversity are, as seen from the notes to Table 4, from different plant associations from the remainder of the data.

Thus Gleason's additional data is consistent with the logarithmic series, with index of diversity ranging from about 3 to about 9, and an average density of population of about 10 plants to the square metre.

Ground vegetation in Tectonia *forest in Java*

O. Arrhenius (1923) gives details of the number of species of ground plants on areas of different sizes in different portions of a *Tectonia* forest in Java. His data is summarized in Table 5.

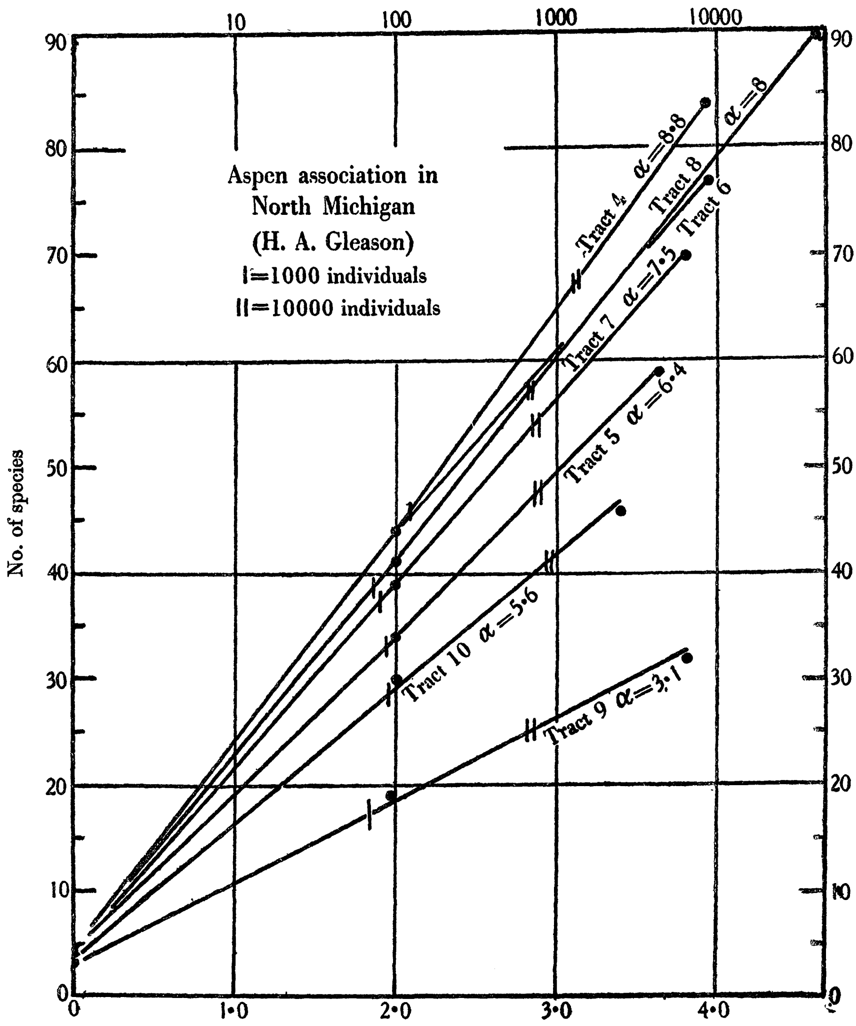


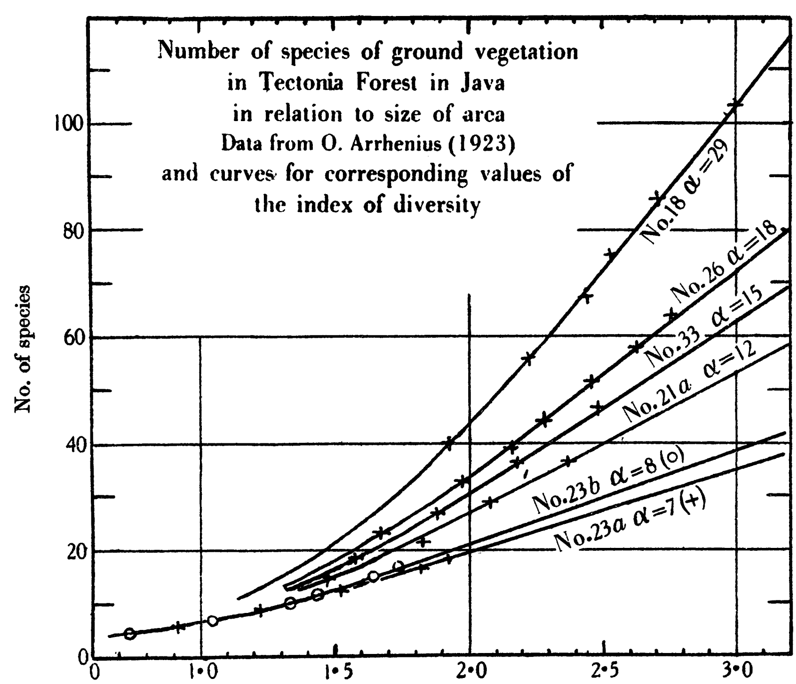
Fig. 4.

When the data of each survey number is plotted graphically the number of species against the log number of square inches in the area, it will be seen from Fig. 5 that each survey number gives an approximately straight line relationship. From this it is possible to estimate closely the index of diversity (α) (Table 5) and to draw the curves for these values on a number of species/log number of individuals basis.

The curves for the relation between number of individuals in the sample and number of species for the corresponding values of α are shown in Fig. 5, and the data for the plants are superimposed on these in the position of best fit.

Table 5. *Ground vegetation in* Tectonia *forest in Java* (after Arrhenius, 1923)

	No. of species observed						
Survey no. ...	18	26	33	21*a*	16*a*	23*a*	23*b*
Area in sq. m.							
100	40	23	18·6	15	15	5·7	4·5
200	56	33	26·4	21	22	8·6	6·7
300	66·8	39	—	—	27	—	—
400	75·0	44	36·9	29	30	12·1	10·0
500	—	—	—	—	—	13·5	11·5
600	86·3	51	—	—	36	—	—
700	—	—	—	37	—	—	—
800	—	—	47·0	—	41·5	16·2	14·6
900	97·0	58	—	—	—	—	—
1000	—	—	—	—	—	18·2	16·7
1200	103·7	64	—	—	49·5	—	—
1400	—	—	—	—	—	—	—
Index of diversity	29	18	15	12	—	7	8



Log no. of individuals

Fig. 5.

It will be seen first of all that the fit is exceptionally close, and secondly, that the position of the first observation in each survey plot, i.e. the number of species on 100 sq. m., to the number of individuals in the corresponding position on the α curve, differs in each plot.

Thus survey number 18 gives 40 species on 100 sq. m., and this on the α curve of 29 corresponds to about 95 individuals. That is to say that theoretically a population with an index of diversity of 29 would give 40 species if a sample of 95 individuals was taken.

The corresponding values for all the areas are as follows:

Survey nos.	No. of individuals corresponding to species on 100 sq. m.
18	95
26	47
33	37
21*a*	30
23*a*	5·5
23*b*	8

Thus the theoretical treatment suggests that there were many more individual plants per square metre in plot 18 than in plot 23. It is curious to note that, in all cases except the two plots 23*a* and 23*b*, the higher value of α is associated with a larger number of individuals per square metre.

Zoological

Distribution of head lice on a number of hosts

Dr P. A. Buxton kindly gave to me for study his records of the number of lice found on the heads of prisoners in the jail at Cannamore, South India (see P. A. Buxton, 1940).

Table 6 shows an analysis of the total number of lice of all stages on the heads of Hindu male prisoners in 1937, 1938 and 1939. The first column is the number of lice per head, and the second the number of heads in which that number of lice occurred. The third column are subtotals of every five groups followed by the total number of lice in the subtotal in brackets. Thus it will be seen that there were 622 heads free from lice; 106 heads

Table 6

No. of lice per head	No. of heads Observed		No. of heads Calculated		No. of lice per head	No. of heads Observed		No. of heads Calculated		No. of lice per head	No. of heads observed	
0	622									51	1	
1	106		107·15		26	4		2·87		53	1	
2	50		52·81		27	6		2·72		54	1	
3	29		34·69		28	2		2·59		57	1	
4	33		25·64		29	4		2·46		58	1	
5	20	238 (525)	22·22	240·51	30	1	17 (468)	2·35	12·98	59	1	
6	14		16·61		31	.		2·24		60	1	7 (392)
7	12		14·03		32	1		2·14		64	1	
8	18		12·10		33	5		2·04		66	1	2 (130)
9	11		10·60		34	.		1·95		71	1	
10	11	66 (521)	9·40	62·74	35	2	8 (267)	1·87	10·24	74	1	
11	3		8·43		36	2		1·79		79	1	
12	10		7·61		37	4		1·72		80	1	4 (304)
13	8		6·93		38	3		1·65		83	1	
14	6		6·34		39	.		1·58		88	1	2 (171)
15	3	30 (386)	5·83	35·13	40	4	13 (494)	1·52	8·26	110	1	
16	6		5·39		41	1		1·46		121	1	
17	7		5·00		42	.		1·41		128	1	
18	4		4·65		43	.		1·36		129	1	
19	7		4·34		44	2		1·31		145	1	
20	7	31 (560)	4·07	23·45	45	3	6 (264)	1·26	6·79	149	1	6 (782)
21	3		3·82		46	1		1·21		188	1	1 (188)
22	4		3·59		47	5		1·17		239	1	
23	4		3·39		48	1		1·13		270	1	2 (509)
24	4		3·20		49	.		1·09		303	1	
25	3	18 (414)	3·03	17·02	50	1	8 (379)	1·05	5·66	385	1	2 (688)

Total number of heads with lice, 461; total no. of lice, 7442.

with one louse each; and 525 lice distributed on the 238 heads which harboured from one to four lice inclusive.

Buxton had already (1938) shown that at various localities, individuals with fewer lice (1–10) were more frequent than those with more lice (10–100 or over 100), but no detailed study of the distribution had been made.

The data for Cannamore, exclusive of the number without lice, is shown graphically by the histogram in Fig. 6, and it will be seen immediately that it gives a curve of the

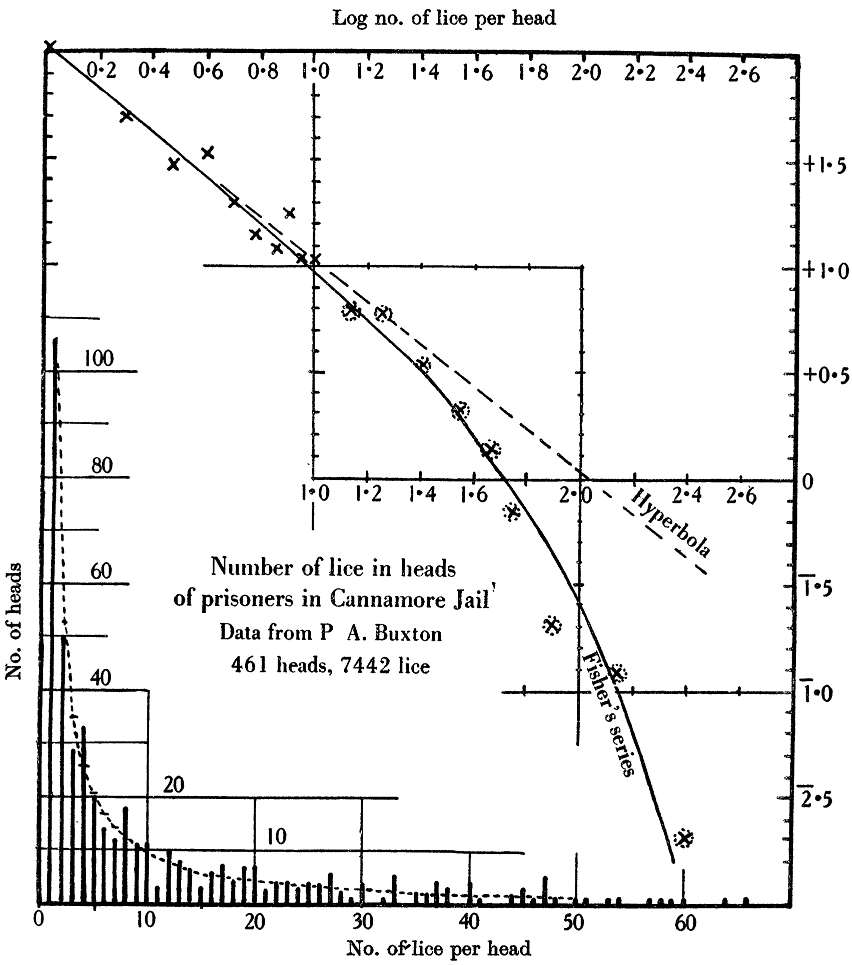


Fig. 6.

'hollow' type closely resembling a hyperbola. The same figure shows (as crosses) the data transformed to logarithmic co-ordinates, and it is then seen that the points do not fall on the straight line required by a hyperbola but agree well with the logarithmic series.

There were altogether 461 infested heads and on these were 7442 lice. From these two figures the corresponding logarithmic series can be calculated. It is found that $x = 0{\cdot}9856$ and $n_1 = 107{\cdot}15$. The very close resemblance of the calculated n_1 to the observed number, 106, is immediately noticed.

The calculated values for the series up to 50 lice per head, together with various sub-totals, is shown in Table 6 and graphically on number and logarithmic basis superimposed on the figures in Fig. 6.

The remarkably close fit cannot be missed, particularly in the log transformation which is particularly accurate for small numbers. Thus, in addition to remarkably close resemblance in the number of heads with one louse already pointed out, the calculated number with 1–5 lice is 240·5 and the observed 238; the calculated number with 1–25 lice is 378·9 and the observed 383. Between 25 and 50 lice per head the fit is not so good, the observed figures being rather larger than the calculated.

There is no doubt that the two series are very close indeed, and it is difficult to believe that this resemblance is entirely accidental.

It may be mentioned that tests were made on the same data with other series. First the negative binomial was tested. This takes into account the number of heads (622) without lice, and with this series the number of heads with one louse was calculated to be 68·4, which is widely different from the observed number.

Secondly, a hyperbolic series was calculated to fit all the values (excluding free heads) up to 100 lice per head. This gave the number of heads with one louse as 86·8, again far from the observed value.

Thus the logarithmic series was closer to the observed than any other yet tested.

British nesting birds

Mr James Fisher has informed me that the most recent estimates of bird populations suggest that in May of any one year there are about 40 million pairs of nesting birds in England and Wales, and that these belong to about 170 species.

If the logarithmic series is applied to these figures we find that $x = 0{\cdot}9999997182$, that $n_1 = 11{\cdot}772$, and that α is the same as n_1 to about the 4th decimal place. From these data the figures in Table 7 have been calculated.

