

**Translation of and commentary on George H. Pólya's (1930) "Eine
Wahrscheinlichkeitsaufgabe in der Pflanzensoziologie"**

by

Alexander M. Kerr

University of Guam Marine Laboratory Technical Report 172

February 2025



ACKNOWLEDGEMENTS

I have benefitted from discussion with my former students Christian (Keiki) Potter (Doctor of Medicine Program, University of Queensland) and Jordan Gault (Ph.D. Program, Institut für Chemie und Biologie des Meeres, Universität Oldenburg).

Dankulu na Saina Ma'åse!

SUMMARY

This report provides an English translation of a paper by George Pólya that is often cited in the literature on biogeography, yet is apparently seldom read in its original language. It concerns how higher taxa are expected to be distributed at random in ecological communities given their species richnesses and that of a joint regional pool. Pólya arrived at an analytical solution of the expected value by considering the problem as the number of colours of balls drawn on average in a multivariate hypergeometric ball-and-urn experiment.

TABLE OF CONTENTS

Acknowledgements	iii
Summary	v
Introduction	1
Translation	2
Discussion	2
Literature cited	5
Translation of Pólya (1930)	7

INTRODUCTION

What follows is an English translation of George Pólya's (1930) paper in German on his mathematical investigation of a problem in plant biogeography and community ecology. It was undertaken as part of a project to calculate the likelihood function of observed genus-species ratios under a multivariate hypergeometric expectation.

George H. Pólya (1897–1985) was a Hungarian mathematician who taught and did research in Switzerland and the USA. He is considered one of the most influential mathematicians of the 20th century. He had broad interests and made fundamental contributions to numerous fields, including those relevant here, such as combinatorics and probability theory. Beginning early in his career, Pólya also wrote important works on mathematical education and problem solving in general for use by students and teachers. Among the many heuristics that he developed was the now classic Pólya urn problem, a model in which collections of objects to be considered statistically were represented as coloured balls in an urn or sack. Draws from the urn are repeated in such a way as to mimic the random sampling process being modelled. This is the case with the problem in the translated paper below.

In mid-February of 1930, Pólya submitted a paper to *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, a quarterly journal of research in the natural sciences, still in publication today, the fifth of eight papers he would publish in that journal. As well, it was one of six papers he would publish that year, including one co-authored with the, at the time, perhaps better-known number-theorist duo from Cambridge, G. H. Hardy and J. E. Littlewood (see bibliography in Chung et al. 1987). The submission was written to develop analytically an emerging idea being contested in the field of biogeography by his contemporaries, primarily the botanists P. Jaccard (1928) and A. Maillefer (1929). Pólya (1930) first reviewed how genus-species ratios in a local region (his "sub-area") are probabilistically drawn from a pool of potential species in a larger region (his "basic" or "basal" area). He modelled this problem conceptually and then mathematically as the expected number of colours of balls drawn without replacement from an urn of balls of several colours. How local communities of organisms are assembled taxonomically from a regional pool of species is still of

interest today (Enquist 2002; Gotelli 2002; Fan et al. 2017; Potter 2019). A review of the history of the study of taxonomic ratios is found in Gotelli (2004).

TRANSLATION

The original article was obtained as a pdf from the journal's website <https://www.ngzh.ch/publikationen/vjs/75/3-4> [Accessed 15 May 2019]. The article will enter the public domain in the USA on 1 January 2026; see Hirtle (2023). I make the translation available now with the understanding that it falls under fair use, for nonprofit educational and research purposes, that it is a factual work (*contra* an artistic one), and that the use has no market effect (the original is freely available on the web). I do not speak German, so I relied heavily on online translators and dictionaries, especially dict.leo.org, which is awesome. I've been unsuccessful in having the final version completely checked by native speakers. Hence, any errors remain my own.

I used Times New Roman to mimic the original serif typeface. The reduction in font size to indicate the short proof on page 216 is likewise preserved. Cited authors are left as in the original, in small caps. Citations are left in the original language, either German, French, or English. Original page numbers and my comments are in Arial typeface in square brackets. Original page numbers are placed at the approximate point in the translated text where the page begins. The original footnotes are placed on the page of the translation in which the in-text superscripts occur. Equations and tables retain their original organisation and are left in their original position relative to the text, in the latter case, centred and at the end of the paper. Parenthetical comments and those in quotes are the original author's. The single *lapsus*, a typesetting error or perhaps loss of resolution in imaging, was left uncorrected, but noted in the translation: in Eq. 9 the left-hand term " g_s " should have been rendered as the expected value, as elsewhere, \bar{g}_s .

