

When Reasonably Stop Sampling? How to Estimate the Gain in Newly Recorded Species According to the Degree of Supplementary Sampling Effort

Jean Béguinot^{1*}

¹Département of Bio Géosciences, Université de Bourgogne, F 21000 – Dijon, France.

Author's contribution

The sole author designed, analyzed and interpreted and prepared the manuscript.

Article Information

DOI: 10.9734/ARRB/2015/18809

Editor(s):

- (1) J. David Puett, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, USA.
(2) George Perry, Dean and Professor of Biology, University of Texas at San Antonio, USA.

Reviewers:

- (1) Anonymous, Hacettepe University, Turkey.
(2) Anonymous, University of Management & Technology, Pakistan.
(3) Anonymous, University of Regina, Canada.

Complete Peer review History: <http://sciencedomain.org/review-history/10089>

Method Article

Received 11th May 2015
Accepted 19th June 2015
Published 8th July 2015

ABSTRACT

Aim: Register new species gradually becomes more difficult as sampling of a community progresses, addressing increasingly rarer species. Thus, although biodiversity assessments would ideally require complete samplings, only partial samplings are ordinarily achieved when species abundances distributions are highly heterogeneous within communities, which is often the case. Then, in the frequent context of partial samplings, answering knowingly whether to continue or stop an ongoing inventory require to tentatively assess the “profitability” of the extra sampling effort. That is, trying to estimate the number of species expected to be newly recorded thanks to a given further increase of the sampling size.

Methods: Such estimate may be conveniently derived on the basis of the recorded numbers f_1, f_2, f_3, f_x , of species already recorded once, twice, three, ... x-times within the ongoing sampling. The derivation involves a Taylor expansion of the species accumulation curve, with the successive derivatives of the species accumulation curve being respectively expressed in terms of the successive recorded values of f_x .

Results: A simple nonparametric estimator of the expected number of newly recorded species is derived as a function of the foreseen additive sampling effort. Depending only upon the directly

*Corresponding author: E-mail: jean-beguिनot@orange.fr;

recorded values of the f_x within the ongoing sampling, this estimator is easy-to-implement and, in particular, does not require recording explicitly the species accumulation curve.

Conclusion: The practical interest of this estimator is to offer a convenient way to gauge the additional sampling effort required for a given increase in sample completeness, thus providing quantitative elements to determine whether further continuing an ongoing sampling looks appropriate or not, in the context of both limited available time expenditure and possible other competing priorities.

Keywords: Partial sampling; non-parametric estimator; biodiversity assessment; centered unbiased estimate.

1. INTRODUCTION

The assessment of the composition of a community of species progressively requires more and more sampling effort to gain a same given increment in newly recorded species. This is because progressively rarer species are expected to be collected in the course of sampling [1,2]. In practice, when the time expenditure allowed for investigation is more or less limited, the consistently growing cost of detection of new species thus poses the question of whether to continue or stop sampling, i.e. when reasonably stop, considering also other competing priorities and limited available time expenditure. In other terms, the question of “cost-effectiveness” may be included among the criteria involved in making the decision whether or not further continuing sampling [3,4]. In this respect, procedures are welcomed, that may provide estimates of the number of newly recorded species expected from a given increase of sampling effort (one common way of appreciation of the sampling effort being typically the sample size, i.e. the number of individuals recorded in the sample [5,6]). One classical type of such procedures involves the appropriate fitting of an empirical function to the recorded species accumulation curve (S.A.C.) and its further extrapolation beyond the current size of the sample [5,7-12]. Yet, this approach requires explicitly the prior recording of the species-accumulation curve (completely or at least partially), which indeed is not usual practice in the field. Hereafter, I propose an alternative approach which, precisely, does not impose the prior recording of the S.A.C. but only requires to consider the directly available numbers f_x of species recorded x -times in the sample under consideration (singletons f_1 , doubletons f_2 , etc...). I thus derive a simple formulation for the increment in number of recorded species, δ_{sp} , that might be expected from a given relative increase $\delta N/N$ of the size N of an ongoing sampling, solely on the basis of the recorded values of the numbers f_x .