Table 7

Pairs per species	No. of species: At	No. of species: Up to	Individual % of total population: Up to
1	11·3	—	—
2	5·7	—	—
3	3·8	—	—
4	2·8	—	—
5	2·3	25·8	—
10	—	33·0	—
100	—	58·5	—
1000	—	84·2	—
10,000	—	136·3	3·28
1 million	—	159·8	24·6
2 million	—	164·9	43·1

It will be seen that there should be, according to the theory, about 11 species represented by only a single pair, about 26 species represented by less than 5 pairs, and about 58 species represented by less than 100 pairs.

At the other end there should be about 160 species represented by one million pairs or less, which leaves 10 species out of the original 170 which must be represented by more than one million pairs.

Unfortunately, the calculations for the rarer birds are difficult to check against observed figures, but Mr Fisher informs me that they are not unlikely results.

In the case of the common birds Fisher (1940, p. 102) has given a list of the estimated numbers of the common birds of England and Wales as shown in Table 8. It will be seen that he mentions 9 species of birds with about a million or more pairs as compared with a calculated figure of 10.

Table 8

	Nos. in millions	
	Nos.	Pairs
(1) Chaffinch	10	5
(2) Blackbird	10	5
(3) Starling	7	3·5
(4) Robin	7	3·5
(5) House sparrow	3	1·5
(6) Hedge sparrow	3	1·5
(7) Thrush	3	1·5
(8) Meadow pipit	3	1·5
(9) Rook	1·75	0·9

It is interesting to note, from the third column of Table 7, that, according to the calculations, all the individuals of species up to 100,000 pairs per species only constitute 3·28% of the total bird population; and that three-quarters of the total bird population is made up of the 10 species with more than a million pairs each. J. Fisher's figures in Table 8 suggest that the 9 commonest species comprise about 24 million pairs out of a total of 40 million or 60% of the population.

Mosquitoes caught in a light trap in ten cities in Iowa, U.S.A.

J. A. Rowe (1942) gives the number of mosquitoes caught by light traps in ten cities in the State of Iowa. The data is summarized in the first two columns of Table 9, and graphically in Fig. 7, where the log of the number of individuals in the sample is plotted against the actual number of species.

Table 9

	No. of individuals	No. of species	Index of diversity
Ruthven	20,239	18	1·95
Des Moines	17,077	20	2·24
Davenport	15,280	18	2·02
Ames	12,504	16	1·81
Muscatine	6,426	16	1·98
Dubuque	6,128	14	1·71
Lansing	5,564	16	2·03
Bluffs	1,756	13	1·90
Sioux City	661	12	2·08
Burlington	595	12	2·13
Total	86,230	28	2·70

In the third column of Table 9 is shown the index of diversity (α) of each sample calculated from the logarithmic series. These are subject to a standard error of from 7·5 to 15% (i.e. approximately 0·15 to 0·3). The smaller error being in the larger samples. It will be seen that they are grouped closely round an average value of 2·0 and none differ significantly from this value.

To Fig. 7 has been added the curve showing the theoretical relationship of species and individuals for $\alpha=1$, 2 and 3, and the closeness of the fit of the different samples for $\alpha=2$ is apparent.

From this we can infer that the extent to which the individuals are diversified into species is more or less the same in all the cities. The small number of species caught in some cities and the larger number caught in others is almost exactly accounted for by the varying number of individuals captured.

When, however, all the ten catches are added together we get 86,230 individuals of 28 species showing an $\alpha = 2{\cdot}7$ which is significantly different from the average of 2·0 for the separate cities. From this we can infer that although the single cities are similar in their diversification, they have not identical populations. The area considered as a whole has a greater variety of species (and hence probably of environment) than any one of its parts.

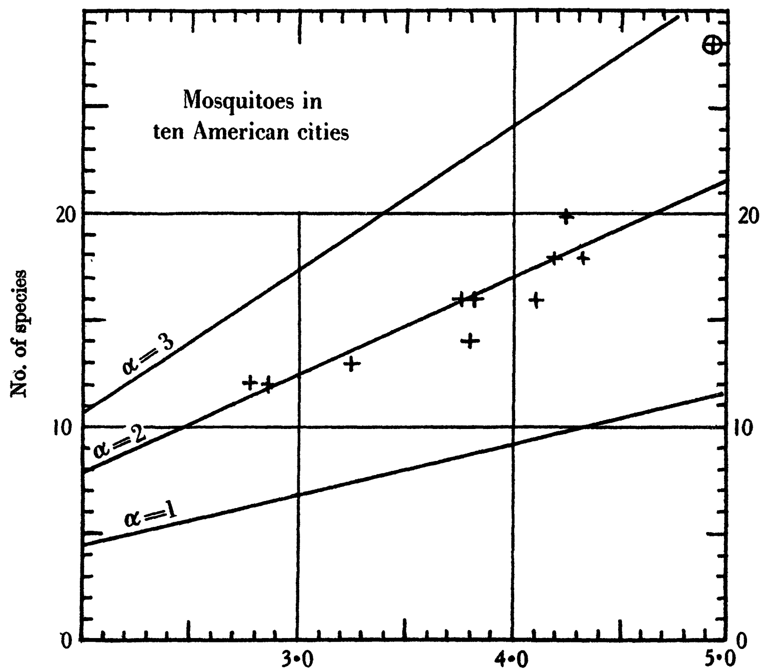


Fig. 7.

Butterflies captured in Malaya by Dr A. S. Corbet

The data obtained by Dr Corbet from a large collection of butterflies made in Malaya and already briefly discussed by him (Fisher *et al.* 1943, p. 42) provides an interesting problem in fitting incomplete data to the logarithmic series.

The collection was made by hand over a period of several years, and the important modification was that after Dr Corbet had caught about 25 specimens of a species he did not make regular attempts to get more. Thus up to about 25 specimens the data are a more or less random sample, but above this figure the number of individuals per species is consistently too low.

The data consisted therefore of three sets of information:

(1) The total number of species with less than 25 individuals per species was 501.

(2) The total number of species at all levels was 620.

(3) The total number of individuals at frequencies below 25 individuals per species was 3306.

The problem is to see if it is possible to obtain from this any values of x and n_1, that will give us the whole series.

From (1) we have $$n_1\left(1+\frac{x}{2}+\frac{x^2}{3}+\frac{x^3}{4}+\dots+\frac{x^{23}}{24}\right)=501.$$

From (2) we have $$n_1\frac{-\log \overline{1-x}}{x}=620.$$

From (3) we have $$n\,(1+x+x^2+x^3+\dots+x^{23})=3306.$$

From each of these equations it is possible to get a relation between n_1 and x. A number of these were calculated and lines through them are shown diagrammatically in Fig. 8.

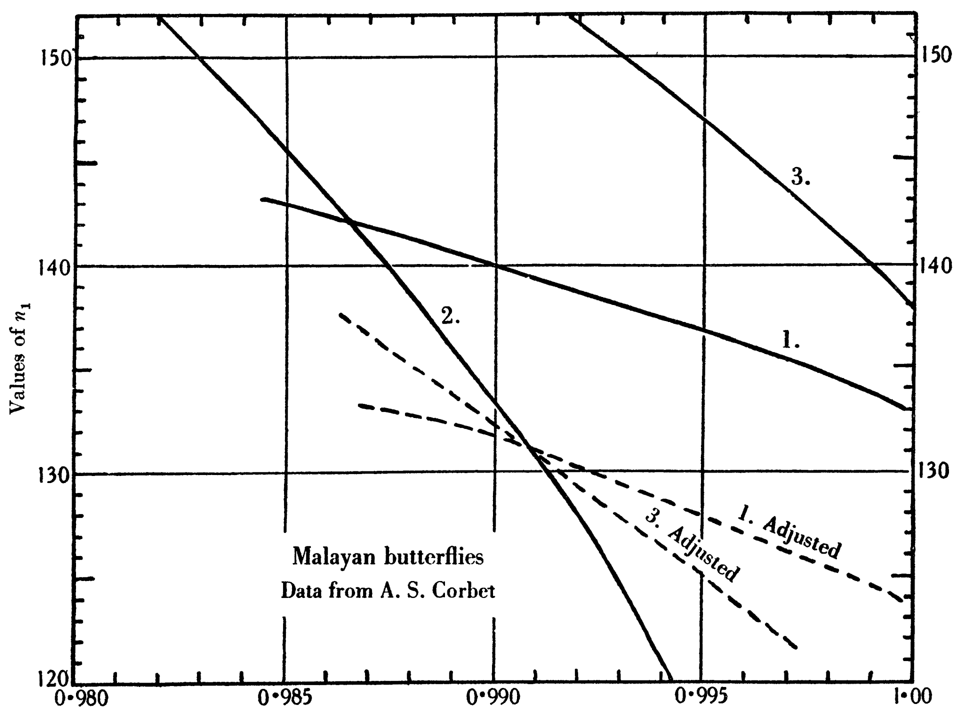


Fig. 8.

It will be seen that it is possible to find values for x and n_1, that will fit any two of the equations, but the three line relationships do not meet at a point and hence there is no values of x and n_1 that will fit simultaneously all three. Either the logarithmic series does not fit or there is some error in the data.

In view of the care taken in the identification the total number of species in the collection, 620, is probably very accurately determined. This means that the line of equation (2) is probably correct.

If the assiduity with which individuals were collected gradually fell off below the 24 individuals per species, instead of at or above this level, it would result in too many species (and hence too many individuals) being included in the 'below 25' category. If we agree that this might have taken place we can make a small reduction in the observed figures in equations (1) and (3).

If we reduce the observed species 'below 25' in equation (1) by 29 (just over 5%), and the observed individuals in equation (3) by 565 (which is just over 1 individual per species) we find that the lines for equations (1) and (3) in Fig. 8 are shifted to new positions as shown by the dotted lines, and that now all three lines meet approximately at the point represented by $n_1 = 131$ and $x = 0{\cdot}991$.

Otherwise the data can be made to fit the theory if we assume that there was a slight overestimation of the number of species and of individuals in the group below 25 individuals per species. The number of species below 25 individuals per species should have been about 472 and the number of individuals about 2841.

This is of course not proof of the fit of the theory to the observed facts, but it is interesting that the difference between the theory and the observed facts is in the direction of, and can be explained by, errors which might easily have occurred, and indeed Dr Corbet tells me that the assumption that assiduity in collecting fell off gradually before the 25 mark was reached is extremely probable.

3. Problems connected with the grouping of species into genera

Up to the present we have discussed the logarithmic series in relation chiefly to individuals grouped into species.

It has long been known that in many groups of plants and animals, the frequency of genera, with different numbers of species forms a 'hollow' curve of the hyperbolic type. The matter has been particularly discussed by Willis in his *Age and Area* (1922) and by Chamberlin (1924). Both writers assumed that the correct solution of the frequency distribution was the hyperbolic series and used this as a basis on which to formulate theories of evolution. Chamberlin unfortunately used also an incorrect method for obtaining the hyperbola to fit his data.

I have examined a number of sets of data of the frequency of genera with different numbers of species, in both animals and plants, and these are discussed below in relation to the logarithmic series.

Zoological

Orthoptera of the world

W. F. Kirby (1904–10) published a catalogue of the Orthoptera of the world based largely on his own studies with the collections at the Natural History Museum in London. When the number of genera with different numbers of species are tabulated the results are as shown in Table 10.

In the Acridiidae (short-horned grasshoppers) he describes 4112 species in 826 genera. This gives, on the basis of the logarithmic series, a calculated n_1 of 289·3, whereas the observed number is 320, about 10% above the calculated. For all genera containing four species or less the calculated number is 613 and the observed 639, which is within 4% of the calculated. The very close fit of the observed and calculated at higher values is best seen from the log scale in Fig. 9. There is no doubt whatever that the fit of the observed point to the calculated curve is very remarkable, and there is no evidence of a good fit to the hyperbola which is shown, transformed to log co-ordinates, on the same diagram.

In the Mantidae (Table 10 and Fig. 10) there were 805 species in 209 genera. This gives the calculated number of genera with one species as 82·27, as compared with the observed figure of 82! The figures for genera with up to 5 species per genus are 167 calculated and 169 observed, and the remarkably close fit at higher values is shown in Fig. 10.

In the Blattidae (Table 10) with 1612 species in 197 genera there is again a very close agreement between 54 observed and 57 calculated for the number of genera with one species. The rest of the fit is almost equally accurate; 127 calculated to 127 observed genera with 1–5 species; 29 calculated to 28 observed with 6–10 species; 14 calculated to 13 observed with 10–15 species, and so on.

In the Forficulidae (Table 10 and Fig. 11) there are much smaller numbers, only 480 species in 52 genera, but again the fit of calculated and observed is close. For genera

Table 10. *Genera and species of Orthoptera of world* (from Kirkby, 1904–10)

No. of species per genus	Acridiidae		Mantidae		Blattidae		Forficulidae	
	Observed	Calculated	Observed	Calculated	Observed	Calculated	Observed	Calculated
1	320	289·3	82	82·27	54	56·80	12	14·3
2	131	134·5	28	36·93	30	27·39	9	6·94
3	86	88·35	27	22·11	16	17·62	5	4·49
4	61	58·11	12	14·88	14	12·75	2	3·26
5	41 (639)	43·21 (613·47)	20 (169)	10·69 (166·88)	13 (127)	9·88 (124·44)	3 (31)	2·53 (31·52)
6	27	33·47	7	8·00	5	7·91	3	2·05
7	21	26·68	4	6·16	6	6·54	—	1·66
8	18	21·70	8	4·84	7	5·52	1	1·44
9	23	17·94	4	3·86	8	4·74	—	1·25
10	17 (106)	15·01 (114·80)	1 (24)	3·12 (25·98)	2 (28)	4·11 (28·82)	— (4)	1·09 (7·49)
11	12	12·68	4	2·55	3	3·61	3	0·96
12	8	10·80	—	2·09	—	3·19	2	0·85
13	9	9·27	2	1·74	4	2·84	1	0·76
14	3	8·00	—	1·45	6	2·54	—	0·69
15	5 (37)	6·94 (47·69)	2 (8)	1·21 (9·04)	— (13)	2·29 (14·47)	1 (7)	0·62 (3·88)
16	4	6·05	1	1·02	2	2·07	—	0·56
17	3	5·30	—	0·86	4	1·88	—	0·52
18	6	4·65	—	0·73	1	1·71	2	0·47
19	2	4·10	1	0·62	2	1·57	—	0·44
20	3 (18)	3·62 (23·72)	1 (3)	0·53 (3·76)	— (9)	1·44 (8·67)	— (2)	0·40 (2·39)
21	1	—	—	—	4	1·32	1	—
22	1	—	—	—	—	1·21	—	—
23	2	—	—	—	—	1·12	—	—
24	1	—	—	—	1	1·04	—	—
25	— (5)	—	—	0·25	1 (6)	0·96 (5·65)	— (1)	—
26	2	—	—	—	4	0·89	—	—
27	—	—	—	—	1	0·83	—	—
28	—	—	—	—	2	0·77	—	—
29	4	—	—	—	—	0·72	1	—
30	— (6)	—	— (0)	—	— (7)	0·67 (3·88)	1 (2)	—
	Also 2 at 31 and 34 and single genera at 35, 36, 38, 41, 43, 51, 54, 58, 72, 75, 103 and 202		Also 2 at 30		Also at 31, 40, 44, 48, 56, 85 and 160		Also at 32, 35 (2), 36 and 47	
	Total: 4112 species in 826 genera		Total: 805 species in 209 genera		Total: 1612 species in 197 genera		Total: 480 species in 52 genera	
	$\alpha = 311{\cdot}2$, $x = 0{\cdot}92964$		$\alpha = 91{\cdot}62$, $x = 0{\cdot}89781$		$\alpha = 58{\cdot}89$, $x = 0{\cdot}96476$		$\alpha = 14{\cdot}83$, $x = 0{\cdot}97003$	

with one species the figures are calculated 14·3 and observed 12, while for genera with up to 5, the calculated is 31·42 and the observed 31! The rest of the results are shown diagrammatically in Fig. 11.

