J. F. Grassle, G. P. Patil, W. Smith, and C. Taillie, (eds.),
Ecological Diversity in Theory and Practice, pp. 51-62.

SPECIES EQUITABILITY: A COMPARATIVE APPROACH

C. TAILLIE

Department of Statistics
The Pennsylvania State University
University Park, Pennsylvania 16802 USA

SUMMARY. The Lorenz curve, familiar in economics, is used to define an equitability ordering on the set of ecological communities. The ordering is only partial in the sense that two communities may not be comparable. Certain of Hill's (1973) ratios are consistent with this ordering but many of the other measures of equitability found in the ecological literature are inconsistent with the ordering.

KEY WORDS. diversity, evenness, species richness, Lorenz curve, gamma model, lognormal model.

1. INTRODUCTION

The concept of species diversity has undergone intensive study in recent years, with the overwhelming conclusion that no single diversity index adequately summarizes community structure. A more complete summarization is possible if, instead of a single index, one uses a parametric family of indices whose members have varying sensitivities to the rare and abundant species. Several such families have been proposed (e.g., Hill, 1973; Patil and Taillie, 1976; Smith and Grassle, 1977). Given such a parametric family, diversity can be plotted against the parameter and the resulting curves compared for the various communities under study. This technique is described in detail in Patil and Taillie (1979) and seems to work reasonably well if only a few communities are to be compared. But with larger numbers of communities it becomes increasingly difficult to disentangle the patterns and to identify groups of communities having similar structure. In this case a more convenient graphical display is possible if diversity is partitioned into a richness and an evenness component, each community then being represented by a single point in the richness-evenness plane. The scatter diagram might then be inspected visually or, perhaps, analyzed formally using techniques of classification or ordination. See Solem (1979) for an example of this approach.

Several methods of measuring evenness have been proposed in the ecological literature, the most popular of which seems to be Pielou's J', defined as the ratio $H' / \log s$ where H' is the Shannon index and s is the number of species in the community. Hill (1973) has defined a double continuum of evenness measures,

$$E_{a,b} = N_a/N_b, \quad -\infty < a,\ b < \infty, \quad a \neq b,$$

where $N_a = (\pi_1^a + \pi_2^a + \cdots + \pi_s^a)^{1/(1-a)}$ and $\pi_1, \pi_2, \cdots, \pi_s$ are the relative abundances of the different species. May (1974) refers to $E_{1,0} = \exp(H')/s$ as the MacArthur-Terborgh index although it appears to be due to Sheldon (1969). Solomon (1979) has proposed an index based upon the ordered relative abundances $\pi_1^\# \geq \pi_2^\# \geq \cdots \geq \pi_s^\#$,

$$I = [2\ \Sigma i\pi_i^\# - 2]/[s-1].$$

This index takes the value 1 for perfectly even communities and values less than 1 for all other communities. Engen (1977, 1979) has introduced the index

$$V = \Sigma\pi_i\ (\log \pi_i)^2 - (\Sigma\pi_i \log \pi_i)^2$$

as an *inverse* measure of evenness. Note that V is the variance of the logged abundances and is large when the abundances are very uneven.

The present note discusses the general concept of equitability (i.e., evenness) and assesses the above-mentioned methods for its measurement. We begin with the purely ordinal problem of ranking communities according to their evenness. Borrowing from the field of economics, a curve (known as the *Lorenz curve*) is associated with each community. A community with greater equitability has its Lorenz curve everywhere above a second, less equitable, community. In general, however, two Lorenz curves may intersect, in which case the communities are not comparable *vis a vis* equitability. We find that Hill's ratios $E_{a,b}$ are consistent with the Lorenz ordering if and only if $b = 0$. Further, all the other proposed "measures of equitability" described above are inconsistent with the Lorenz ordering. While these may be useful descriptive parameters for some purposes, in our opinion it is inappropriate to refer to them as measures of equitability.

The paper concludes with a discussion of equitability within the context of abundance models. Any two gamma models are comparable with respect to the Lorenz ordering and the exponent k in a gamma model is an equitability measure. Similarly, any two lognormal models are comparable and the logarithmic variance σ^2 is an *inverse* measure of equitability.

