Williams, C. B. 1953. The relative abundance of different species in a wild animal population. *Journal of Animal Ecology.* 22 (1), pp. 14-31.

The Relative Abundance of Different Species in a Wild Animal Population
Author(s): C. B. Williams
Source: *Journal of Animal Ecology*, Vol. 22, No. 1 (May, 1953), pp. 14-31
Published by: British Ecological Society
Stable URL: https://www.jstor.org/stable/1687

THE RELATIVE ABUNDANCE OF DIFFERENT SPECIES IN A WILD ANIMAL POPULATION

By C. B. WILLIAMS
Rothamsted Experimental Station, Harpenden

(With 9 Figures in the Text)

In a series of papers published during the past six years I have discussed the frequency distribution of the relative abundance of different species of animals in random samples from wild populations.

From data obtained, chiefly by collecting Lepidoptera in a light-trap, it was found that there was an orderly distribution in such samples, and in practically all cases the number of species represented by one individual was greater than the number represented by two, the number with two greater than that with three, and so on. The distribution in the sample, on an arithmetic scale of number of individuals per species, was in the form of a 'hollow curve'.

Many different mathematical formulae can be represented graphically by curves of this type, and it is unfortunately often difficult to distinguish between them, especially with variable biological data. This difficulty is greater in small samples, but less in large samples.

It was found that the logarithmic series (which is a special form of the negative binomial) gave a very good fit to the data obtained from most of our trapping experiments, and also to evidence from other sources. This series has the form

$$\alpha x; \quad \alpha x^2/2; \quad \alpha x^3/3; \quad \alpha x^4/4; \quad \text{etc.};$$

where the successive terms are the number of species represented by 1, 2, 3, 4, etc. individuals. The x is a number less than unity which is constant for any one sample from a population; but it varies according to the size of the sample, being larger in large samples and gradually approaching unity in very large samples. The α, on the other hand, is a constant for all samples from the same population; it is a property of the population and not of the sample, and we have called it the 'Index of Diversity' (Fisher, Corbet & Williams 1943).

Since the number of species represented by one individual each is αx it follows that, with increasing size of sample, this must increase as x increases until in large samples it approaches very close to α. Thus, in populations with a frequency distribution based on the log-series, α is not only a measure of diversity, but is also the upper limit to the theoretical number of species which can be represented by one individual either in a very large sample or in the population itself.

There is no doubt that the log-series has given a very close fit to observed data in many samples of animal populations, and that many deductions from it, such as the straight-line relation between number of species and the logarithm of the number of individuals in samples of different sizes, have also been found to be correct within the limit of normal sampling. Normal samples, however, usually consist of a few thousand individuals, or perhaps occasionally a few hundred thousand, while populations, at least in insects, have to be reckoned in millions; so there is always considerable extrapolation necessary in arguing from the structure of a sample to the structure of the population.

Theoretically the log-series is one in which the form of the series is not altered by the process of sampling. If a population is arranged in a log-series, then any random sample from it is also in the form of a log-series with the same α but a lower value of x. Thus if any sample was a mathematically exact fit to log-series, we would be correct in inferring a similar distribution in the population sampled. This mathematical exactitude is, however, just what can never be obtained in biological observations.

In 1948 F. W. Preston suggested that the frequency distribution of an animal population might be a log-normal distribution, and not a logarithmic series. This distribution (which is a continuous curve and not a discontinuous integer series) is the normal Gaussian curve when the dimensions of the variate are expressed in logarithmic classes instead of on an arithmetic scale. Preston further suggested that the frequency distribution found in a small sample from a population arranged in a log-normal pattern is a 'truncate log-normal', and that this, when shown on an arithmetic scale, may very closely resemble a logarithmic series distribution.

It is therefore of considerable interest to see how these two alternative suggestions can be separated, if the biological data at our disposal are sufficiently accurate to support one or the other, and what are the implications of accepting either one or the other. There remains, of course, always the likelihood that neither is the really correct solution.

In the first place attention should be drawn to the fact that the log-normal distribution has three constants or parameters, while the logarithmic series has only two. In the case therefore of data that are a possible fit to either, the curve with three constants would automatically give a better fit to the data without explaining more. For a given number of individuals and a given number of species, there is only one possible log-series, but many log-normals. A logarithmic series is defined by the total number of individuals and the total number of species, but to define a log-normal curve a third quantity is necessary—one involving the 'standard deviation' of the distribution. The form of the log-normal curve is determined by the total number of species (the area of the curve) and by the standard deviation: the position of its median is determined by the number of individuals.

It has already been pointed out that, if in the logarithmic distribution a series of samples of increasing size is taken from the same population, the number of species with one individual (n_1) is equal to αx; it is small with very small samples, increases rapidly at first, and then more and more slowly till it approaches to the value of α. It cannot, except by chance error, surpass this value, nor does it theoretically show any fall as the sample size is further increased so as to include the whole population.

In the case of the log-normal population, the number of species represented by one individual will start small in small samples, rise at first rapidly, then more slowly, as sample size increases, till it reaches a peak when the sample is of such a size that about half the species in the population have been obtained; after this the number will fall at first slowly, then more rapidly and finally very slowly again, till it becomes very small in the whole population.

A second difference is that with the logarithmic series the number of species represented in a series of samples of increasing size is, except in small samples, always proportional to the logarithm of the size of the sample; i.e. there is graphically a straight-line relation between the number of species and the logarithm of the number of individuals. With the log-normal the number of species in samples of increasing size at first increases slowly, then takes up a straight-line relation very similar to the logarithmic series, but in larger samples, when over half of all the species in the population are represented, the rate of increase of species falls off so that the curve for the relation

Table 1. *Number of species of Macro-lepidoptera with different numbers of individuals, caught in a light trap at Harpenden in the year 1935*

(Total catch = 6814 individuals representing 197 species; an average of 34·6 individuals per species. On the assumption of the log series $x = 0{\cdot}994$, $\alpha = 38$.)

ON ARITHMETIC SCALE

Individuals	Species	Individuals	Species	Individuals	Species	Individuals	Species
1	37	11	2	21	4	31	—
2	22	12	4	22	1	32	—
3	12	13	2	23	1	33	2
4	12	14	3	24	—	34	2
5	11	15	2	25	1	35	—
6	11	16	2	26	—	36	—
7	6	17	4	27	—	37	—
8	4	18	2	28	2	38	1
9	3	19	—	29	2	39	1
10	5	20	4	30	—	40	3

and also at 42(2), 48(2), 51, 52, 53, 58, 61, 64(2), 69, 73, 75, 83, 87, 88, 105, 115, 131, 139, 173, 200, 223, 232, 294, 323, 603 and 1799

IN GEOMETRIC $\times 3$ CLASSES

I	II	III	IV	V	VI	VII	VIII
37	46	48	37	19	8	1	1

between number of species and log-number of individuals is sigmoid. It would be necessary, however, to have a sample large enough to contain at least 80% and perhaps 90% of the species in the population before the sigmoid nature of the curve would be obvious graphically.

