‘Modified Gower’

Anderson MJ, Ellingsen KE, McArdle BH 2006, *Ecol Lett* 9: 683-693 used Czekanowski’s mean character difference (above) as their preferred distance measure after a specific transformation of the original counts, advocated for its interpretable properties, namely: $y^′ = \log(y) + 1$, unless $y = 0$, when $y^′ = 0$. Choice of the base for the logarithm explicitly scales how much weight the counts get in relation to the presence/absence structure. For example, base 2 gives the step from 0 (absence) to 1 (individual) the same weight as the step from 1 to 2, or from 2 to 4, or 4 to 8 etc. Base 10 gives 0 to 1 the same weight as 1 to 10, or 10 to 100 etc. Thus high bases give more weight to the presence/ absence structure. Thus, this work mainly concerns an added transformation choice rather than a new resemblance measure, but it is convenient to bundle the transformation with Czekanowski’s measure into a single coefficient, which the authors called *modified Gower* (though note that it avoids one of the defining, and usually problematic, features of the Gower coefficient S_{19}, below – that of standardising each species by its range of values across the samples). It is important to stress that the transform applies only to genuine counts (without other initial standardising/transforming). For densities, biomass, cover etc., the logic breaks down: y values can be less than 1, for which the transformed $y^′$ can be <0. Thus high densities give positive values for $y^′$ but low densities can give negative $y^′$ and an even lower density (absence) will give $y^′ = 0$ – the transform is not monotonic! To avoid this, any y values in (0,1) are initially rounded down or up to 0 or 1 before computation but this changes the number of perceived absences. Unless you are clear about the implications, the safest course is to use Modified Gower only for real counts – for which it is designed!

Revision #4
Created 6 June 2024 22:55:28 by Arden
Updated 15 January 2025 01:42:07 by Abby Miller