DISCUSSION

Pólya (1930) was more straightforward to translate than anticipated. Because of the paper's author, a giant of 20th-century mathematics, its unfamiliar topic, probability

theory, and its publication nearly a century ago, I was initially worried that the text would correspondingly be an alien landscape of discipline- and era-specific terms. I had previously encountered a similar issue in translating a portion of a German-language book, Hubert von Ludwig's (1888–1892) *Die Seewalzen* (see Kerr 2013). Hence, I was relieved to find that Prof. Pólya's writing was a model of clarity. Recent developments in online translation also likely played a role in the ease of translation. Still, one wonders if his lucid style was due in large measure to his strong interests in mathematical teaching and the importance he lay in conveying to students a compelling conceptual model of statistical problems. If so, then perhaps this also explains his avoidance of specialist terminology and giant leaps in the derivations. In fact, he was later to say,

If the proof starts from axioms, distinguishes several cases, and takes thirteen lines in the text book ... it may give the youngsters the impression that mathematics consists in proving the most obvious things in the least obvious way. (Pólya 1968: 129)

He would go on to finish two volumes on the pedagogy of problem solving and yes: amusingly, which is to say with charm and affection, employ the word 'youngsters' throughout.

A second proof

Although invariably provided without a source (see e.g., www.brainyquote.com), he is also often quoted as having said, 'it is better to solve one problem in five ways...'. In that spirit, then, I offer another solution to the main problem in the translated paper. Pólya (1930) obtained the expected value, the mean, of the distribution of the number of colours drawn hypergeometrically from an urn (as the number of genera of species occurring in a local area). Here, he made use of the probability generating function and then took its first derivative, a standard method for obtaining the moments of a probability distribution. There is a simpler method, however, involving the sums of random indicator variables, whose asymptotic theory was to be developed later (e.g., Volkova 1996) and so, perhaps, was not available or did not occur to him at the time. Alternatively, perhaps he felt that his derivation was the more fundamental or fruitful approach.

Briefly, using the method of 'random indicators' outlined by numerous authors, e.g., succinctly by Nicolas (2015), we can find the average number of colours of balls $E(m)$ of multivariate hypergeometric draws of n balls from an urn of N balls with N_i balls of $i = 1$ to M colours. First, let $I_i = 1$ when the i -th colour is drawn at least once and $I_i = 0$ otherwise. Then, the number of colours drawn is $m = I_1 + I_2 + \dots + I_M$ and through a rule called the 'linearity of expectation', which states that the expected value of a sum of random variables is the sum of their individual expected values,

$$E(m) = E(I_1) + E(I_2) + \dots + E(I_M).$$

Since each I_i are independent, albeit non-identically distributed outcomes of a Bernoulli trial, their expected value is just $E(I_i) = Pr(I_i = 1)$, which happens to be quicker to calculate as its one complement or

$$\begin{aligned} E(I_i) &= Pr(I_i = 1) \\ &= 1 - Pr(I_i = 0) \\ &= 1 - \frac{\binom{N - N_i}{n}}{\binom{N}{n}}. \end{aligned}$$

Now, summing over all M of these expectations, we get

$$E(m) = M - \sum_{i=1}^M \frac{\binom{N - N_i}{n}}{\binom{N}{n}}. \quad (1)$$

This, reassuringly, is the same as Pólya's (1930) result (his Eq. 6) from deriving and differentiating the probability generating function. \square

Alexander M. Kerr
alexander.kerr@aya.yale.edu
Marine Laboratory
University of Guam
21 February 2025

