

RESEARCH

Open Access

Evenness indices once again: critical analysis of properties

Tarald O Kvålseth

Abstract

Various properties have been advocated for biological evenness indices, with some properties being clearly desirable while others appear questionable. With a focus on such properties, this paper makes a distinction between properties that are clearly necessary and those that appear to be unnecessary or even inappropriate. Based on Euclidean distances as a criterion, conditions are introduced in order for an index to provide valid, true, and realistic representations of the evenness characteristic (attribute) from species abundance distributions. Without such value-validity property, it is argued that a measure or index provides only limited information about the evenness and results in misleading interpretations and evenness comparisons and incorrect results and conclusions. Among the overabundant variety of evenness indices, each of which is typically derived by rescaling a diversity measure to the interval from 0 to 1 and thereby controlling or adjusting for the species richness, most are found to lack the value-validity property and some lack the property of strict Schur-concavity. The most popular entropy-based index reveals an especially poor performance with a substantial overstatement of the evenness characteristic or a large positive value bias. One evenness index emerges as the preferred one, satisfying all properties and conditions. This index is based directly on Euclidean distances between relevant species abundance distributions and has an intuitively meaningful interpretation in terms of relative distances between distributions. The value validity of the indices is assessed by using a recently introduced probability distribution and from the use of computer-generated distributions with randomly varying species richness and probability (proportion) components.

Keywords: Biological evenness; Evenness indices; Value validity; Biodiversity

Introduction

There has become an embarrassment of riches for indices used to measure the evenness or uniformity of species distributions in biology. A researcher seeking an evenness index to use in a particular study is faced with a bewildering choice. Also, there does not appear to be any general agreement as to which index is to be preferred over others. Extensive reviews of such indices and their properties have been provided by Smith and Wilson (1996) and Tuomisto (2012). Table 1 summarizes proposed evenness indices that take on values over the interval from 0 to 1 (including a new index E_{11}).

Evenness indices are typically functions of some diversity measure and the number of individuals in a species sample or collection, called species richness and denoted by S , i.e.,

$$\text{Evenness index} = f(\text{Diversity}, S) \in [0, 1] \quad (1)$$

where the function f is generally such that the evenness index ranges in value from 0 to 1. While there is a plethora of measures of species diversity and no consensus about which measure is “best” (e.g., Magurran 2004, Ch. 4; Grassle et al. 1979; Tuomisto 2012), the present paper is confined to the measurement of species evenness with the focus on the properties of such measures or indices.

A variety of properties have been imposed on evenness indices (see, e.g., Smith and Wilson 1996; Taillie 1979; Engen 1979; Tuomisto 2012; Gosselin 2001; Ricotta 2004). Some of these appear to be entirely appropriate and generally agreed upon, while others seem to be unnecessary and even unreasonable. This paper starts off with a discussion of properties that are considered to be necessary for any acceptable evenness index, followed by comments about some dubious properties. One of the properties introduced as necessary is the *value-validity* property, which basically requires that an evenness index

Correspondence: kvals001@umn.edu
Department of Mechanical Engineering, University of Minnesota,
Minneapolis, MN 55455, USA

Table 1 Proposed evenness indices varying over the interval from 0 to 1 and based on the species abundance probabilities (proportions) p_1, \dots, p_S and species richness S

Designation	Formula	Reference	Notes
E_1	$H/\log S = -\sum_{i=1}^S p_i \log p_i / \log S$	Pielou (1966)	
E_2	$(e^H - 1)/(S - 1)$	Heip (1974)	a
E_3	$(1 - \sum_{i=1}^S p_i^2)/(1 - 1/S)$	Smith and Wilson (1996)	
E_4	$-\log \sum_{i=1}^S p_i^2 / \log S$	Smith and Wilson (1996)	
E_5	$(1/\sum_{i=1}^S p_i^2 - 1)/(e^H - 1)$	Alatalo (1981)	a
E_6	$\left(1 - \sqrt{\sum_{i=1}^S p_i^2}\right) / \left(1 - \sqrt{1/S}\right)$	Pielou (1969)	
E_7	$(1/\sum_{i=1}^S p_i^2 - 1)/(S - 1)$	Kvålseth (1991)	
E_8	$\left(\sum_{i=1}^S \min\{p_i, 1/S\} - 1/S\right) / (1 - 1/S)$	Bulla (1994)	
E_9	$1 - \left[\left(\sum_{i=1}^S p_i^2 - 1\right) / (S - 1)\right]^{1/2}$	Williams (1977)	b
E_{10}	$2 \sum_{i=1}^S (i-1)p_i / (S-1)$	Solomon (1979)	c
E_{11}	$\left(1 - \sum_{i=1}^S p_i/i\right) / \left[1 - (1/S) \sum_{i=1}^S 1/i\right]$	New	c

Notes: a. The H stands for the Shannon (1948) entropy defined for E_1 and with base-e (natural) logarithms (E_1 and E_4 are indifferent as to which logarithm is used).
 b. Engen (1979) attributed this index to F.M. Williams (1977) in an unpublished manuscript.
 c. The p_i 's are here in descending order ($p_1 \geq p_2 \geq \dots \geq p_S$).

takes on values that are all reasonable with respect to an acceptable criterion. It is argued that this property, which nearly all proposed evenness indices lack, is necessary in order for comparisons between differences in evenness values to be valid. An index lacking this property may cause misleading results and incorrect conclusions.

The value-validity requirement of an index is based on the recently introduced *lambda distribution* (Kvålseth 2011) and a criterion involving Euclidean distances between species abundance distributions. The assessment of whether

or not an index meets this requirement is then done analytically using the lambda distribution and empirically using computer-generated random species abundance distributions. It is also proved that some well-known indices are not valid since they lack the essential strict Schur-concavity property and hence do not preserve the Lorenz order.

Properties of evenness indices

Let $E(P_S)$, or simply E , represent the value of a generic evenness index for the species abundance distribution

$P_S = (p_1, \dots, p_S)$ where p_i is the probability (proportion) of the i th species, with $p_i \geq 0$ for $i = 1, \dots, S$ and $\sum_{i=1}^S p_i = 1$.

If each p_i is based on a sample or collection as opposed to unknown population probabilities, then $p_i = n_i/N$ where n_i is the frequency of individuals in species i , $i = 1, \dots, S$, and $N = \sum_{i=1}^S n_i$ is the sample size. The extreme distributions are then

$$P_S^1 = \left(\frac{1}{S}, \dots, \frac{1}{S}\right), P_S^0 = (1, 0, \dots, 0) \tag{2}$$

for which E is expected to take on its extremal values for any given species richness S . In the case of a censused community or collection of species, the most uneven species abundance distribution becomes $(1 - (S - 1)/N, 1/N, \dots, 1/N)$ and the most even distribution equals P_S^1 in (2) only when N/S is an integer. However, when the sample size N is large, these two extreme distributions become effectively equivalent to those in (2).

Based on this notation, the various potential properties of an evenness index E will be discussed, starting with an outline and explanation of those properties that are suggested as being indeed necessary. Subsequent sections of this paper will discuss the *value-validity* property necessary for making appropriate comparisons and interpretations of E -values.

Necessary properties

(P1) *Continuity*: E is a continuous function of each of the p_i , $i = 1, \dots, S$, for any given S .

(P2) *Symmetry*: E is (permutation) symmetric in p_1, \dots, p_S so that the value of E is invariant with respect to all permutations of the p_i 's.

(P3) *Normalization*: E takes on values between 0 and 1.