2. DERIVING A NON-PARAMETRIC POINT-ESTIMATOR OF THE NUMBER OF NEWLY RECORDED SPECIES VERSUS SUPPLEMENTARY SAMPLING EFFORT

Consider the progressive sampling of a community of species, providing a steadily growing number $R_{(N)}$ of recorded species with increasing sampling effort, i.e. with increasing sample size N . Let $f_{1(N)}$, $f_{2(N)}$, $f_{3(N)}$, ..., $f_{x(N)}$, be the numbers of species recorded respectively one, two, three, ..., x -times, within this sample of given size N . A bi-univocal relationship may be derived algebraically between each of the successive derivatives $\partial^x R_{(N)} / \partial N^x$ of the theoretical species-accumulation curve $R(N)$ and the expected value of each $f_{x(N)}$ (the details of the derivation of this relationship are provided in the Appendix):

$$f_{x(N)} = (-1)^{(x-1)} C_{N,x} [\partial^x R_{(N)} / \partial N^x] \quad (1)$$

with $[\partial^x R_{(N)} / \partial N^x]$ as the x^{th} derivative of $R_{(N)}$ with respect to N , at point N and $C_{N,x} = N! / (N-x)! / x!$

Accordingly, the Taylor expansion of the equation governing the theoretical accumulation curve $R_{(N)}$ yields:

$$R_{(N+\delta N)} - R_{(N)} = \delta_{sp} = \sum_x [\partial^x R_{(N)} / \partial N^x] (\delta N)^x / x!$$

with δ_{sp} designing the increment in number of recorded species and \sum_x designing the operation summation upon x , from $x = 1$. According to equation (1):

$$\delta_{sp} = \sum_x (-1)^{(x-1)} f_{x(N)} (\delta N)^x / x! / C_{N,x} \quad (2)$$

In practice, the expansion may generally be limited to the few first terms (see below for practical advice regarding the truncation of the Taylor expansion), so that x remains well below N in practice and thus, $(x! C_{N,x}) \approx N^x$. An estimator of the expected gain in number of recorded species resulting from a given supplementary sampling effort ($\delta N/N$) may thus be written as:

$$\bar{d}_{sp} \approx \sum_x (-1)^{(x-1)} f_{x(N)} (\bar{\delta}N/N)^x \quad (3)$$

The formally similar expression for incidence data is:

$$\bar{d}_{sp} \approx \sum_x (-1)^{(x-1)} Q_{x(N)} (\bar{\delta}N/N)^x \quad (3 \text{ bis})$$

where $Q_1, Q_2, \dots, Q_x, \dots$ are the numbers of uniques, duplicates, x^{th} , that is, respectively, the numbers of species recorded in 1, 2, ..., x, quadrats or investigated sites within the whole set of sampled quadrats or sites under consideration.

Thus defined, \bar{d}_{sp} is a *non-parametric* estimator from two points of view: (i) primarily, the derivation of equations (1, 2, 3) implies no restriction relative to the particular shape of the theoretical accumulation curve $R(N)$ and (ii) as this derivation is essentially algebraic, the values taken by \bar{d}_{sp} are expected to be centered estimates and, in this respect, the estimator is structurally unbiased.

The expression of \bar{d}_{sp} highlights the two causes which combine to make the gain of newly recorded species increasingly difficult as sampling progresses, namely:

- i. The dependence of \bar{d}_{sp} on the *relative* (rather than absolute) increase of sampling effort ($\bar{\delta}N/N$);
- ii. The trend for f_1, f_2, \dots to finally steadily decrease when reaching high values of N .

