Journal of Micropalaeontology, **23:** 61–65.

On confidence intervals from micropalaeontological counts

STEFAN A. REVETS
Koninklijk Belgisch Instituut voor Natuurwetenschappen, Departement Paleontologie, Vautierstraat 29, B-1000, Brussel, Belgium
(e-mail: s.revets@tiscali.be).

ABSTRACT – In micropalaeontology, statistics resulting from counts are rarely reported with error margins or confidence intervals. A short discussion of taxon proportions and of Fisher's α diversity index is augmented with tables and graphs to assist the researcher in calculating confidence intervals. Any results, discussions or further calculations based on counts can then be put on a more secure statistical footing. *J. Micropalaeontol*. **23**(1): 61–65, May 2004.

INTRODUCTION

Quantitative studies in micropalaeontology continue to develop, with new and powerful multivariate statistical methods being tried and adopted. However, confidence intervals are rarely, if ever, calculated, let alone referred to. All too often, numbers from counts are presented as being both accurate and precise and used as such in more complex methods. This omission is detrimental as it makes it very difficult to gauge significance and reliability of any outcome of such statistical operations.

Although the statistics of sampling are very well established, tables and graphs of immediate use to the micropalaeontologist are not as readily available as one might wish. This article deals only with the numbers obtained from the picking tray or slide traverse. The problems of dealing adequately with the sampling of a population as it occurs out there in the world are different from the problems of correctly measuring and representing the taxa present in the physical samples taken. Therefore, the term 'population' is used here in the statistical sense, referring to the specimens in the collected rock sample or in the (washed) residue, not to the biological population out in the field.

It is relevant and important to have a reliable and correct measure of the precision with which one counts and calculates taxon presence and proportions. When working with extant organisms and their heterogeneous distribution, the measure of 'laboratory precision' has a direct bearing on the nature and, particularly, the intensity of any sampling program in the field. In marked contrast, many, if not most, micropalaeontologists work with fossil material from outcrop or drillcore. They simply cannot afford the luxury of dealing with such ecological considerations and issues of replicate sampling. Nevertheless, palaeoecological interpretations have to be made from whatever material could be recovered. Such interpretations necessarily rely on the abundances of any taxa encountered and statistically justified confidence intervals are, therefore, important. Anyone confronted with such problems will hopefully benefit from the following discussion, tables and graphs.

CONFIDENCE INTERVALS ON PROPORTIONS

A general problem which has greatly exercised the minds of researchers and statisticians alike is the minimal size of a sample required to answer a particular question. The micropalaeontologist, in particular, is confronted with the question of how many, or few, specimens to pick from a rock sample or washed residue and still retain confidence in the detection level and possibly proportions calculated of the taxa in the sample. Phleger (1960) applied the results obtained by Dryden (1931) and proposed the now commonly used 300 specimens as a sufficient and practical number with which to determine relative abundances of foraminiferal species. In a different context, Shaw (1964) also used the Binomial Distribution and arrived at the same kind of answers. Dennison & Hay (1967) applied Shaw's analysis to the problem of the size of sample area and provided a very useful graph (calculated and refigured here in Fig. 1). They also discussed the wider applicability of their results, including amongst others Dryden's and Phleger's work.

Their plot shows the number of specimens required if one is prepared to overlook with a probability P a species making up a particular proportion in the population. For example, the graph in Figure 1 shows that if one is prepared to miss out 1 in 20 (p = 0.05) species which make up a proportion of 0.01 of the population, then one needs to count 300 specimens.

What these analyses and associated graphs do not provide is a measure of the precision of these proportions. There is a substantial difference between finding a taxon making up a particular proportion in a sample and calculating that proportion. Many, if not most, taxa are rare (Fisher *et al*., 1943; Buzas *et al*., 1982) and may well play an important role in ecology, palaeoecology and any assessment or characterization of environments (Cao *et al*., 1998; Cao & Larsen, 2001). It is, therefore, important and relevant to be able to justify statistically any measurement of occurrence of such rare taxa.