To differentiate between the two alternative theories by either of these methods, a series of samples of different sizes must be obtained, and the larger samples must be taken under conditions which are identical with those of the smaller samples. This is not always practical. A third, more promising, method of distinction is a study of the form of frequency distribution within a sample when the numbers of individuals per species are grouped together in classes on a geometrical scale instead of in an arithmetic scale. To do this Preston made what he called a series of 'octaves', each class being twice the size of the previous, by adding all species with from 1 to 2 individuals in octave I, 2 to 4 individuals in octave II, 4 to 8 individuals in octave III, 8 to 16 individuals in octave IV, etc.

This system, however, necessitated splitting the numbers of species with 1, 2, 4, 8, etc., individuals into those below the integer, which go into the lower class, and those above the integer which go into the higher. The word 'octave' is also unfortunate as its derivation and general use in music is based on the eight notes into which each octave is divided, and not on the fact that the frequency of vibration is doubled.

If instead of this series we define the first class to include all values between 0·5 and 1·5, and (since 1·5 is 3 times 0·5) the second class all values between 1·5 and 4·5, and the third from 4·5 to 13·5, etc., we get a 'three times' geometric classification* into

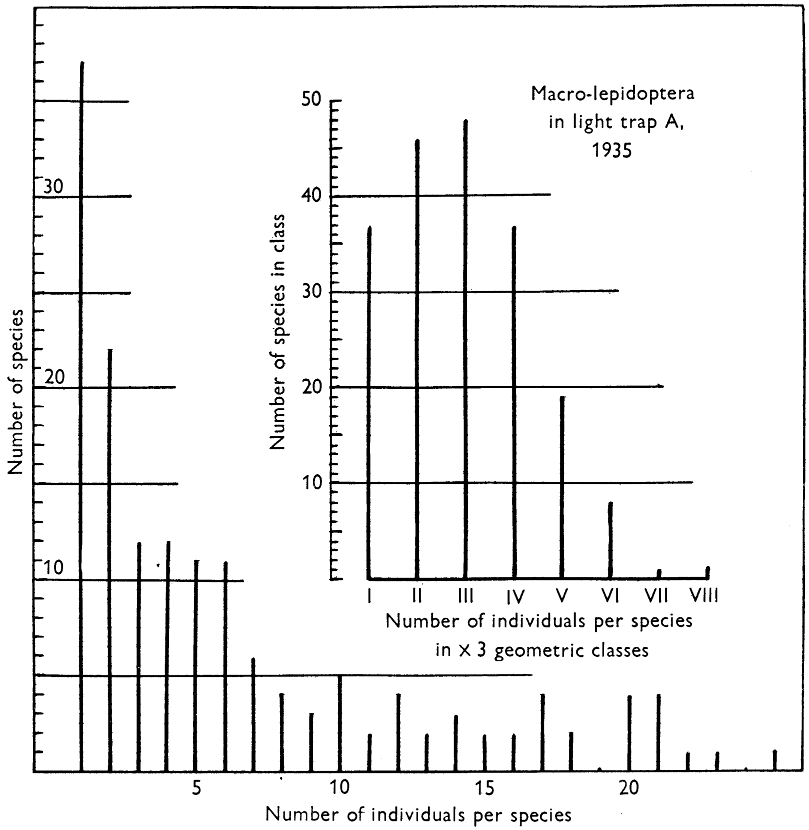

Fig. 1. No. of species of Macro-lepidoptera with different numbers of individuals captured in a light trap (A) at Rothamsted in the year 1935. First with the numbers of individuals per species on an arithmetic scale, and second in × 3 geometric classes as explained in the text. Data in Table 1.

which the original integer observations fit very simply and require no splitting. The integers in the classes are thus

Class			Class	
I	1		VI	122–364
II	2–4	inclusive	VII	365–1,093
III	5–13	,,	VIII	1,094–3,280
IV	14–40	,,	IX	3,281–9,841
V	41–121	,,	X	9,842–29,524

To give an example, the number of species of moths represented by 1, 2, 3, 4, etc., individuals trapped in a light trap at Rothamsted during the year 1935 is shown in Table 1 and in Fig. 1, at first on an arithmetic integer scale and then on a × 3 geometric classification.

* If for the material in hand this grouping is too close, a × 5 classification can be made with the dividing lines at 0·5, 2·5, 12·5, 62·5, etc., i.e. with the integers 1–2, 3–12, 13–62, etc. Or any larger *odd* number can be used for multiplication.

It will be seen that on an arithmetic scale these data give a 'hollow curve' somewhat resembling a hyperbola, but when arranged in $\times 3$ classes there is a curve resembling a truncate normal distribution with the peak between classes II and III.

If calculations are made for different sized samples taken from a population arranged in a logarithmic series the number of species in the $\times 3$ classes are as shown in Table 2 and in Fig. 2. The distributions are similar for all samples with the same average number of individuals per species—that is to say also with the same value of x.

Table 2. *Theoretical frequency distribution of species, in $\times 3$ classes, in samples of different sizes taken from a population arranged in a logarithmic series with $\alpha = 100$**

Size of sample				Number of species in class			
Individuals	Species	Average individuals per species	x	I (1)	II (2–4)	III (5–13)	IV (14–40)
43	37	1·16	0·3	**30**	5·60	0·065	—
100	69	1·44	0·5	**50**	18·23	1·09	—
233	120	1·94	0·7	**70**	41·94	8·32	—
900	230	3·91	0·9	**90**	81·28	47·91	10·94
1,900	300	6·34	0·95	**95**	94·07	72·61	33·50
2,173	306	6·95	0·956	**95·6**	**95·6**	—	—
3,746	360	10·5	0·974	97·4	**99·77**	88·56	58·65
9,900	461	21·47	0·990	99·0	**105·36**	101·00	86·05
19,900	531	37·48	0·995	99·5	**106·95**	105·27	97·25
99,900	691	144·6	0·999	99·9	108·04	**108·50**	107·18
Inf.†	Inf.	Inf.	1·000	100	108·33	109·68	109·84

* For other values of α the numbers in the above table must be multiplied by $\alpha/100$.
† Hyperbolic or harmonic series.