2. THE LORENZ CURVE

Overall, 2.8 percent of the property owners [in Brazil] control 66 percent of the privately held land.

Joseph B. Treaster
The Atlantic Monthly
August, 1979, p. 8

Questions involving equitability arise frequently in the social sciences (e.g., income inequality, tax burden, racial balance) and there is a vast literature on the subject. Pertinent references include Alker (1970), Dalton (1920), Hart (1975), Lorenz (1905), and Theil (1967). Typically these problems involve some "commodity" distributed amongst a population of units and the evenness (or unevenness) of the distribution is expressed by comparing population shares with commodity shares. The quotation at the beginning of this section illustrates the point.

The purpose of this paper is to suggest that this method can form the basis for an index-free definition of species equitability. Here the population units are species while the individual organisms comprise the "commodity." Now let us gradually accumulate individuals starting with members of the most abundant species. The *Lorenz curve* is obtained by plotting cumulative proportions of individuals as abscissae (X) against corresponding cumulative proportions of species as ordinates (Y). Formally, letting

$$\pi_1^\# \geq \pi_2^\# \geq \cdots \geq \pi_s^\#$$

be the ordered relative abundances, the Lorenz curve is the polygonal path joining the successive points

$$P_0 = (0,0),$$

$$P_1 = (\pi_1^{\#},\ 1/s),$$
$$P_2 = (\pi_1^{\#} + \pi_2^{\#},\ 2/s),$$
$$P_3 = (\pi_1^{\#} + \pi_2^{\#} + \pi_3^{\#},\ 3/s)$$
$$\vdots$$
$$P_s = (\pi_1^{\#} + \pi_2^{\#} + \cdots + \pi_s^{\#},\ s/s) \equiv (1,1).$$

Figure 1 displays a Lorenz curve for a hypothetical 5-species community. The 45° line in this figure represents the norm of perfect equality, achieved only by completely even communities. With this in mind, we may define an *intrinsic equitability ordering* in the following manner. Let C and C' be two communities. Then C is said to have greater equitability than C' (written $C \overset{e}{\geq} C'$) provided the Lorenz curve of C is everywhere above that of C'; on the other hand, if the two Lorenz curves cross one another, then the two communities are not intrinsically comparable.

It may be noticed that the Lorenz curve is very similar to the intrinsic diversity profile as defined by Patil and Taillie (1979a, Section 6). The latter uses as ordinate the cumulative *number* of species whilst the ordinate of the Lorenz curve is

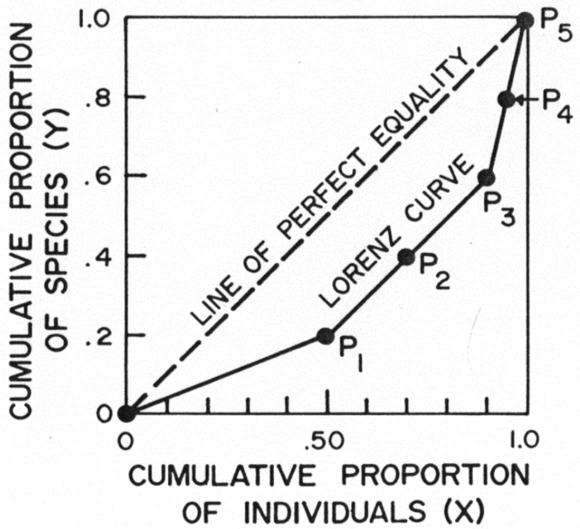

FIG. 1: Lorenz curve for the hypothetical five-species community with relative abundances 0.5, 0.2, 0.2, 0.05, 0.05.

the cumulative *proportion* of species. This has the important consequence that, when comparisons are restricted to communities with the same number of species, the intrinsic equitability ordering is identical to the intrinsic diversity ordering (cf. Solomon, 1979; Patil and Taillie, 1979a,b). This is as it should be since there is no essential difference between diversity and equitability when species richness is held constant.