(P4) *Schur-concavity*: E is strictly Schur-concave. In order to explain this property, the concept of majorization needs to be first defined. The rather vague notion that the components of $P_S = (p_1, \dots, p_S)$ are "less spread out", "more nearly equal", or "more even" than are the components of the distribution $Q_S = (q_1, \dots, q_S)$ can be expressed as " P_S being majorized by Q_S " and denoted by $P_S < Q_S$. Thus, if the p_i ($i = 1, \dots, S$) are ordered such that

$$p_1 \geq p_2 \geq \dots \geq p_S \tag{3}$$

with Q_S being similarly ordered, then, by definition,

$$P_S < Q_S \text{ if } \sum_{j=1}^i p_j \leq \sum_{j=1}^i q_j, \quad i = 1, \dots, S-1 \tag{4}$$

and with $\sum_{i=1}^S p_i = \sum_{i=1}^S q_i = 1$. Then, by definition, E is strictly Schur-concave if

$$P_S < Q_S \text{ implies } E(P_S) > E(Q_S) \tag{5}$$

and P_S is not a permutation of Q_S (Marshall et al. 2011, pp. 8, 80).

As a simple numerical example, consider the two distributions $P_4 = (0.40, 0.30, 0.20, 0.10)$ and $Q_4 = (0.70, 0.15, 0.10, 0.05)$ where the components have been ordered as in (3). It is readily apparent from (4) that P_4 is majorized by Q_4 , i.e., $P_4 < Q_4$. Then, since the evenness index E is required to be strictly Schur-concave, $E(P_4) > E(Q_4)$ from (5). Clearly, the components of P_4 are "more evenly distributed" or "more nearly equal" than those of Q_4 and the evenness based on P_4 should be greater than that based on Q_4 . This is precisely what Property P4 requires.

(P5) *Value Validity*: E takes on values that are all valid representations of the true extent of the species evenness characteristic.

This last property, which is considered necessary for making appropriate evenness difference comparisons, is explained in detail in subsequent sections of the paper.

Comment on property P1

The continuity requirement ensures that, for any given or fixed S , small changes in some of the p_i 's cause only a small change in the value of E . However, since E may be a function of both the p_i ($i = 1, \dots, S$) and S , a change in S and hence changes in the p_i 's will necessarily result in a discontinuous change in E . In fact, if S is very small, the addition of one more species may produce a substantial jump in the value of E . Such discontinuity with varying S has been discussed by, for instance, Routledge (1983), Jost (2010), and Ricotta (2004).

Other implications from property P4

The strict Schur-concavity Property P4 of E has some important implications. Since, for any $P_S = (p_1, \dots, p_S)$, it is readily apparent from the definition in (4) that $P_S^1 < P_S < P_S^0$ for the P_S^1 and P_S^0 in (2), it then follows from (5) that

$$E(P_S^0) \leq E(P_S) \leq E(P_S^1) \tag{6}$$

with the inequalities being strict if P_S is not a permutation of P_S^0 or P_S^1 . This can be expressed as the following sub-properties:

(P4a) For any given species richness S , E attains its minimum and maximum values for the species abundance distributions P_S^0 and P_S^1 in (2), respectively.

Another consequence of Property P4 is that E satisfies the *principle of transfers*, which, within the context of income distributions, was introduced by Dalton (1920). That is, if $p_i < p_j$ in the species abundance distribution $P_S = (p_1, \dots, p_S)$ and if an amount δ is transferred from p_j to p_i , with $\delta < p_j - p_i$, then the value of E increases. This can be seen to follow from (3)-(5). See also Marshall et al. (2011), pp. 6-8. That is, a sub-property of E is as follows:

(P4b) E satisfies the transfer property. As a simple example, consider the four-species distribution $Q_4 = (0.30, 0.15, 0.50, 0.05)$ with $\delta = 0.20$ being transferred from p_3 to p_4 to produce the distribution $P_4 = (0.30, 0.15, 0.30, 0.25)$. It is then clear from (3)-(4) that P_4 is majorized by Q_4 , so that, from (5) and the strict Schur-concavity of E , $E(P_4) > E(Q_4)$.

A third consequence of Property P4 may be stated in terms of Lorenz curves as follows (see Marshall et al. 2011, pp. 5–6, 713–715):

(P4c) E preserves the Lorenz order. As an explanation, the Lorenz curve for a distribution $P_S = (p_1, \dots, p_S)$ is typically based on the cumulative probabilities $F_{(i)} = \sum_{j=1}^i p_{(j)}$

after the p_i 's have been ordered such that $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(S)}$. The Lorenz curve (after Lorenz 1905) is then obtained by joining the consecutive points $(i/S, F_{(i)})$ for $i = 0, 1, \dots, S$ and $F_{(0)} = 0$ by successive line segments to connect the origin $(0, 0)$ with the point $(1, 1)$. If the Lorenz curve for P_S is nowhere below and does not coincide everywhere with the Lorenz curve for the distribution $Q_S = (q_1, \dots, q_S)$, this is equivalent to the majorization in (3)-(4) so that, from (5), $E(P_S) > E(Q_S)$. The Lorenz curve for the uniform or completely even distribution P_S^1 in (2) is the straight line between the two points $(0, 0)$ and $(1, 1)$ so that the value $E(P_S)$ increases as the Lorenz curve for P_S gets closer to this straight line.

Introductory comments on value validity

The motivation behind the need for an evenness index to have value validity can perhaps be best illustrated by means of some simple numerical examples. Consider, for instance, the species abundance distributions $P_2 = (0.75, 0.25)$ and $Q_5 = (0.60, 0.10, 0.10, 0.10, 0.10)$ and the Pielou index E_1 in Table 1 for which $E_1(P_2) = 0.81$ and $E_1(Q_5) = 0.76$. Both of these results would seem to indicate a “high” degree of evenness, with P_2 indicating a more even distribution than Q_5 of about 7%. However, the components of each of these two distributions are equally far from the corresponding components of the respective pairs of extreme distributions P_2^0 and P_2^1 and Q_5^0 and Q_5^1 defined in (2) so that the only reasonable evenness values should be 0.50 for both P_2 and Q_5 . Furthermore, if, say, $\sqrt{E_1}$ or more generally E_1^α ($\alpha > 0$) were to be considered as alternative indices, which can be shown to have the same Properties P1-P4 as E_1 , the results could be even more unreasonable such as $\sqrt{E_1(Q_5)} = 0.87$.

Of course, the purpose of using an evenness index E is to compare the degree or extent of evenness of different species abundance distributions. Thus, in simplified notation, if e_1, e_2, \dots denote values of the index E for different species distributions, the various comparisons of potential interest may be defined as follows:

$$\text{Size (order) comparisons : } e_1 > e_2 \tag{7a}$$

$$\text{Difference comparisons : } e_1 - e_2 > e_3 - e_4 \tag{7b}$$

$$\text{Proportional difference comparisons : } e_1 - e_2 = c(e_3 - e_4) \tag{7c}$$

where c is some constant. Even the above interpretation of $E_1(0.75, 0.25) = 0.81$ as indicating a “high” degree of evenness is basically a difference comparison as in (7b), with $0.81 - 0 > 1 - 0.81$, i.e., the 0.81 value is much further from the minimum E_1 -value of 0 than it is from the maximum E_1 -value of 1. Similarly, the above statement that $E_1(0.75, 0.25) = 0.81$ exceeds $E_1(0.60, 0.10, \dots, 0.10)$ by about 7% is equivalent to $e_1 = 0.81$, $e_2 = e_3 = 0.76$, $e_4 = 0$, and $c = 0.07$ in (7c).

