Biodiversity Indexes

Índices de biodiversidad

(http://en.wikipedia.org/wiki/Diversity_index)

... Simpson index

The Simpson index was introduced in 1949 by Edward H. Simpson to measure the degree of concentration when individuals are classified into types. The same index was rediscovered by Orris C. Herfindahl in 1950. The square root of the index had already been introduced in 1945 by the economist Albert O. Hirschman. As a result, the same measure is usually known as the Simpson index in ecology, and as the Herfindahl index or the Herfindahl-Hirschman index (HHI) in economics.

The measure equals the probability that two entities taken at random from the dataset of interest represent the same type. It equals:

$$\lambda = \sum_{i=1}^{R} p_i^2$$

This also equals the weighted arithmetic mean of the proportional abundances p_i of the types of interest, with the proportional abundances themselves being used as the weights. Proportional abundances are by definition constrained to values between zero and unity, but their weighted arithmetic mean, and hence λ, can never be smaller than $1/S$, which is reached when all types are equally abundant.

By comparing the equation used to calculate λ with the equations used to calculate true diversity, it can be seen that $1/\lambda$ equals ${}^2D$, i.e. true diversity as calculated with $q = 2$. The original Simpson's index hence equals the corresponding basic sum.

... Índice de Simpson

El índice de Simpson fue propuesto en 1949 por Edward H. Simpson para medir el grado de concentración cuando los individuos se clasifican en tipos. Orris C. Herfindahl definió el mismo índice otra vez en 1950. El economista Albert O. Hirschman ya había introducido la raíz cuadrada del índice en 1945. Por ello el mismo índice es conocido normalmente como índice de Simpson en ecología e índice de Herfindahl o de Herfindahl-Hirschmann (IHH) en economía.

Expresa la probabilidad de que dos entidades tomadas al azar en la muestra pertenezcan al mismo tipo. Su expresión es:

$$\lambda = \sum_{i=1}^{R} p_i^2$$

También es igual a la media ponderada de las abundancias relativas p_i de los tipos estudiados, con las propias abundancias relativas haciendo de factores de ponderación. Las abundancias relativas están por definición acotadas entre los valores cero y uno, pero su media ponderada (y por ello también λ) nunca puede ser menor que 1/S, valor obtenido cuando todos los tipos están igualmente representados.

Al comparar la ecuación usada para calcular λ con las ecuaciones empleadas para hallar la diversidad real, se puede comprobar que $1/\lambda$ es igual a ${}^2D$, esto es, la diversidad real calculada cuando q=2. El índice original de Simpson es por lo tanto igual a la suma básica correspondiente.

The interpretation of λ as the probability that two entities taken at random from the dataset of interest represent the same type assumes that the first entity is replaced to the dataset before taking the second entity. If the dataset is very large, sampling without replacement gives approximately the same result, but in small datasets the difference can be substantial. If the dataset is small, and sampling without replacement is assumed, the probability of obtaining the same type with both random draws is:

$$l = \frac{\sum_{i=1}^{R} n_i(n_i - 1)}{N(N-1)}$$

where n_i is the number of entities belonging to the ith type and N is the total number of entities in the dataset. This form of the Simpson index is also known as the Hunter-Gaston index in microbiology.

Since mean proportional abundance of the types increases with decreasing number of types and increasing abundance of the most abundant type, λ obtains small values in datasets of high diversity and large values in datasets of low diversity. This is counterintuitive behavior for a diversity index, so often such transformations of λ that increase with increasing diversity have been used instead. The most popular of such indices have been the inverse Simpson index (1/λ) and the Gini-Simpson index (1 - λ). Both of these have also been called the Simpson index in the ecological literature, so care is needed to avoid accidentally comparing the different indices as if they were the same.

<u>Inverse Simpson index</u>

The inverse Simpson index equals:

$$1/\lambda = \frac{1}{\sum_{i=1}^{R} p_i^2} = {}^2D$$

This simply equals true diversity of order 2, i.e. the effective number of types that is obtained when the weighted arithmetic mean is used to quantify average proportional abundance of types in the dataset of interest.