LITERATURE CITED

- Chung, K. L., R. P. Boas, D. H. Lehmer, D. Schattschneider, R. C. Read, M. M. Schiffer and A. H. Schoenfeld. 1987. Obituary: George Pólya. *Bulletin of the London Mathematical Society* 96: 559–608.
- Enquist, B. J., Haskell, J. P., & Tiffney, B. H. 2002. General patterns of taxonomic and biomass partitioning in extant and fossil plant communities. *Nature* 419: 610–613.
- Fan, C., L. Tan, C. Zhang, X. Zhao and K. von Gadow. 2017. Analysing taxonomic structures and local ecological processes in temperate forests in North Eastern China. *BMC Ecology* 73: 33, 1–11. DOI 10.1186/s12898-017-0143-y.
- Gotelli, N. J. 2002. Biodiversity in the scales. *Nature* 419: 575–576.
- Gotelli, N. J. 2004. Assembly rules. Pages 1027–1034 in M. V. Lomolino, D. F. Sax, and J. H. Brown (eds.), *Foundations of Biogeography: Classic Papers with Commentaries*. Chicago: University of Chicago Press. xx + 1291 p.
- Hirtle, P. B. 2023. *Copyright Term and the Public Domain in the United States*. Ithaca, NY: Cornell University Library. 11 p.
- Jaccard, P. 1928. Die statistisch-floristische methode als grund-lage der pflanzensozio-logie. Pages 165–202 in E. Abderhalden (ed.), *Handbuch Biologischer Arbeitsmethoden, Abteilung XI, Teil 5*. Berlin-Wien: Urban & Schwarzenberg. 38 p.
- Kerr, A. M. 2013. Herbert [sic!] Ludwig's (1889–1892) 'Introduction' to Die Seewalzen. *University of Guam Marine Laboratory Technical Report* 146: 1–46.
- Maillefer, A. 1929. Le coefficient générique de P. Jaccard et sa signification. *Mémoires de la Société Vaudoise des Sciences Naturelles* 3: 113–183.
- Nicolas, A. 2015. Answer to 'Urn problem: probability of drawing balls of k unique colors'. <https://math.stackexchange.com/q/1175360/> [Accessed 12 Feb 20].
- Pólya, G. 1930. Eine Wahrscheinlichkeitsaufgabe in der Pflanzensoziologie. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* 75: 211–219.
- Pólya, G. 1968. *Mathematical Discovery: On Understanding, Learning and Teaching Problem Solving. Volume 2*. New York: Wiley. 191 p.

Potter, C. T. 2019. *Taxonomic Partitioning of Coral Reefs in Micronesia*. M.Sc. thesis, University of Guam. 64 p.

Volkova, A. Y. 1996. A refinement of the central limit theorem for sums of independent random indicators. *Theory of Probability and its Applications* 40: 791–794.

Translation of G. H. Pólya (1930. Eine Wahrscheinlichkeitsaufgabe in der Pflanzensoziologie. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* 75: 211–219)

A Probabilistic Problem in Phytosociology

by

G. PÓLYA (Zürich)

(Manuscript received on June 14, 1930.)

1. Background to the problem. In his basic statistical floristic investigations, P. JACCARD introduced the concept¹ of generic coefficients ("coefficient générique"): In an area there exists a total of s species comprised of g genera, wherein JACCARD defined $100g/s$ as the generic coefficient of the area concerned.

In a recent interesting communication,² A. MAILLEFER has taken a very noteworthy approach to precompute the generic coefficient probabilistically. The MAILLEFER approach is, it seems to me, most understandable if one sets the task as follows:

It is well known how many species are present in a larger basic area and how these species are distributed among the genera represented; It is also known how many species are present in a subarea of the basic area. What is required, based on these data, is to calculate (at least approximately) the number of genera in said subarea.

Let's introduce terms. Let

S be the number of species in the basal area,
 G is the number of genera in the basal area,
 s is the number of species in the sub-area,
 g is the number of genera in the sub-area.

[p. 212]

g is thus unknown and sought, whereas S , G , s are known, and it is known how many of the G genera are represented by only one species in the basic area, how many by two, how many by three species, etc.

Of course, if the s species of the sub-area were chosen among the S species of the basal area "on purpose" (for example, by an enemy of mathematical methods), then a prediction of g would be futile, but the species of the subspecies has been chosen by nature, and it seems to me that the following working hypothesis can not be discarded a priori: nature has randomly selected the species of the sub-area, as far as the division into genera is concerned, so we pre-specify the given task as follows: determine the probable or average number of genera under the assumption that the species of the sub-area have been selected among those of the basic area "by chance".

We can make a random selection ourselves as follows: we represent each species by a ball, we throw the balls together in an "urn" (in a sack or such) and draw the required number blindly. We want the balls, which represent species of the same genus, and provided with the

¹ See the summary P. JACCARD, Die statistisch-floristische Methode als Grundlage der Pflanzensoziologie; *Abderhalden's Handbuch der biolog. Arbeitsmethoden*, Abt. XI, Teil 5, p. 165-202.