It will be seen that for small samples the curve is of the 'hollow' type even on this geometric classification. When, however, we reach a sample with $x = 0{\cdot}956$ (or just under seven individuals per species on an average in the sample), the number of species in class II becomes equal to class I. With just over 100 individuals per species class III becomes greater than class II. Any samples larger than this give a curve with a peak which gradually moves further along the class ordinate. When the sample is infinite x becomes 1 and the log-series becomes a hyperbolic or harmonic series. In this case the frequency distribution is 100 species (i.e. α) in class I, rising to 108·3 in class II and then becoming constant just below 110 in all further classes. Thus one of the characteristics of the log-series is that the number of species in the peak class cannot (except by accidental error) be more than 10% above the number of species with one individual however large the sample.

We thus have in animal populations three different ways of distinguishing between the frequency distribution of species and individuals based on the logarithmic series and on the log-normal:

(1) From changes in the number of species with one individual (n_1) in samples of increasing size from the same population.

(2) From the rate of increase of number of species in relation to sample size with very large samples.

(3) From the distribution of species in $\times 3$ (or any other geometric) grouping of classes of abundance.

All these three methods are most useful with large samples, as the resemblance between the two theories is so close in small samples as to make them practically indistinguishable. The third method has the advantage of requiring only a single large sample.

It is next necessary to examine the data available to see how this can be analysed to show any of the above differences.

The most complete series of samples from an animal population that are available are those taken from wild mixed insect populations by means of light traps.

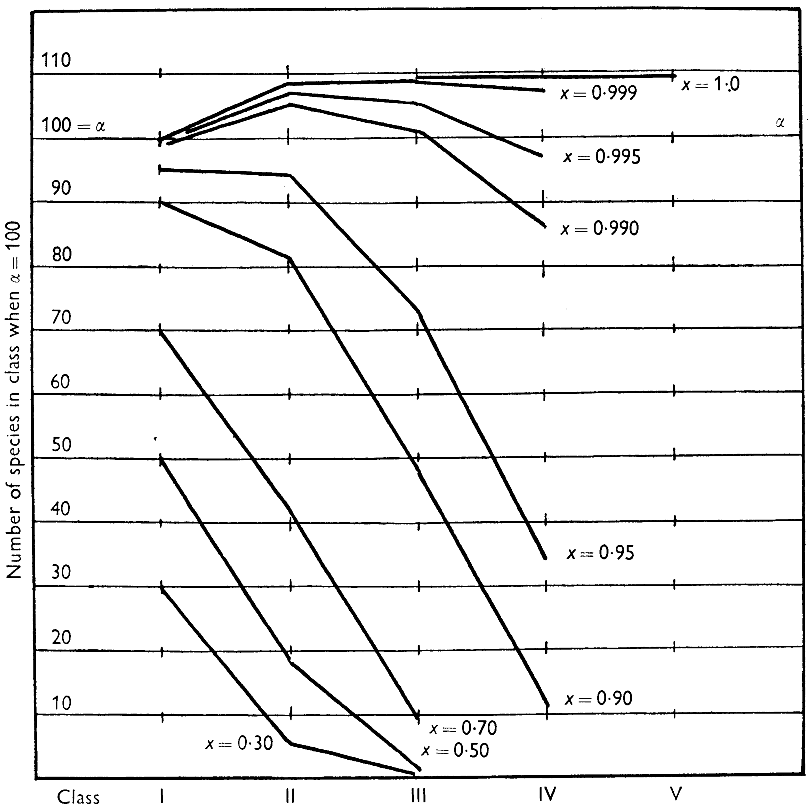

Fig. 2. The theoretical relation between the sizes of × 3 class groupings of species with different numbers of individuals, in samples of increasing size (and hence increasing values of x) taken from a population arranged in a logarithmic series with the index of diversity = 100. Data in Table 2.

In the course of 8 years trapping by this method at Rothamsted Experimental Station, Harpenden (about 25 miles north of London) nearly 100,000 Macro-lepidoptera were captured and identified. A total of nearly 350 species were represented from the families Noctuidae (Agrotidae), Geometridae, Sphingidae, Bombycidae, Arctiidae, Lithosiidae and a few other related families. Of this total nearly 33,000, representing 285 species, were taken in a single trap (A) which was in continuous use in one place for two complete periods of 4 years (Williams 1939, for first 4 years).

A summary of the data from this trap is shown in Table 3 and in Fig. 3. The table gives the number of individuals and of species, and the distribution of species in

2-2

Table 3. *Frequency distribution (in × 3 classes) of Macro-lepidoptera caught in a single light trap (trap A) at Rothamsted Experimental Station during 1933–37 and 1946–50*

Period	Individuals	Species	Average individuals per species	Number of species with different numbers of individuals (geometric classes × 3): I (1)	II (2–4)	III (5–13)	IV (14–40)	V (41–121)	VI (122–364)	VII (365–1093)	VIII (1094–3280)	IX (3281–9841)
				One-eighth of year (Av.)								
1933	440	89·6	4·9	**36·0**	27·3	19·3	7·0	0·1	—	—	—	—
1934	409	86·8	4·7	**34·3**	27·8	16·4	8·3	0·3	—	—	—	—
1935	855	107·1	8·0	**39·5**	34·4	21·1	9·0	2·0	0·9	0·3	—	—
1936	250	67·6	3·7	**33·6**	19·4	10·9	3·6	0·1	—	—	—	—
	Average of 4 years			**35·8**	27·2	16·8	7·0	0·6	0·2	0·1	—	—
				Whole year								
1933	3,454	173	20·0	32	**54**	36	28	18	5	—	—	—
1934	3,276	168	19·5	33	**42**	39	35	12	7	—	—	—
1935	6,530	191	34·2	37	**47**	45	36	18	6	1	1	—
1936	1,961	154	12·7	**54**	34	35	21	8	2	—	—	—
	Average of 4 years			39·0	**44·3**	38·8	30·0	14·0	5·0	0·3	0·3	—
1946	3,124	156	20·0	39	**43**	40	17	14	1	2	—	—
1947	3,786	184	20·6	45	**53**	36	28	17	4	1	—	—
1948	3,771	187	20·2	47	**54**	43	22	18	1	2	—	—
1949	6,107	191	32·0	48	**49**	**49**	23	16	3	2	1	—
	Average of 4 years			44·7	**49·8**	42·0	22·5	16·3	2·3	1·8	0·3	—
	Average of 8 years			41·9	**47·0**	40·4	26·3	15·2	3·6	1·0	0·3	—
				Series of years								
1933–36 (4 years)	15,221	234	65·0	34	41	**47**	46	34	24	7	1	—
1946–49 (4 years)	16,972	254	66·8	45	43	**57**	52	29	21	5	1	1
	Average of two periods			39·5	42	**52**	49	31·5	22·5	6	1	0·5
All 8 years	32,853	285	115·3	38	41	48	**67**	43	28	14	5	1

geometric $\times 3$ classes for periods varying from $\frac{1}{8}$ of a year (i.e. every 8th day throughout the year) to the total catch for 8 years.