3. PRACTICAL USE OF THE PROPOSED ESTIMATOR OF THE NUMBER OF NEWLY RECORDED SPECIES

Figs. 1 to 4 provide examples of estimates of the number \bar{d}_{sp} of supplementary recorded species expected from varied relative increases of the sampling size $\bar{\delta}N/N$ (%) as a function of the number of terms retained in the development of \bar{d}_{sp} before truncation of the Taylor expansion. As expected, reliable estimates are obtained for a limited range of increasing sampling efforts: Usually $\bar{\delta}N/N$ not in excess of 50%, exceptionally 75%.

Four land snails communities from the calcareous Burgundy Cost are considered:

"Combe Lavaux": $f_1 = 11, f_2 = 1, f_3 = 1, f_4 = 0, f_5 = 0, f_6 = 0, f_7 = 0, f_8 = 0$.

"Nolay": $f_1 = 8, f_2 = 3, f_3 = 2, f_4 = 4, f_5 = 1, f_6 = 1, f_7 = 1, f_8 = 0$.

"Cersot": $f_1 = 10, f_2 = 1, f_3 = 2, f_4 = 0, f_5 = 3, f_6 = 1, f_7 = 0, f_8 = 3$.

"Val des Choues": $f_1 = 4, f_2 = 0, f_3 = 2, f_4 = 5, f_5 = 2, f_6 = 5, f_7 = 0, f_8 = 1$.

Strictly speaking, equations (3) and (3bis) are based on theoretically infinite Taylor expansions. Yet, in practice, the Taylor expansion may be truncated, that is limited to the beginning of the expansion. Making such a truncation without compromising the reliability of the Taylor expansion is allowed due to the decrease of $(\bar{\delta}N/N)^x$ with growing values of x , when $\bar{\delta}N/N < 1$; this decrease being all the more rapid that $\bar{\delta}N/N$ is lesser than 1. The appropriate position of the truncation thus depends, in particular, upon the foreseen level of supplementary sampling effort ($\bar{\delta}N/N$): all other things being equal, the greater the projected increase of the sampling effort, the greater will be the number of terms to be retained before allowing truncation of the Taylor expansion. More specifically, the number of terms to be retained before truncation, so as to insure reliable expectations for \bar{d}_{sp} , is objectively determined when the stabilisation of \bar{d}_{sp} with growing numbers of terms in the expansion is actually obtained. This is clearly highlighted in the examples given in Figs.1 to 4: the four diagrams show how the estimation of the expected number of newly recorded species, \bar{d}_{sp} , is progressively stabilizing with an increasing number of terms considered in the Taylor expansion before truncation. Thus, for additive sampling efforts, $\bar{\delta}N/N$, up to $\approx 50\%$, the convergence and stability of \bar{d}_{sp} is generally satisfactorily obtained when 3 or 4 terms are retained before truncation of the Taylor development.

For substantially greater additive sampling efforts ($\bar{\delta}N/N > 50\%$), the diagrams show that a rapidly growing number of terms may be needed before allowing truncation. Thus, $\bar{\delta}N/N = 75\%$ may require retaining at least 6 terms (example "Val des Choues"). For still greater additive effort, $\bar{\delta}N/N$ up to 100%, retaining 8 terms at least before truncation is just sufficient in the case of "Nolay" and larger numbers of terms would obviously be needed for "Cersot" and "Val des Choues". Accordingly, in general practice, the procedure proposed above should appropriately address foreseen additional sampling efforts not in excess of 100%.

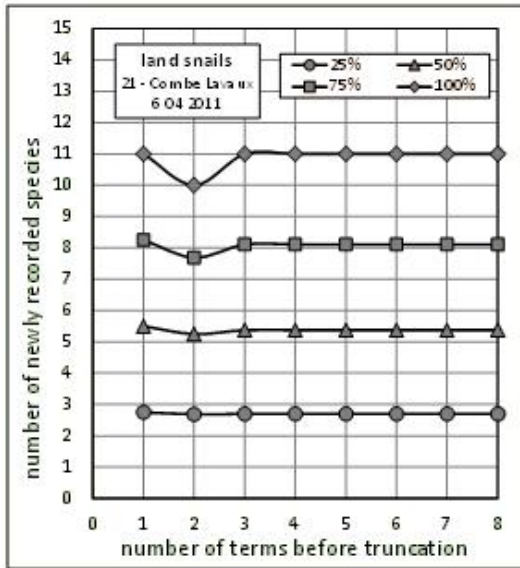


Fig. 1. Estimates of the number δ_{sp} of supplementary recorded species expected from varied relative increases of the sampling size $\delta N/N$ (%), as a function of the number of terms retained in the development of δ_{sp} before truncation of the Taylor expansion. Land snails community from "Combe Lavaux" National Park (Burgundy)