When making such comparisons, it is essential that they are not limited to the evenness index itself, but that they provide true representations of the attribute (characteristic) being measured. Otherwise, the results obtained may be invalid and misleading, which is not uncommon in the published literature. What is needed of the index E is that it takes on values throughout its range that are all accurate, true, or valid representations of the evenness attribute. That is, E has to have *value validity* (Property P5). In measurement theory, the term *validity* of a measurement procedure means that it does in fact measure what it purports to measure. While several different types of validity have been defined (see, e.g., Hand 2004, pp. 129–134), *value validity* as used here applies specifically to the requirement that all potential numerical values of a measure have to be appropriate or reasonable with respect to some generally acceptable criterion as explained in the next subsection of this paper.

The importance of a concept such as value validity does not appear to have been the subject of rigorous attention in the numerous publications involving evenness measurement and indices. Some have commented that certain evenness indices “tend to strongly overestimate evenness” or “give unreasonably high values of evenness” (Bulla 1994, pp. 167, 169). The results from computer-generated combinations of species frequencies by Fager (1972) implied that indices such as Pielou’s E_1 and Smith and Wilson’s E_3 in Table 1 tended to overstate evenness and had “strongly skewed distributions...” indicating that those indices “would give poor discrimination over much of the range of samples (p. 301).” Molinari (1989) did consider the importance of reasonable intermediate values of an evenness index, focusing on the case of two-species communities, and Smith and Wilson (1996) as well as Jost (2010) emphasized the need to consider intuitively reasonable intermediate index values. However, no specific value-validity conditions appear to have been formulated.

Conditions for value validity

In order to establish value-validity conditions on E (Property P5), it will be convenient to use the following *lambda distribution* recently introduced by this author (Kvålseth 2011, 2014):

$$P_S^\lambda = \left(1-\lambda + \frac{\lambda}{S}, \frac{\lambda}{S}, \dots, \frac{\lambda}{S}\right), 0 \leq \lambda \leq 1 \tag{8}$$

where λ is a parameter and S denotes the species richness. The λ is basically an evenness or uniformity parameter with $\lambda = 1$ for the completely even (uniform) distribution P_S^1 in (2) and $\lambda = 0$ for the degenerate distribution P_S^0 in (2). Note also that each component of P_S^λ can be considered the weighted mean of components of P_S^0 and P_S^1 , i.e.,

$$P_S^\lambda = \lambda P_S^1 + (1-\lambda)P_S^0 \tag{9}$$

Since the strict Schur-concavity of E ensures that the inequality in (6) holds for all distributions $P_S = (p_1, \dots, p_S)$, it also holds for P_S^λ and all S and λ . Therefore, for any given P_S , there exists one, and only one, value of λ such that

$$E(P_S) = E(P_S^\lambda) \text{ for any given } P_S \text{ and one unique } \lambda\text{-value} \tag{10}$$

Of course, for any given S , there may be any number of different P_S -distributions for which the value of E would be the same and hence for which (10) would apply. Because of (10), the value validity of E and appropriate conditions can be considered in terms of the P_S^1 -distribution in (8).

Consider now that each distribution P_S is a point or vector in S -dimensional Euclidean space with Cartesian coordinates p_1, \dots, p_S . The Euclidean distance between the points P_S and Q_S is given by $d(P_S, Q_S) = \left[\sum_{i=1}^S (p_i - q_i)^2\right]^{1/2}$.

A most logical requirement for value validity based on Euclidean distances may then be formulated as the following equality between distance ratios:

$$\frac{|E(P_S^\lambda) - E(P_S^1)|}{|E(P_S^0) - E(P_S^1)|} = \frac{d(P_S^\lambda, P_S^1)}{d(P_S^0, P_S^1)} \tag{11}$$

for the $P_S^0, P_S^1, P_S^\lambda$ defined in (2) and (8). Since the right-hand side ratio in (11) can be seen to equal $1 - \lambda$ and from the Schur-concavity Property P4, specifically (6), the relationship in (11) can equivalently be expressed as

$$E(P_S^\lambda) = \lambda E(P_S^1) + (1-\lambda)E(P_S^0), 0 \leq \lambda \leq 1 \tag{12}$$

$$= \lambda \text{ for } E(P_S^1) = 1 \text{ and } E(P_S^0) = 0 \tag{13}$$

The distance-based condition in (12) is also a logical implication from (9). That is, the evenness value in (12)

consists of a weighted mean equivalent to that of the underlying species distributions in (9).

The condition in (11) and hence in (12)-(13) applies specifically to the lambda distribution P_S^λ in (8). However, because of (10), it seems entirely reasonable to also require that (11) should generally provide a good (close) approximation when P_S^λ is replaced by any species distribution $P_S = (p_1, \dots, p_S)$. Therefore, for any distribution P_S , the approximation corresponding to (13) becomes

$$E(P_S) \approx 1 - \frac{d(P_S, P_S^1)}{d(P_S^0, P_S^1)} \tag{14}$$

so that, with $E(P_S^0) = 0$ and $E(P_S^1) = 1$, (13) and (14) become conditions for value validity of the evenness index E .

Further property comments

Distance criterion

The Euclidean distance is the metric being used as the basic criterion for the value-validity conditions in (11)-(14). Other metric distance functions could have been considered such as alternative members of the Minkowski family

$$d_\alpha(P_S, Q_S) = \left[\sum_{i=1}^S |p_i - q_i|^\alpha\right]^{1/\alpha} \text{ for parameter } \alpha \geq 1, \text{ with}$$

$\alpha = 2$ being the Euclidean metric and $\alpha = 1$ being the so-called city-block or Manhattan metric (e.g., Upton and Cook 2008, pp. 118–119). While it is easily seen that, if any member of this distance family is used in (11), the expressions in (12)-(13) would still apply, but the right-hand side of (14) would generally differ for different values of α . The Euclidean distance is used here since it is the standard or ordinary measure of distance used in mathematics and the various sciences. The use of any other alternative distance metric or function would seem to require some particular justification or explanation.

Also, it is of importance to note that the right-hand side of (14) based on the Euclidean distance is strictly Schur-concave whereas, had it been based on the city-block distance, the corresponding expression would have been Schur-concave, but not strictly Schur-concave. This follows from the fact that (a) the Euclidean distance between P_S and P_S^1 is strictly Schur-convex while the city-block distance is Schur-convex, but not strictly so (Kvålseth 2011), and (b) the right-hand side of (14) is a strictly decreasing function of those two alternative distance metrics (Marshall et al. 2011, pp. 88–91, 138–139).

Bounds on E

For a censused community with species abundance frequencies n_1, \dots, n_S and total number of individuals

$N = \sum_{i=1}^S n_i$, the most uneven species abundance distribution becomes $P_S^{0+} = (1-(S-1)/N, 1/N, \dots, 1/N)$ and the most even distribution P_S^{1-} equals P_S^1 in (2) only if N/S is an integer. For these extremal distributions, none of the evenness indices in Table 1 can attain the lower and upper bounds of 0 and 1 since $E_i(P_S^{0+}) > 0$ and $E_i(P_S^{1-}) < 1$. Of course, when N is large, P_S^{0+} and P_S^{1-} become effectively equal to the distributions in (2) and $E_i(P_S^{0+}) \approx 0$ and $E_i(P_S^{1-}) \approx 1$.

As a simple example, consider $S = 5$ and $N = 98$. Then, from the above definition, $P_5^{0+} = (1-4/98, 1/98, \dots, 1/98)$ for the frequencies $n_i = 94, 1, \dots, 1$ and $P_5^{1-} = (20/98, 20/98, 20/98, 19/98, 19/98)$. For, say, the Williams' E_9 in Table 1, the values for these two extreme distributions are found to be $E_9(P_5^{0+}) = 0.05$ and $E_9(P_5^{1-}) = 0.99$ as compared to $E_9(P_5^0) = 0$ and $E_9(P_5^1) = 1$ for the P_5^0 and P_5^1 in (2).