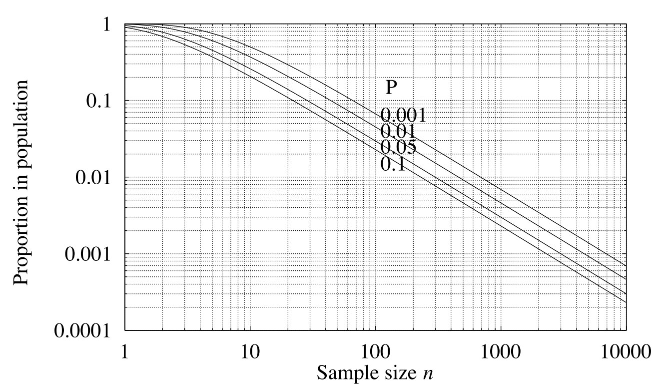

Fig. 1. Estimating sample size. For a number of probabilities of failing to find a taxon, the curves show how many specimens must be counted to detect a taxon present at a particular proportion in the population samples.

Finding specimens of rare species in a sample is a good example of a discontinuous stochastic process known as the Poisson Process – well described by the Poisson Distribution (Poisson, 1837; Student, 1907). There are benefits in deriving this distribution from discontinuous stochastic processes (e.g. Hald, 1952). A very important, but all too often overlooked, result is that the Poisson Distribution describes a process in its own right and should not be seen as a limiting case of the Binomial Distribution.

Thanks to the properties of the Poisson Distribution (some of which are discussed in the Appendix), confidence intervals on any observed number of occurrences can be derived and calculated. The value and the power of using a Poisson Distribution lies in the fact that it allows one to deal directly with the number of occurrences. This is in marked contrast to the Binomial Distribution which needs both the number of occurrences and the total number of events.

With the Poisson Distribution, it is a straightforward matter to derive the confidence limits (equations 4) and then calculate the confidence interval on the proportion of the species of interest. Rather than have the reader refer to χ^2 tables to calculate confidence intervals, Table 1 lists the lower and upper limits of the 0.99, 0.95 and 0.90 confidence intervals for number of specimens between 0 and 50.

An example will clarify the use of this table. Suppose that six specimens of a taxon are found in a picking tray. From Table 1 one reads that the 0.95 confidence interval for six specimens lies between 2.2 and 13.1. The proportion and the 0.95 confidence limits of this proportion are simply these numbers divided by the total number of specimens counted or picked from the tray. If 100 specimens were found in total, the species would make up 0.06 on average of the population and this proportion would lie anywhere between 0.02 and 0.13. Similarly, if 300 specimens had yielded the six specimens, the proportion would be 0.02 with the 0.95 confidence interval stretching from 0.007 to 0.044.

Inspection of the table shows the asymmetric nature of the confidence intervals, especially for the smaller values. This asymmetry is a clear indication of the non-Gaussian nature of the distribution and argues against the use of normal approximations (Garwood, 1936; Hald, 1952). With the table calculated here, there is no need to resort to approximations: the values presented are exact.

A related, but different, problem is posed if one wishes to know if two taxa make up different proportions in the population. How much of a difference in counts is necessary for one to conclude that the taxa make up different proportions? Or indeed, how much of a difference in counts may one allow for the taxa to occur in the same proportions?

Once again, thanks to the properties of the Poisson Distribution, these critical numbers can be calculated exactly. The calculations are straightforward but tedious (see equation 9), therefore, the results are listed here in Table 2. The table shows for counts from 0 to 50 of a taxon A how many specimens of a taxon B have to be found for B to be deemed more abundant than A, at the usual confidence levels of 0.90, 0.95 and 0.99.

Revisiting the previous example with six specimens of species A, how many specimens does one need of species B before it can be said to be more abundant (and make up a larger proportion of the population)? Looking under the entries for six in Table 2 shows that at least 16 specimens of species B are required (at a probability of 0.95). These numbers can then be translated to proportions by dividing them by the total number of specimens counted.

If, amongst the counts, there are 10 specimens of species A and 20 of species B, one looks up the entry for 10 to find that at least 21 specimens are needed for B to be more abundant (at 0.95