Coccidae of the world

Two catalogues of the Coccidae of the world have been examined. Fernald (1903) recognized 1439 species in 166 genera. MacGillivray (1921) increased the number of known species to 1762, but more than doubled the genera to 352. The frequency of genera of different size according to each author is shown in Table 11, and diagrammatically on the log scale in Fig. 12.

It will be seen that in both cases the observed number of genera with one species is well above the calculated: in the case of MacGillivray's classification it is nearly 50% above. Genera with 2, 3, 4 and 5 species are, however, in MacGillivray's case all below

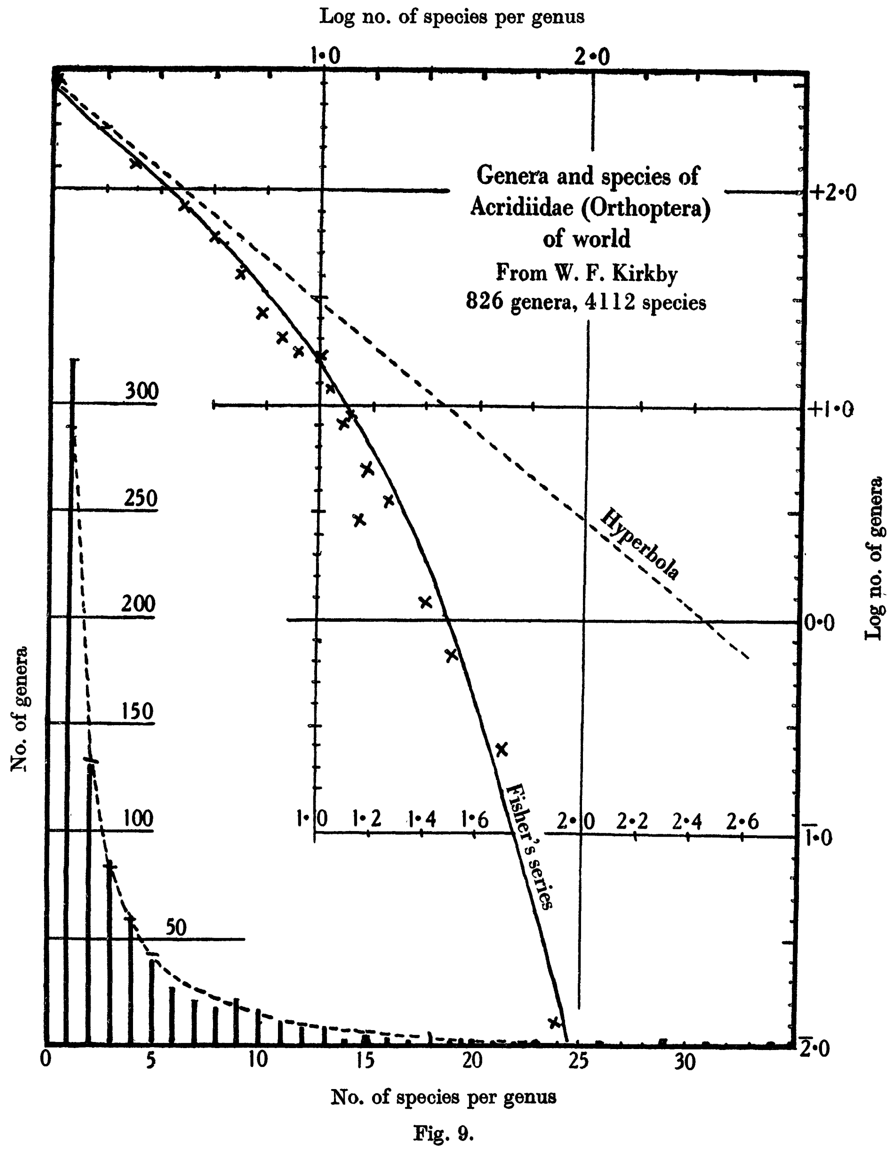


Fig. 9.

the calculated, so that the difference on all genera with up to 5 species is only 282 observed to 259 calculated, an excess of about 9%. It will be seen from Fig. 12 that although the points are variable and more off the theoretical curve than in the case of Orthoptera, yet the logarithmic series does generally represent the form of the curve, and is much closer to the truth than is the hyperbola.

British insects

In 1904 Beare and Donisthorpe published a catalogue of the British Coleoptera. They recognized 3268 species in 804 genera (Table 12, 1st column). Calculations from these two figures give the series shown in the same table and diagrammatically in Fig. 13. It will be seen that the observed number of genera with one species is about 20% above the calculated, but that all the other genera with from 2 to 10 species are below so that the

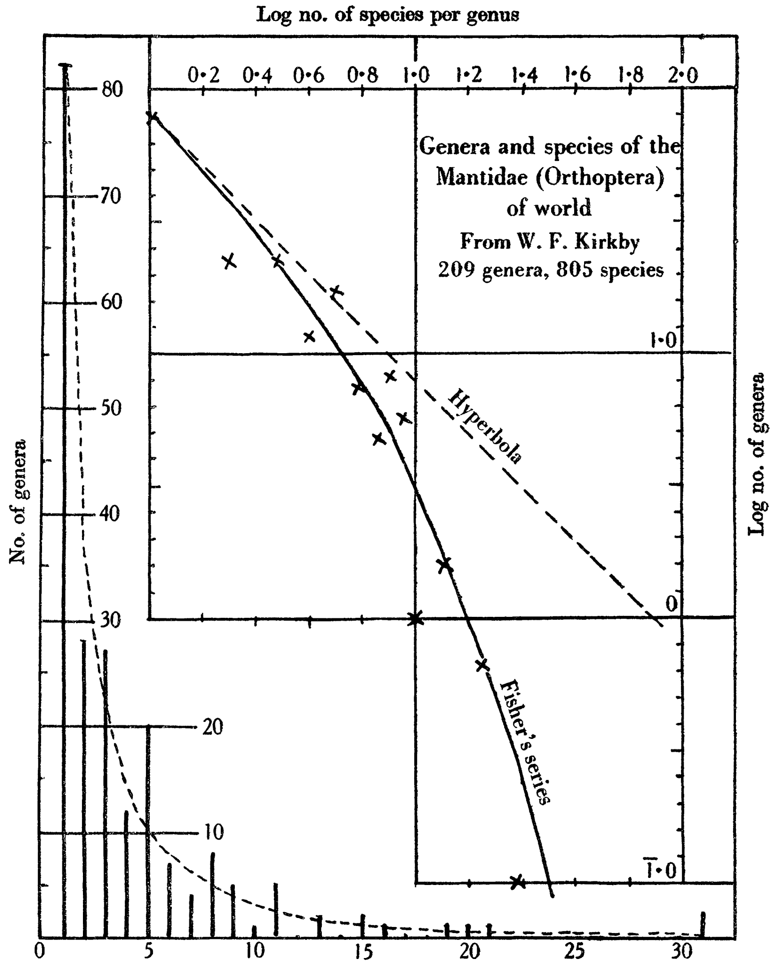


Fig. 10.

calculated number of genera with up to 10 species, 733, is almost identical with the calculated 734·4. The general close resemblance of the observed and calculated figures is well shown on the logarithmic curve in Fig. 13.

In 1857 Stainton published a list of the British Lepidoptera. He recognized 1902 species in 499 genera. This gives a calculated n_1 of 197·4 as against an observed 243 (Table 12, 2nd column); but again, as in the Coleoptera, every other value from 2 to 10 species per genus the observed is below the calculated, so that the total observed (468) and calculated (461) figures to 10 species per genus are very close indeed. The results are shown graphically in Fig. 14.

In the last two columns of Table 12 and in Fig. 15 is shown a comparison of the figures for Macrolepidoptera (excluding butterflies), according to Stainton's 1857 catalogue and the most recent classification by Edelsten in the 1939 edition of South's *Moths of the British Isles*. Stainton recognized 703 species in 249 genera; Edelsten has 713 species in 357 genera. It will be noted that an increase of only 10 species has been accompanied by an increase of over 100 genera. It will be seen that the calculated number of genera with one species is in both cases below the observed, but again the following genera in the

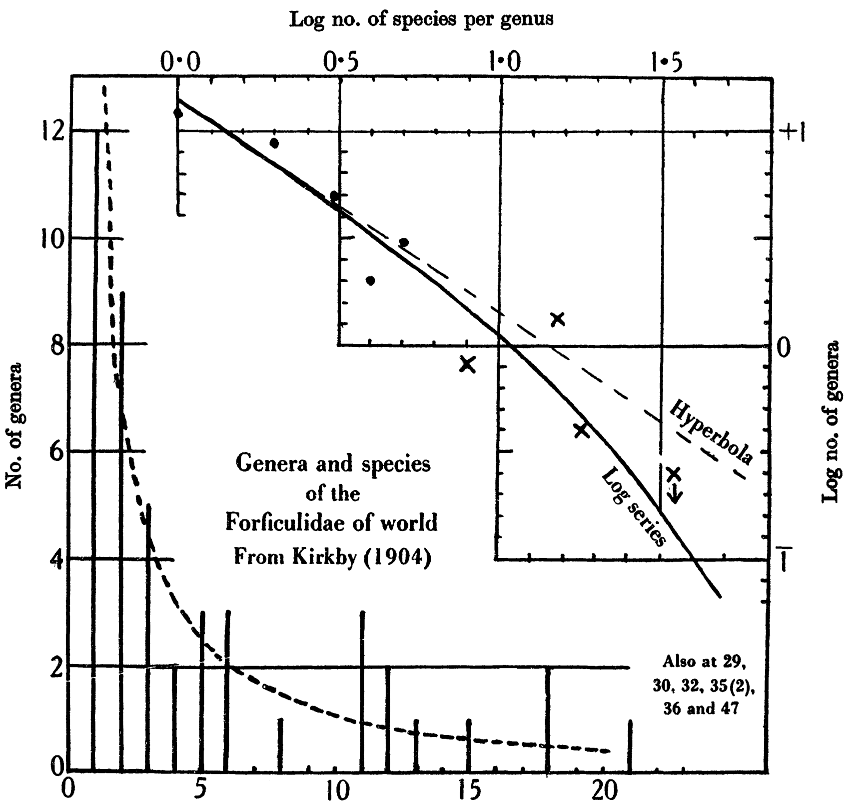


Fig. 11.

series are rather below so that either up to 5 or up to 10 species per genus there is a very close agreement. The results are shown graphically, in a logarithmic scale only, in Fig. 15, when it will be seen that Stainton's classification is more regularly consistent with the calculated series than is the more recent classification of Edelsten. The possible meaning of this will be discussed later.

In 1896 Edwards published a monograph containing families of the British Homoptera in the Cicadina (excluding Psyllidae); he recognized 271 species in 48 genera. Quite recently Dr C. G. Butler has given me a MSS. list showing the most recent classification of the same group, based, I understand, on discussions between himself and Mr China.

The list includes 360 species in 73 genera (Table 13 and Fig. 16). Both series agree fairly well with the calculated, though in each case the observed n_1 is greater than the calculated; the difference, however, is smaller than in the previous groups. If all genera up to 5 species per genus are considered the agreement is very close; 34 calculated to 32 observed for Edwards, and 54 calculated to 53 observed for Butler and China. Again the logarithmic graphs in Fig. 16 show how much closer the observed results are to the logarithmic series than to the hyperbola.

Table 11. *Frequency of genera with different number of species in the scale insects (Coccidae) of the world*

No. of species per genus	MacGillivray		Fernald	
	Observed	Calculated	Observed	Calculated
1	181	122·98	65	46·90
2	42	57·20	18	22·69
3	22	35·47	16	14·63
4	21	24·73	8	10·61
5	16 (282)	18·41 (258·79)	5 (112)	8·22 (103·05)
6	6	14·28	5	6·62
7	5	11·39	7	5·49
8	7	9·27	3	4·65
9	5	7·66	3	4·00
10	4 (27)	6·42 (49·02)	2 (20)	3·48 (24·24)
11	6	5·43	1	3·06
12	5	4·63	5	2·71
13	4	3·97	—	2·42
14	2	3·43	1	2·18
15	6 (23)	2·98 (20·44)	2 (9)	1·97 (12·34)
16	—	2·60	1	1·78
17	1	2·28	2	1·62
18	1	2·00	2	1·48
19	—	1·76	1	1·36
20	2 (4)	1·56 (10·20)	1 (7)	1·25 (7·49)
21	1	1·38	—	1·15
22	—	1·23	1	1·06
23	—	1·09	—	0·98
24	2	0·97	—	0·91
25	1 (4)	0·87 (5·54)	1 (2)	0·85 (4·95)
26	—	—	—	—
27	2	—	1	—
28	2	—	2	—
29	—	—	1	—
30	4 (8)	—	1 (5)	—
	Also at 31, 60, 70 and 200		Also at 33, 34, 37, 57, 59, 62 (3), 70, 79 and 99	
	Total: 1762 species in 352 genera		Total: 1439 species in 166 genera	
	$\alpha = 132{\cdot}2$, $x = 0{\cdot}93024$		$\alpha = 48{\cdot}47$, $x = 0{\cdot}96741$	

British birds

In 1941 Witherby, Jourdain, Ticehurst and Tucker published, in the last volume of their handbook, a check list of the British Birds. They recognize 198 genera containing 424 species, or 420 species and subspecies. Table 14 and Fig. 17 shows the calculated and observed results for both these sets of figures. In each case there is no doubt as to the general close fit of the logarithmic series to the observed data, but again the observed n_1 is above the calculated. Again also if we take all genera up to 5 species per genus the fit is very close indeed; 192 observed to 186 calculated for species only, and 177 observed to 175 calculated if both species and subspecies are included.