It will be seen that the peak number of species is in class I for the $\frac{1}{8}$ year (average number of individuals per species 4–8); and the peak is in class II for whole years

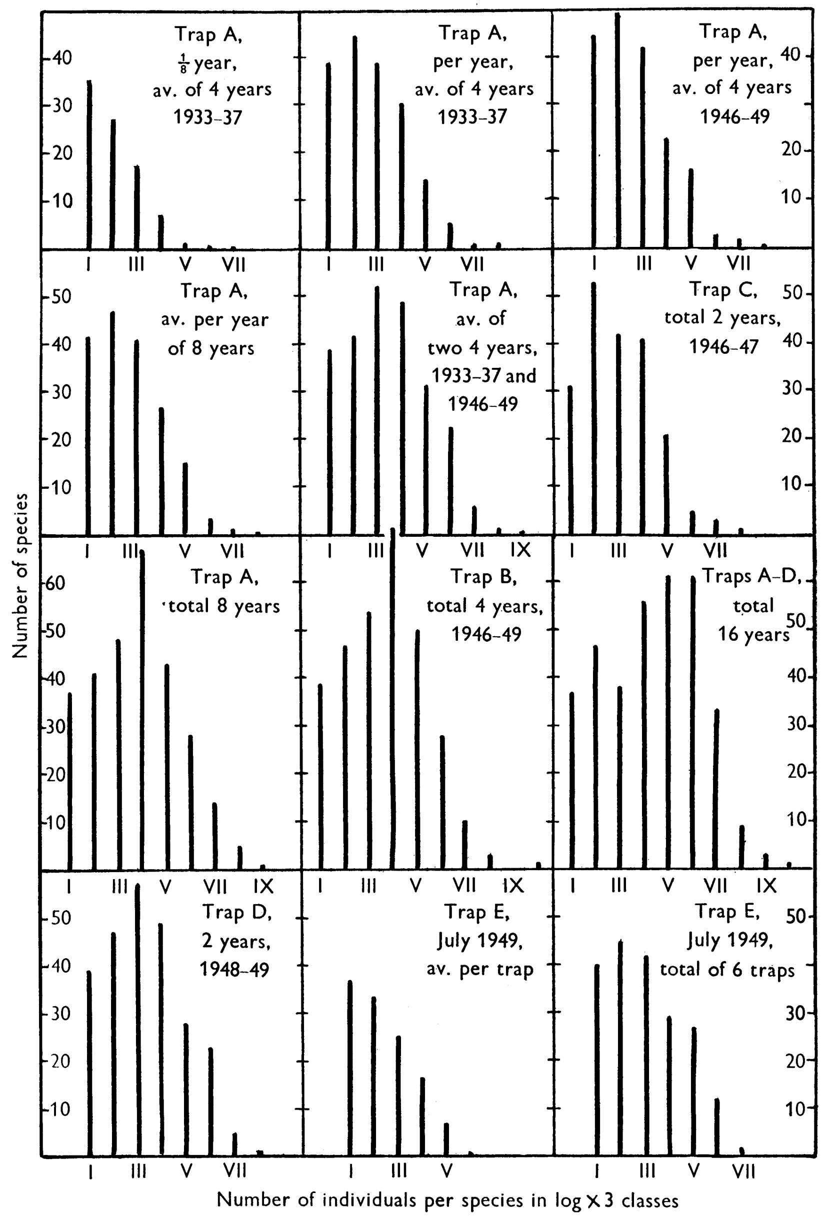

Fig. 3. Relation between number of species and numbers of individuals (the latter in $\times 3$ geometric classes) in catches of Macro-lepidoptera in light traps at Rothamsted. Original data in Tables 3 and 4.

(average individuals per species 19–34) except for the year 1936 when the peak was in class I and the average number of individuals per species only 12·7.

In the two 4-year periods, with an average number of individuals per species of 65–67, the peak is in class III and in the total of all 8 years, with an average of 115 individuals per species, the peak is in class IV. Although there is not very much difference between these results and the relation between peak, class and average individuals per species as shown in Table 2 for the logarithmic series, there is a slight tendency for the peak to be in a higher class than that expected. The height of the peak is, however, in the larger samples (4–8 years) distinctly more than 10% above class I, thus departing from this characteristic of the log-series. A study of the changes in numbers in class I on the contrary shows a slight rise (36–42) from the average for

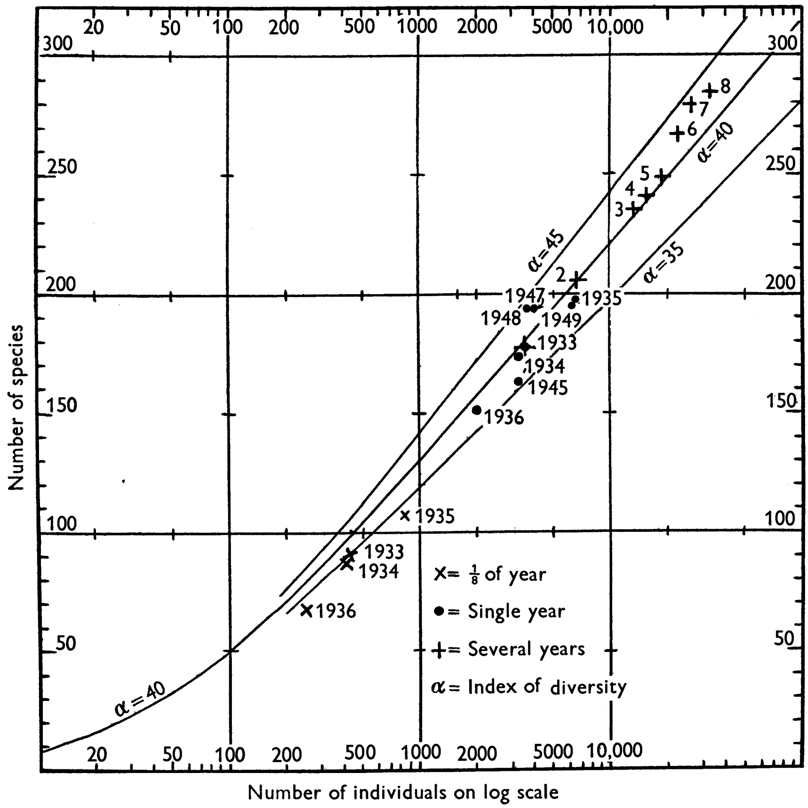

Fig. 4. The relation between increasing size of samples and increasing number of species in catches of Macro-lepidoptera in a light trap (A) at Rothamsted in periods varying from $\frac{1}{8}$ year to 8 years during 1933–36 and 1946–49.