n	0.99		0.95		0.90	
0	0.0000	5.2983	0.0000	3.6888	0.0000	2.9957
1	0.0050	7.4301	0.0253	5.5716	0.0512	4.7438
2	0.1034	9.2737	0.2422	7.2246	0.3553	6.2957
3	0.3378	10.9774	0.6186	8.7672	0.8176	7.7536
4	0.6722	12.5940	1.0898	10.2415	1.3663	9.1535
5	1.0779	14.1497	1.6234	11.6683	1.9701	10.5130
6	1.5369	15.6596	2.2018	13.0594	2.6130	11.8424
7	2.0373	17.1335	2.8143	14.4226	3.2853	13.1481
8	2.5711	18.5782	3.4538	15.7631	3.9808	14.4346
9	3.1324	19.9984	4.1153	17.0848	4.6952	15.7052
10	3.7169	21.3978	4.7953	18.3903	5.4254	16.9622
11	4.3213	22.7792	5.4911	19.6820	6.1690	18.2075
12	4.9431	24.1449	6.2005	20.9615	6.9242	19.4425
13	5.5801	25.4966	6.9219	22.2304	7.6895	20.6685
14	6.2306	26.8359	7.6539	23.4896	8.4639	21.8864
15	6.8933	28.1640	8.3953	24.7402	9.2463	23.0971
16	7.5670	29.4819	9.1453	25.9830	10.0359	24.3011
17	8.2506	30.7905	9.9031	27.2186	10.8321	25.4992
18	8.9433	32.0907	10.6679	28.4477	11.6343	26.6917
19	9.6444	33.3829	11.4392	29.6708	12.4419	27.8792
20	10.3532	34.6680	12.2165	30.8883	13.2546	29.0620
21	11.0692	35.9462	12.9993	32.1007	14.0720	30.2404
22	11.7918	37.2182	13.7872	33.3082	14.8937	31.4148
23	12.5206	38.4843	14.5800	34.5112	15.7195	32.5853
24	13.2553	39.7449	15.3772	35.7101	16.5490	33.7524
25	13.9953	41.0004	16.1786	36.9049	17.3821	34.9160
26	14.7405	42.2509	16.9840	38.0960	18.2185	36.0766
27	15.4906	43.4968	17.7931	39.2835	19.0581	37.2341
28	16.2452	44.7384	18.6058	40.4678	19.9006	38.3889
29	17.0041	45.9758	19.4217	41.6488	20.7459	39.5409
30	17.7672	47.2093	20.2408	42.8268	21.5939	40.6905
31	18.5342	48.4390	21.0630	44.0020	22.4445	41.8376
32	19.3048	49.6652	21.8879	45.1744	23.2974	42.9824
33	20.0791	50.8879	22.7156	46.3442	24.1526	44.1250
34	20.8567	52.1074	23.5459	47.5115	25.0101	45.2656
35	21.6375	53.3238	24.3787	48.6765	25.8696	46.4041
36	22.4215	54.5371	25.2139	49.8391	26.7311	47.5407
37	23.2084	55.7476	26.0514	50.9996	27.5946	48.6754
38	23.9982	56.9554	26.8910	52.1579	28.4599	49.8084
39	24.7908	58.1605	27.7328	53.3142	29.3269	50.9397
40	25.5859	59.3630	28.5765	54.4686	30.1957	52.0693
41	26.3836	60.5631	29.4223	55.6211	31.0661	53.1974
42	27.1838	61.7608	30.2699	56.7718	31.9381	54.3239
43	27.9863	62.9562	31.1193	57.9207	32.8116	55.4490
44	28.7911	64.1494	31.9704	59.0679	33.6866	56.5726
45	29.5981	65.3405	32.8233	60.2135	34.5630	57.6948
46	30.4072	66.5295	33.6777	61.3575	35.4407	58.8158
47	31.2184	67.7165	34.5338	62.5000	36.3198	59.9354
48	32.0316	68.9015	35.3914	63.6410	37.2002	61.0538
49	32.8467	70.0847	36.2504	64.7806	38.0819	62.1710
50	33.6637	71.2660	37.1109	65.9187	38.9647	63.2870

Table 1. Confidence intervals for the Poisson Distribution. The list shows for an observed number of occurrences *n* the lower and upper limits at probabilities 0.99, 0.95 and 0.90.

n	0.99	0.95	0.90
0	8	5	4
1	11	7	6
2	12	9	7
3	14	10	9
4	16	12	10
5	17	13	12
6	19	15	13
7	21	16	14
8	22	18	16
9	24	19	17
10	25	21	18
11	27	22	20
12	28	23	21
13	30	25	22
14	31	26	23
15	32	27	25
16	34	28	26
17	35	30	27
18	37	31	28
19	38	32	29
20	39	33	31
21	41	35	32
22	42	36	33
23	43	37	34
24	45	38	35
25	46	40	37
26	47	41	38
27	49	42	39
28	50	43	40
29	51	45	41
30	53	46	42
31	54	47	44
32	55	48	45
33	56	49	46
34	58	51	47
35	59	52	48
36	60	53	49
37	62	54	51
38	63	55	52
39	64	57	53
40	66	58	54
41	67	59	55
42	68	60	56
43	69	61	57
44	70	62	58
45	72	64	60
46	73	65	61
47	74	66	62
48	75	67	63
49	76	68	64
50	77	69	65

Table 2. Critical values for the discrimination of two Poisson-distributed samples. Given a number of specimens *n* for one species, the list shows the minimal number required for another species to occur with a higher abundance, at confidence probabilities of 0.99, 0.95 and 0.90.

probability): therefore, one has to conclude that both taxa occur with the same abundance.