En la interpretación de λ como la probabilidad de que dos entidades tomadas al azar de la muestra sean del mismo tipo, se asume que la primera entidad es repuesta en la muestra antes de extraer la segunda. Si la muestra es muy grande el muestreo sin reposición da aproximadamente el mismo resultado, pero en muestras pequeñas la diferencia puede ser importante. Si la muestra es pequeña y no se hace reposición, la probabilidad de extraer el mismo tipo en las dos extracciones aleatorias es:

$$l = \frac{\sum_{i=1}^{R} n_i(n_i - 1)}{N(N-1)}$$

donde n_i es el número de entidades que pertenecen al i-ésimo tipo y N es el número total de entidades en la muestra. Esta expresión del índice es llamada en microbiología índice de Hunter-Gaston.

Puesto que la abundancia relativa media de los tipos aumenta cuantos menos tipos haya y cuanto mayor sea la abundancia del tipo modal, λ da valores pequeños en muestras diversas y valores altos en muestras poco diversas. Esto va en contra de lo que intuitivamente se espera de un índice de diversidad, por lo que a menudo se han usado transformaciones de λ que crecen al crecer la diversidad. Las más utilizadas de dichas transformaciones han sido el índice inverso de Simpson (1/λ) y el índice de Gini-Simpson (1 – λ). Ambos se han denominado en ocasiones en textos de ecología índice de Simpson, lo que debe tenerse en cuenta para no comparar de forma accidental dos índices que no son en realidad el mismo.

<u>Índice inverso de Simpson</u>

El índice inverso de Simpson es igual a:

$$1/\lambda = \frac{1}{\sum_{i=1}^{R} p_i^2} = {}^2D$$

Esto es sencillamente la diversidad real de orden 2, esto es, el número efectivo de tipos que se obtiene cuando se usa la media ponderada para cuantificar la abundancia relativa media de los tipos en la muestra estudiada.

Gini-Simpson index

The original Simpson index λ equals the probability that two entities taken at random from the dataset of interest (with replacement) represent the same type. Its transformation $1 - \lambda$ therefore equals the probability that the two entities represent different types. This measure is also known in ecology as the probability of interspecific encounter (*PEI*) and the Gini-Simpson index. It can be expressed as a transformation of true diversity of order 2:

$$1 - \lambda = 1 - \sum_{i=1}^{R} p_i^2 = 1 - 1/{}^2D$$

The Gibbs-Martin index of sociology, psychology and management studies, which is also known as the Blau index, is the same measure as the Gini-Simpson index.

Berger-Parker index

The Berger-Parker index equals the maximum p_i value in the dataset, i.e. the proportional abundance of the most abundant type. This corresponds to the weighted generalized mean of the p_i values when q approaches infinity, and hence equals the inverse of true diversity of order infinity ($1/{}^\infty D$)...

Índice de Gini-Simpson

El índice original de Simpson λ es igual a la probabilidad de que dos entidades tomadas al azar de la muestra (con reposición) sean del mismo tipo. Su transformación $1-\lambda$ es por ello la probabilidad de que ambas entidades sean distintas. Esto es lo que se llama en ecología probabilidad de un encuentro interespecífico (*PEI*) o índice de Gini-Simpson. Expresado como transformación de la diversidad real de orden 2:

$$1 - \lambda = 1 - \sum_{i=1}^{R} p_i^2 = 1 - 1/{}^2D$$

Se corresponde con el índice de Gibbs-Martin, también conocido como índice de Blau, y que es usado en sociología, psicología y en gestión de empresas.

Índice de Berger-Parker

El índice de Berger-Parker es igual al valor máximo p_i en el conjunto, esto es, la abundancia relativa del tipo más común. Esto corresponde a la media ponderada generalizada de los valores p_i cuando q tiende a infinito, y por ello es igual al inverso de la diversidad real de orden infinito ($1/{}^\infty D$)...