$\frac{1}{8}$ year to the average for 1 year, but only a slight indication of a fall after this to 39·5 for 4 years and 38 for 8 years. The increase in number of species with increased sample size is shown graphically in Fig. 4, and it will be seen that so far from falling below the number expected by the log-series in higher values, there is actually an excess. In other words, the diversity of the sample on the basis of the log-series is increasing.

There is, however, a disturbing factor from the pure theory in the way the samples have been taken. The larger samples have been taken over a longer period than the smaller (except in the case of the $\frac{1}{8}$-year samples) and are thus not strictly comparable. There is no doubt that this accounts for the steady increase in diversity, as the population sampled during 8 years by one trap will not give the same result as a trap catching eight times as many insects in one year. This increase in diversity will also affect the changing values of n_1 (or class I) as already discussed above; a slightly higher class I would be expected in the larger samples, and so the small fall which has been found is more significant than it would otherwise be.

In Fig. 5 the numbers of species in the different $\times 3$ classes in trap A for different size samples have been shown one on top of the other with the position of class I changed (logarithmically) to allow for the increased size of samples as shown on a horizontal scale below the diagram. For example, if the sample size were increased three times, the position of class I would move one class interval to the left. It will be

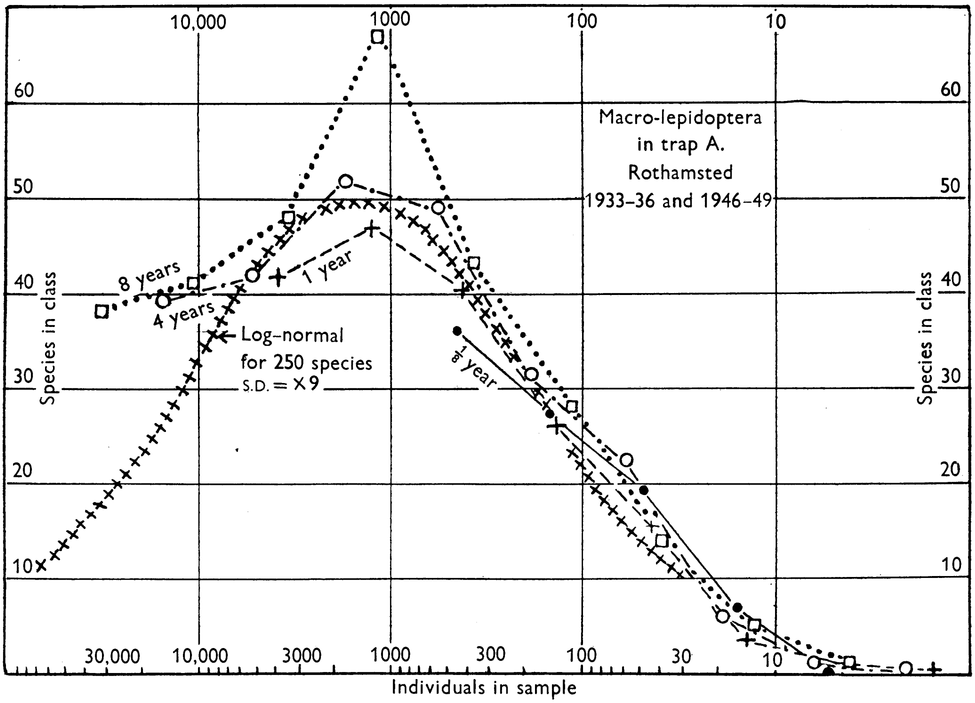

Fig. 5. Frequency distribution of species with different numbers of individuals (in $\times 3$ geometric classes) of Macro-lepidoptera caught in a light trap (A) at Rothamsted in different periods from $\frac{1}{8}$ year to 8 years. The resulting distributions arranged so that the position of class I varies according to the size of the sample on a horizontal scale reading from right to left. The superimposing of the curves illustrates Preston's theory of sampling from a log-normal distribution.

seen that the whole pattern tends to form a continuous curve, resembling Preston's suggestion that with a log-normal population samples of different sizes are truncate log-normals represented by moving the zero line on the original population curve. Grundy's work shows that this is approximately correct except for very small samples.

Superimposed on these curves is a log-normal calculated to fit 250 species with a standard deviation equal to two $\times 3$ groups (i.e. $\times 9$). It will be seen how closely the observed results fit to this log-normal distribution, with the exception of the total of 8 years. It is questionable whether in this case the two periods of 4 years, separated by force of circumstance by a gap of 9 years, should really be combined into a single sample. The figure is not meant to demonstrate mathematical identity, but to show a trend in observed biological data which is approximately paralleled in the theory of sampling from a log-normal population.

Table 4. *Frequency distribution of abundance of Macro-lepidoptera in various light traps at Rothamsted Experimental Station*

Period	Individuals	Species	Average individuals per species	Number of species with different numbers of individuals. Geometric classes (× 3): I (1)	II (2–4)	III (5–13)	IV (14–40)	V (41–121)	VI (122–364)	VII (365–1093)	VIII (1094–3280)	IX (3281–9841)	X (9842–29,542)
					Trap B								
4 years 1946–49	35,428	304	116·5	39	47	54	**72**	50	28	10	3	0	1
					Trap C								
2 years 1946–47	5,972	197	30·3	31	**53**	42	41	21	5	3	1	—	—
					Trap D								
2 years 1948–49	12,456	249	50·0	39	47	**57**	49	28	23	5	1	—	—
					All four traps all years								
16 Trap years	87,400	346	252·6	37	47	38	56	**61**	**61**	33	9	3	1
					Traps E 1–6. Month of July 1949 only, in woodland								
Average per trap	1,230	119·7	10·3	**36·8**	33·7	25·8	16·2	7·0	0·2	—	—	—	—
Total of six traps	7,378	197	37·5	40	**45**	42	29	27	12	2	—	—	—
					Trap R (mercury vapour), August 1950								
11 days in garden	3,349	106	31·6	24	**31**	24	14	7	2	3	—	—	—
15 days in woodland	5,232	84	62·3	**24**	19	**23**	11	5	1	—	—	1	—

To overcome the difficulty of comparing samples taken over different periods a trapping experiment was made during the month of July 1949, in which six traps were working simultaneously in a small woodland at Harpenden (traps E 1–6). The results showing the average per trap and the total of all six traps are shown in Table 4 and Fig. 3. The average number of individuals per species was 10·3 in the single trap, and the peak in class I, while in all six traps together the peak was in class II with an average of 37·5 individuals per species. The log-series (see Table 2) would require a peak between classes I and II in the first case and between II and III in the second, so that these figures do not clash with the observed. The total of all six traps is still too small to provide a strict criterion.

