A Generalized Diversity Index

A. J. Baczkowski
Department of Statistics, University of Leeds
Leeds LS2 9JT, UK
sta6ajb@leeds.ac.uk

1. Introduction

Suppose a population consists of s species with ordered relative abundances $\pi_1, \pi_2, \dots, \pi_s$. Good (1953) suggested measuring diversity using an index of the form

$$H(\alpha, \beta) = \sum_{i=1}^{s} \pi_i^{\alpha} \{-\ln(\pi_i)\}^{\beta},$$

defined for non-negative integer values of α and β. This attempted to give a general diversity measure which included as special cases both $H(1,1)$, Shannon's (1948) index, and $H(2,0)$, related to Simpson's (1949) index. Other diversity indices can be expressed in terms of $H(\alpha, \beta)$; for example, Hill's (1973) index, N_k, satisfies $N_k = H(k,0)^{1/(1-k)}$.

Baczkowski *et al.* (1997, 1998) further generalized Good's index by allowing (α, β) to take values in the real plane, and determined the range of values (α, β) for which $H(\alpha, \beta)$ satisfies two simple key properties of Pielou (1975, p.7), namely:
P1: for fixed s, the index increases as the relative abundances become more equal,
P2: for equal relative abundances, the index is an increasing function of s.
The index $H(\alpha, \beta)$ satisfies both properties P1 and P2 for a finite region $R_1 \subset \{0 < \alpha \leq 1, \beta \geq 0\}$. For the region $R_2 = \{\alpha > 1, \beta \leq 0\}$ an "inverting" transformation such as $1 - H(\alpha, \beta)$ or $1/H(\alpha, \beta)$ is necessary.

In practice a sample of size n is available, of which n_i are observed belonging to species i. The relative abundance of species i can be estimated using $p_i = n_i/n$ and the generalized diversity index estimated by $h(\alpha, \beta)$. The moments of $h(\alpha, \beta)$ are available and can be used to suggest suitable approximating distributions; see Baczkowski *et al.* (2000). Applications include simple tests of hypothesis for the diversity index.

It is found that the generalized index highlights different aspects of diversity in different regions of the (α, β) plane. This may be more useful than using a single diversity measure. Thus Shannon's index has been used to register PET and MR medical images, see for example, Studholme *et al.* (1996); using other parts of the (α, β) plane would allow different aspects of the two images to influence optimal registration.

2. Pooling property of the generalized index

Routledge (1979) suggested N_2 as the "best" measure of diversity. However, Gadagkar (1989) presented numerical evidence showing that N_2 does not possess a desirable pooling property, as do Shannon's index and Simpson's $\lambda = 1 - H(2,0)$, namely:

P3: if H_1 and H_2 are the values of a diversity index for two communities and H_P the diversity index for the two communites when pooled together, it is desirable that $H_P \geq \frac{1}{2}(H_1 + H_2)$.

Schur-concaveness of $H(\alpha, \beta)$ allows proof that it satisfies property P3 for $(\alpha, \beta) \in R_1$. However, for $(\alpha, \beta) \in R_2$ the transform $1/H(\alpha, \beta)$ is unsuitable whereas the complement transformation is acceptable. This conclusion generalizes the numerical work of Gadagkar. The conclusions can be readily extended to the case of pooling several communities.

REFERENCES

Baczkowski, A.J., Joanes, D.N. and Shamia, G.M. (1997). Properties of a generalized diversity index. *J.Theor.Biol.*, **188**, 207-213.

Baczkowski, A.J., Joanes, D.N. and Shamia, G.M. (1998). Range of validity of α and β for a generalized diversity index $H(\alpha, \beta)$ due to Good. *Math. Bioscience*, **148:2**, 115-128.

Baczkowski, A.J., Joanes, D.N. and Shamia, G. (2000). The distribution of a generalized diversity index due to Good. *Environmental and Ecological Statistics*, **7**, 329-342.

Gadagkar, R. (1989). An undesirable property of Hill's diversity index N_2. *Oecologia*, **80**, 140-141.

Good, I.J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, **40**, 237-264.

Hill, M.O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology*, **54**, 427-432.

Pielou, E.C. (1975). *Ecological Diversity. Wiley.*

Routledge, R.D. (1979). Diversity index: which ones are admissible? *J.Theor.Biol.*, **76**, 503-515.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell Sys.Tech.J.*, **27**, 379-423.

Simpson, E.H. (1949). Measurement of diversity. *Nature*, **163**, 688.

Studholme, C., Hill, D.L.G., Wong, J., Maisey, M.N., and Hawkes, D.J. (1996). Registration measures for automated 3D alignment of PET and intensity distorted MR images. In *Proceedings in Image Fusion and Shape Variability Techniques*, (eds. K.V. Mardia, C.A. Gill, I.L. Dryden), Leeds University Press, 186-193.

RESUME

Quelques propriétés d'un index de diversité sont presentée. La région dans laquelle l'index a une propriété de groupe de Gadagkar est montrée.