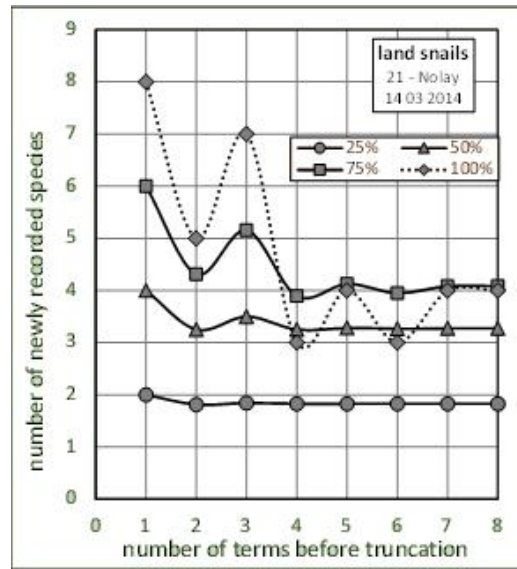


Fig. 2. Estimates of the number δ_{sp} of supplementary recorded species expected from varied relative increases of the sampling size $\delta N/N$ (%), as a function of the number of terms retained in the development of δ_{sp} before truncation of the Taylor expansion. Land snails community from Nolay (Burgundy)

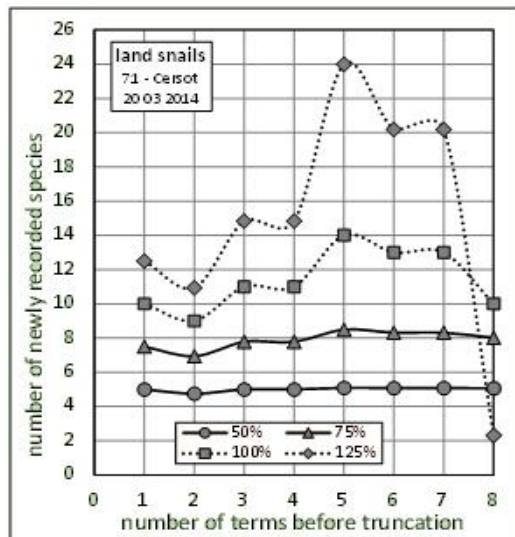


Fig. 3. Estimates of the number δ_{sp} of supplementary recorded species expected from varied relative increases of the sampling size $\delta N/N$ (%), as a function of the number of terms retained in the development of δ_{sp} before truncation of the Taylor expansion. Land snails community from Cersot (Burgundy)

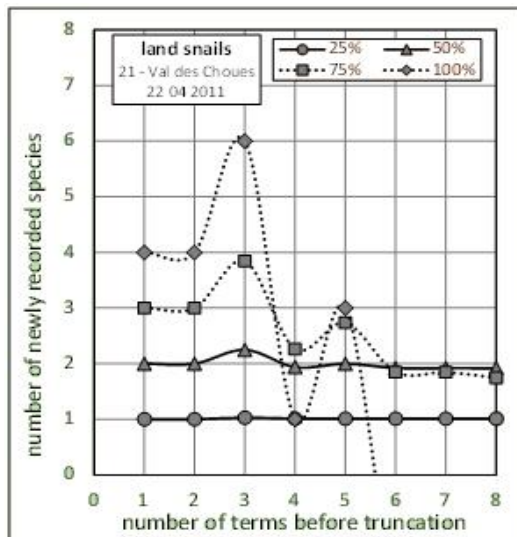


Fig. 4. Estimates of the number δ_{sp} of supplementary recorded species expected from varied relative increases of the sampling size $\delta N/N$ (%), as a function of the number of terms retained in the development of δ_{sp} before truncation of the Taylor expansion. Land snails community from "Val des Choues" (Burgundy)