During the period 1946–50 several other traps (B, C and D) were run in different localities in the same neighbourhood as shown in Table 4 and Fig. 3. Trap B over 4 years caught 35,000 moths with an average of 116·5 individuals per species and a peak in class IV (above that expected by the log-series).

Table 5. *Macro-lepidoptera caught in mercury-vapour light trap* (*G*) *at Rothamsted during the nine months March to November 1951*

Class	Individuals per species	NOCTUIDAE	GEOMETRIDAE	Other families	Total
I	1	14	12	5	31
II	2–4	18	13	**14**	45
III	5–12	25	21	9	55
IV	13–40	30	**22**	**13**	**65**
V	41–121	**32**	5	6	43
VI	122–364	12	1	4	17
VII	365–1,093	4	—	—	4
VIII	1,094–3,280	4	—	—	4
IX	3,281–9,841	1	—	—	1
Total species		140	74	57	265
Total individuals		23,425	1,258	1,617	26,300
Average individuals per species		167·3	17·0	13·7	99·3

The total of all four traps in a total of sixteen trap-years caught 87,000 moths with an average of 253 individuals per species and a peak between classes V and VI. Too much emphasis must not, however, be laid on this grand total owing to the long period of time and the widely different conditions under which the figures were obtained.

Both in trap B (4 years) and in the total of all traps the increase of the peak above the value of class I is far beyond the limit of 10% required by the log-series. The rise is from 39 to 72 in trap B and from 37 to 61 in all four traps.

In order to get a still larger sample in a shorter time two mercury vapour ultraviolet light traps were started in March 1951 (after very short trials in 1950) and were run continuously during 1951 under the supervision of Mr M. Hosni; one (F) inside the woodland already referred to and one (G) just outside in a locality where no trapping had previously been done. At the end of 9 months (March to November) in the trap outside the wood 26,300 Macro-lepidoptera had been captured belonging to 265 species. This is a much greater number than had ever before been captured in so short a time, and is indeed 60% more individuals and 25 more species than were caught in the whole of the first 4 years trapping with trap A. The results on a $\times 3$ geometric classification are shown in Table 5 and Fig. 6.

These results support again the idea that large samples depart from the logarithmic series in the direction of the log-normal, as in each family subdivision, and in the total, the peak class has a value about double that of class I, but even with this large catch the distribution is not symmetrical about the peak.

Apart from light-trap catches, very large samples of insect populations which are randomized for abundance, and in which all or very nearly all of the species have been identified, are not often recorded; but some examples are given by Palmèn (1944) of great masses of Coleoptera washed in from the sea and piled up in long rows on the southern shores of Finland. They cannot be strictly said to have been taken from one particular association, and we do not know for certain if they have been washed out to sea by flooded rivers or have been brought down on the surface of the ocean by heavy rains or some similar cause. Over 970 species of beetles were identified by Palmèn in nine such aggregations. For each species in a large random sample from each of the aggregations he gives the actual number of individuals if it is below 50;

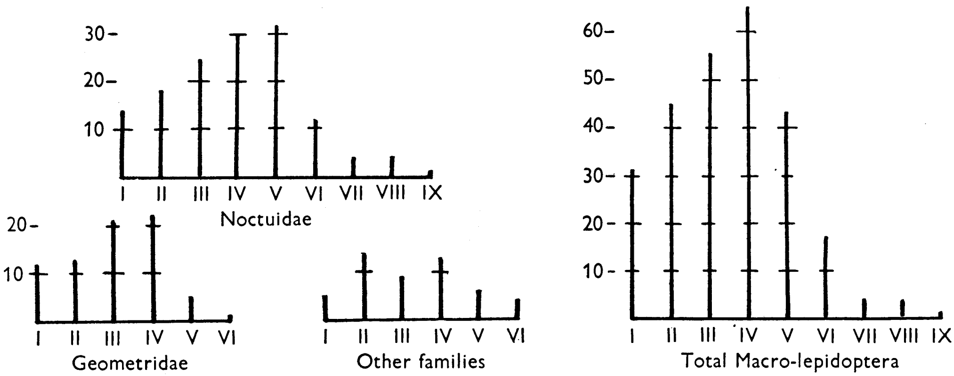

Fig. 6. Distribution of numbers of species with different numbers of individuals (the latter in $\times 3$ geometric classes) as shown by captures of Macro-lepidoptera in a light trap (G) at Rothamsted from March to November 1951. Based on 26,300 individuals belonging to 265 species. Data in Table 5.

higher numbers are grouped into over 50, over 100, over 200, over 300, over 500, over 1000, and 'infinite'. It is thus not possible to get the total number of individuals in a sample or the average number of individuals per species, as by far the greater proportion of individuals in any population is made up from the few most abundant species. The structure of two samples, with the actual number of species with up to 50 individuals, is given in Table 6 and Fig. 7.

The first sample (Palmèn's Pa 1) was taken on 12–13 June 1939, and contained 393 different species. It will be seen that in this sample there is not a true 'hollow curve' as the number of species with two individuals is greater than the number with one. This is the only random sample of insects that I have so far traced in which this condition is found. The frequency distribution on a $\times$ 3 scale shows a peak in class II (2–4 individuals), and might easily be a portion of a log-normal distribution. Class II is more than 100% above class I and so the distribution is definitely not of the logarithmic series type.

The second sample (Palmèn's Pa 3) was taken from an aggregation on 4–5 July 1939, and contained 466 species. In this the number of species represented by one individual is larger than that by 2, but not as different as is required by the logarithmic series, in which the number with 2 must always be less than half the number with one. The classification into $\times 3$ logarithmic classes gives a maximum in class II, as before, with class II over 60% above class I, and a close resemblance to a log-normal distribution.

Table 6. *Numbers of species of Coleoptera with different numbers of individuals found by Palmèn (1944) in drifts along the sea-shore in Finland*

(The figures in parentheses are sub-totals for the geometric $\times 3$ classes.)

