Oecologia (Berl.) 25, 187–191 (1976)

Oecologia

A Trophic Diversity Index for Presence-Absence Food Data

Carlos M. Herrera

Estación Biológica de Doñana, Paraguay 1-2, Sevilla-12, Spain

Summary. An index to assess trophic diversity from presence-absence food data is proposed. The index is computed according to the expression $D = -\sum_{i=1}^{s} \lg \hat{p}_i$, where the $\hat{p}_i$'s are the frequencies of occurrence of the various prey categories. The upper and lower limits of D are derived. A test of the reliability of D was carried out by comparing D and H (Shannon's information function) values obtained from a set of twenty-three food analyses from vertebrate animals. Results show that, although a significant correlation exists between D and H, only a small fraction of H-variation is explained by D-variation. D contains two kinds of information, one referred to 'species richness' and another relative to the degree of between-samples heterogeneity. The former is shared in common with H and this presumably explains the fairly weak correlation found between both measures.

Introduction

Food habits data for many animal species are frequently given in the literature as lists of presence-absence records of the diverse prey categories in discrete sampling units, e.g. stomachs, pellets, faeces, etc. Results are given as 'percentages of occurrence' of the various prey categories, obtained by dividing the total number of presences by the total number of discrete samples. These percentages do not reflect the contribution of each prey type to the diet but rather the relative frequency of capture. Trophic diversity is not ussually assessed from data of this kind since the expressions commonly used (e.g. Shannon's information measurement) are not applicable to this semiquantitative information. I propose below a simple, very easily computed diversity index which allows to obtain a measurement of trophic diversity on the basis of presence-absence data. The term 'trophic diversity' is used here to denote the diversity of the diet of a predator species (Hurtubia, 1973; Herrera, 1974; among others). So defined, trophic diversity may be taken as a measure of food niche breadth (Levins, 1968).

Description of the Index

Let us assume to be:

p_i = probability of occurrence of i-th food category in a sampling unit drawn at random from the study population.

$i = 1, 2, \ldots, s$ = total number of qualitative prey categories.

Obviously, $\sum_{i=1}^{s} p_i$ is not necessarily equal to unity.

The p_i's are unknown in practice, but unbiased estimators of them, $\hat{p}_i$'s, are given by

$$\hat{p}_i = \frac{ni}{N}$$

where

n_i = number of occurrences of the i-th prey type.

N = total number of discrete samples under consideration or sample size.

Then, the $\hat{p}_i$'s are the observed 'percentages of occurrence' referred to unity.

An estimate of the probability of occurrence of all the s prey types together in a single sampling unit drawn at random is the product of all the individual probabilities:

$$P = \hat{p}_1 \cdot \hat{p}_2 \cdot \ldots \hat{p}_s = \prod_{i=1}^{s} \hat{p}_i. \tag{1}$$

In the context of food studies, P would be the estimated probability of finding in the same food-sample (e.g. stomach, pellet) all the prey types involved. As P increases, the trophic diversity of the predator will follow the opposite trend, since P is a measurement of *certainty* and, as such, opposed to diversity.

By applying logarithms in (1), we get:

$$\lg P = \lg \left(\prod_{i=1}^{s} \hat{p}_i \right) = \sum_{i=1}^{s} \lg \hat{p}_i \tag{2}$$

and by changing the sign so as to measure diversity and not certainty, the index finally becomes:

$$D = -\lg P = -\sum_{i=1}^{s} \lg \hat{p}_i. \tag{3}$$

It is worth noting the external similitude between expression (3) and Shannon's formula but it must be noted that the $\hat{p}_i$'s are not equivalent in both expressions.