4. AN ADDITIONAL SUGGESTION: CONVERTING THE ESTIMATED GAIN IN NUMBER OF NEWLY RECORDED SPECIES IN TERM OF RELATIVE IMPROVEMENT OF THE DEGREE OF COMPLETENESS OF SAMPLING

Beyond the absolute gain, in term of the number of newly recorded species δ_{sp} that may be expected from a given supplementary effort of sampling, it may be also interesting to evaluate the corresponding *relative* gain in term of the increment of sampling exhaustivity, δ_{sp}/S (S being the expected total richness). Using the improved Chao estimator 'iChao' [13] for the expected number of missing species, the ratio δ_{sp}/S can be expressed in term of the foreseen supplementary sampling effort ($\delta N/N$) by considering the recorded values of the *four* first $f_{x(N)}$:

$$\delta_{sp}/S = [f_1 \cdot (\delta N/N) - f_2 \cdot (\delta N/N)^2 + f_3 \cdot (\delta N/N)^3 - f_4 \cdot (\delta N/N)^4] / [R + iChao(f_1, f_2, f_3, f_4)] \quad (4)$$

with R as the number of species already recorded during the ongoing sampling.

5. CONCLUSION

A nonparametric estimator of the gain in number of recorded species expected from a given supplementary sampling effort is newly derived. As such, this estimator may help making a decision on a rational basis, when the question is raised whether it is appropriate or not to continue an ongoing sampling, taking into account the expected additional cost of enlarged sampling. Basically, estimates of the extra sampling effort needed for a given gain of sampling completeness would require knowing exactly the shape of the species accumulation curve, especially its prolongation beyond the actual sample size. But, most often, the species accumulation curve is not actually recorded and, *a fortiori*, its prolongation is out of direct empirical reach. Yet, the possibility to extrapolate the species accumulation curve, on the basis of the recorded values of the f_x and using Taylor expansion procedure, leads to a convenient alternative solution. As exemplified by a series of examples, the procedure proposed above provides a rapid, easy to run tool to forecast the expected supplementary sampling effort needed for a given increment in sample completeness. As such, this is a way to better manage field

programs in the context of limited available time expenditure.

ACKNOWLEDGMENTS

Three anonymous reviewers are gratefully acknowledged for their useful comments on a previous version of the manuscript.

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. May RM. Patterns of species abundance and diversity", in Cody ML & Diamond JM, Ecology and Evolution of Communities. The Belknap Press of Harvard University, Cambridge Massachusetts; 1975.
2. Scharff N, Coddington JA, Griswold CE, Hormiga G, De Place Bjorn. When to quit? Estimating spider species richness in a northern European deciduous forest. The Journal of Arachnology. 2003;31:246-273.
3. Azevedo GHF, Faleiro BT, Magalhaes IL, Benedetti AR, Oliveira U, Pena-Barbosa JP, Santos M, Vilela F, De Maria M, Santos AJ. Effectiveness of sampling methods and further sampling for accessing spider diversity: A case study in a Brazilian Atlantic rainforest fragment. Insect Conservation and Diversity; 2013. DOI: 10.1111/icad.12061.
4. Cao Y, Larsen DP, Hughes RM. Evaluating sampling sufficiency in fish assemblage surveys: A similarity-based approach. Canadian Journal Fisheries & Aquatic Sciences. 2001;58:1782-1793.
5. Basset Y, Novotny V. Species richness of insect herbivore communities on Ficus in Papua New Guinea. Biological Journal of the Linnean Society. 1999;67:477-499.
6. Moreno CE, Halffter G. On the measure of sampling effort used in species accumulation curves. Journal of Applied Ecology. 2001;38(2):487-490.
7. Thompson GG, Withers PC, Pianka ER, Thompson SA. Assessing biodiversity with species accumulation curves; inventories of small reptiles by pit-trapping in Western Australia. Austral Ecology. 2003;28:361-383.