How then does the Lorenz curve effect equitability comparisons when species richness varies? Consider two communities which are replicates of one another in the sense that they have the same relative abundance vectors, but no species in common. It seems plausible that combining these communities should give a community with the same evenness but twice the richness as either of the original communities. In other words,

$$C : \pi_1, \pi_2, \cdots, \pi_s$$

and $$C' : \pi_1/2, \pi_1/2, \pi_2/2, \pi_2/2, \cdots, \pi_s/2, \pi_s/2$$

are expected to have the same evenness. The Lorenz curves of C and C' are indeed the same; the effect of replication is simply to insert "extra" vertices such as P_2 in Figure 1. Now suppose we wanted to compare the evenness of two communities with, say, 3 and 5 species, respectively. We could replicate the 3-species community 5 times and the 5-species community 3 times and then carry out a *diversity* comparison between the pair of resulting 15-species communities. The Lorenz curves do all this automatically.

This interesting replication property was first pointed out by Hill (1973) in his discussion of the $E_{a,b}$ measures. Hill failed to notice the connection with the Lorenz curve however.

Note: In the social sciences literature, the Lorenz curve is usually taken to be the locus of points (1-Y, 1-X). Both definitions produce the same Lorenz ordering but the definition used above simplifies comparisons with the intrinsic diversity profile.

3. EQUITABILITY INDICES

A proposed measure of equitability E is said to be consistent with the Lorenz ordering provided $E(C) \geq E(C')$ whenever $C \overset{e}{\geq} C'$. The considerations of Section 2 lead to the following criteria for consistency: (i) replication should not change the value of E and (ii) E should be consistent with the intrinsic diversity ordering when restricted to communities with the same number of species. See Solomon (1979) or Patil and Taillie (1979a,b) for circumstances under which (ii) is satisfied.

The indices described in Section 1 can now be assessed against these criteria. The measure J' violates the principle of replication and is therefore inconsistent with the Lorenz ordering. Solomon's I also violates the replication principle, but a simple modification,

$$I' = [2\ \Sigma i\pi_i^{\#} - 1]/s,$$

is consistent with the Lorenz ordering. In fact I' is twice the area under the Lorenz curve and is often used as a measure of income equality under the name "Gini index" (cf. Alker, 1970, p. 200). The measure V, although it satisfies the replication principle, is known to be inconsistent with the Lorenz ordering even in the reverse sense (Dasgupta, Sen, and Starrett, 1973). Finally, we have already pointed out that Hill's ratios $E_{a,b}$ are consistent with the Lorenz ordering if and only if $b = 0$. The tedious but straightforward proof is omitted.

A fairly wide class of measures of unevenness has the form

$$U_\Psi = \frac{1}{s} \sum_{i=1}^{s} \Psi(s\pi_i) , \qquad (1)$$

where Ψ is an arbitrary convex function. The functional form of U_Ψ assures that the principle of replication is satisfied and convexity of Ψ guarantees property (ii) above (in the reverse sense). Measures of evenness would result if Ψ were concave, but the resulting measures generally take on negative values and it seems preferable to adopt some decreasing transformation of U_Ψ as the equitability measure. If we use absolute abundances λ_i instead of relative abundances, then $\pi_i = \lambda_i/\Sigma\lambda_i$ and $s\pi_i = \lambda_i/\bar{\lambda}$ where $\bar{\lambda}$ is the mean absolute abundance. Taking

$$\Psi(t) = \Psi_\gamma(t) = \frac{t^{\gamma+1} - t}{\gamma(\gamma+1)} , \qquad -\infty < \gamma < \infty,$$

equation (1) becomes

$$U_\gamma = \frac{s^{-1}\Sigma(\lambda_i/\bar{\lambda})^{\gamma+1} - 1}{\gamma(\gamma+1)} .$$