Thus we can say that, in the Zoological Kingdom, the logarithmic series closely represents the frequency of genera with different numbers of species, except there is a tendency for the calculated number of genera with one species to be below the observed.

Botanical

Flowering plants of the world

J. C. Willis in his *Age and Area* (1922) has already drawn attention to the fact that the frequency of the number of genera of plants with different numbers of species forms a 'hollow curve'.

The first column of Table 15 and Fig. 18 show data obtained by taking the first one thousand genera of flowering plants and ferns from the same author's *Dictionary of*

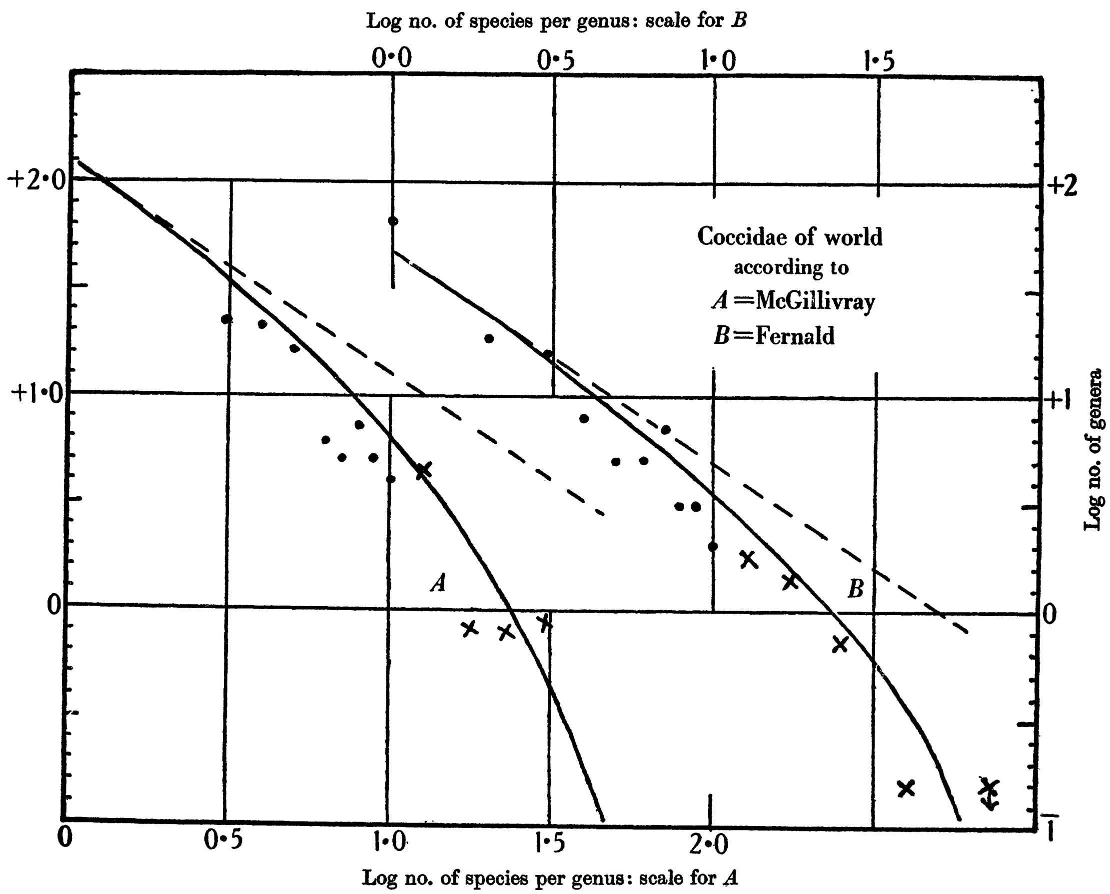


Fig. 12.

Flowering Plants and Ferns (1925 ed.). The number of species in each genus is given accurately when small, but the larger numbers are unfortunately only approximate, as will be seen from the fact that there are 54 genera mentioned with 10 species, but only 4 with 9, and 3 with 11; again at 20 species per genus there are 37 genera listed, but none at either 19 or 21. The figures will, however, probably serve their purpose, as the higher numbers can be averaged.

In all there were 15,473 species in the 1000 genera. This gives by calculation $\alpha = 238{\cdot}9$, $x = 0{\cdot}98475$ and n_1 (the number of genera with one species) as 235·34. The observed figure for the latter was 309, so that once again the observed n_1 is higher than the calculated.

The observed n_2 is also a little larger than the calculated, but after that the fit becomes good and the number of genera with 6–10 species is 136 observed and 137 calculated.

Fig. 18 shows that there are also rather too many genera with large numbers of species than the calculations predict. In fact, for the first time one might suggest that a straight line (shown dotted) might fit the observed results as well as the log series. The hyperbola is however definitely ruled out as in all other cases.

Table 12. *Frequency of genera with different number of species in British insects*

No. of species per genus	British Coleoptera (Beare & Donisthorpe, 1904)		British Lepidoptera (Stainton, 1857)		British Macrolepidoptera (excl. Butterflies): From Stainton (1857)		British Macrolepidoptera (excl. Butterflies): From South & Edelsten (1939)	
	Observed	Calculated	Observed	Calculated	Observed	Calculated	Observed	Calculated
1	391	308·5	243	197·4	132	112·13	239	203·4
2	126	139·7	75	88·47	40	47·12	51	72·68
3	75	84·35	49	52·85	25	26·40	25	34·62
4	48	57·28	35	35·52	12	16·64	12	18·56
5	29 (669)	41·50 (631·33)	21 (423)	25·47 (399·71)	13 (222)	11·19 (213·48)	8 (335)	10·61 (339·87)
6	20	31·32	17	19·02	8	7·84	3	6·32
7	18	24·31	5	14·31	3	5·64	4	3·87
8	10	19·26	9	11·46	3	4·15	3	2·42
9	10	15·51	8	9·13	3	3·10	2	1·54
10	6 (64)	12·65 (103·06)	6 (45)	7·36 (61·28)	— (17)	2·35 (23·08)	3 (15)	0·99 (15·14)
11	14	10·40	4	6·00	3	1·79	1	0·64
12	7	8·63	4	4·93	—	1·38	—	0·42
13	6	7·22	2	4·08	1	1·07	1	0·28
14	4	6·07	1	3·39	—	0·84	—	0·18
15	4 (35)	5·13 (37·45)	4 (15)	2·84 (21·24)	2 (6)	0·66 (5·74)	1 (3)	0·12 (1·64)
16	5	4·36	—	2·38	—	0·52	—	0·08
17	4	3·71	1	2·01	—	0·41	2	0·06
18	—	3·18	1	1·70	1	0·32	1	0·04
19	4	2·73	1	1·45	—	0·26	—	0·03
20	— (13)	2·34 (16·32)	— (3)	1·23 (8·77)	— (1)	0·21 (1·72)	— (3)	0·02 (0·23)
21	3	—	—	—	—	—	—	—
22	1	—	1	—	—	—	—	—
23	—	—	1	—	1	—	—	—
24	2	—	—	—	—	—	—	—
25	1 (7)	1·14	— (2)	—	— (1)	—	— (0)	—
26	1	—	1	—	1	—	—	—
27	—	—	1	—	—	—	—	—
28	1	—	1	—	—	—	—	—
29	—	—	—	—	—	—	—	—
30	1 (3)	0·58	— (3)	—	— (1)	0·024	— (0)	—
	Also at 31, 33, 35, 37 (2), 39, 40, 47, 53, 64, 77 and 155		Also at 32, 39, 40, 41, 43, 45, 53 and 101		And one at 40		And one at 43	
	Total: 3268 species in 804 genera		Total: 1902 species in 499 genera		Total: 703 species in 249 genera		Total: 713 species in 357 genera	
	$\alpha = 340·7$, $x = 0·90559$		$\alpha = 220·3$, $x = 0·89620$		$\alpha = 133·4$, $x = 0·840506$		$\alpha = 284·59$, $x = 0·7147$	

Two examples of the relation of genera and species in the British flowering plants are shown in the last two columns of Table 15 and in Fig. 19. They are from the *Floras* of Bentham & Hooker (1906 ed.), who recognized 1251 species in 479 genera, and of Babington (1922 ed.), who recognized 1671 species in 511 genera.

It will be seen, particularly from the diagrams, that the fit to the logarithmic series is in general good, except that once more the observed numbers of genera with one species is higher than the calculated. The hyperbolic series is ruled out, and there is no suggestion that any other straight line would pass through the points on the logarithmic transformation.

Grouping of genera into families

If we consider the frequency of the different number of genera in higher groups such as families we once more get data which fit in closely with the logarithmic series.

Table 16 and Figs. 20 and 21 show series of the number of families with different number of genera in the flowering plants of the world, in the British flora and the flora of Colorado.

The data for the plants of the world from Rendle (1925) is shown in the first column of

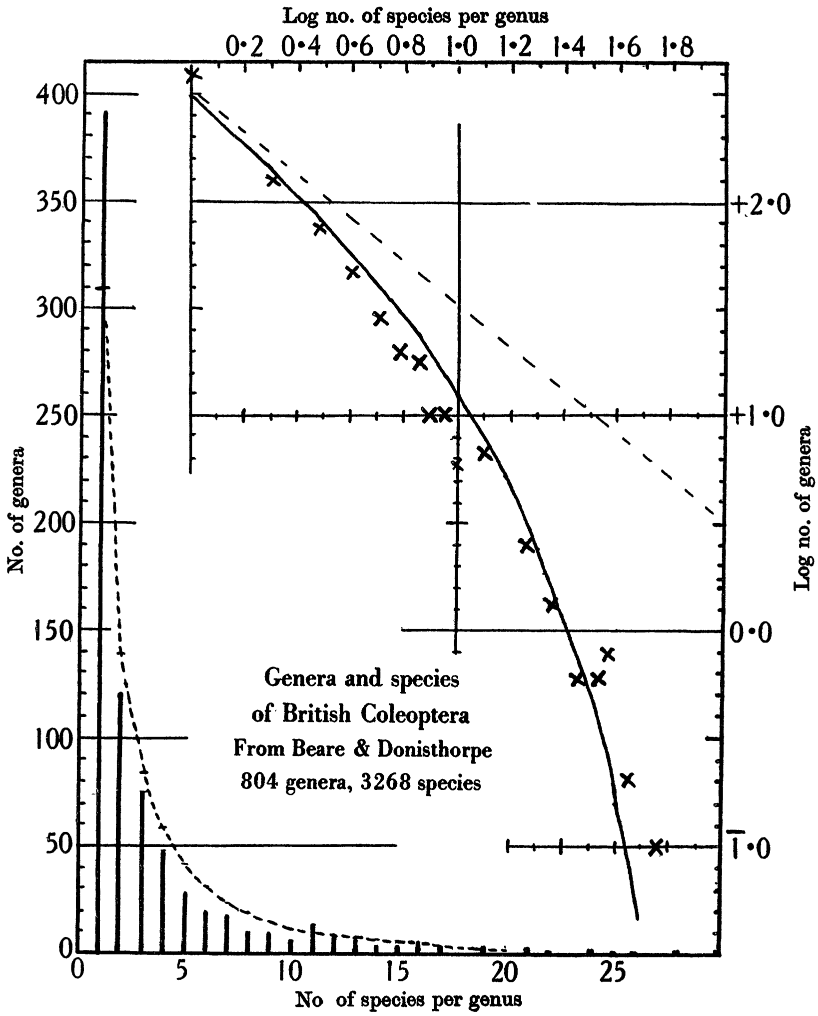


Fig. 13.

Table 16 and in Fig. 20. It is of interest, as it is the first case we have dealt with here where the observed n_1 is distinctly below the calculated. In fact, all except one of the first five figures are below the calculated. To balance this there is, as will be best seen from the diagram, a slight excess of observed genera with 10–100 species. On the whole, however, the log series is not an unreasonable fit. The two sets of data for the British flora are shown in the second and third columns of Table 16 and in Fig. 21. In each case the fit is good in view of the irregularity of the points, and it is most certainly unlikely that any other simple curve would give a better fit. In each case, unlike Rendle's data for the world, there is a slight excess of observed families with a single genera over the calculated.

The data for Colorado (Fig. 20 and the last column in Table 16) are again a very close fit to the logarithmic series; n_1 is observed to be 46 and calculated 43; n_2 is rather above the calculated, but this is balanced by n_2, n_4 and n_5 all being below. The higher observed values, as will be best seen from the diagram, are a very close fit to calculated figures.

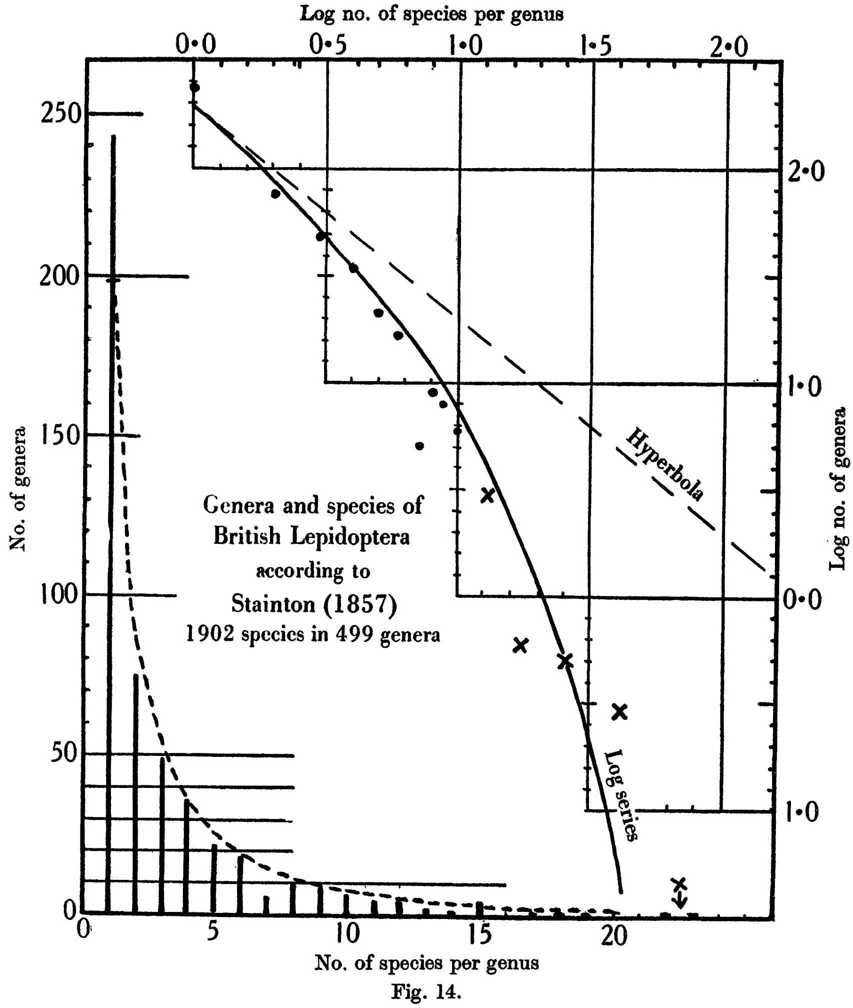


Fig. 14.