Some (e.g., Fager 1972) have suggested that an evenness index should be able to attain the values 0 and 1 for any S and N while others do not impose such requirement as long as an index takes on values near 0 and 1 when the species distribution is as uneven or even as possible for any given S and N (e.g., Smith and Wilson 1996). Of course, an index E with $E(P_S^{0+}) > 0$ and $E(P_S^{1-}) < 1$ can easily be rescaled to the interval $[0,1]$ by defining

$$E^*(P_S) = \frac{E(P_S) - E(P_S^{0+})}{E(P_S^{1-}) - E(P_S^{0+})} \tag{15}$$

In the above example, with $S = 5$ and $N = 98$, consider $n_i = 20, 5, 60, 10, 3$ for which the value of E_9 in Table 1 becomes $E_9(P_5) = 0.46$ whereas, from (15), $E_9^*(P_5) = (0.46 - 0.05) / (0.99 - 0.05) = 0.44$.

However, there seems to be no particular basis for imposing such a rescaling requirement on an evenness index. In fact, on intuitive grounds, it seems unreasonable. To suggest that the above distribution $P_5^{1-} = (20/98, 20/98, 20/98, 19/98, 19/98)$ indicates complete evenness is simply not appropriate. Rather, this P_5^{1-} is as even as it can be when $S = 5$ and $N = 98$. Similarly, the above distribution $P_5^{0+} = (94/98, 1/98, \dots, 1/98)$ is as uneven as possible for given $S = 5$ and $N = 98$, but it should not be considered to imply perfect unevenness. The above value, $E_9(P_5^{0+}) = 0.05$ rather than zero, would seem to provide a reasonable representation of the evenness.

Independence of S

It has been argued strongly by some that an evenness index E should be independent of the species richness S

(e.g., Peet 1975; Hill 1973; Heip et al. 1998). Such a requirement would eliminate all indices in Table 1 for an insufficient reason. It is probably inevitable that any E with the above Properties P1-P5 will to some extent depend on S . This dependence has also been discussed more recently by Gosselin (2006) and Jost (2010).

It can certainly be argued that the dependence of E on S is not by itself an undesirable property. The purpose of an evenness index E is to measure how evenly (uniformly) the relative abundances $p_i (i = 1, \dots, S)$ are distributed across the S different species, irrespective of the value of S . The reason for normalizing (rescaling) an evenness index such that the values of E all fall within the interval from 0 to 1 (or nearly so) is simply to be able to properly compare values of E for species distributions with differing S -values. That is, the normalized E controls or adjusts for S . Without such control or adjustment of S , an evenness index would be a confounded representation of species richness S and the form of the species distribution. For such an index, one would not be able to tell whether different index values were due to differences in S -values or differences in the form of species distributions or both. An evenness index $E \in [0, 1]$ is only supposed to measure the form of the species distribution, thereby requiring the control (adjustment) for S , and thus making E a function of both p_1, \dots, p_S and S .

In statistical analysis of categorical data, an evenness index such as E_3 in Table 1 has been used as a measure of variation and referred to as the *index of qualitative variation (IQV)* (e.g., Mueller and Schuessler 1977, pp. 175–181; Reynolds 1984, pp. 61–64). For such measurement, the fact that the normalized $E_3 \in [0, 1]$ depends on the number of categories S does not appear to be much of an issue of concern, while papers on biological evenness seem to have made the dependence on S into an unnecessary issue. Just as it is for a measure of qualitative variation, inclusion of S into the formulation of $E \in [0, 1]$, besides the ease of interpretation, is precisely for the purpose of obtaining an evenness index that can be used to compare the evenness of different species distributions of differing S .

One type of situation in which there is reason for genuine concern about E depending on species richness is when the relative species frequencies $p_i = n_i/N (i = 1,$

$\dots, S)$ and sample size $N = \sum_{i=1}^S n_i$ are based on random

observations from a population of species and when the computed (sample) value $E(n_1/N, \dots, n_S/N)$ is used to make statistical inferences about the corresponding unknown population index E_p . Such statistical inferences would include (a) the use of the sample value of E as an

appropriate estimate of E_p and (b) the construction of confidence intervals for E_p . However, the necessary conditions for making such statistical inferences are rarely met in biological studies. It is typical for such sampling data that neither the total number of species S_p in the defined population nor all the specific types of species are known, with $S < S_p$ and S depending to some extent on the sample size N . It is then most prudent to confine the evenness measurement to the sample itself.

Replication principle

A measurement property that has also been referred to as the “replication property” or “replication principle” was introduced by Dalton (1920) for the measurement of economic inequality. Thus, if (x_1, \dots, x_n) denotes the set of incomes of n individuals and if another set of identical incomes for n other individuals are combined to produce the set of incomes $(x_1, x_1, x_2, x_2, \dots, x_n, x_n)$ for $2n$ individuals, then a measure of income inequality should have the same value for $(x_1, x_1, \dots, x_n, x_n)$ as for (x_1, \dots, x_n) . Any number of such replications should produce the same result. Dalton (1920) called this the “principle of proportionate additions to persons”, while others (e.g., Cowell 2011, pp. 63–64) have referred to it as the “principle of population”. The term “replication” property or principle, as also used for biological evenness measures (e.g., Taillie 1979; Tuomisto 2012), would seem to be most descriptive and will be used here.

However, it may perhaps be questionable if such a replication property should be considered necessary or even desirable, especially when the underlying data are categorical such as biological species frequencies (counts) as opposed to quantitative data such as personal incomes. Even when measuring income inequality, some feel that the desirability of this property is debatable and not self-evident (Cowell 2011, pp. 63–64), while others do not even mention this property when discussing inequality measures (Bellù and Liberati 2006; Marshall et al. 2011, Sec 13F).

For quantitative data, it does at least make realistic sense to consider a replication $(x_1, x_1, \dots, x_n, x_n)$. However, for categorical data such as a set of biological species abundance frequencies (counts) (n_1, n_2, \dots, n_S) , a replication $(n_1, n_1, n_2, n_2, \dots, n_S, n_S)$ would seem to be an unrealistic and intuitively meaningless concept. Each n_i has to be identified with a specific species (or category) within a set of mutually exclusive and exhaustive set of species of size S . If, say, (n_1, \dots, n_S) are the frequencies for a set of S different types of apple trees in one geographic area and if exactly the same frequencies are observed for the same set of S apple trees in an adjacent area, it would make no sense to determine an evenness value based on the data $(n_1, n_1, \dots, n_S, n_S)$ for the two combined areas. Such determination should be based on $(2n_1, \dots, 2n_S)$. Otherwise, one would be “comparing apples and oranges”.

One situation in which it does make sense to consider replication would be when a set of S different species with frequencies (n_1, \dots, n_S) are split into males and females, resulting in the species-sex frequencies $(n_1/2, n_1/2, \dots, n_S/2, n_S/2)$. The S “species” categories have been split into $2S$ “species-sex” categories. These are two different sets of categories for which there would seem to be no basis for requiring that the evenness should be the same.

Jost (2010) has also argued against the requirement that an evenness index should possess the replication property. In fact, he argues that replication invariance would be an undesirable property for an evenness index $E \in [0, 1]$.