The upper and lower limits of D derive easily by rearranging (1) and obtaining the corresponding limits to P:

$$P = \prod_{i=1}^{s} \hat{p}_i = \frac{1}{N^s} \prod_{i=1}^{s} n_i. \tag{4}$$

The trivial, zero lower limit of P must be neglected, since any $n_i = 0$ would imply a food category which must be rejected since it makes no contribution to the diet. The actual lower limit is reached when all the n_i's are equal to unity, then:

$$P = \frac{1}{N^s}. \tag{5}$$

The upper limit is obviously set by unity and is reached when all the n_i's are equal to N. Then, we get:

$$\frac{1}{N^s} \leqq P \leqq 1. \tag{6}$$

Consequently, limits of D are set by the inequality

$$0 \leqq D \leqq s \lg N \tag{7}$$

which derives easily from (6) by applying logarithms and changing the direction of the signs. As the number of qualitative categories, s, increases, upper limit of D increases linearly. Sample size (N) affects the upper limit to a lesser extent.

A Comparison of Results Obtained by Applying Both '*D*' and Shannon's Information Function

In order to test the reliability of the proposed index D, a comparison of results obtained by using both D and Shannon's information function (H) has been carried out. Twenty-three food analyses were gathered from the literature (Valverde, 1967; Höglund and Lansgren, 1968; Mysterud and Hagen, 1969; Rothkopf, 1970) and from unpublished material. The data are from a variety of vertebrate animals: fishes (1 analysis), amphibians (1), lizards (3), snakes (6), birds of prey (9), and other birds (3). In seven analyses, the sampling units were raptor pellets and in the remaining were stomach contents. All of them gave separate analyses for each sampling unit, then D and H values could be computed simultaneously. D values were obtained according to expression (3) above. H was computed according to

$$H = -\sum_{i=1}^{s} q_i \lg q_i$$

where

$i = 1, 2, \ldots, s =$ number of qualitative food categories, the same as for D calculations.

$q_i =$ frequency of i-th food category with respect to the total number of prey items in the analysis. Then, $\sum_{i=1}^{s} q_i = 1$.

Each food analysis was characterized by the D and H values, number of food categories (s), number of sampling units (N) and total number of prey items. Natural logarithms were used throughout.

Spearman's rank correlation coefficient (Siegel, 1956) between D and H was found to be statistically significant ($r_s = 0.594$, $P < 0.005$, $n = 23$), and Figure 1 illustrates the relation found between both measurements in the sample studied. Thus D may be a useful tool in calculating effective trophic diversity from presence-absence data alone since the results obtained are in good agreement with the corresponding H values.

To determine the main sources of D variation, multiple linear regression analyses were carried out. Taking D as the dependent variable and N and s as

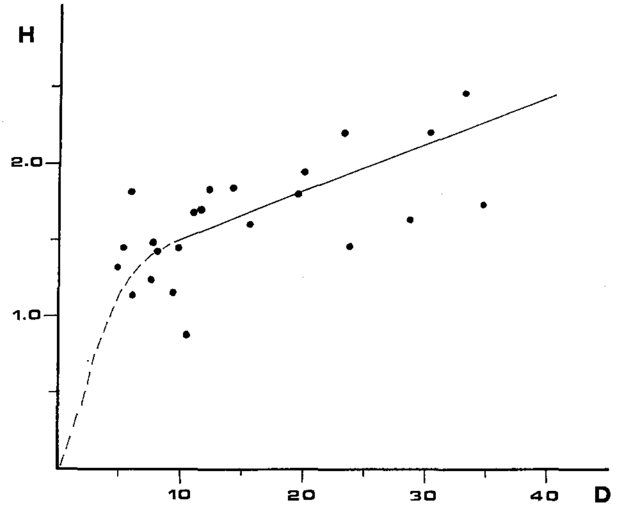

Fig. 1. Graph showing the relation found between H (Shannon's information function) and D (the index proposed here) in twenty-three food analyses from vertebrate animals, and line fitted by eye. Correlation is statistically significant ($r_s = 0.59$, $P < 0.005$)

independent ones, in a first step the variation of s explains 81.01 % of the variation of D, whereas s and N together account for 97.37 % of the observed variation. Taking N alone, the amount of D variation explained by this variable is 28.49 %, thus reflecting a relative independence of D with respect to sample size. A similar analysis taking H as dependent variable and s and N as independent ones showed that s explained 28.35 % of the variation of H, the incorporation of N in a second step increasing it only up to 28.53 %. Number of prey items alone account for 0.90 % of the variation of H.