² Le coefficient générique de P. JACCARD et sa signification, *Mém. de la Soc. Vaudoise des Sciences Naturelles*, vol. 3 (1929) p. 113-183.

same color, and distinguish the different genera by different colors. There are thus S spheres that carry G different colors. When we draw blindly from these spheres s , we find more some times, and at other times we find fewer different colors on the drawn balls, and if we repeat this experiment with the balls enough times, we can find the average number of different colors on s blindly-drawn balls, which, if our working hypothesis is correct, also may be considered as the average number of genera in subregions with s species. That is why we call, with MAILLEFER, this average number determined by urn experiments the probable genus number for s species; we denote this probable genus number by g , or more specifically to emphasize the dependence on the species number s , with \bar{g}_s . This number, which we transmit not only by urn experiments, as we shall see later, but also by probability calculus, seems to me to be a not unreasonable solution to the problem given above. [p. 213]

MAILLEFER adopted Switzerland as the basic area and determined the expected genus number \bar{g}_s for several species numbers through urn experiments. (He did not experiment with colored balls, but with written notes, which, evidently came to the same conclusion.) MAILLEFER has compared the expected genus numbers determined by urn experiments with numerous real numbers obtained by floristic-statistical observations, both in parts of Switzerland, as in areas outside of Switzerland. Without being able to claim a competence which does not concern me in botanical questions, may I say that, in this respect, on the basis of the material summarized by MAILLEFER the agreement seems to me to be quite a good one, so that the real genus numbers should be regarded as having been explained by the expected genus numbers, at least for the most part.

In the following I want to solve the still outstanding mathematical task, whose solution MAILLEFER has also called desirable³: Calculate the probabilistic numbers \bar{g}_s determined by probability experiments.

I want to give the exact solution to the mathematical problem in Section 2 and derive from it in Section 3 an approximate formula that is well-suited for smaller values of s . Finally, in Section 4, I compile some numerical data.

It would be possible to pursue the solution in different directions, but I have refrained from doing so, in particular because, as a non-botanist, I can not judge whether the delimitation of species and genera possesses the degree of firmness which is generally desirable in the delineation of features in statistical studies.

2. Solution to the probabilistic problem. So we have been led to the following task: An urn contains S balls that carry G different colors. Let k_1 be balls of the first, k_2 ; from the second, .. k_G of the last G -th color present, so that

(1)

$$k_1 + k_2 + k_3 + \dots + k_G = S$$

[p. 214]

³ *Op. cit.*, p. 119. I will, in complete agreement with MAILLEFER, understand the task in the way he calls "manière b)".

We are looking for the mean number or mathematical expectation g of the different colors that appear on the balls taken from the urn at the same time.

Let us first look for the probability that exactly g colors appear on the balls taken from the urn. The number of possible cases is obviously $\binom{S}{g}$. The number of favorable cases, we call them A_{sg} is the sum of all products

$$\binom{k_1}{r_1} \binom{k_2}{r_2} \dots \binom{k_G}{r_G} = S$$

of the kind that

$$r_1 + r_2 + \dots + r_G = s$$

$$0 \leq r_1 \leq k_1, 0 \leq r_2 \leq k_2, \dots, 0 \leq r_G \leq k_G$$

and among the numbers r_1, r_2, \dots, r_G exactly g are different from 0. We get A_{sg} most conveniently as coefficients in a generating function:

(2)

$$f(x, y) = \prod_{v=1}^G \left\{ 1 + \binom{k_v}{1} xy + \binom{k_v}{2} x^2 y + \dots + \binom{k_v}{k_v} x^{k_v} y \right\}$$

$$= \prod_{v=1}^G \{ 1 + [(1+x)^{k_v} - 1]y \}$$

$$= \sum_s \sum_g A_{sg} x^s y^g$$

The desired probability is the quotient $A_{sg}/\binom{S}{g}$ and the desired average number of colors is

(3)

$$\bar{g}_s = \binom{S}{s}^{-1} \sum_g g A_{sg}.$$

According to (2), we have,

$$\frac{\partial f}{\partial y} = \sum_s \sum_g A_{sg} x^s y^{g-1}$$

and therefore, according to (3),

(4)

$$\left(\frac{\partial f}{\partial y} \right)_{y=1} = \sum_s \sum_g A_{sg} x^s = \sum_g \binom{S}{g} \bar{g}_s x^s.$$