Any proposed condition such as the one involving replication also has to make some intuitive sense when simply looking at the data, both in terms of frequencies n_i and proportions (relative frequencies) $p_i = n_i/N$ for $i = 1, \dots, S$ and $N = \sum_{i=1}^S n_i$. For instance, consider two sets of data: frequencies (80, 20) or proportions $P_2 = (0.80, 0.20)$ versus the replicated (80, 80, 20, 20) or $Q_4 = (0.40, 0.40, 0.10, 0.10)$. It would be hard to justify the proposition that both data sets reveal the same degree of evenness, especially when considering the P_2 versus the Q_4 . Clearly, Q_4 indicates greater evenness than P_2 since the components of Q_4 differ less than those of P_2 .

As an example of a replication-invariant evenness index producing such intuitively unreasonable values, consider

$$E_T = \frac{1}{S} \left(2 \sum_{i=1}^S ip_i - 1 \right) \in \left[\frac{1}{S}, 1 \right] \tag{16}$$

where p_1, \dots, p_S are ordered as in (3) (Taillie 1979, p. 56). This E_T represents a slight modification of Solomon’s E_{10} in Table 1. Because of its lower bound $1/S$, an interpretation such as $E_T(1, 0) = 0.50$ showing more than twice the evenness of $E_T(1, 0, 0, 0) = 0.20$ makes no intuitive sense. Similarly, in terms of frequencies n_1, \dots, n_S , with $N = \sum_{i=1}^S n_i$, it is found from (16) with $p_i = n_i/N$ ($i = 1, \dots, S$) that

$$E_T(N-S+1, 1, \dots, 1) = \frac{1}{S} + \frac{S-1}{N}$$

with such unreasonable results as $E_T(99, 1) = 0.51$ and $E_T(96, 1, 1, 1) = 0.24$. By comparison, from the E_9 in Table 1, $E_9(99, 1) = 0.02 < E_9(96, 1, 1, 1) = 0.05$, an intuitively reasonable result.

From the definition of the Lorenz curve under Property P4c, it is readily seen that the Lorenz curve of, say,

$P_2 = (0.80, 0.20)$ and $Q_4 = (0.40, 0.40, 0.10, 0.10)$ coincide (one curve with two more points than the other), a result that holds for all replication situations. However, considering the preceding arguments, such a Lorenz-curve fact would not be sufficient to argue for a proposition that an evenness index should be invariant with respect to replication.

Assessment of proposed indices

Based on the above discussion of appropriate properties for an evenness index, the validity of various proposed indices may now be assessed. All of the indices listed in Table 1 can be seen to meet the requirements for continuity, symmetry, and normalization, i.e., Properties P1-P3. However, as will be explained next, some of those indices are not strictly Schur-concave (Property P4) and most do not satisfy the conditions for value validity (Property P5).

Schur-concavity

Alatalo’s index E_5 in Table 1 is not Schur-concave. This is rather apparent since the denominator of E_5 is strictly Schur-concave because (a) the entropy H is strictly Schur-concave and (b) a strictly increasing function of a strictly Schur-concave function is strictly Schur-concave (Marshall et al. 2011, pp. 89, 101). Also, as an example, consider the distributions $P_3 = (0.80, 0.10, 0.10)$ and $Q_3 = (0.80, 0.19, 0.01)$ for which $E_5(P_3) = 0.58$ exceeds $E_5(Q_3) = 0.67$ even though P_3 is majorized by Q_3 so that E_5 cannot be Schur-concave (see (3)-(5)). Consequently, E_5 is not a valid evenness index.

Bulla’s index E_8 in Table 1 is Schur-concave, but it is not strictly Schur-concave as required by Property P4. A proof of the Schur-concavity of E_8 is given below in Appendix A. The fact that this index is not strictly Schur-concave can be verified by the following counterexample: $E_8(0.80, 0.15, 0.05) = E_8(0.80, 0.10, 0.10) = 0.30$ even though $(0.80, 0.10, 0.10)$ is majorized by $(0.80, 0.15, 0.05)$ as can be seen from (3)-(4). Consequently, from (5), E_8 cannot be strictly Schur-concave.

Among other proposed evenness indices, Smith and Wilson (1996) introduced one as a strictly decreasing function of the variance V of the logged frequencies,

$$i.e., V = \sum_{i=1}^S \left[\log n_i - \sum_{i=1}^S (\log n_i) / S \right]^2 / S.$$

However, since V is not Schur-convex and a decreasing function of V is therefore not Schur-concave (Marshall et al. 2011, pp. 89, 561), the index proposed by Smith and Wilson (1996) is not Schur-concave and is therefore not a valid evenness index.

Hill (1973) proposed a family of evenness indices as

$$E_{H\alpha\beta} = \frac{N_\alpha}{N_\beta}, N_\alpha = \left(\sum_{i=1}^S p_i^\alpha \right)^{1/(1-\alpha)}, N_\beta = \left(\sum_{i=1}^S p_i^\beta \right)^{1/(1-\beta)} \tag{17}$$

with no restriction on the parameters α and β . However, in order for a member of this family to be strictly Schur-concave, it is necessary that $\alpha > 0$ and $\beta < 0$. These parameter ranges do not include the specific member $E_{H2,1}$ for $\alpha = 2$ and $\beta = 1$ in (17) mentioned by Hill (1973) and of which the E_5 in Table 1 is a slight modification. Also, $E_{H\alpha\beta}$ in (17) is undefined for $\beta < 0$ unless it is assumed that all p_i are positive. A proof of these parameter restrictions is given below in Appendix B.

Hill (1997) also proposed the evenness index

$$\left(\frac{N_3}{N_2} \right)^2 = \frac{\left(\sum_{i=1}^S p_i^2 \right)^2}{\sum_{i=1}^S p_i^3}$$

where N_2 and N_3 are defined from (17). This index is replication invariant and relatively independent of S , but, as can be inferred from the preceding paragraph, it cannot be Schur-concave since it is a ratio of Schur-convex functions. The fact that this index is not Schur-concave and therefore not a valid evenness index is also evident from, for instance, the two distributions $(0.6, 0.4)$ and $(0.99, 0.01)$ for which this Hill’s index takes on the respective values 0.97 and 0.99. This is clearly an absurd result since the components of $(0.6, 0.4)$ are more even than those of $(0.99, 0.01)$, i.e., $(0.6, 0.4)$ is majorized by $(0.99, 0.01)$ (see (3)-(4)).

Ricotta (2004) made an interesting attempt to formulate an evenness index by using a fuzzy set theory approach. With the p_i ’s ordered as in (3), this index is defined as

$$E_R = 1 - \sum_{i=1}^S \left(\frac{p_i - p_{i+1}}{i} \right)$$

where $p_{S+1} = 0$. However, this index is not Schur-concave (nor is it Schur-convex) and is therefore not a valid evenness index. This fact can be proved by using a counterexample such as the two species distributions $P_3 = (0.6, 0.2, 0.2)$ and $Q_3 = (0.6, 0.3, 0.1)$ where Q_3 is derived from P_3 by means of the transfer of 0.1 from p_3 to p_2 . From the definition in (4), it is clear that P_3 is majorized by Q_3 , but $E_R(P_3) = 0.53 < E_R(Q_3) = 0.57$ so that, from (5), E_R is not Schur-concave.