4. Index of diversity in different families

Since the index of diversity, α, is a measure of the extent to which the units under discussion are collected into larger groups, it might be interesting to see to what extent the process of grouping species into genera, or splitting genera into species, has gone on in different families of plants and animals.

It must be recalled that a higher value of α means that the same number of units are divided into more groups, i.e. in the present case it means that the same number of species will be classified into more genera, with of course fewer species per genus, unless the 'sample' is larger. A low value of α means fewer genera with a higher average number of species per genus.

The same α value for species with different numbers of units means that the increase in the number of genera can be accounted for by the sampling processes involved in selecting a larger number of units. For the same index of diversity a large number of

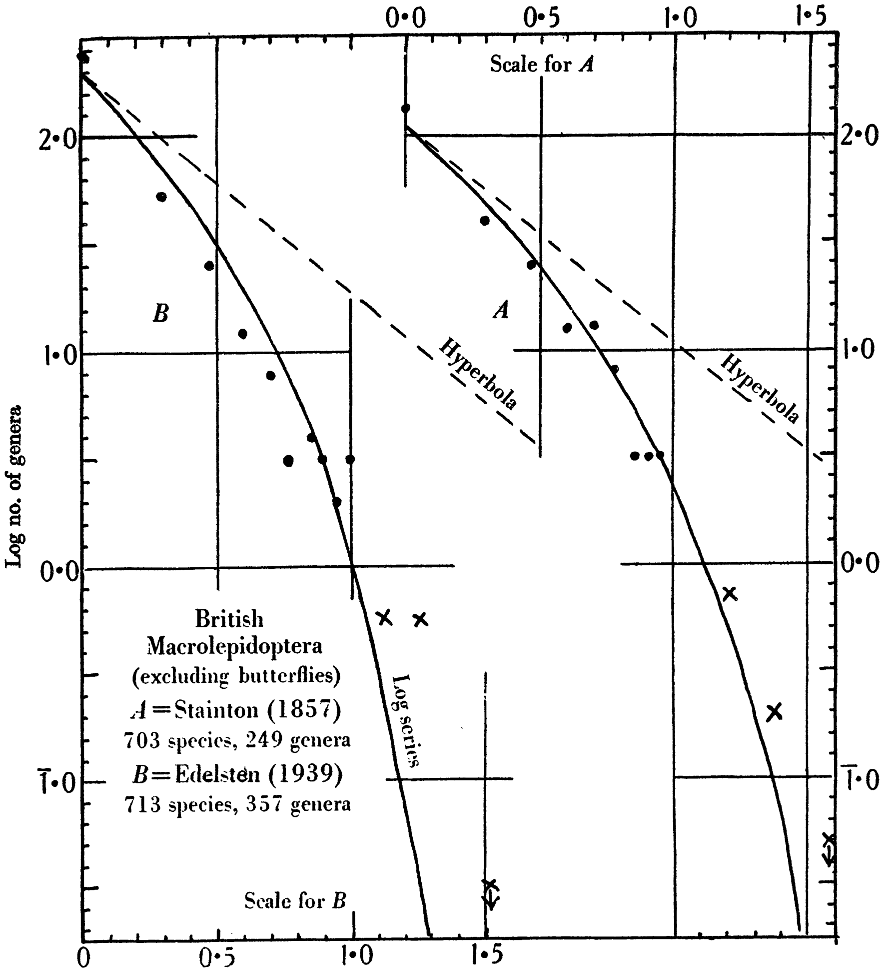


Fig. 15.

species will be divided into 'proportionally' fewer genera with more species in each, than will be the case with a small number of species.

Fig. 22 shows the log number of species and the number of genera in 35 subfamilies of the Orthoptera of the world from Kirkby's catalogue. On the same diagram are lines showing the expected relationships for different values of α. It will be seen that index of diversity varies in the different subfamilies from below 1 to about 120; and that in general the families with more genera and species have the higher index. Subfamilies with under

Table 13. *Frequency of genera with different numbers of species of British Cicadina according to Edwards* (1894–6) *and Butler & China* (1942 *MSS.*)

No. of species per genus	Edwards (1896) Observed	Edwards (1896) Calculated	Butler & China (1942) Observed	Butler & China (1942) Calculated
1	19	15·94	33	25·7
2	9	7·50	10	11·93
3	2	4·73	5	7·39
4	2	3·32	3	5·15
5	— (32)	2·50 (33·99)	2 (53)	3·82 (53·99)
6	2	1·96	4	2·96
7	2	1·58	3	2·36
8	2	1·30	2	1·91
9	2	1·09	2	1·58
10	— (8)	0·92 (6·85)	— (11)	1·32 (10·13)
11	1	0·79	—	1·12
12	2	0·68	1	0·95
13	—	0·59	1	0·81
14	—	0·52	—	0·70
15	1 (4)	0·45 (3·03)	2 (4)	0·61 (4·19)
16	—	0·40	—	0·53
17	—	0·36	—	0·46
18	1	0·32	—	0·41
19	—	0·28	—	0·36
20	— (1)	0·25 (1·61)	1 (1)	0·32 (2·08)

Edwards (1896): Also at 26 and 56. Total: 271 species in 48 genera. $\alpha = 16{\cdot}94$, $x = 0{\cdot}94117$

Butler & China (1942): Also at 21 (2), 33 and 41. Total 360 species in 73 genera. $\alpha = 27{\cdot}65$, $x = 0{\cdot}92867$

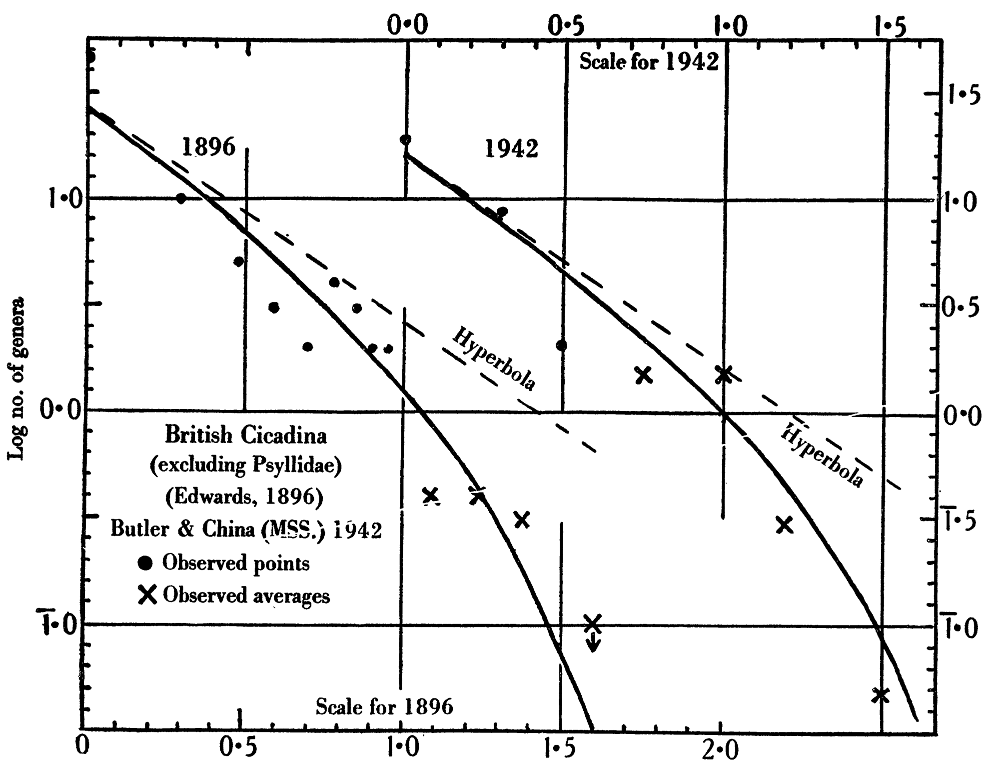


Fig. 16.

Table 14. *Frequency of genera with different numbers of species in British Birds according to Witherby, Jourdain, Ticehurst and Tucker (vol. 5), 1941*

No. of species in genus	Species only: Observed	Species only: Calculated	Species and subspecies: Observed	Species and subspecies: Calculated
1	121	107·8	107	93·08
2	34	40·21	32	38·68
3	12	19·99	17	21·43
4	10	11·18	19	13·36
5	15 (192)	6·67 (185·85)	2 (177)	8·88 (175·43)
6	3	4·15	2	6·15
7	5	2·65	3	4·38
8	1	2·17	2	3·18
9	3	1·15	3	2·35
10	1 (13)	0·77 (10·89)	1 (11)	1·76 (17·82)
11	1	0·45	5	1·33
12	1	0·35	1	1·01
13	1	0·26	4	0·78
14	—	0·17	—	0·60
15	— (3)	0·12 (1·35)	1 (11)	0·46 (4·18)
	Total: 424 species in 198 genera		Total: 520 species and subspecies in 198 genera	
	$\alpha=144{\cdot}56$, $x=0{\cdot}7457$		$\alpha=112{\cdot}01$, $x=0{\cdot}8310$	

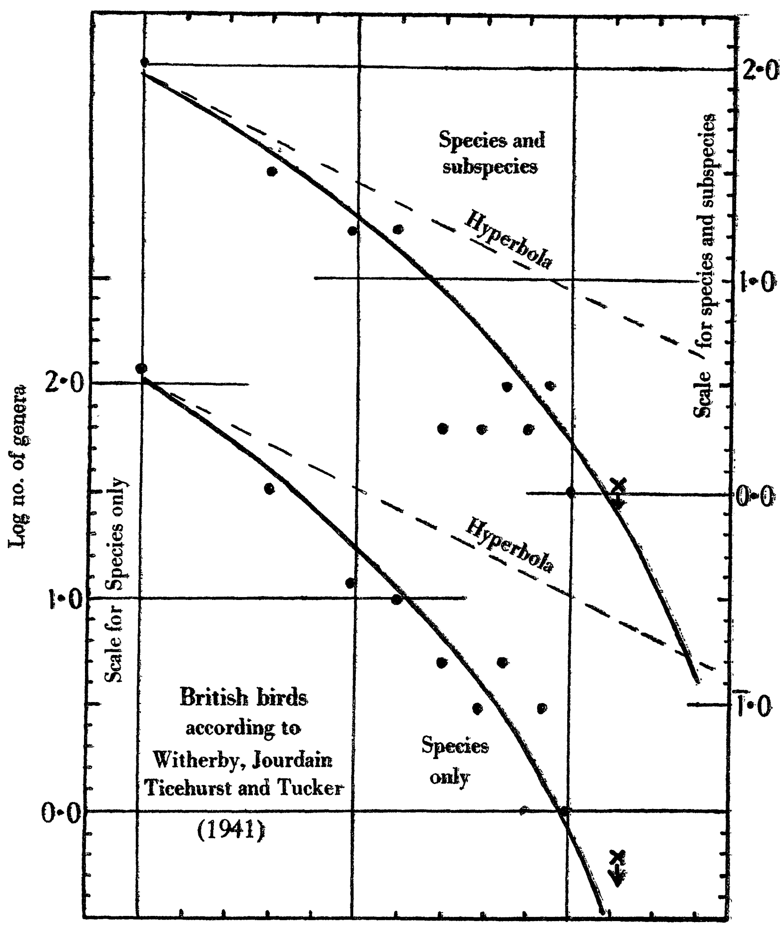


Fig. 17.

100 species nearly all have indices below 10; subfamilies with between 100 and 1000 species have indices from 5 to 50, and the only subfamily with over 1000 species has an index of about 120.

The general interpretation is that the larger subfamilies have more genera and fewer species per genus than would be expected if they were merely larger samples for a population of a similar make up to those from which the smaller families are drawn. On the other hand, some quite large families, e.g. the Phyllodrominae with 463 species and 26 genera, have a relatively low index, as compared with the Truxalinae with 641 species and 141 genera.

Table 15. *Frequency of genera and species of flowering plants*

No. of species per genus	Plants and ferns of world (Willis, 1000 genera)		British flowering plants Babington, 1922 ed.)		British flowering plants (Bentham & Hooker, 1906 ed.)	
	Observed	Calculated	Observed	Calculated	Observed	Calculated
1	309	235·34	255	217·75	256	231·3
2	126	115·9	93	94·69	84	94·3
3	77	76·08	45	54·89	51	51·3
4	37	56·18	37	35·96	22	31·3
5	51 (600)	44·27 (527·77)	20 (450)	24·91 (428·20)	21 (434)	20·4 (428·6)
6	32	36·33	14	18·05	8	13·9
7	18	30·67	8	13·46	5	9·69
8	28	26·42	3	10·23	5	6·91
9	4	23·14	5	7·91	5	5·01
10	54 (136)	20·50 (137·06)	4 (34)	6·19 (55·84)	6 (29)	3·67 (39·18)
11	3	18·35	7	4·90	4	2·72
12	30	16·58	3	3·90	3	2·03
13	2	15·06	4	3·13	1	1·53
14	2	13·77	2	2·53	1	1·16
15	32 (69)	12·66 (76·42)	1 (17)	2·05 (16·51)	3 (12)	0·88 (8·32)
16	3	—	—	1·67	1	0·67
17	2	—	1	1·37	1	0·52
18	2	—	—	1·13	—	0·40
19	0	—	—	0·93	—	0·31
20	37 (44)	8·79	— (1)	0·78 (5·88)	— (2)	0·24 (2·14)
21	0	—	1	—	1	0·18
22	1	—	—	—	—	0·14
23	0	—	—	—	—	0·11
24	2	—	—	—	—	0·09
25	25 (28)	—	1 (2)	—	— (1)	0·07 (0·59)
	Also at 30 (16), 35 (7), 36, 40 (11), 45 (11), 50 (11), 55 (3), 60 (7), 65 (2), 70 (3), 75 (5), 80 (4), 85 (2), 90 (4), 100 (5), 105, 110, 115, 120 (7), 125 (3), 130, 140, 150 (2), 160 (2), 180 (4), 190, 220, 250, 260, 320, 325, 500, 550		Also at 29, 31 (2), 48, 74 and 97		Also one at 47	
	Total: 15,473 species in 1000 genera		Total: 1671 species in 511 genera		Total: 1251 species in 479 genera	
	$\alpha = 238{\cdot}9$, $x = 0{\cdot}98479$		$\alpha = 250{\cdot}4$, $x = 0{\cdot}8696$		$\alpha = 283{\cdot}8$, $x = 0{\cdot}81509$	

Fig. 23 shows the same data for 171 families of dicotyledonous plants of the world from Rendle (1925). The range for species and genera is much greater. Four families have only a single species, and of course a single genus, while at the other end the Compositae have about 1000 genera and 23,000 species.