Starting from (2), we can calculate the function (4) differently:

$$\frac{\partial f}{\partial y} = f \frac{\partial \log f}{\partial y} = f(x, y) \sum_{v=1}^G \frac{(1+x)^{k_v} - 1}{1 + [(1+x)^{k_v} - 1]y}$$

[p. 215]

(5)

$$\begin{aligned}
\left(\frac{\partial f}{\partial y}\right)_{y=1} &= f(x, 1) \sum_{v=1}^G \frac{(1+x)^{k_v} - 1}{(1+x)^{k_v}} \\
&= (1+x)^S \sum_{v=1}^G [1 - (1+x)^{-k_v}] \\
&= G(1+x)^S - \sum_{v=1}^G (1+x)^{-k_v};
\end{aligned}$$

To determine $f(x, 1)$, we used (1). The comparison of the coefficient of x^s in (4) and (5) gives the desired average number:

(6)

$$\bar{g}_s = G - \sum_{v=1}^G \binom{S - k_v}{s} / \binom{S}{s} = G - \sum_{v=1}^G \frac{(S - k_v)! (S - s)!}{(S - k_v - s)! S!}.$$

By the consideration of $\left(\frac{\partial^2 f}{\partial y^2}\right)_{y=1}$, we could use a similar formula for the mean deviation of the number of colors (genus number) g from the mean value \bar{g}_s .

3. Discussion and approximation formulas. In order to solve our problem, we must know how the S species present in the basal area are distributed among the genera represented by G . In particular, we must know the number of genera represented in the basal area by one species only; this number is denoted by H_1 . Similarly, let H_2 be the number of genera, those in the basal area by two, H_3 the number of those that are represented by three species, and so on. So we need all these frequencies

$$H_1, H_2, H_3, H_4, \dots$$

which indicate the distribution of the species according to the number of species native to them in the area. They form a table that we could call a "ranked genus table". Such tables were dealt with by J. C. WILLIS; he also examined the genus sizes within families, the sizes of endemic genera, etc.⁴ At the end of this work, the reader will find the genus table of Swiss flora used by MAILLEFER (Table I); it is of a structure similar [p. 216] to the genus tables considered by WILLIS; H_1 , the number of monotypic genera within Switzerland is the largest, and generally H_k decreases as k grows, initially fairly smoothly and regularly.

From the definition of the frequencies H_k , it follows without further ado that the sums extend over all existing values of k as

(7)

$$\sum_k H_k = G,$$

⁴ J.C. WILLIS, *Age and Area*, Cambridge 1922. For a mathematical treatment, see in particular G. UDNY YULE, A mathematical theory of evolution, based on the conclusions of Dr. J.C. WILLIS, *Phil. Transactions R. S. London, Ser. B*. Vol. 213 (1924).

(8)

$$\sum_k kH_k = S.$$

Formula (6) can be rewritten after the introduction of H_k as

(9)

$$g_s = G - \sum_k H_k \frac{(S-k)!(S-s)!}{(S-k-s)!S!}.$$

[sic; printed in the original article as g_s , but what is meant here is \bar{g}_s]

By slightly transforming the above, one obtains the following expression for the expected generic coefficient:

(10)

$$\frac{100\bar{g}_s}{s} = 100 \sum_k H_k \frac{1}{s} \left[1 - \left(1 - \frac{s}{S}\right) \left(1 - \frac{s}{S-1}\right) \dots \left(1 - \frac{s}{S-k+1}\right) \right].$$

By reason of this expression, one can show that the average generic coefficient, $100\bar{g}_s/s$, decreases steadily, from 100 to $100G/S$, as s , the number of species in the subarea, increases from 1 to S .

The proof is based on the following general remark: If the polynomial $f(x)$ has all positive roots, of which a is the smallest, and $f(0) > 0$ then the second derivative $f''(x)$ is left to a does not vanish, positive in the interval $0 < x < a$, so that the curve $y = f(x)$ seen from below is convex, and therefore $(f(x) - f(0))/x$, as the directional coefficient of a chord, increases steadily when x from 0 to a . The application to the polynomial whose roots are $S - k + 1, \dots, S - 1, S$, and which assumes the value 1 for $x = 0$, yields that the general term of the sum on the right in (10) decreases steadily as s increases steadily from 0 to $S - k + 1$. From here to S , however, when s grows in an integer, the term coincides with the decreasing function s^{-1} .