From Table 2 it can also be seen that a count of 300 specimens is just sufficient to differentiate at the 0.95 confidence level between a 0.01 and 0.04 proportion in a population: a count of three differs significantly from 12 (at 0.95 confidence level), which, divided by 300, translates into differentiating a 1% from a 4% proportion.

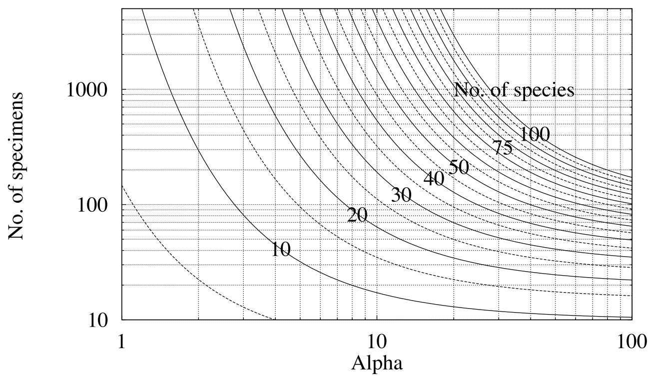

Fig. 2. Fisher's α diversity index. From the number of specimens laid out on the *y*-axis the α index can be read off on the *x*-axis using the curve with the number of species found. The successive curves increase with a step of five species from 5 to 100.

CONFIDENCE INTERVALS ON FISHER'S ALPHA INDEX

This diversity index was proposed by Fisher as a natural extension of the Poisson Distribution (Fisher *et al.*, 1943). In a great many cases, the distribution of species in function of the numbers of specimens found can be described as a negative binomial distribution, i.e. many taxa are represented by only a few specimens while only a very few occur in great numbers. The *α* index characterizes with a single number the shape of this distribution: it gives the number of species represented by a single specimen in a sample, whatever its size. Unfortunately, the index is an awkward function of two variables – number of specimens and number of species – and direct calculation is not possible. Various kinds of graphs can be made up to show the relation: Figure 2 is such a graph, which differs in portrayal from the one used most often in micropalaeontological circles since its introduction and excellent discussion by Murray (1968, 1973).

The nature of the index is counterintuitive: it seems odd that the number of taxa represented by a single specimen should remain constant, regardless of the number of specimens collected. In practice, an increase in the value of the index is often encountered. In the original article by Fisher *et al.*, Williams showed increases in *α* (from 31.38 to 40.24) when continuing to collect lepidoptera over a four-year period – and that with specimen numbers running up to 15 000. Murray (1968) also found an increase in *α* when going on picking foraminifera from 100 to about 600 specimens. Nevertheless, he concluded that the variation he had come across was not sufficiently large to cause concern.

A rerun of Murray's counting experiment on two different samples failed to reproduce this increase. The samples used were the Sands of Wemmel, Belgium (Eocene, Lutetian) and from Antikephalina Bay, Paros, Greece (Recent). The samples had been washed on a 63 µm sieve and portions randomly scattered on a picking tray. Table 3 shows the results: these should allay any concern about the validity of the α index as a consistent measure of diversity.

Wemmel Sands, Belgium			Paros, Greece		
N	*S*	α	*N*	*S*	α
142	29	11.0 ± 3.0	192	45	18.6 ± 4.0
228	32	10.5 ± 2.5	344	52	16.5 ± 3.0
353	37	10.5 ± 2.0	409	58	18.5 ± 2.9
556	42	10.6 ± 1.9	502	62	18.6 ± 2.9
757	42	9.6 ± 1.5	620	70	20.2 ± 2.8

Table 3. Cumulative counts of specimens (*N*) and species (*S*) with the resulting Fisher's α index from a sample of the Eocene Sands of Wemmel and a sample of the Recent of Paros.

When Fisher proposed this index, he also investigated the effects of variation. His calculation of the variance of the index is of particular interest. As one would expect from the functions defining the index, the formula for the variance is somewhat involved and its relation to the parameters making it up – number of species and number of specimens – is far from transparent. However, by adopting the coefficient of variation (standard deviation divided by the mean), a highly informative graph can be drawn by plotting the coefficient of variation against the number of specimens used for a selection of values of *α*, as shown in Figure 3.