SAMPLE I

Individuals per species	Species	Individuals per species	Species	Individuals per species	Species
1	72 (72)	13	5 (90)	28	1
2	91	14	4	29	1
3	47	15	0	39	1 (29)
4	23 (161)	16	6	50	2
5	30	17	1	> 50	14
6	19	18	3	> 100	5
7	8	19	3	> 200	2
8	12	20	3	> 300	1
9	3	21	2	> 500	8
10	2	22	1	'∞'	9
11	4	25	1	Total species	393
12	7	26	2		

SAMPLE II

Individuals per species	Species	Individuals per species	Species	Individuals per species	Species
1	111 (111)	15	5	32	1
2	87	16	4	33	1
3	59	17	5	34	1
4	35 (181)	18	1	36	2 (37)
5	29	19	1	49	1
6	20	20	0	50	1
7	6	21	2	> 50	11
8	10	22	1	> 100	9
9	17	23	1	> 200	4
10	3	24	1	> 500	1
11	3	25	1	> 1000	1
12	7	29	2	'∞'	10
13	4 (99)	30	1	Total species	466
14	5	31	2		

Since the general principles of the structure of populations are not likely to be confined to any one group of animals, it is interesting to see to what extent the results of surveys and censuses made by ornithologists will support or extend those based on insect populations. Table 7 shows two sets of data on bird population from the United States; in each case the original numbers have been reclassified on a $\times 3$ basis. The figures in both these cases are of breeding pairs. The average number of pairs per species is 4·2 in the first with a peak in class I, and 180 in the second with a rather flat

Table 7. *Data for bird populations, (1) from Stewart & Aldrich (1949), (2) from Saunders (1936)*

	Total species	Total pairs	×3 class I	II	III	IV	V	VI	VII	VIII
(1) Appalachian Mountains	24	100	**10**	6	6	2	—	—	—	—
(2) Quaker Run Valley	79	14,348	2	9	12	**15**	**14**	**16**	8	3

peak ranging from classes IV to VI. The small Appalachian sample is close to a logarithmic series, as one calculated to fit 100 pairs in 24 species would give the first term 9 and the second 4·05 as compared with observed values 10 and 4. The position

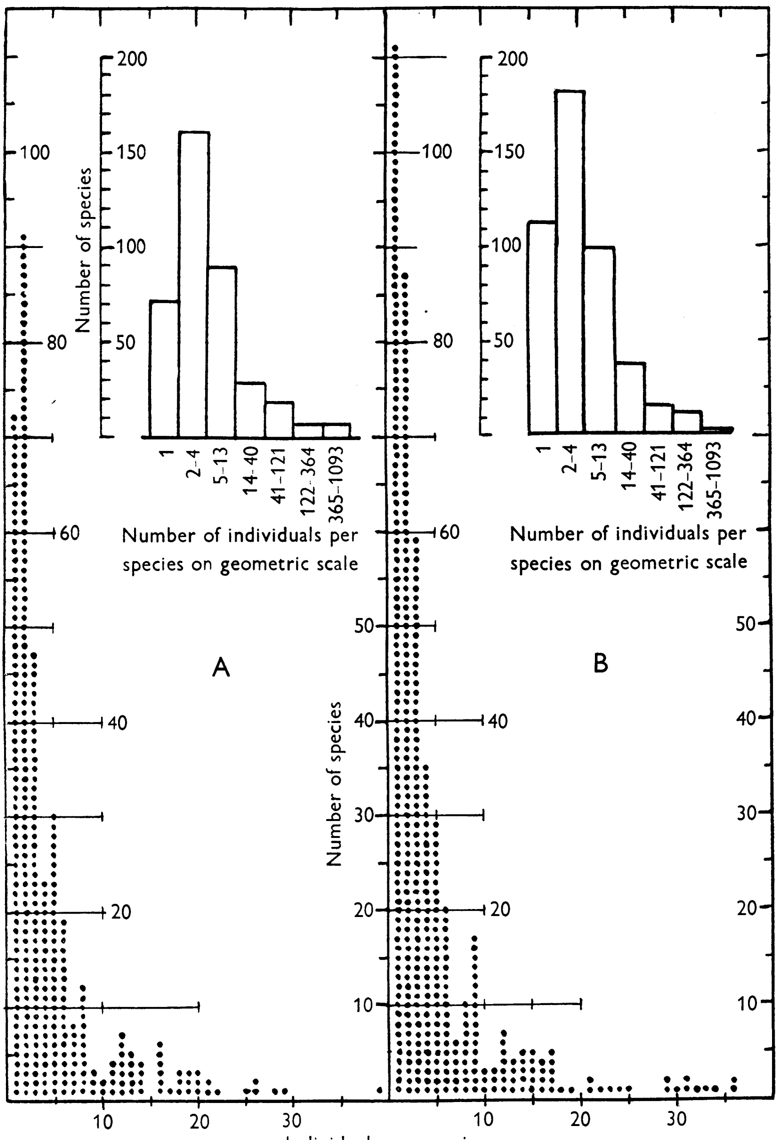

Fig. 7. Distribution of numbers of species of Coleoptera with different numbers of individuals in two samples taken by Palmèn from great drifts of insects found on the sea shore in Finland in 1939. The number of individuals is shown both on an arithmetic scale (up to 40) and on a geometric $\times 3$ scale (up to 1093 individuals per species). Data in Table 6.

of the peak for the log-series has not been calculated for such high average numbers of units per group as that found in the second sample, but the height of the peak is eight times as large as class I, which is impossible for the log-series, but quite possible for a log-normal distribution.

To get a still greater 'sample' I asked Dr James Fisher if he could group all the known nesting land-birds in England and Wales in categories of abundance in powers of 10, i.e. whether the average number of individuals present in a breeding season

Table 8. *Estimate by Dr James Fisher of number of species of nesting land-birds in England and Wales with different levels of abundance*

Class	Number of individuals per species	Number of species
I	1–10	7
II	10–100	9
III	100–1000	22
IV	1000–10,000	42
V	10,000–100,000	32
VI	100,100–1 million	16
VII	1–10 million	12
VIII	over 10 million	2*
	Total	143

* Just above 10 million.

was 1–10, 10–100, 100–1000, etc. He has done this for me and the details have been published (Fisher 1952). A summary of the number of species in each group is given in Table 8.

The results, shown graphically in Fig. 8, give a very close approximation to a log-normal distribution with a mean near the top of class IV at about 8000 individuals per species, and a standard deviation of $10^{1\cdot4}$. It bears no resemblance to a logarithmic

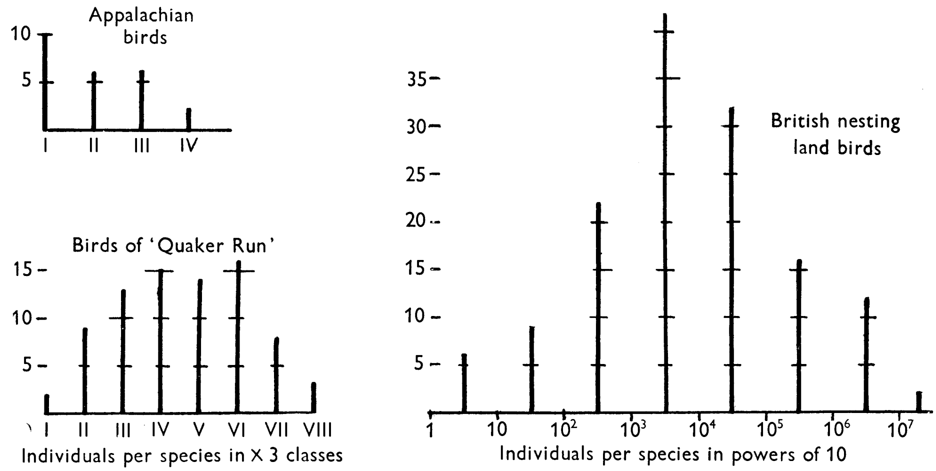

Fig. 8. Distribution of numbers of species of birds with different numbers of individuals in two samples of American birds (with the number of individuals on a $\times 3$ geometric scale), and in an estimate of the total population of 144 nesting landbirds in England and birds in England and Wales, with the numbers of individuals on a $\times 10$ geometric scale. For data see Tables 7 and 8.

