Modelling dependence in ecological community data

Shirley Pledger[a]

[a]Department of Mathematics, Statistics and Operations Research
Victoria University of Wellington
Wellington, New Zealand
shirley.pledger@vuw.ac.nz

Keywords: community structure and dynamics; multispecies models

Abstract: Species abundance data often occur as a species by sample (n by p) matrix of counts. Such a matrix may be used to provide ordination by multidimensional scaling or by correspondence analysis, with their associated low-dimensional plots. Traditional methods of analysis are mathematical, using distance metrics or matrix decomposition. Alternative likelihood-based methods using finite mixtures yield low-dimensional plots which are usually similar to those from the traditional methods, but have the added advantage of providing model comparisons and statistical decision-making (Pledger and Arnold, 2014). However, a major concern for ecologists is the assumption of independence in the model-based methods. The previous Poisson-based models are now modified to use the negative binomial distribution as the building block. Making the negative binomial dispersion parameter species-specific helps to account for the spatial clustering frequently observed within species. This talk briefly outlines the model-building, then compares the Poisson and negative binomial results using ecological community data. Possible model extensions to allow for dependence in the data caused by competition between species will be suggested, and discussion will be welcomed. All models are wrong, but useful ones are appreciated.

References
Pledger, S. and Arnold, R. (2014) Multivariate methods using mixtures: Correspondence analysis, scaling and pattern-detection. *Computational Statistics and Data Analysis*, 71:241-261.