8. Ugland I, Gray JS, Ellingsen KE. The species –accumulation curve and estimation of species richness. *Journal of Animal Ecology*. 2003;72:888-897.
9. Colwell RKC, Mao CX, Chang J. Interpolating, extrapolating and comparing incidence-based species accumulation curves. *Ecology*. 2004;85(10):2717-2727.
10. Dove ADM, Cribb TH. Species accumulation curves and their applications in parasite ecology. *Trends in Parasitology*. 2006;22(12):568-574.
11. Mourthé I. How much effort should be employed for surveying a low-diversity Amazonian mammal assemblage using line-transects? *Zoologia*. 2013;30(1):119-124.
12. Gómez-Anaya JA, Novelo-Gutiérrez R, Ramírez A, Arce-Pérez R. Using empirical field data of aquatic insects to infer a cut-off slope value in asymptotic models to assess inventories completeness. *Revista Mexicana Biodiversidad*. 2014;85(1):218-227.
13. Chiu CH, Wang YT, Walther BA, Chao A. An improved nonparametric lower bound of species richness via a modified good-turing frequency formula. *Biometrics*. 2014;70:671-682.

APPENDIX

Derivation of the relationship between the successive derivatives of an accumulation curve and the number of species recorded x -times in a sample

Consider a community of species containing an unknown total number 'S' of species.

Let R be the number of recorded species in a partial sampling of this community, comprising N individuals and let p_i be the probability of occurrence (assimilated to the relative abundance) of species 'i' within this community. The number Δ of missed species (unrecorded in the sample) is

$$\Delta = S - R.$$

The estimated number Δ of species that escape recording during sampling of the community is a decreasing function $\Delta_{(N)}$ of the sample size N, which depends on the particular distribution of species abundances which govern the series of probabilities p_i :

$$\Delta_{(N)} = \sum_i (1-p_i)^N \quad (A1)$$

with \sum_i as the operation summation extended to the totality of the 'S' species 'i' (either *recorded* or *not*)
The number f_x of species recorded x times in the sample, is then, according to the binomial distribution:

$$f_x = [N!/x!(N-x)!] \sum_i [(1-p_i)^{N-x} p_i^x] = C_{N,x} \sum_i (1-p_i)^{N-x} p_i^x \quad (A2)$$

We shall now derive the relationship between the successive derivatives of $R_{(N)}$, the theoretical species accumulation curve and the expected values for the series of ' f_x '.

According to equation (A2):

$$\begin{aligned} \blacktriangleright f_1 &= N \sum_i [(1-p_i)^{N-1} p_i] = N \sum_i [(1-p_i)^{N-1} (1 - (1-p_i))] = N \sum_i [(1-p_i)^{N-1}] - N \sum_i [(1-p_i)^{N-1} (1-p_i)] \\ &= N \sum_i [(1-p_i)^{N-1}] - N \sum_i [(1-p_i)^N]. \end{aligned}$$

Then, according to equation (A1) it comes: $f_1 = N (\Delta_{(N-1)} - \Delta_{(N)}) = -N (\Delta_{(N)} - \Delta_{(N-1)})$

$$= -N (\partial \Delta_{(N)} / \partial N) = -N \Delta'_{(N)}$$

where $\Delta'_{(N)}$ is the first derivative of $\Delta_{(N)}$ with respect to N. Thus:

$$f_1 = -N \Delta'_{(N)} \quad (= -C_{N,1} \Delta'_{(N)}) \quad (A3)$$

Similarly :