series. The total number of birds included in the census is about 63 million, giving an arithmetic mean of about 440,000 individuals per species, with a peak on the $\times 3$ classification near the top of class IX (3281–9840). Fig. 9 shows the same data on 'probability' paper showing the very close relation of the observed figures to the straight line required by the log-normal distribution, with a mean at 8000, and a standard deviation of $10^{1\cdot4}$.

SUMMARY AND DISCUSSION

Thus we see that a survey of large samples taken from mixed animal populations tends to support the idea that the distribution of species with different numbers of individuals is nearer to the log-normal than to the log-series form. This suggestion was first made by Preston (1948), and has also been independently suggested to me by Mr W. S. Volkers, a student of statistical biology at the University of Utrecht in Holland.

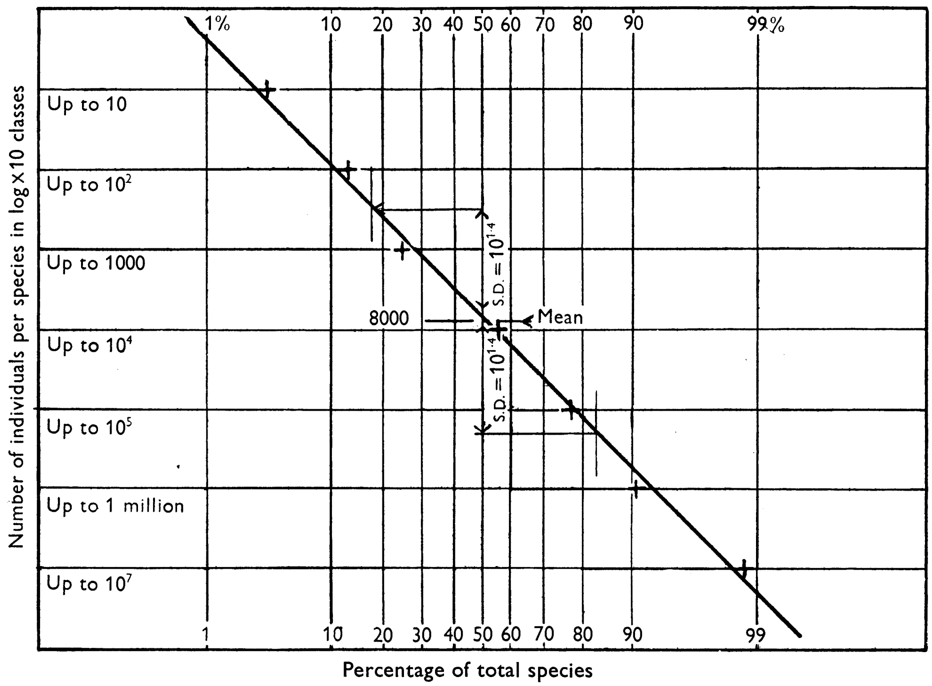

Fig. 9. James Fisher's data (Table 8) on the relative abundance of British nesting land-birds plotted on log-probability paper, so as to show the relation between the accumulated percentage total of species up to each level of abundance. The recorded points are very close to the straight line which shows a mean at $10^{3\cdot9}$ (=approx. 8000) individuals per species, and standard deviation of $10^{\pm1\cdot4}$.

The mathematical problems relating to the question of sampling from a log-normal population, and of fitting a log-normal distribution to data, have been recently taken up by Mr P. Grundy of the Statistical Department at Rothamsted. The problem is made more complex by the fact that most biological data are incomplete, as it is not possible to say how many species present in the population were not represented in the sample. His work is in the process of publication (Grundy 1951).

The main differences in conception, apart from the increased mathematical complexity, between the log-series and the log-normal are:

1. The former postulates no limit to the numbers either of species or of individuals in the population from which the samples are taken, while the log-normal implies a finite number of species, although the number of individuals is theoretically unlimited.

2. Three parameters are necessary to define a log-normal, whereas two are sufficient for a log-series. As a result to the total number of individuals and of species in the

sample, which is sufficient to define a log-series, must be added a third constant fixing in some way the standard deviation of the curve, before a log-normal can be fitted to given data.

3. With increasing size of sample from a log-series population, the relation between the number of species and the log-number of individuals tends to a straight line (Fig. 4), but with a log-normal population the curve is sigmoid and flattens out as it approaches the limit of the total number of species in the population which is being sampled.

4. There is with the log-normal a decline in the number of species with one individual (n_1, or class I) after the sample is large enough to include about 50% of the species in the population, which does not occur if the distribution is of the log-series form.

5. There is in the log-normal distributed population the gradually increasing difference between the number of species with one individual (n_1 or class I), and the number in the peak class as sample size increases. This difference is limited to 10% in a log-series population.

It is evident that many more examples of large random samples of animal populations, with accurate counts and determination of species, are required before the general form or forms of distribution of the species, and hence the pattern or balance of the populations, can be determined.

REFERENCES

Fisher, J. (1952). Bird numbers; a discussion on the breeding population of inland birds of England and Wales. *S. East. Nat.* **57**, 1–10.

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.

Grundy, P. (1951). The expected frequencies in a sample of an animal population in which the abundances of species are log-normally distributed. Part I. *Biometrika*, **38**, 427–34.

Palmèn, E. (1944). Die anemohydrochore Ausbreitung der Insekten als zoogeographischer Faktor. *Ann. Soc. Zool.-Bot. Vanamo*, **10** (1), 1–262.

Preston, F. W. (1948). The commonness, and rarity, of species. *Ecology*, **29**, 254–83.

Saunders, A. A. (1936). Ecology of the birds of Quaker Run Valley, Alleghany State Park, New York. *Handb. N.Y. St. Mus.* **16**, 123–4.

Stewart, R. E. & Aldrich, J. W. (1949). Breeding bird populations in the spruce region of the Central Appalachians. *Ecology*, **30**, 75–82.

Williams, C. B. (1939). An analysis of four years captures of insects in a light trap. Part I. General survey; sex proportion; phenology; and time of flight. *Trans. R. Ent. Soc. Lond.* **89**, 79–131.

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

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