The numerical evaluation of formula (10) is quite laborious. One obtains an approximation formula useful for small values of s by roughly replacing the product [p. 217]

$$\left(1 - \frac{s}{S}\right) \left(1 - \frac{s}{S-1}\right) \dots \left(1 - \frac{s}{S-k+1}\right)$$

appearing in the main term of the sum in (10) by the k th power

$$\left(1 - \frac{s}{S}\right)^k$$

and this by its first four terms

$$1 - \frac{k}{S}s + \frac{k(k-1)}{2S^2}s^2 - \frac{k(k-1)(k-2)}{6S^3}s^3$$

The result is the approximate formula

(11)

$$\begin{aligned} \frac{100\bar{g}_s}{s} &\approx 100 \sum_k H_k \left[\frac{k}{S} - \frac{k(k-1)}{2S^2}s + \frac{k(k-1)(k-2)}{6S^3}s^2 \right] \\ &\approx 100 - bs + cs^2. \end{aligned}$$

We use (8) and set

(12)

$$b = \frac{100}{2S^2} \sum_k H_k k(k-1), c = \frac{100}{6S^3} \sum_k H_k k(k-1)(k-2).$$

Obviously, the approximate formula (11) is useful at most until s remains below the abscissa of the minimum, that is, as long as

(13)

$$s < \frac{b}{2c}.$$

A similar but somewhat more accurate approximation formula is obtained by applying STIRLING's formula to the terms found in (9) and then developing it by decreasing powers from S to S^{-3} inclusive. A longer calculation, which I suppress here, yields

(14)

$$\frac{100\bar{g}_s}{s} \approx 100 \left(1 - \frac{1}{6S^2}\right) - \left[b \left(1 - \frac{1}{S}\right) + 3c \right] (s-1) + c(s^2 - 1).$$

This approximation formula is also useful at most to its minimum.

4. Numerical example. Formula (10) allows us to compute the expected generic coefficient $\frac{100\bar{g}_s}{s}$, starting from any known area, for a subarea of any given species number s . The exact formula [p. 218] (10), if s is sufficiently small, is to be replaced by the much more convenient approximation formula (14).

I have carried out the calculation, for Switzerland as a basal area, with the data used by MAILLEFER, which are summarized in Table I. The rather tedious exact formula (10) was used only for the species number $s = 200$.⁵ By inserting the data from Table I into (12), one can easily obtain the numerical values for b and c (which are the "factorial moments" familiar to the statistician) and by substituting the values of b , c and S in (14) gives the approximate formula

(15)

$$\frac{100\bar{g}_s}{s} \approx 100.27 - 0.2674s + 0.001329s^2.$$

The minimum of the right side is $s = 100.5$; the formula (15) actually remains useful for approximately $s < 90$. In Table II, which follows at the end of this work, I present the values that I based on formulas (10) and (15), in the third column headed "theoretical" opposite those obtained by MAILLEFER through urn experiments, and those in the second column headed "observed". The agreement is completely satisfactory. Since then, MAILLEFER has successfully compared his figures with an extensive plant geographic data, the communication of the preceding calculations seemed to me of some interest, both for the probability calculation and for the study of plant distributions, in which study the importance of statistical recordings and numerical relationships have just been pointed out in the basic work of P. JACCARD.

⁵ The calculation was carried out by Mr. E. MOECKLIN, Assistant.

Finally, I would like to thank my dear colleague, Professor P. JACCARD, for his pointed reference to the discussion just dealt with.
[p. 219]

Table I.

Summary table of Swiss flora according to MAILLEFER⁶
 H_k = number of genera represented by k species.

k	H_k	k	H_k	k	H_k	k	H_k
1	331	11	8	21	1	31	1
2	133	12	3	22	2		
3	56	13	6	23	3	34	1
4	36	14	4	24	1		
5	27	15	1	25	0	74	1
6	20	16	2	26	1		
7	15	17	4	27	0	85	1
8	8	18	2	28	1		
9	7	19	4	29	2		
10	7	20	5	30	1		

Table II.

Probable generic coefficients, determined through urn experiments and calculation.
 s = number of species in the subarea.

s	Observed	Theoretical
5	99.7	98.9
10	97.9	97.7
20	95.3	95.5
30	93.4	93.5
40	91.8	91.7
50	89.6	90.2
60	88.4	89.2
80	87.1	87.6
200	72.1	71.7

⁶ *Op. cit.*, p. 118. Note the remarks on p. 117.