Putting $\gamma = 1$, we obtain

$$2U_1 = s^{-1}\ \Sigma(\lambda_i/\bar{\lambda})^2 - 1,$$

which is the squared coefficient of variation of the absolute abundances. It follows that the reciprocal of the squared coefficient of variation is a valid measure of equitability. Putting $\gamma = -1$ results in

$$U_{-1} = -s^{-1} \Sigma \log (\lambda_i/\bar{\lambda})$$
$$= -s^{-1} \Sigma \log (\lambda_i) + \log(\bar{\lambda}).$$

Hence $\exp(U_{-1})$ is the ratio AM/GM of the arithmetic and geometric means of the absolute abundance. Again, the reciprocal GM/AM is a valid equitability measure. Finally, $\gamma = -2$ gives

$$2U_{-2} + 1 = s^{-1} \Sigma(\lambda_i/\bar{\lambda})^{-1}.$$

which is the ratio AM/HM of the arithmetic and harmonic means of the absolute abundances. Consequently, HM/AM may be employed as an equitability measure.

4. EQUITABILITY AND ABUNDANCE MODELS

The abundance patterns in large species assemblages can usually be adequately described by a continuous distribution, that is, the number of species with (absolute) abundances greater than some positive value λ is the integral

$$\int_{\lambda}^{\infty} f(t)\ dt.$$

Since $\int_0^{\infty} f(t)dt$ gives the total number of species, the ordinate of the Lorenz curve is

$$Y = \int_{\lambda}^{\infty} f(t)\ dt/\int_0^{\infty} f(t)dt. \qquad (2)$$

The abscissa corresponding to this ordinate is

$$X = \int_{\lambda}^{\infty} t\ f(t)dt/\int_0^{\infty} tf(t)dt. \qquad (3)$$

As λ varies from 0 to ∞, equations (2) and (3) give a parametric representation of the Lorenz curve.

The two most popular abundance models are surely the gamma and lognormal, the former being fashionable amongst statisticians and the latter amongst ecologists. The species density function for the gamma model is

$$f(t) = s\, c^k t^{k-1} e^{-ct}/\Gamma(k)\ , \quad c,\ k > 0,$$

and equations (3) and (4) simplify to

$$Y = \int_{c\lambda}^{\infty} t^{k-1}\, e^{-t}\, dt/\Gamma(k)$$

and

$$X = \int_{c\lambda}^{\infty} t^k e^{-t} dt/\Gamma(k+1).$$

Lorenz curves for the gamma model are displayed in Figure 2 for several values of k. The Lorenz curve does not depend upon the scale parameter c. It may be noticed that these curves never cross one another and steadily increase as k increases. It follows that any two gamma models are comparable with respect to the Lorenz ordering and the exponent k may be described as a complete equitability parameter for the gamma family. A formal proof of this fact is quite difficult but follows along the lines of Patil and Taillie (1979a, Section 7). Notice also that, for the gamma distribution, k is reciprocal of the squared coefficient of variation which we have already observed to be a valid measure of evenness.

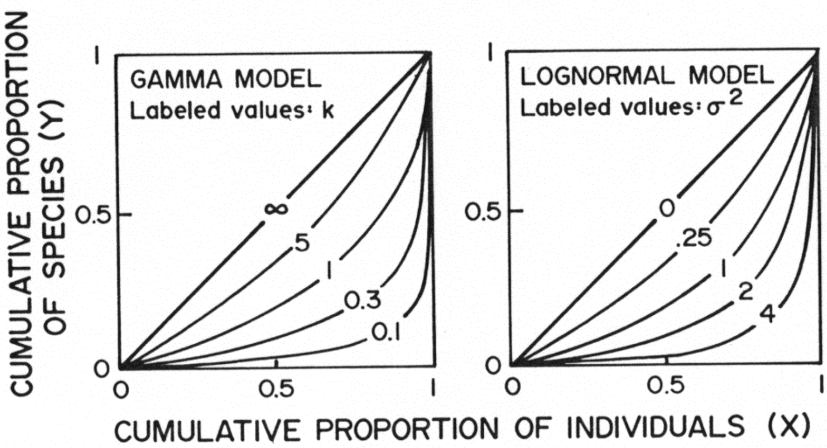

FIG. 2: Lorenz curves for the gamma and lognormal models.