It transpires from this graph that the coefficient of variation at counts of 300 specimens is about 0.1, or, that the standard deviation amounts to about one-tenth (0.092–0.145 to be exact) of *α*, and that for a very wide range of values of *α* (100–2). This means that the 95% confidence interval of *α* when calculated from 300 counted specimens, is about 20% of its value (i.e. 0.1 × 1.96 which is near enough 0.2, or 20%), a substantial amount. It takes roughly an order of magnitude more in specimens to be counted, about 3000, to reduce this confidence interval to plus or minus 10% of the value of *α*. These surprisingly large intervals should be borne in mind when attempting to compare and contrast samples.

MATHEMATICS APPENDIX

The Poisson Distribution

The Poisson Distribution was proposed by Poisson (1837) and independently by 'Student' (pseudonym of W. S. Gosset, 1907). It is a function defined by a single parameter *m*. This remarkable and valuable property contributes substantially to the power of this function in many of its applications. In contrast, all other distributions are functions determined by two or more parameters. The probability distribution is defined as

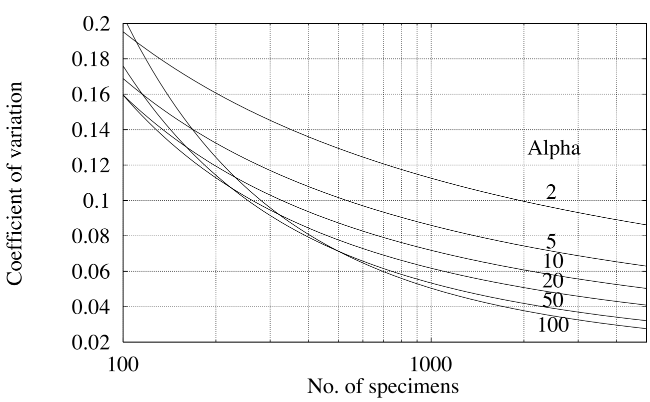

Fig. 3. Variation coefficient of Fisher's α diversity index. A plot of the coefficient of variation of α in function of the number of specimens used to calculate α, and that for a selection of values of α.

$$p(x) = e^{-m}\frac{m^x}{x!}, \qquad x = 0,1,2... \tag{1}$$

Confidence intervals. The cumulative distribution function can be shown to be related to the χ^2 distribution (Deming, 1950; Hald, 1952):

$$P(x) = 1 - P(\chi^2 < 2m), \qquad f = 2(x+1) \tag{2}$$

This relation gives the means with which to calculate confidence intervals for the parameter *m* of the Poisson Distribution:

$$P\{\underline{\mathrm{m}}(x_0) \le m \le \overline{\mathrm{m}}(x_0)\} = P_2 - P_1 \tag{3}$$

yielding

$$\begin{cases} \overline{\mathrm{m}} = \frac{1}{2}\chi^2_{(1-P_1)}, & f = 2(x_0+1) \\ \underline{\mathrm{m}} = \frac{1}{2}\chi^2_{(1-P_2)}, & f = 2x_0 \end{cases} \tag{4}$$

Since the χ^2 distribution is defined through a Γ-function, the values of the confidence intervals can be calculated directly. Table 1 lists the results of such calculations carried out with the Mathematica program.

Critical values. To find out if two observations come from two Poisson Distributions with the same or with different means relies on the addition theorem for the Poisson Distribution. This easily proved theorem states that if x_1, x_2, …, x_n are stochastically independent and Poisson-distributed with means m_1, m_2, …, m_n then the sum $x = x_1+x_2+...+x_n$ will be Poisson-distributed with mean $m = m_1+m_2+...+m_n$.

With the null hypothesis that the two means are equal, the probability of observing n_1, n_2 is simply

$$p(n_1+n_2) = e^{-2m}\frac{(2m)^{n_1+n_2}}{(n_1+n_2)!} \tag{5}$$

and, therefore, the conditional probability becomes

$$p(n_1|n_1+n_2) = \frac{p(n_1, n_2)}{p(n_1+n_2)} = \frac{(n_1+n_2)!}{n_1!n_2!}\left(\frac{1}{2}\right)^{n_1+n_2} \tag{6}$$