Although a significant correlation between H and D has been demonstrated, only about 36 % of the variation of H is accounted for by the variation of D. Presumably, this correlation comes from the fact that both indices contain an amount of information on 'species richness' which must be roughly the same for both measures. This is supported by the significant correlation existing between each index and lg s: $r_s = 0.561$ and $r_s = 0.844$ for H and D, respectively. However, each index has another information content which is of different nature for D and H. H contains an amount of information on the relative distribution of prey items among categories ('evenness', Pielou, 1969; Peet, 1974) which account for a large part of H-variation and which is lacking in D owing to the particular kind of data on which it is based. An attempt of correlating the values of evenness ($H/\lg s$) with the corresponding $D/s \lg N$ values was unsuccessful ($r_s = 0.204$, $P > 0.05$) as it would be expected a priori. On the other hand, D depend not only on the 'species richness' component but also on the aggregation patterns of prey items relative to sampling units or, in other words, on the degree of individual differences in feeding habits when data are referred to stomachs contents. If all the individuals of a population feed alike (that is, with equal food preferences), for a given H-value and, consequently, a fixed qualitative 'species richness', the D-

value will be low. Inversely, populations whose members show appreciable individual differences in food choice or exhibit opportunistic feeding habits will give high D-values for a given H-value. Then, D-values reflect not only the variety of prey items in the diet of a population but also the variety of ways in which population members are utilizing them. All these conclusions follow easily from expression (3): higher $\hat{p}_i$'s values, corresponding to greater uniformity in food utilization, lead to lower D-values, and conversely.

When sampling units under study are raptor pellets instead stomachs, 'heterogeneity' information would simply reflect the amount of temporal heterogeneity in food availability and/or food preferences faced by a single individual.

To conlude, it must be noted that although H and D measure definitely separate things, a combination of both measures (e.g. the ratio D/H) would be useful to assess complementary aspects of the feeding ecology of a population in those instances in which sufficient information were available as to compute both indices simultaneously. High D/H values would imply high levels of inter-sample heterogeneity, and conversely.

Acknowledgements. I gratefully acknowledge to J. A. Hernando for the supply of unpublished food data of fishes and to Dr. F. Alvarez for kindly checking the English version of the manuscript. F. Amores was largely responsible for the emergence of the ideas reported here and I am indebted to him. Dr. R. Margalef kindly read the manuscript and the comments of a reviewer were most useful in improving the paper. The author was supported by a grant from the Consejo Superior de Investigaciones Científicas, Spain.

References

Herrera, C. M.: Trophic diversity of the Barn Owl *Tyto alba* in continental Western Europe. Ornis Scand. **5**, 181–191 (1974)
Hurtubia, J.: Trophic diversity measurement in sympatric predatory species. Ecology **54**, 885–890 (1973)
Höglund, N. H., Lansgren, E.: The Great Grey Owl and its prey in Sweden. Viltrevy **5**, 363–421 (1968)
Levins, R.: Evolution in changing environments. Princeton: Princeton Univ. Press 1968
Mysterud, I., Hagen, Y.: The food of the Ural Owl (*Strix uralensis* Pall.) in Norway. Nytt. Mag. Zool. **17**, 165–167 (1969)
Peet, R. K.: The measurement of species diversity. Ann. Rev. Ecol. Syst **5**, 285–307 (1974)
Pielou, E. C.: An introduction to mathematical ecology. New York: Wiley 1969
Rothkopf, D.: Eine Analyse von Gewöllen der Schleiereule *Tyto alba* aus der Eifel. Bonn. zool. Beitr. **21**, 63–82 (1970)
Siegel, S.: Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill 1956
Valverde, J. A.: Estructura de una comunidad de vertebrados terrestres. Monog. Est. Biol. Doñana **1**, 1–218 (1967)

Received December 1, 1975