Presence/ Absence similarities

There are numerous similarity measures defined for simple species lists, i.e. when the data consist only of presence (1) or absence (0) of each species in each sample. Any similarity defined between samples 1 and 2 must then be a combination of only four numbers: a, the number of species present in both samples; b, the number present in 1 but absent from 2; c, the number absent in 1 but present in 2; d, the number absent from both. Clearly, the coefficient must be symmetric in b and c, and the more biologically useful coefficients are also not a function of joint absences, d. There still remain a large number of options, of which PRIMER 7 calculates the following:

$S_1 = 100 \frac{a+d}{a+b+c+d} \text{\hspace{30mm} simple matching;} $

$S_2 = 100 \frac{a+d}{a+2b+2c+d} \text{\hspace{28mm} Rogers \& Tanimoto;} $

$S_5 = 25 \left[\frac{a}{a+b} + \frac{a}{a+c} + \frac{d}{b+d} + \frac{d}{c+d} \right] \text{;} $

$S_6 = 100 \frac{a}{\sqrt{(a+b)(a+c)}} \times \frac{d}{\sqrt{(b+d)(c+d)}} \text{;} $

$S_7 = 100 \frac{a}{a+b+c} \text{\hspace{34mm} Jaccard;} $

$S_8 = 100 \frac{2a}{2a+b+c} \text{\hspace{33mm} Sørensen;} $

$S_{11} = 100 \frac{a}{a+b+c+d} \text{\hspace{30mm} Russell \& Rao;} $

$S_{13} = 50 \left[\frac{a}{a+b} + \frac{a}{a+c} \right] \text{\hspace{25mm} Kulczynski (P/A);} $

$S_{14} = 100 \frac{a}{\sqrt{(a+b)(a+c)}} \text{\hspace{26mm} Ochiai (P/A);} $

$S_{26} = 100 \frac{a+(d/2)}{a+b+c+d} \text{\hspace{31mm} Faith;} $

A quantitative matrix input to one of these calculations will automatically be reduced to a simple array of 1’s and 0’s before computation. The most frequently met of the presence/absence measures are Sørensen, which is Bray-Curtis calculated on P/A data, and Jaccard – the definition shows how alike they are. In fact they are monotonically related (as one increases, so does the other), so the procedures in PRIMER which are based only on rank values of the coefficients (i.e. most of them: nMDS, ANOSIM, BEST, RELATE etc, in our largely non-parametric approach to resemblance matrix analysis) will give exactly the same outcome for these two coefficients.

Revision #29
Created 7 June 2024 01:25:43 by Arden
Updated 15 January 2025 02:26:28 by Abby Miller