Table 2 Values of the evenness indices in Table 1 for the lambda distribution P_S^λ in (8) with different λ and S values

S	λ	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}
2	0.2	0.47	0.38	0.36	0.29	0.57	0.61	0.22	0.20	0.20	0.20	0.20
2	0.5	0.81	0.75	0.75	0.68	0.79	0.72	0.60	0.50	0.50	0.50	0.50
2	0.8	0.97	0.96	0.96	0.94	0.96	0.95	0.92	0.80	0.80	0.80	0.80
5	0.2	0.41	0.23	0.36	0.21	0.43	0.28	0.10	0.20	0.20	0.20	0.20
5	0.5	0.76	0.60	0.75	0.57	0.62	0.66	0.38	0.50	0.50	0.50	0.50
5	0.8	0.96	0.92	0.96	0.91	0.90	0.94	0.83	0.80	0.80	0.80	0.80
10	0.2	0.38	0.15	0.36	0.17	0.35	0.26	0.05	0.20	0.20	0.20	0.20
10	0.5	0.73	0.48	0.75	0.49	0.48	0.63	0.23	0.50	0.50	0.50	0.50
10	0.8	0.94	0.87	0.96	0.87	0.81	0.92	0.71	0.80	0.80	0.80	0.80
20	0.2	0.35	0.10	0.36	0.14	0.28	0.24	0.03	0.20	0.20	0.20	0.20
20	0.5	0.70	0.37	0.75	0.42	0.35	0.60	0.13	0.50	0.50	0.50	0.50
20	0.8	0.93	0.80	0.96	0.81	0.68	0.91	0.55	0.80	0.80	0.80	0.80
30	0.2	0.34	0.07	0.36	0.13	0.25	0.24	0.02	0.20	0.20	0.20	0.20
30	0.5	0.68	0.32	0.75	0.38	0.29	0.58	0.09	0.50	0.50	0.50	0.50
30	0.8	0.92	0.76	0.96	0.77	0.58	0.89	0.44	0.80	0.80	0.80	0.80

Value validity (condition (13))

Of the evenness indices identified so far in this paper, most of them lack the value validity property as defined and discussed above. Those indices cannot therefore be used to make valid difference comparisons as in (7b)-(7c) and their numerical values cannot be used as valid indicators of the true extent of evenness from a data set. Only “larger than” comparisons as in (7a) may be valid for those indices. Among the indices defined in Table 1, only E_8, E_9, E_{10} , and E_{11} can be seen to meet the condition in (13) for value validity. However, as proved above, Bulla’s E_8 , while being Schur-concave, lacks the strict Schur-concavity property. As a demonstration of such lack of value validity, consider the results in Table 2 giving the values of the measures in Table 1 for the lambda distribution P_S^λ in (8) with some different values of λ and S .

It is clearly indicated by these results that, except for E_8, \dots, E_{11} , the various indices violate the condition in (13) that $E(P_S^\lambda) = \lambda$, some more so than others. For some of the indices, their values for the distribution P_S^λ exceed the λ -values, i.e., they overstate the extent of evenness, while such bias is reversed for other indices, depending upon the values of λ and S . Although these results are based on the specific distribution P_S^λ in (8), they have definite implications for species distributions $P_S = (p_1, \dots, p_S)$ in general because of the relationship in (10).

It is then apparent from Table 2 together with (10) that the general tendency for each of the E_1, \dots, E_6 is to overstate the degree of evenness, especially when the species richness S is small. The indices E_1, E_3 , and E_6

appear to consistently overstate the evenness. In terms of the *value bias* of an evenness index E (VBE) (as distinct from statistical bias of an estimator) defined as

$$VBE(P_S^\lambda) = E(P_S^\lambda) - \lambda \tag{18}$$

E_1, E_3 , and E_6 appear to have such a consistent positive bias for all P_S^λ and hence generally for all $P_S = (p_1, \dots, p_S)$ from (10). Interestingly, such positive bias for E_3 does not depend explicitly on S since, from (18) and the expression for E_3 in Table 1, this bias is found to be $VBE_3(P_S^\lambda) = \lambda(2-\lambda) - \lambda = \lambda(1-\lambda)$, which also is clearly maximal for $\lambda = 0.5$. Note also in Table 2 for $S > 2$ the extremely large negative value bias for the index E_7 proposed by this author (Kvålseth 1991), making it entirely invalid for the difference comparisons in (7b)-(7c).

For the evenness indices E_9, E_{10} , and E_{11} , which all have Properties P1-P4 and which meet the condition in (13), it remains to be determined if those indices also meet the value-validity condition in (14). This is considered next.

Value validity (condition (14))

The Williams’ E_9 meets the condition in (14) exactly since

$$\begin{aligned} E_9 &= 1 - \frac{d(P_S, P_S^1)}{d(P_S^0, P_S^1)} \\ &= 1 - \left(\frac{\sum_{i=1}^S (p_i - 1/S)^2}{(1-1/S)^2 + (S-1)(1/S)^2} \right)^{1/2} \\ &= 1 - \left(\frac{S \sum_{i=1}^S p_i^2 - 1}{S-1} \right)^{1/2} \end{aligned} \tag{19}$$

which equals E_9 in Table 1. The E_9 is also equivalent to the *coefficient of nominal variation* (CNV) proposed by Kvålseth (1995) as a measure of variation for nominal categorical data. It can also be expressed as a linear function of the standard deviation s_S of p_1, \dots, p_S (with divisor S), i.e.,

$$E_9 = CNV = 1 - \frac{S}{\sqrt{S-1}} s_S$$

This is so rather obvious an evenness index that undoubtedly others have also thought of it.

The determination of whether or not Solomon’s E_{10} and this author’s E_{11} comply with the requirement in (14) then becomes a comparison of E_{10} and E_{11} with the ‘gold standard’ E_9 for different distributions $P_S = (p_1, \dots, p_S)$. The E_{11} is presented as an interesting alternative to E_{10} , both of which are based on the ordering $p_1 \geq p_2 \geq$

... ≥ p_S. The E₁₁ is based on the complement of the homogeneity (dominance, concentration) measure $\sum_{i=1}^S p_i/i$ proposed by Kvålseth (1993) and normed to the [0, 1] -interval. For large S, the denominator of E₁₁ can more conveniently be computed as 1 - (logS + 0.5772)/S + 1/2S², which is found to be accurate to four decimal places when S > 11 (see, e.g., Knopp 1990, p. 538).

In order to obtain the necessary data to determine if E₁₀ and E₁₁ comply with the requirement in (14), i.e., determine the extent to which E₁₀ and E₁₁ approximate E₉, randomly generated distributions P_S = (p₁, ..., p_S) were used. For each such P_S, the value of S was first generated as a random integer between 2 and 30 (inclusive) and then each p_i was generated in descending order (p₁ ≥ p₂ ≥ ... ≥ p_S) as random numbers (to 6 decimal places) within the following intervals:

$$\begin{aligned} \frac{1}{S} &\leq p_1 \leq 1 \\ \frac{1-p_1}{S-1} &\leq p_2 \leq \min \{p_1, 1-p_1\} \\ &\vdots \\ \frac{1-\sum_{j=1}^{S-2} p_j}{S-(S-2)} &\leq p_{S-1} \leq \min \left\{ p_{S-2}, 1-\sum_{j=1}^{S-2} p_j \right\} \\ p_S &= 1-\sum_{j=1}^{S-1} p_j \end{aligned}$$

For each of the 25 such computer-generated distributions P_S, the values of E₉, E₁₀, and E₁₁ were computed as were the values of Pielou's E₁ in Table 1. Some (five) generated data sets were excluded since the index values were all nearly equal to 0 or 1. The normalized entropy E₁ was included, although it clearly violates (13), since this entropy-based index appears to be the most popular evenness index. The results are summarized in Table 3.

As expected from the results in Table 2 together with (10), Table 3 shows that E₁ substantially and consistently violates the value-validity condition in (14) since its values generally differ considerably from those of the 'gold standard' E₉. The overstatement (positive value bias) of evenness by Pielou's E₁ appears to be as large as up to about 150 percent. By comparison, the approximation in (14) is much better for E₁₀ and E₁₁, i.e., the values of E₁₀ and E₁₁ in Table 3 are generally much closer to the corresponding values of E₉ than are those of E₁. This is especially the case for E₁₁.