To test the hypothesis that $m_1 = m_2$ against the alternative of $m_1 > m_2$, one requires the probability

$$P(n \ge n_1|n_1+n_2) = \left(\frac{1}{2}\right)^{n_1+n_2}\sum_{n=n_1}^{n_1+n_2}\binom{n_1+n_2}{n} \tag{7}$$

$$= 1 - P\left(F < \frac{n_1}{n_2+1}\right), \qquad f_1 = 2(n_2+1),\ f_2 = 2n_1 \tag{8}$$

If this probability is less than or equal to the significance level α, then reject the null hypothesis that $m_1 = m_2$. Alternatively, the equation can be transformed to

$$\frac{n_1}{n_2 + 1} \geq F_{1-\alpha}(2(n_2 + 1), 2n_1) \tag{9}$$

which allows for the calculation from the F-Distribution of the critical values for which the means of the parent Poisson Distributions differ. The critical values at significance levels 0.99, 0.95 and 0.90 are listed for n_2 values between 0 and 50 in Table 2.

Fisher's α

Fisher derived the α index by extending the Poisson Distribution by considering the effects of the parameter (the mean) having a distribution of its own. He chose the Euler function (now usually known as the T-function), and arrived at the expression

$$f(n) = \frac{(k + n - 1)!}{n!(k - 1)!} \frac{p^n}{(1 + p)^{k+n}} \tag{10}$$

and setting the constant factor in the denominator $(k-1)! = \alpha$, $k = 0$ and $p/(p+1) = x$, simplified it to

$$\frac{\alpha}{n} x^n \tag{11}$$

Through summations, Fisher calculated the expected number of species S and the number of individuals N to obtain

$$S = -\alpha \ln(1 - x), \qquad N = \frac{\alpha x}{1 - x} \tag{12}$$

By eliminating x from these two formulae the equation becomes the well-known

$$S = \alpha \ln\left(1 + \frac{N}{\alpha}\right) \tag{13}$$

Fisher also derived a formula for the variance of α

$$s_\alpha^2 = \frac{\alpha^3 \left\{ (N + \alpha)^2 \ln\left(\frac{2N + \alpha}{N + \alpha}\right) - \alpha N \right\}}{(SN + S\alpha - \alpha N)^2} \tag{14}$$

The coefficient of variation is defined as the ratio between standard deviation and mean: applying this to the variance formula for α yields

$$c_\alpha = \frac{\sqrt{\alpha \left\{ (N + \alpha)^2 \ln\left(\frac{2N + \alpha}{N + \alpha}\right) - \alpha N \right\}}}{SN + S\alpha - \alpha N} \tag{15}$$

Manuscript received 7 April 2003
Manuscript accepted 4 December 2003

REFERENCES

Buzas, M.A., Koch, S.J., Culver, S.J. & Sohl, N.F. 1982. On the distribution of species occurrence. *Paleobiology,* **8**: 142–150.

Cao, Y. & Larsen, D.P. 2001. Rare species in multivariate analysis for bioassessment: some considerations. *Journal of the North American Benthological Society,* **20**: 144–153.

Cao, Y., Williams, D.D. & Williams, N.E. 1998. How important are rare species in aquatic community ecology and bioassessment? *Limnology and Oceanography,* **43**: 1403–1409.

Deming, W.E. 1950. *Some Theory of Sampling.* Dover Publications, New York.

Dennison, J.M. & Hay, W.W. 1967. Estimating the needed sampling area for subaquatic ecological studies. *Journal of Paleontology,* **41**: 706–708.

Dryden, A.L. 1931. Accuracy in percentage representations of heavy mineral frequencies. *Proceedings of the National Academy of Sciences,* **17**: 233–238.

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. *Journal of Animal Ecology,* **12**: 42–57.

Garwood, F. 1936. Fiducial limits for the Poisson Distribution. *Biometrika,* **28**: 437–442.

Hald, A. 1952. *Statistical Theory with Engineering Applications.* Wiley & Sons, New York.

Murray, J.W. 1968. Living foraminifers of lagoons and estuaries. *Micropaleontology,* **14**: 83–96.

Murray, J.W. 1973. *Distribution and ecology of living benthic foraminiferids.* Heinemann, London.

Phleger, F.B. 1960. *Ecology and distribution of Recent Foraminifera.* John Hopkins University Press, Baltimore.

Poisson, S.D. 1837. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, precedées des règles générales du calcul des probabilités,* Paris.

Shaw, A.E. 1964. *Time in Stratigraphy.* McGraw-Hill, New York.

Student 1907. On the error of counting with a haemacytometer. *Biometrika,* **5**: 351–360.