$$\begin{aligned} \blacktriangleright f_2 &= C_{N,2} \sum_i [(1-p_i)^{N-2} p_i^2] \quad \text{according to equation (A2)} \\ &= C_{N,2} \sum_i [(1-p_i)^{N-2} (1 - (1-p_i^2))] = C_{N,2} [\sum_i [(1-p_i)^{N-2}] - \sum_i [(1-p_i)^{N-2} (1-p_i^2)]] \\ &= C_{N,2} [\sum_i [(1-p_i)^{N-2}] - \sum_i [(1-p_i)^{N-2} (1-p_i)(1+p_i)]] = C_{N,2} [\sum_i [(1-p_i)^{N-2}] - \sum_i [(1-p_i)^{N-1} (1+p_i)]] \\ &= C_{N,2} [(\Delta_{(N-2)} - \Delta_{(N-1)}) - f_1/N] \quad \text{according to equations (A1) and (A2)} \\ &= C_{N,2} [-\Delta'_{(N-1)} - f_1/N] = C_{N,2} [-\Delta'_{(N-1)} + \Delta'_{(N)}] \quad \text{since } f_1 = -N \Delta'_{(N)} \quad (\text{cf. equation (A3)}). \\ &= C_{N,2} [(\partial \Delta'_{(N)} / \partial N)] = [N(N-1)/2] (\partial^2 \Delta_{(N)} / \partial N^2) = [N(N-1)/2] \Delta''_{(N)} \end{aligned}$$

where $\Delta''_{(N)}$ is the second derivative of $\Delta_{(N)}$ with respect to N. Thus :

$$f_2 = [N(N-1)/2] \Delta''_{(N)} = C_{N,2} \Delta''_{(N)} \quad (A4)$$

$$\begin{aligned} \blacktriangleright f_3 &= C_{N,3} [\sum_i (1-p_i)^{N-3} p_i^3] \text{ which, by the same process, yields:} \\ &= C_{N,3} [\sum_i (1-p_i)^{N-3} - \sum_i (1-p_i)^{N-2} - \sum_i [(1-p_i)^{N-2} p_i] - \sum_i [(1-p_i)^{N-2} p_i^2]] \\ &= C_{N,3} [(\Delta_{(N-3)} - \Delta_{(N-2)}) - f_1^*/(N-1) - 2 f_2/(N(N-1))] \text{ according to equations (A1) and (A2)} \end{aligned}$$

where f_1^* is the number of singletons that would be recorded in a sample of size $(N - 1)$ instead of N . According to equations (A3) & (A4):

$$f_1^* = -(N-1) \Delta'_{(N-1)} = -C_{N-1,1} \Delta'_{(N-1)} \text{ and } f_2 = [N(N-1)/2] \Delta''_{(N)} = C_{N-1,2} \Delta''_{(N)} \quad (A5)$$

where $\Delta'_{(N-1)}$ is the k^{th} derivate of $\Delta_{(N)}$ with respect to N , at point $(N-1)$. Then,

$$\begin{aligned} f_3 &= C_{N,3} [(\Delta_{(N-3)} - \Delta_{(N-2)}) + \Delta'_{(N-1)} - \Delta''_{(N)}] = C_{N,3} [-\Delta'_{(N-2)} + \Delta'_{(N-1)} - \Delta''_{(N)}] \\ &= C_{N,3} [\Delta''_{(N-1)} - \Delta''_{(N)}] = C_{N,3} [-\partial \Delta''_{(N)}/\partial N] = C_{N,3} [-\partial^3 \Delta_{(N)}/\partial N^3] = C_{N,3} \Delta'''_{(N)} \end{aligned}$$

where $\Delta'''_{(N)}$ is the third derivative of $\Delta_{(N)}$ with respect to N . Thus :

$$f_3 = -C_{N,3} \Delta'''_{(N)} \quad (A6)$$

Now, generalising for the number f_x of species recorded x times in the sample:

$$\begin{aligned} \blacktriangleright f_x &= C_{N,x} \sum_i [(1-p_i)^{N-x} p_i^x] \text{ according to equation (A2),} \\ &= C_{N,x} \sum_i [(1-p_i)^{N-x} (1 - (1 - p_i^x))] = C_{N,x} [\sum_i (1-p_i)^{N-x} - \sum_i [(1-p_i)^{N-x} (1 - p_i^x)]] \\ &= C_{N,x} [\sum_i (1-p_i)^{N-x} - \sum_i [(1-p_i)^{N-x} (1 - p_i)(\sum_j p_i^j)]] \end{aligned}$$

with \sum_j as the summation from $j = 0$ to $j = x-1$. It comes:

$$\begin{aligned} f_x &= C_{N,x} [\sum_i (1-p_i)^{N-x} - \sum_i [(1-p_i)^{N-x+1} (\sum_j p_i^j)]] \\ &= C_{N,x} [\sum_i (1-p_i)^{N-x} - \sum_i (1-p_i)^{N-x+1} - \sum_k [(\sum_i (1-p_i)^{N-x+1} p_i^k)]] \end{aligned}$$

with \sum_k as the summation from $k = 1$ to $k = x-1$; that is:

$$f_x = C_{N,x} [(\Delta_{(N-x)} - \Delta_{(N-x+1)}) - \sum_k (f_k^*/C_{(N-x+1+k),k})] \text{ according to equations (A1) and (A2)}$$

where $C_{(N-x+1+k),k} = (N-x+1+k)!/k!(N-x+1)!$ and f_k^* is the number of species recorded k times during a sampling of size $(N-x+1+k)$ (instead of size N).

The same demonstration, which yields previously the expression of f_1^* above (equation (A5)), applies for the f_k^* (with k up to $x-1$) and gives:

$$f_k^* = (-1)^k (C_{(N-x+1+k),k}) \Delta^{(k)}_{(N-x+1+k)} \quad (A7)$$

where $\Delta^{(k)}_{(N-x+1+k)}$ is the k^{th} derivate of $\Delta_{(N)}$ with respect to N , at point $(N-x+1+k)$. Then,

$$f_x = [N!/x!(N-x)!] [(\Delta_{(N-x)} - \Delta_{(N-x+1)}) - \sum_k ((-1)^k \Delta^{(k)}_{(N-x+1+k)})]$$

which finally yields :

$$f_x = [N!/x!(N-x)!] [(-1)^x \partial \Delta^{(x-1)}_{(N)}/\partial N] = [N!/x!(N-x)!] [(-1)^x \partial^x \Delta_{(N)}/\partial N^x]. \text{ That is :}$$

$$f_x = (-1)^x C_{N,x} \Delta^{(x)}_{(N)} = (-1)^x C_{N,x} [\partial^x \Delta_{(N)}/\partial N^x] \quad (A8)$$

where $[\partial^x \Delta_{(N)}/\partial N^x]$ is the x^{th} derivative of $\Delta_{(N)}$ with respect to N , at point N .

Now, the number of recorded species $R_{(N)}$ is equal to the total species richness S minus the expected number of missed species $\Delta_{(N)}$. Then it comes:

$$f_x = (-1)^{(x-1)} C_{N,x} [\partial^x R_{(N)} / \partial N^x] \quad (A9)$$

with $[\partial^x R_{(N)} / \partial N^x]$ as the x^{th} derivative of $R_{(N)}$ with respect to N , at point N and $C_{N,x} = N! / (N-x)! / x!$

A relationship is thus derived between the series of expected values of f_x (the numbers of species recorded x -times) and the series of successive derivatives $[\partial^x R_{(N)} / \partial N^x]$ of the species accumulation curve $R_{(N)}$: $f_1 = C_{N,1} [\partial R_{(N)} / \partial N]$; $f_2 = -C_{N,2} [\partial^2 R_{(N)} / \partial N^2]$; $f_3 = C_{N,3} [\partial^3 R_{(N)} / \partial N^3]$; etc...

© 2015 Béguinot; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
 The peer review history for this paper can be accessed here:
<http://sciencedomain.org/review-history/10089>