In the case of Solomon's E₁₀, the data in Table 3 indicate that values of E₁₀ tend to be systematically and sometimes considerably smaller than those of E₉. For data Sets 3, 6, 12, 14, and 18, the E₁₀ values are half or less of the E₉ values. If E₉ is used to predict E₁₀ as $\hat{E}_{10} =$

Table 3 Values of indices E₁, E₉, E₁₀, and E₁₁ defined in Table 1 for species distributions P_S = (p₁, ..., p_S) with randomly generated S and p_i(i = 1, ..., S)

Data set	S	E ₉	E ₁₀	E ₁₁	E ₁
1	15	0.27	0.22	0.25	0.44
2	7	0.88	0.81	0.87	0.98
3	18	0.45	0.24	0.44	0.53
4	2	0.15	0.15	0.15	0.39
5	4	0.29	0.22	0.26	0.51
6	12	0.35	0.11	0.28	0.42
7	17	0.92	0.91	0.92	0.99
8	14	0.32	0.30	0.31	0.51
9	8	0.85	0.77	0.84	0.97
10	17	0.63	0.56	0.61	0.82
11	29	0.84	0.84	0.84	0.95
12	5	0.19	0.10	0.15	0.33
13	30	0.38	0.30	0.35	0.53
14	5	0.32	0.18	0.28	0.49
15	27	0.82	0.69	0.79	0.92
16	2	0.29	0.29	0.29	0.59
17	5	0.52	0.40	0.50	0.74
18	30	0.31	0.15	0.24	0.39
19	10	0.57	0.42	0.53	0.75
20	26	0.95	0.91	0.94	0.99
21	19	0.67	0.61	0.66	0.84
22	19	0.77	0.76	0.76	0.91
23	17	0.46	0.40	0.44	0.65
24	12	0.74	0.70	0.73	0.91
25	2	0.50	0.50	0.50	0.81

E₉, then the coefficient of determination (when properly computed; see Kvålseth 1985) is found from Table 3 to be R² = 1 - Σ(E₁₀ - E₉)²/Σ(E₁₀ - \hat{E}_{10})² = 0.86. Similarly, the root mean square difference (RMSD) between the values of E₁₀ and E₉ is found to be 0.10. When similarly comparing E₁₁ with E₉, it turns out that R² = 0.99 and RMSD (E₁₁, E₉) = 0.03. That is, nearly 100% of the total variation of E₁₁ (about its mean) is explained by the 'fitted' model $\hat{E}_{11} = E_9$ as compared to 86% in the case of E₁₀. Consequently, E₁₁ cannot be rejected for lacking value validity since it complies with the requirement in (13) and seems to have a high degree of approximation to E₉ as required by (14). However, in the case of Solomon's E₁₀, while it complies with (13), one could certainly question its value validity since its approximation to E₉ is rather marginal.

Concluding comments

As advocated in this paper and about which most researchers seem to agree, an evenness index E should take

on values over the interval from 0 to 1, with $E(P_S^0) = 0$ and $E(P_S^1) = 1$ for the distributions in (2). These fixed bounds are preferable when comparing the evenness of different species collections or communities with differing species richness S . Some indices have been proposed, however, for which the bounds depend on S such as $E(P_S^0) = 1/S$ for indices E_T in (16) by Taillie (1979) and E_C by Camargo (1993) defined as

$$E_C = 1 - \frac{\sum_{i=1}^{S-1} \sum_{j=i+1}^S |p_i - p_j|}{S}$$

The E_T is seen to be a simple transformation of E_{10} in Table 1, i.e., $E_T = [(S - 1)E_{10} + 1]/S$. The E_C index has been referred to as Camargo's evenness index (e.g., Magurran 2004, p. 118) although it is the same as the E_T proposed earlier by Taillie (1979). The fact that $E_C = E_T$ follows directly from the equality

$$\sum_{i=1}^{S-1} \sum_{j=i+1}^S |p_i - p_j| = \sum_{i=1}^S (S + 1 - 2i)p_i$$

where p_1, \dots, p_S on the right-hand side are ordered as in (3).

Since all the indices in Table 1 can be seen to have Properties P1-P4, with the exception of Alatalo's E_5 and Bulla's E_8 that are not strictly Schur-concave (Property P4), those indices can reasonably be expected to provide valid size (order) comparisons as in (7a). That is, for an evenness index E with Properties P1-P4 and for which, say, $E(P_S) = 0.90$ and $E(Q_R) = 0.60$, one can reasonably say that the species with abundance distribution P_S has higher evenness than that with abundance distribution Q_R , but nothing more can be said with validity. One cannot credibly say that one has "considerably higher" evenness than the other or that the 0.90 value shows a "very high" degree of evenness even though the 0.90 value is near the top end of the $[0, 1]$ -interval over which E is defined. Such interpretations and the types of difference comparisons in (7b)-(7c) require that E meets conditions (13)-(14) for value validity, conditions which are only clearly satisfied by Williams' E_9 and the new index E_{11} in Table 1.

Of the two indices E_9 and E_{11} with the Properties P1-P5, there seems to be no particular reason for preferring E_{11} over E_9 . The E_9 also has an intuitively appealing geometric interpretation in terms of the Euclidean distances between points in S -dimensional space in (19), i.e., E_9 is the relative extent to which the distance between P_S^1 in (2) and the species abundance distribution (point) $P_S = (p_1, \dots, p_S)$ is less than the distance between P_S^1 and P_S^0 defined in (2). Alternatively, since P_S is majorized by P_S^0 and since the

distance $d(P_S, P_S^1)$ can be shown to be strictly Schur-concave, the E_9 in (19) can be expressed as

$$E_9 = 1 - \frac{d(P_S, P_S^1)}{\max_{P_S} d(P_S, P_S^1)} \tag{20}$$

that is, E_9 is the relative extent to which the distance between an observed P_S and the P_S^1 is less than the maximum such distance over all S -species distributions. For instance, from the data by Magurran (2004), pp. 226–227 of the abundances of 16 species of dung beetles around Bangalore, India, (19) or (20) gives $E_9(P_{16}) = 0.48$, which means that the distance between the dung beetle abundance distributions P_{16} and P_{16}^1 is 48% less than its maximum possible value for 16-species distributions. If, for these data with $S = 16$ and total number of dung beetles $N = 1745$, one uses $P_{16}^{0+} = (1730/1745, 1/1745, \dots, 1/1745)$ and $P_{16}^{1-} = (110/1745, 109/1745, \dots, 109/1745)$ instead of the P_{16}^0 and P_{16}^1 in (2) as discussed above, one obtains the distances $d(P_{16}, P_{16}^{1-}) = 0.5034$ and $d(P_{16}^{0+}, P_{16}^{1-}) = 0.9588$ so that, from (19), $E_9(P_{16}) = 1 - 0.5034/0.9588 = 0.47$.

In order to avoid misrepresentations and incorrect interpretations by using evenness indices lacking important properties, the conclusion from the analysis in this paper seems clear: E_9 is the index of choice. Having all of the necessary properties (Properties P1-P5), E_9 is a most informative index for which all of the comparisons in (7a)-(7c) can reasonably be considered valid.