It will be seen that the same distribution is found as in the Orthoptera: nearly all families with less than 100 species have α values below 10; with 100–1000 species the indices range from below 1 up to 50; with 1000 to 10,000 species there are few with α

below 10 and the upper limit reaches 100; while both families with over 10,000 species have indices over 100.

Again the conclusion is that the larger families have many more genera than would be expected by simple increases in the number of species; and that therefore they differ in structure from the smaller families. On the other hand, there are resemblances and differences which are worth noting. For example, the Piperaceae, with over 1300 species

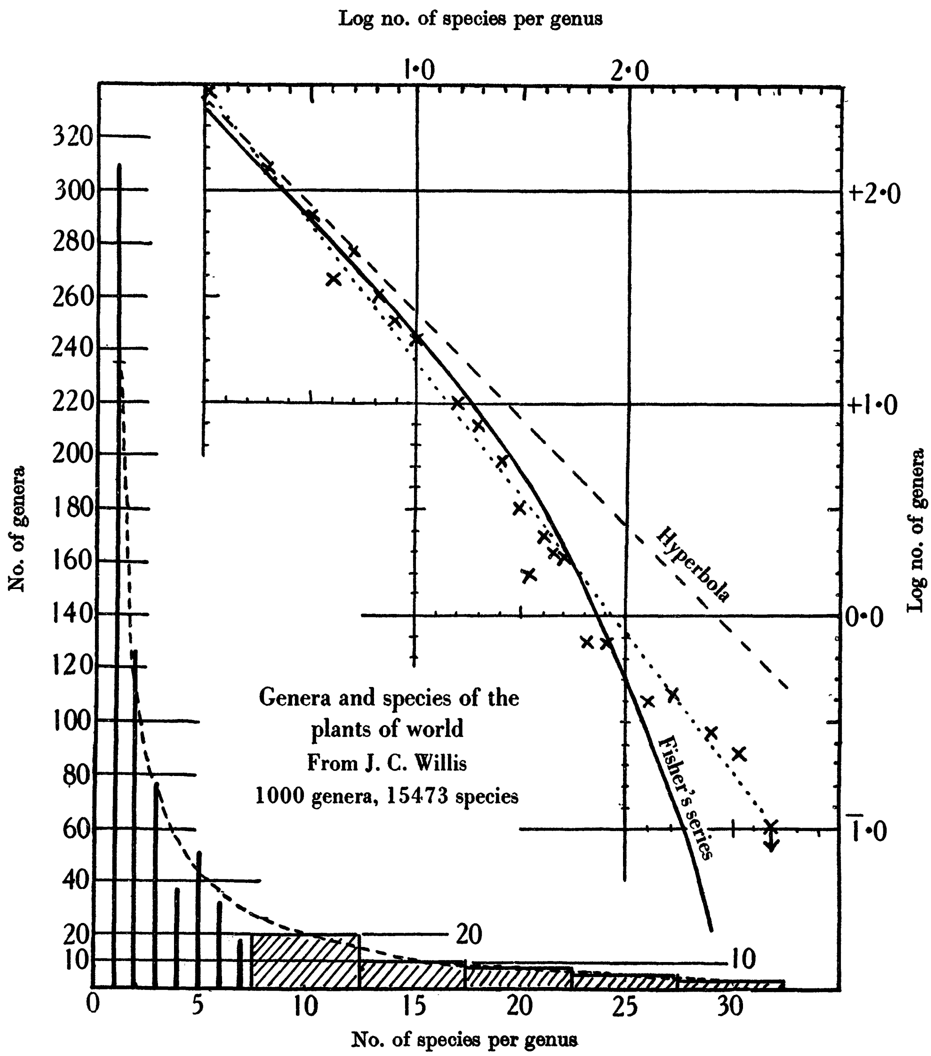


Fig. 18.

and only 9 genera, have an α value of just over 1, which is almost identical with the Sonnerabiaceae which has only 12 species and 3 genera. On the other hand, the Acanthaceae with just over 2000 species but with 220 genera have an index of 64, which makes a remarkable contrast with the Piperaceae. This point will be discussed later.

As final evidence that the principle applies also to local floras, Fig. 24 shows the results for families of the British flowering plants as given by Bentham & Hooker (1906). Below

10 species per family the error of calculation of α is very high, but again it will be seen that, above this number, the families with more species have higher values for the index of diversity.

Frequency of occurrence of different values of α

Fig. 25 shows diagrammatically the frequency of occurrence of different α values in the dicotyledons of the world (Rendle, 1925). The most frequent value for a family is less

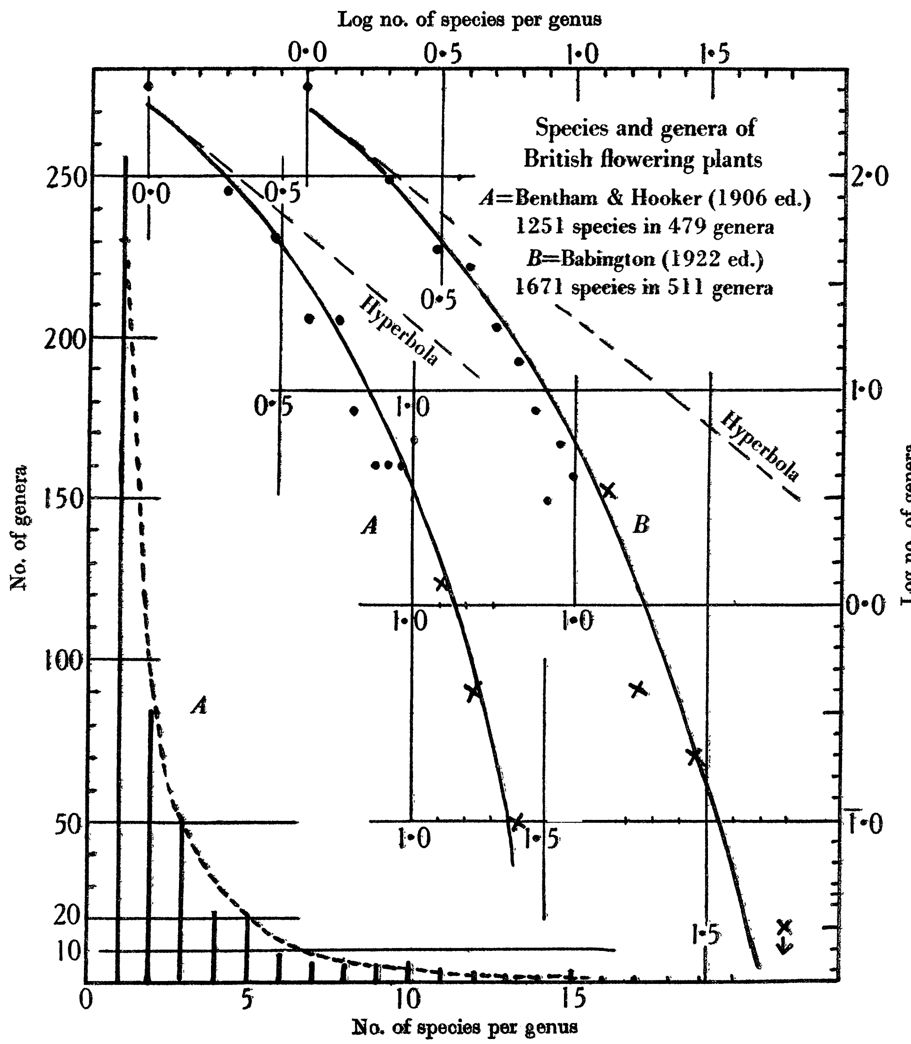


Fig. 19.

than unity; values of 1–2 and 2–3 occur almost equally often, but after that there is a steady fall off. Out of 171 families of dicotyledons, 27 have indices below unity, 22 between 1 and 2, and 23 between 2 and 3.

The α value of all the families together, i.e. of the whole group of dicotyledons, will be the sum of all the values for the different families (see p. 3) and is approximately 1980.

Classification of families of dicotyledons on index of diversity

The problem of the variation in index of diversity in different families appears to be of possible importance in the evolution of the group, as it is only partly explained by the different size of the families.

Other points such as the degree of study of the groups—or the idiosyncrasies of the particular authority who drew up the classification, whether he be a 'splitter' or a 'lumper will also come in.

Table 16. *Frequency of families of flowering plants with different numbers of genera*

No. of genera per family	Plants of world (Rendle)		British plants (Bentham & Hooker)		British plants (Babington)		Plants of Colorado (Rydberg)	
	Observed	Calculated	Observed	Calculated	Observed	Calculated	Observed	Calculated
1	19	31·08	32	29·3	36	32·55	46	42·88
2	7	15·48	17	13·8	19	15·25	36	20·09
3	14	10·28	9	8·6	13	9·51	7	12·55
4	5	7·67	8	6·1	3	6·68	7	8·82
5	5 (50)	6·11 (70·62)	3 (69)	4·6 (62·4)	4 (75)	5·00 (68·99)	4 (100)	6·61 (90·95)
6	7	5·07	2	3·6	3	3·90	3	5·16
7	4	4·33	1	2·9	1	3·13	3	4·15
8	3	3·77	—	2·4	—	2·57	3	3·40
9	7	3·34	3	2·0	1	2·14	3	2·83
10	5 (26)	2·99 (19·50)	1 (7)	1·7 (12·6)	3 (8)	1·80 (13·54)	— (12)	2·39 (17·93)
11	6	2·70	1	1·4	1	1·53	1	2·04
12	3	2·47	—	1·2	1	1·32	2	1·75
13	3	2·27	1	1·1	1	1·14	—	1·51
14	1	2·10	2	0·92	—	0·99	1	1·32
15	1 (14)	1·95 (11·49)	— (4)	0·81 (5·43)	1 (4)	0·86 (5·84)	— (4)	1·15 (7·77)
16	2	1·83	2	0·71	—	0·76	2	1·01
17	3	1·71	1	0·63	—	0·67	1	0·89
18	5	1·61	2	0·56	5	0·59	1	0·79
19	—	1·52	—	0·50	—	0·52	—	0·70
20	3 (13)	1·44 (8·11)	— (5)	0·44 (2·84)	— (5)	0·47 (3·01)	1 (5)	0·62 (4·02)
21	2	1·36	—	0·40	—	—	1	—
22	4	1·29	—	0·35	—	—	—	—
23	2	1·23	—	0·32	—	—	—	—
24	1	1·18	—	0·29	—	—	—	—
25	2 (11)	1·13 (6·19)	— (0)	0·26 (1·62)	— (0)	—	— (1)	—
	Also at 26, 28, 30 (2), 32 (3), 33, 38, 40 (4), 41 (2), 45 (4), 47, 50 (3), 58, 60 (2), 61, 62, 63, 66, 70 (4), 73, 74, 75, 80 (2), 85, 88, 90, 100 (2), 110, 120, 130, 150, 170 (2), 200, 205, 220, 270, 280, 380, 550, 1000		Also one at 27		Also at 30, 37, 46 and 47		Also at 27 (2), 36, 70 and 80	
	Total: 7452 genera in 172 families		Total: 479 genera in 89 families		Total: 511 genera in 96 families		Total: 682 genera in 127 families	
	$\alpha=31{\cdot}2$, $x=0{\cdot}995830$		$\alpha=31{\cdot}19$, $x=0{\cdot}93887$		$\alpha=34{\cdot}76$, $x=0{\cdot}93631$		$\alpha=45{\cdot}76$, $x=0{\cdot}93712$	

To show the differences and to stimulate discussion I have prepared Table 17 which shows all the families of dicotyledons, according to Rendle's classification, grouped according to

(1) the value of the index of diversity,

(2) the number of genera in the family.

Two sets of questions emerge from this table. First, what are the reasons that cause two families to be in the same group? Is it accident, or amount and quality of study, or some question of geographical distribution, or the age of the group in evolution?

Table 17. *Classification of the families of dicotyledonous plants according to value of* α *and number of genera*

Value of α	No. of genera: 1	2–5	6–10	11–20	21–50	51–100
Less than 1	Tropaeolaceae Callitrichaceae Nepenthaceae Garryaceae Myrtaceae Casuarinaceae Calycanthaceae Hydrostachyaceae Clethraceae Symplocaceae	Hippocastanaceae Aceraceae Balsaminaceae Elatinaceae Salicaceae Hydnoraceae Chloranthaceae Droseraceae Tamaricaceae Agnifoliaceae Meophrastraceae Nolanaceae Lentibulariaceae Plantaginaceae Stylidiaceae	Begoniaceae Fagaceae Piperaceae Ebenaceae			
1		Sarraceniaceae Frankeniaceae Caricaceae Margraviaceae Elaeagnaceae Sonneratiaceae Salvadoraceae Globulariaceae Martyniaceae	Betulaceae Aristolochiaceae Resedaceae Cistaceae Buxaceae Oxalidaceae Polygalaceae Haloragaceae Valerianaceae	Dilleniaceae Geraniaceae		
2			Juglandaceae Magnoliaceae Berberidaceae Nymphaeaceae Linaceae Staphyleaceae Vitaceae Pittosporiaceae Connaraceae Cornaceae Plumbaginaceae Styracaceae Dipsacaceae	Myristicaceae Passifloraceae Loasaceae Burseraceae Polemoniaceae Orobanchaceae Caprifoliaceae Goodeniaceae		
3				Ulmaceae Aizoaceae Violaceae Dipterocarpaceae Crassulaceae Combretaceae		
4			Pyrolaceae	Portulacaceae Ternstroemiaceae	Loranthaceae Cactaceae Primulaceae Oleaceae	