For the lognormal model, the species density function is

$$tf(t) = s(2\pi\sigma^2)^{-\frac{1}{2}} \exp\{-(\log(t) - \mu)^2/2\sigma^2\},$$

where $\sigma^2 > 0$ and $-\infty < \mu < \infty$. The ordinate and abscissa of the Lorenz curve are

$$Y = 1 - \Phi(z)$$

and $$X = 1 - \Phi(z - \sigma)$$

where Φ denotes the standard normal cumulative distribution function and $z = (\log(t) - \mu)/\sigma$. Lorenz curves are plotted in Figure 2 for various values of σ^2. The different curves never cross one another and are steadily decreasing as σ^2 increases. Consequently, any two lognormal models are comparable with respect to the Lorenz ordering and $1/\sigma^2$ is a complete equitability parameter for the lognormal family.

REFERENCES

Alker, H. R. (1970). Measuring inequality. In *The Quantitative Analysis of Social Problems*, E. R. Tufte, ed. Addison-Wesley, Reading, Massachusetts. 191-211.

Dalton, H. (1920). The measurement of the inequality of incomes. *Economic Journal*, 30, 348-361.

Dasgupta, P., Sen, A., and Starrett, A. (1973). Notes on the measurement of inequality. *Journal of Economic Theory*, 6, 180-187.

Engen, S. (1977). Exponential and logarithmic species area curves. *American Naturalist*, 111, 591-594.

Engen, S. (1979). Some basic concepts of ecological equitability. In *Ecological Diversity in Theory and Practice*, J. F. Grassle, G. P. Patil, W. K. Smith, and C. Taillie, eds. International Co-operative Publishing House, Fairland, Maryland.

Hart, P. E. (1975). Moment distributions in economics: an exposition. *Journal of the Royal Statistical Society, Series A*, 138, 423-434.

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

Lorenz, M. C. (1905). Methods of measuring the concentration of wealth. *Journal of the American Statistical Association*, 9, 209-219.

May, R. M. (1974). General Introduction to *Ecological Stability*, M. B. Usher and M. H. Williams, eds. Chapman and Hall, London. 1-14.

Patil, G. P. and Taillie, C. (1976). Ecological diversity: concepts, indices, and applications. *Proceedings of the 9th International Biometric Conference*, Boston, Massachusetts. 383-411.

Patil, G. P. and Taillie, C. (1979). A study of diversity profiles and orderings for a bird community in the vicinity of Colstrip, Montana. In *Contemporary Quantitative Ecology and Related Ecometrics*, G. P. Patil and M. L. Rosenzweig, eds. International Co-operative Publishing House, Fairland, Maryland.

Patil, G. P. and Taillie, C. (1979a). An overview of diversity. In *Ecological Diversity in Theory and Practice*, J. F. Grassle, G. P. Patil, W. K. Smith, and C. Taillie, eds. International Co-operative Publishing House, Fairland, Maryland.

Patil, G. P. and Taillie, C. (1979b). Diversity as a concept and its measurement. *Journal of the American Statistical Association* (in revision).

Sheldon, A. L. (1969). Equitability indices: dependence on the species count. *Ecology*, 50, 466-467.

Smith, W. K. and Grassle, J. F. (1977). Sampling properties of a family of diversity indices. *Biometrics*, 33, 283-292.

Solem, J. (1979). A comparison of species diversity indices in Trichoptera communities. In *Ecological Diversity in Theory and Practice*, J. F. Grassle, G. P. Patil, W. K. Smith, and C. Taillie, eds. International Co-operative Publishing House, Fairland, Maryland.

Solomon, D. L. (1979). A comparative approach to species diversity. In *Ecological Diversity in Theory and Practice*, J. F. Grassle, G. P. Patil, W. K. Smith, and C. Taillie, eds. International Co-operative Publishing House, Fairland, Maryland.

Theil, H. (1967). *Economics and Information Theory*. North-Holland Publishing Co., Amsterdam.

[*Received October* 1979]