Appendices

Appendix A

In order to prove that Bulla's E_8 in Table 1 is Schur-concave, but not strictly Schur-concave, consider the function

$$F_\alpha = \sum_{i=1}^S \left(\frac{p_i^\alpha + (1/S)^\alpha}{2} \right)^{1/\alpha} = \sum_{i=1}^S m_{\alpha i}$$

where α is some arbitrary parameter. This F_α is simply the sum of arithmetic means $m_{\alpha i}$ of order α and of which E_8 is the particular limiting member $F_{-\infty}$ as α goes to $-\infty$. By taking the partial derivatives of $m_{\alpha i}$ with respect to p_i it is found that

$$\frac{\partial^2 m_{\alpha i}}{\partial p_i^2} = \left(\frac{\alpha-1}{4} \right) m_{\alpha i}^{1-2\alpha} \left(\frac{1}{S} \right)^\alpha p_i^{\alpha-2}$$

which is negative for $\alpha < 1$ so that $m_{\alpha i}$ is a strictly concave function of p_i for $\alpha < 1$ and fixed S . Thus, being a sum of strictly concave functions, F_α is strictly Schur-concave for $\alpha < 1$ (Marshall et al. 2011, p. 92). However, in the limit as $\alpha \rightarrow -\infty$, when $(F_\alpha - 1/S)/(1 - 1/S)$

becomes Bulla's E_8 , the strict Schur-concavity does not hold as demonstrated by the earlier counterexample.

Appendix B

As a proof of the parameter restrictions on Hill's family of evenness indices in (17), consider the partial derivative of $N_\alpha^* = \log N_\alpha$ (which is the entropy family of Rényi 1961) with respect to p_i i.e.,

$$\frac{\partial N_\alpha^*}{\partial p_i} = \left(\frac{\alpha}{1-\alpha} \right) \left(\sum_{j=1}^S p_j^\alpha \right)^{-1} p_i^{\alpha-1}$$

which shows that, if p_1, \dots, p_S are ordered as in (3), $\partial N_\alpha^* / \partial p_i$ is strictly increasing in $i = 1, \dots, S$ (when $p_i > p_{i+1}$ for all i) for $\alpha > 0$ and strictly decreasing if $\alpha < 0$. Therefore, N_α^* and hence N_α are strictly Schur-concave for $\alpha > 0$ and strictly Schur-convex for $\alpha < 0$ (Marshall et al. 2011, pp. 84, 89). Consequently, since the ratio of a strictly Schur-concave function to a strictly Schur-convex function is strictly Schur-concave (Marshall et al. 2011, pp. 89, 145), the $E_{H\alpha\beta}$ in (17) is strictly Schur-concave for $\alpha > 0$ and $\beta < 0$. It can also be proved that this result holds when either $\alpha = 0$ or $\beta = 0$.

Competing interests

The author declares that he has no competing interests.

Acknowledgment

The author thanks an anonymous reviewer for helpful comments.

Received: 19 December 2014 Accepted: 24 March 2015

Published online: 16 May 2015

References

- Alatalo RV (1981) Problems in the measurement of evenness in ecology. *Oikos* 37:199–204
- Bellù LG, Liberati P (2006) Policy impacts on inequality: simple inequality measures. Food and Agriculture Organization (FAO) of the United Nations, Rome, Italy
- Bulla L (1994) An index of evenness and its associated diversity measure. *Oikos* 70:167–171
- Camargo JA (1993) Must dominance increase with the number of subordinate species in competitive interactions? *J Theor Biol* 161:537–542
- Cowell FA (2011) *Measuring inequality*, 3rd edn. Oxford University Press, Oxford
- Dalton H (1920) Measurement of the inequality of incomes. *Econ J* 30:348–361
- Engen S (1979) Some basic concepts of ecological equitability. In: Grassle JF, Patil GP, Smith W, Taillie C (eds) *Ecological diversity in theory and practice*. International Co-operative Publishing House, Fairland, Maryland, USA, pp 37–50
- Fager EW (1972) Diversity: a sampling study. *Am Nat* 106:293–310
- Gosselin F (2001) Lorenz partial order: the best known logical framework to define evenness indices. *Commun Ecol* 2:197–207
- Gosselin F (2006) An assessment of the dependence of evenness indices on species richness. *J Theor Biol* 242:591–597
- Grassle JF, Patil GP, Smith W, Taillie C (eds) (1979) *Ecological Diversity in Theory and Practice*. International Co-operative Publishing House, Fairland, Maryland, USA
- Hand DJ (2004) *Measurement Theory and Practice*. Wiley, Chichester, West Sussex, UK
- Heip C (1974) A new index measuring evenness. *J Mar Biol Ass UK* 54:555–557
- Heip CHR, Herman PMJ, Soetaert K (1998) Indices of diversity and evenness. *Océanis* 24:61–87
- Hill MO (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427–432
- Hill MO (1997) An evenness statistic based on the abundance-weighted variance of species proportions. *Oikos* 79:413–416

- Just L (2010) The relation between evenness and diversity. *Diversity* 2:207–232
- Knopp K (1990) *Theory and Application of Infinite Series*. Dover Publications, New York
- Kvålseth TO (1985) Cautionary note about R^2 . *Am Stat* 39:279–285
- Kvålseth TO (1991) Note on biological diversity, evenness, and homogeneity. *Oikos* 62:123–127
- Kvålseth TO (1993) A measure of homogeneity for nominal categorical data. *Percept Mot Skills* 76:1129–1130
- Kvålseth TO (1995) Coefficients of variation for nominal and ordinal categorical data. *Percept Mot Skills* 80:843–847
- Kvålseth TO (2011) The lambda distribution and its applications to categorical summary measures. *Adv Appl Stat* 24:83–106
- Kvålseth TO (2014) Entropy evaluation based on value validity. *Entropy* 16:4855–4873
- Lorenz MO (1905) Methods for measuring concentration of wealth. *J Am Stat Assoc* 9:209–219
- Magurran AE (2004) *Measuring Biological Diversity*. Blackwell Publishing, Oxford, UK
- Marshall AW, Olkin I, Arnold BC (2011) *Inequalities: Theory of Majorization and its Applications*, 2nd edn. Springer, New York
- Molinari J (1989) A calibrated index for the measurement of evenness. *Oikos* 56:319–326
- Mueller JH, Schuessler KF (1977) *Statistical Reasoning in Sociology*, 3rd edn. Houghton Mifflin Co., Boston
- Peet RK (1975) Relative diversity indices. *Ecology* 56:496–498
- Pielou EC (1966) The measurement of diversity in different types of biological collections. *J Theor Biol* 13:131–144
- Pielou EC (1969) *An Introduction to Mathematical Ecology*. Wiley, New York
- Rényi A (1961) On measures of entropy and information. *Proc Fourth Berkeley Symposium on Math Stat Prob* 1960, Vol I, University of California Press, Berkeley, CA, pp. 547–561
- Reynolds HT (1984) *Analysis of Nominal Data*, 2nd edn. Sage, Beverly Hills, CA
- Ricotta C (2004) A recipe for unconventional evenness measures. *Acta Bioteoretica* 52:95–104
- Routledge RD (1983) Evenness indices: are any admissible? *Oikos* 40:149–151
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:279–423,623–656
- Smith B, Wilson JB (1996) A consumer's guide to evenness indices. *Oikos* 76:70–82
- Solomon DL (1979) A comparative approach to species diversity. In: Grassle JF, Patil GR, Smith W, Taillie C (eds) *Ecological diversity in theory and practice*. International Co-operative Publishing House, Fairland, Maryland, USA, pp 29–35
- Taillie C (1979) Species equitability: a comparative approach. In: Grassle JF, Patil GR, Smith W, Taillie C (eds) *Ecological diversity in theory and practice*. International Co-operative Publishing House, Fairland, Maryland, USA, pp 51–61
- Tuomisto H (2012) An updated consumer's guide to evenness and related indices. *Oikos* 121:1203–1218
- Upton G, Cook I (2008) *Oxford Dictionary of Statistics*. Oxford University Press, Oxford, UK

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com