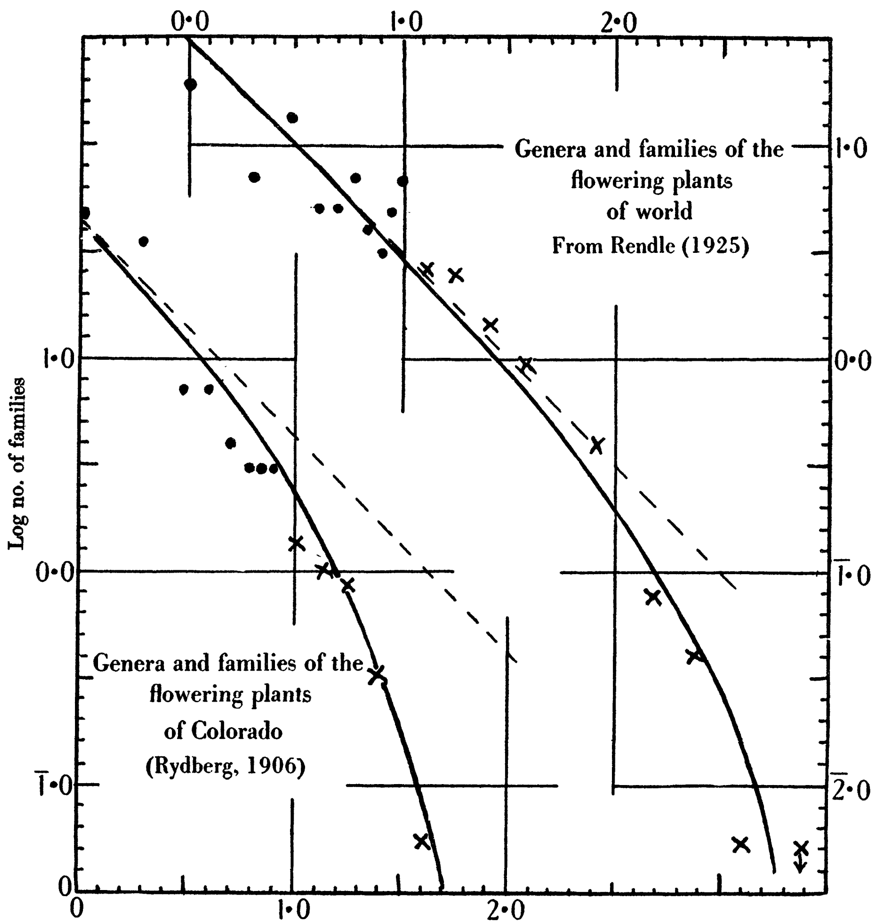
The following eleven Families with less than 10 species are omitted owing to large error in calculating α:

Julianiaceae
Cannabinaceae
Cynomoniaceae
Sauruaceae
Ceratophyllaceae
Bixaceae
Datiscaeae
Empetraceae
Punicaceae
Hippuridaceae
Adoxaceae

					No. of genera		
					101–200	201–500	Over 500
5	Rafflesiaceae Diapensiaceae	Ochnaceae Hydrophyllaceae	Ranunculaceae Lythraceae Epacridaceae				
6–7		Balanophoraceae Nyctaginaceae Lecythidaceae	Santalaceae Polygonaceae Papaveraceae Bombacaceae Amonaceae Myrsinaceae Loganiaceae				
8–10	Lardizabalaceae	Hamamelidaceae	Phytolaccaceae Monimiaceae Lauraceae Capparidaceae Guttiferae Zygophyllaceae Podostemonaceae Thymelaeaceae Onagraceae Sapotaceae Convolvulaceae				
11–15			Urticaceae Amarantaceae Tiliaceae Malvaceae Simarubaceae Meliaceae Celastraceae Rhamnaceae Campanulaceae	Moraceae Proteaceae Myrtaceae			
16–20				Caryophyllaceae Annonaceae Sterculiaceae Malpighiaceae Rosaceae Araliaceae Ericaceae Gentianaceae Boraginaceae Solanaceae			
21–50				Chenopodiaceae Menispermaceae Flacourtiaceae Cucurbitaceae Anacardiaceae Saxifragaceae Verbenaceae Gesneriaceae	Rutaceae Sapindaceae Melastomaceae Apocynaceae Labiatae Bignoniaceae	Scrophulariaceae	
51–100					Cruciferae	Acanthaceae Euphorbiaceae Rubiaceae Umbelliferae Asclepiadaceae	
101–200						Leguminosae	
Over 200							Compositae

3-2

Secondly, why do two families with the same index of diversity have widely differing numbers of genera, or two families with the same number of genera have widely differing values of α? Again much the same questions arise, and I am not botanist enough to contribute greatly to the discussion. I am, however, continuing this line of study with data from different groups of insects.



Log no. of genera per family

Fig. 20.

5. General discussion

We have in this paper discussed the application of the logarithmic series and the conception of the index of diversity as a measure of the extent to which groups are divided up into units, or units are classified into groups. We have shown that calculations based on this theory give a close fit, and sometimes a remarkably close fit, to observed facts in the frequency distribution of plants in different associations and in samples of different sizes; in the distribution of lice on the heads of their hosts; in the distribution of different species of mosquitoes in American cities; and in the number of butterflies of different species captured in Malaya. In a previous paper (Fisher *et al.* 1943) it has also been shown to apply well to numbers of Lepidoptera and of Capsidae (Heteroptera) caught in a light trap at Harpenden.

All these are problems of the distribution of individuals into groups, and usually into species; and they conform more or less to the original conception of sampling of a mixed population from which Fisher deduced the logarithmic distribution. We have also shown here that the same series explains—or at least will fit well—the frequency of genera with different numbers of species in several different groups of animals and plants, and also

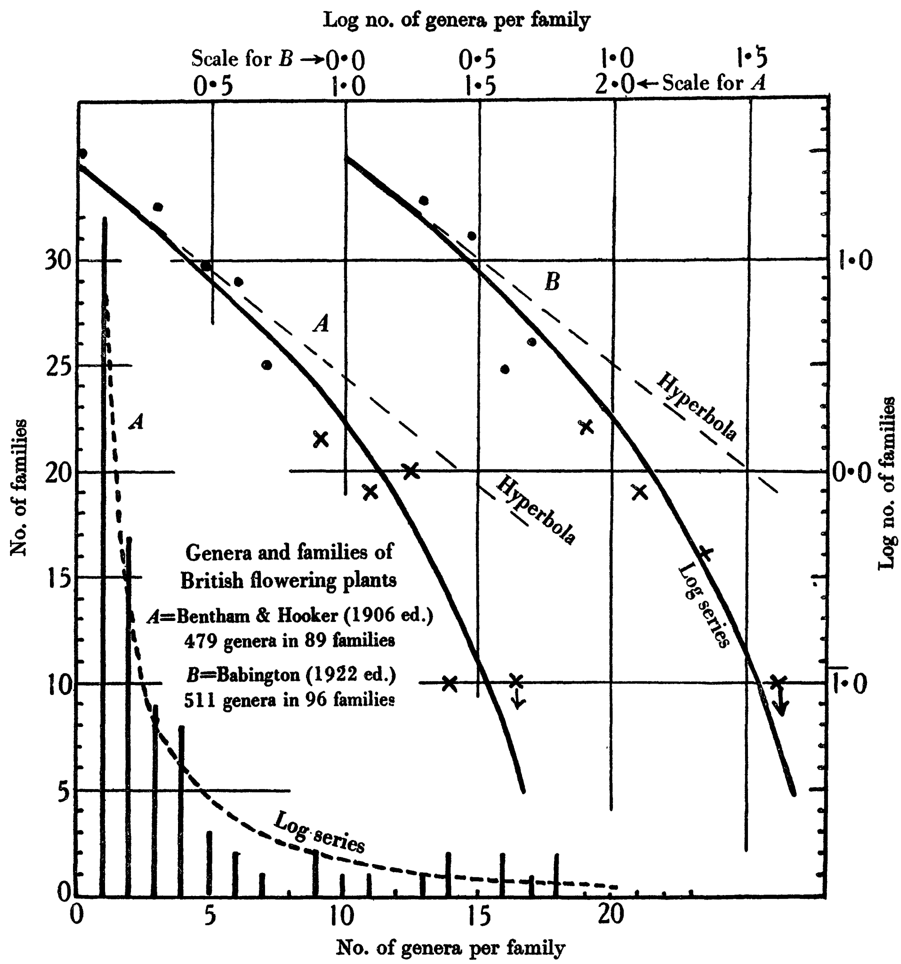


Fig. 21.

the frequency of families or subfamilies with different numbers of genera. The fit of the facts to the logarithmic series is very much closer than to the hyperbolic series previously proposed by several writers.

If we allow that the close fit is anything more than a series of coincidences we must admit that there is a fundamental law and order in the frequency of species with different numbers of individuals; in the frequency of genera with different numbers of species; and in the frequency of families with different numbers of genera; and further, the law and order must be approximately the same for each.

In my own somewhat limited experience of systematic or taxonomic work on insects, I have always held the opinion that the 'species' of good systematists were about 90% biological realities and 10% groupings of convenience. I have never had the same feeling of confidence in the systematist's 'genera' which I would have classified as perhaps 20% biological reality and 80% personal opinion and convenience. If, however, the grouping of individuals into species, as classified by systematists, is subject to the same mathematical laws as the grouping of species into genera, it seems to imply that the biological

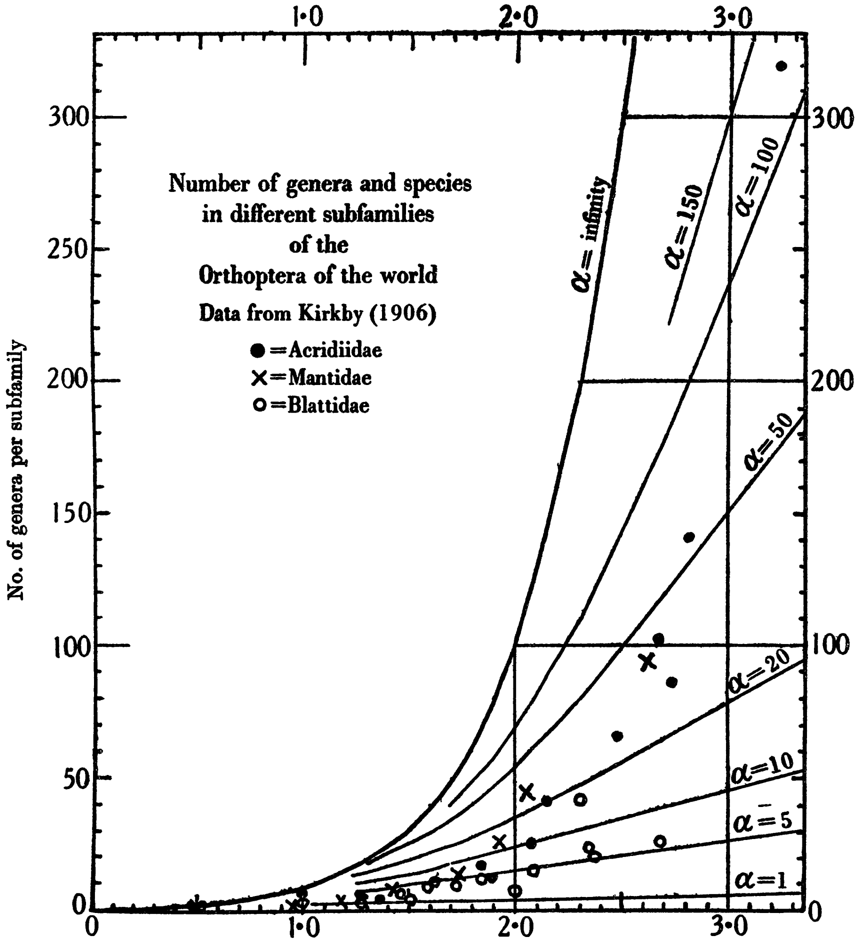


Fig. 22.

reality of genera is of the same order of reality as that of the systematic 'species'. Unless, of course, the logarithmic series is so universal that it can include in its scope the idiosyncrasies of individuals and opinions based on quite insufficient evidence.

It is important, I think, at this stage to distinguish between the real biological species or genus, and the estimate made by the systematist from the material available; much in the same way as the statistician distinguishes the 'parameter', or real truth towards which he is striving from the 'statistic' or estimate of the truth that he determines from his sample. It is also important to note that the logarithmic series gives no information as to whether a systematist is right or wrong in classifying a number of species into many or into a few genera. But once the number of genera has been fixed, we can see immediately if the division has been carried out consistently. One authority may be a 'splitter' and the other a 'lumper' of species or genera; yet both, or neither, may be consistent as tested by the log series.

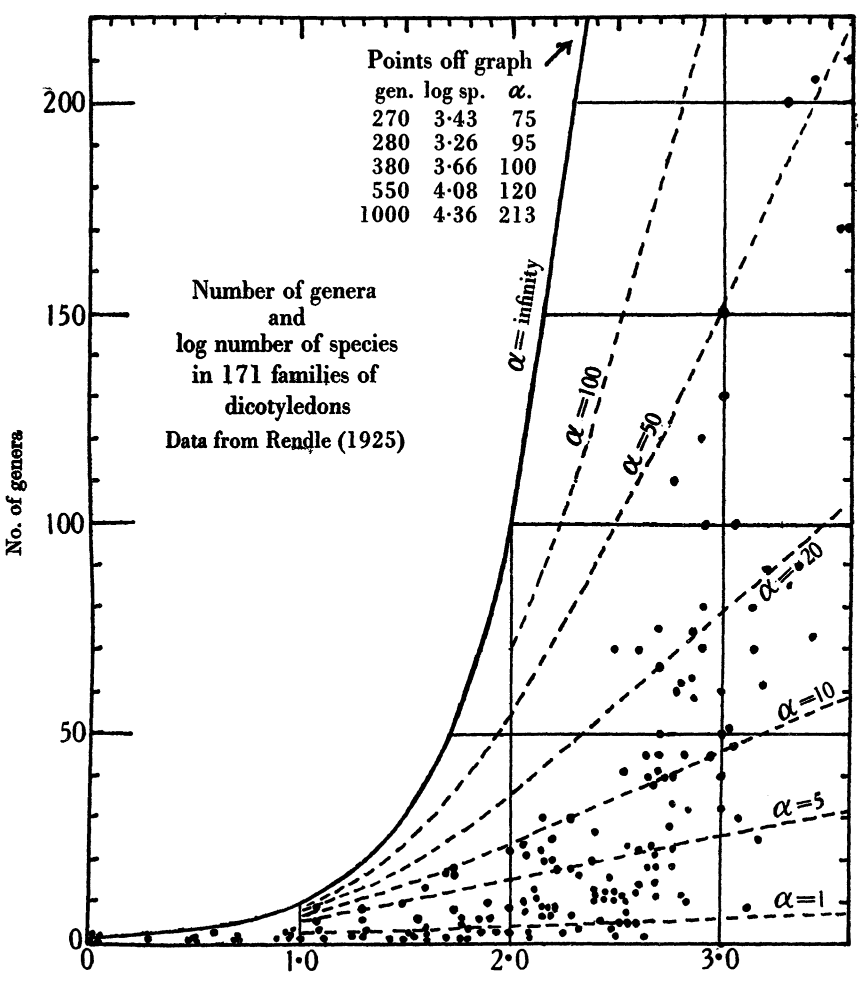
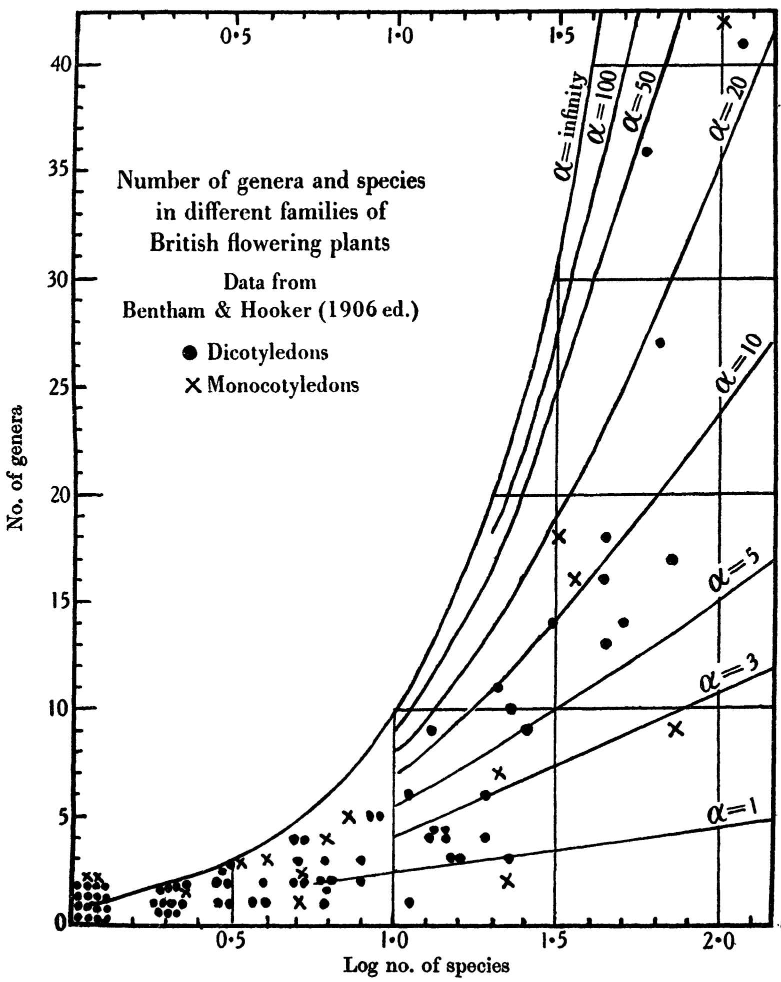


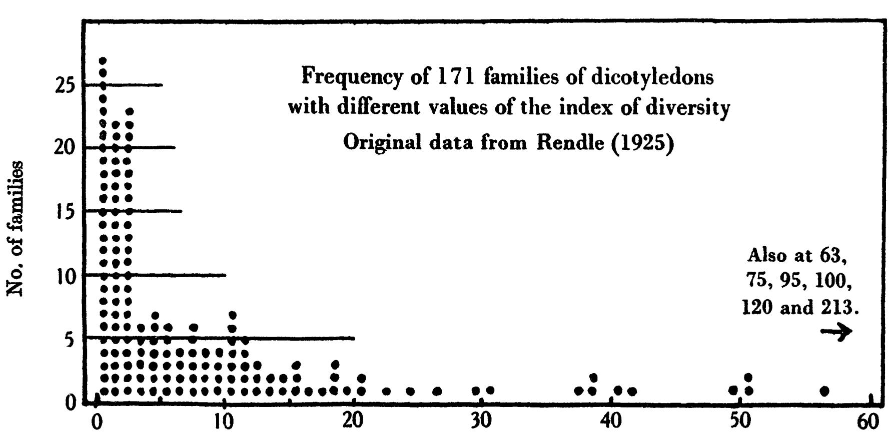
Fig. 23.

In this connexion we should notice how very close indeed to the logarithmic series Kirkby has quite unconsciously come in his classification of some of the families of Orthoptera. The Mantidae are the most remarkable with 805 species classified in 209 genera. This gives the calculated number of genera with one species as 82·27 and the observed 82; the calculated number of genera with 1–5 species as 166·88 and the observed 169; and the calculated number of genera with 6–10 species as 25·98 and the observed 24.



Number of genera and species
in different families of
British flowering plants
Data from
Bentham & Hooker (1906 ed.)
● Dicotyledons
× Monocotyledons
No. of genera
Log no. of species
α = infinity
α = 100
α = 50
α = 20
α = 10
α = 5
α = 3
α = 1

Fig. 24.



Frequency of 171 families of dicotyledons
with different values of the index of diversity
Original data from Rendle (1925)
Also at 63, 75, 95, 100, 120 and 213.
No. of families
Value of index of diversity, α

Fig. 25.

It seems to me possible that in a relatively small group revised largely by one man, and particularly in the days when the multiplicity of species (at least in the insects) had not grown beyond its conception as a whole by one brain, the classification is more likely to be psychologically consistent, and hence mathematically consistent, than in a larger group in which many authorities have had their say to different extents in different parts.

In other words, I would suggest that the basic natural classification would be consistent mathematically with some series, such as the one under consideration, and that departures from consistency would be due to errors or incomplete knowledge on the part of the person or persons responsible for the grouping.

In this connexion it is interesting to note how close to the logarithmic series are the numbers of men with different numbers of lice on their heads. This is a purely mechanical problem, with no question of doubt about the identity of either individual lice or individual heads; and with 7442 lice or 461 heads the calculated number of heads with one louse is 107·15 and the observed 106; and the calculated number of heads with 1–10 lice was 303·25 and the observed 304.

There is, however, one general inconsistency that must be noted. In the present paper calculated and observed series are given for 22 sets of data. In only one of them is the observed n_1 definitely below the calculated; in two more it is slightly below; in two almost identical; and in 17 cases the observed n_1 is above the calculated. In the paper on light-trap captures in Harpenden already published (Fisher *et al.* 1943), there were given 27 sets of figures, in 19 of which the observed n_1 was above the calculated and in only 8 below. In some as yet unpublished analyses of light-trap captures of Lepidoptera in U.S.A. I have calculated eleven series of frequency distributions and they all give calculated n_1 below the observed; sometimes the difference is small but in others quite large; in one case the calculated n_1 is 10 and the observed 19.

There seems therefore to be a general tendency for the logarithmic series to under-estimate the number of groups with one individual, as compared with the observed facts. Either the theory is incomplete or there is some biased error in the observations producing too many very small groups. In some cases groups with 2, 3 and 4 individuals are also under-estimated, but more usually they give observed results below the calculated so that the figures for groups with 1–5 units are very much nearer than calculated results. In the case of individuals and species it means that there are too many extremely rare species, and this might be partly explained by the fact that in general systematists are more interested in rare species than in common ones and will take more trouble over the description of a 'new species' than over the identification of a second species of one already known.

In the case of species grouped into genera it means that there are too many monotypic genera. This might be partly explained by an unfortunate desire on the part of a systematist to describe new genera, but this would not produce the result if the process of splitting was applied equally to large and to small genera.

There is undoubtedly need for further investigation both by testing of new observations and by further examination of the mathematical theory in the light of the present results.

Turning to the conception of the index of diversity we see that it can be applied with interesting and stimulating results to such problems as the number of species of plants on areas of different sizes, and can be used to give an estimate of the number of individual plants on a given area, which can be used as a check against other deductions. Various writers have suggested empirical formulae connecting the number of species on quadrats of different sizes, and some of them have pointed out the approximate straight-line relationship between the logarithm of the area and the number of species over a part of its range. Fisher's index of diversity is, however, the first theory to give an accurate estimate for both small and large samples, the first to be based on a logical foundation, and the first to link up botanical data on quadrats of various sizes with zoological data on samples containing different numbers of animals. It marks, I think, the beginning of a new mathematical study of mixed populations as opposed to the study of the changes in numbers of a single species—perhaps in relation to one or two parasites—which has occupied the attention of mathematically minded biologists, or biologically minded mathematicians, in recent years.

6. Index of diversity used as a comparison of different populations

It is interesting to note that in addition to the points discussed above in connexion with the index of diversity, this mathematical conception gives up a new possibility in the comparison of the fauna or flora of different areas. If we have two areas of land, say two islands, or two counties or two countries, then according to their ecological and geographical relationships there will be more or fewer species common to the two areas. If the two areas are close together and similar ecologically there will be many species common; but of course, owing to the accident of sampling, there will always be some in each area that do not occur in the other. At the other end of the scale, with complete difference between the two areas there would be no species in common.

Now the index of diversity gives us the possibility of calculating—with certain assumptions and under certain conditions—the number of species that would be common to two areas if they were of the same population only differing in size.

The assumption necessary is that the density of population is the same for both; and as will be seen later the areas must be of different sizes. The mathematical treatment is as follows:

If there are two localities with areas A and B and with a and b species respectively in the group of plants and animals under consideration; then if the two are samples of the same population the value of α will be the same for each, and the same for the two together.

Let x be the theoretical total number of different species found on the two areas together on the supposition of identity. Then

(1) The increase in numbers of species by adding B to A

$$=x-a=\log_e \frac{A+B}{A}.$$

(2) The increase in numbers of species by adding A to B

$$=x-b=\log_e \frac{A+B}{A},$$

hence

$$\frac{x-a}{x-b}=\frac{\log (A+B)/A}{\log (A+B)/B}$$

(whether logs to base e or to base 10 are used). From this x can be calculated and the number of species common to both areas is of course $a+b-x$.

It is interesting to note that the above method only gives x if the two areas are of different sizes. Expressed graphically we are testing the fit of three points to a curve of a particular form; and if two points are identical the method of test fails.

As a practical example we may take the two Channel Islands of Guernsey, 24 square miles, having 804 flowering plants and ferns, and Alderney, 3 square miles, with 519 species (Marquand, 1901).

On the assumption of identity of flora

$$x-804=\alpha \log_e \tfrac{27}{24}=\log_e 1\cdot125, \quad x-519=\alpha \log_e \tfrac{27}{3}=\log_e 9\cdot00,$$

whence $x=820\cdot126$.

Otherwise on the theory of identity of flora there should be 820 species in the two islands combined, or 503 species common to both.

The actual number given in Marquand's *Flora* as common to both is 480 or 95·4% of the calculated number. This indicates a close degree of relationship.

When, however, the method was tested to see the relationship between Wicken Fen and Scolt Head Island—each with an area just over a square mile and with 185 species of plants—it broke down as explained above, owing to the similarity of the two areas. In fact, the total number of species common to the two, on the assumption of identity is $185-0\cdot69\alpha$, and so is dependent on the value of the index of diversity, which cannot be found from the data given. It could, however, be found if we knew the number of species on a definite fraction of either of the two localities. Further work along these lines is now being carried out.

7. SUMMARY

The paper describes the application of a logarithmic series to a number of problems of the division of individuals into species and of species into genera, the range of which is best seen from the table of contents.

The series, first suggested by R. A. Fisher in this connexion, is

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

where n_1 is the number of groups with one unit and x is constant less than unity. Unlike the hyperbolic series, which had previously been considered to apply to some of the cases discussed, the logarithmic series is convergent: both the number of groups (e.g. species) and the number of units (e.g. individuals) can be summed. When several samples are taken from a population containing a number of species it is found that the ratio n_1/x is constant and, as it is therefore a characteristic of the population, it has been called the index of diversity.

The logarithmic series is found to fit extremely well to a large number of frequency series drawn from insects, birds, butterflies and plants, except that there is a slight tendency for the calculated n_1 to be below the observed.

It also fits well, sometimes extremely well, to the number of genera with different numbers of species in standard classifications of groups of both animals and plants.

The conception of the index of diversity is applied to problems of the number of species of plants on different areas, and to the comparison of floras of different areas with interesting results.

A classification is given of the 171 families of dicotyledons according to their index of diversity to stimulate a discussion as to which may be the factors which bring about differences and resemblance in this Index. In general, the families with large numbers of species and of genera have large index of diversity, but there may be a very big range of index in families of approximately the same size.

REFERENCES

Arrhenius, O. (1923). Statistical investigations in the constitution of plant associations. *Ecology*, **4**, 68–73.

Babington, C. C. (1922). *Manual of British Botany*, 10th ed.

Beare, T. H. & Donisthorpe, H. St J. K. (1904). *A Catalogue of the British Coleoptera*. London.

Bentham, G. & Hooker, J. D. (1906). *Handbook of the British Flora*, 6th ed.

Blackman, G. E. (1935). A study by statistical methods of the distribution of species in a Grassland Association. *Ann. Bot., Lond.*, **49**, 749–78.

Buxton, P. A. (1938). Studies on populations of headlice (*Pediculus humanus capitis*). II. *Parasitology*, **30**, 100.

Buxton, P. A. (1940). Studies on populations of headlice. III. Material from South India. *Parasitology*, **32**, 296.

Chamberlin, J. C. (1924). Concerning the hollow curve of distribution. *Amer. Nat.* **58**, 350–74.

Edwards, J. (1894–6). *The Hemiptera-Homoptera* (Cicadina *and* Psyllina) *of the British Isles*. London.

Fernald, M. E. (1903). *A Catalogue of the Coccidae of the World*. Hatch Exp. Station, Amherst, Mass. Bull. 88.

Fisher, J. (1940). *Watching Birds*, p. 102. London: Penguin Books.

Fisher, R. A., Corbet, A. S. & Williams, C. B. (1943). The relation between the number of individuals and the number of species 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**, 158–62.

Gleason, H. A. (1925). Species and area. *Ecology*, **6**, 66–74.

Kirkby, W. F. (1904–10). *A Synonymic Catalogue of the Orthoptera*. London: British Museum.

MacGillivray, A. D. (1921). *The Coccidae*. Urbana, Illinois.

Marquand, E. D. (1901). *The Flora of Guernsey and the Lesser Channel Islands*. London.

Rendle, A. B. (1925). *The Classification of the Flowering Plants*. **2**. Dicotyledons. London.

Rowe, J. A. (1942). Mosquito light trap catches from ten American cities, 1940. *Iowa St. Coll. J. Sci.* **16** (4), 487–518.

Rydberg, P. A. (1906). *The Flora of Colorado*. Exp. Station, Fort Collins, Col.

Stainton, H. T. (1857). *A Manual of British Butterflies and Moths*. London.

South, R. & Edelsten, H. M. (1939). *The Moths of the British Isles*, new edition.

Willis, J. C. (1922). *Age and Area*. Cambridge.

Willis, J. C. (1925). *A Dictionary of Flowering Plants and Ferns*, 5th ed. Cambridge.

Witherby, H. F., Jourdain, F. C. R., Ticehurst, N. F. & Tucker, B. W. (1941). *The Handbook of British Birds*, vol. 5.