Biodiversity measures and the role of species similarity

Christina Cobbold
University of Glasgow, Boyd Orr Centre for Population and Ecosystem Health

University of Glasgow

Quantifying diversity

- Richness: How many species?

Species	Site 1	Site 2	Site 3
	1	1	10000
	1	10	1
	1	10	1
	10	1	1

- All sites have 4 species – how do we distinguish them?

University of Glasgow

Quantifying diversity

- Richness: How many species?

Species	Site 1	Site 2	Site 3
	1	1	10000
	1	10	1
	1	10	1
	10	1	1

- All sites have 4 species – how do we distinguish them?
- *Answer:* Include relative contribution of a species to the community

- Diversity captures richness, eveness and dominance

Hill numbers and the spectrum of view points

- Let p_i be the relative abundance of the ith species. $\mathbf{p}=(p_1, p_{2\ldots} p_s)$. The diversity of order q of the community is exp (Rényi entropy):

$$ {}^{q}D(p) = \begin{cases} \left(\sum_{i=1}^{S} p_i^q \right)^{1/(1-q)} & \text{if } q \neq 1, \infty \\ 1/(p_1^{p_1} p_2^{p_2} \cdots p_S^{p_S}) & \text{if } q = 1 \\ 1/\max_{1 \le i \le S} p_i & \text{if } q = \infty \end{cases} $$

University of Glasgow

Jost (2006) Oikos

Why take the exponential of entropy?

- Consider a continent of 1 million equally abundant species. A meteorite hits the continent wiping out 50% of the species. What happens to diversity?

- *Expected answer:* Diversity drops by 50%

University of Glasgow

Jost (2006) Oikos

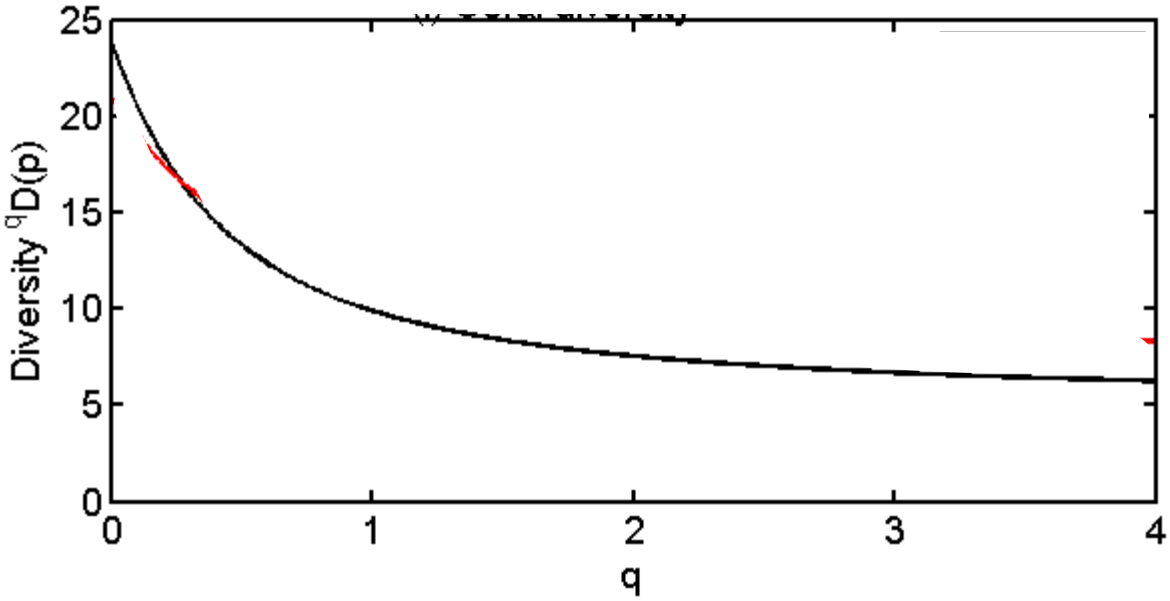
Why take the exponential of entropy?

- Consider a continent of 1 million equally abundant species. A meteorite hits the continent wiping out 50% of the species. What happens to diversity?

- *Expected answer:* Diversity drops by 50%
- BUT… Shannon entropy – drops by just 5%!
- Entropies are not effective numbers, Hill numbers are.

University of Glasgow

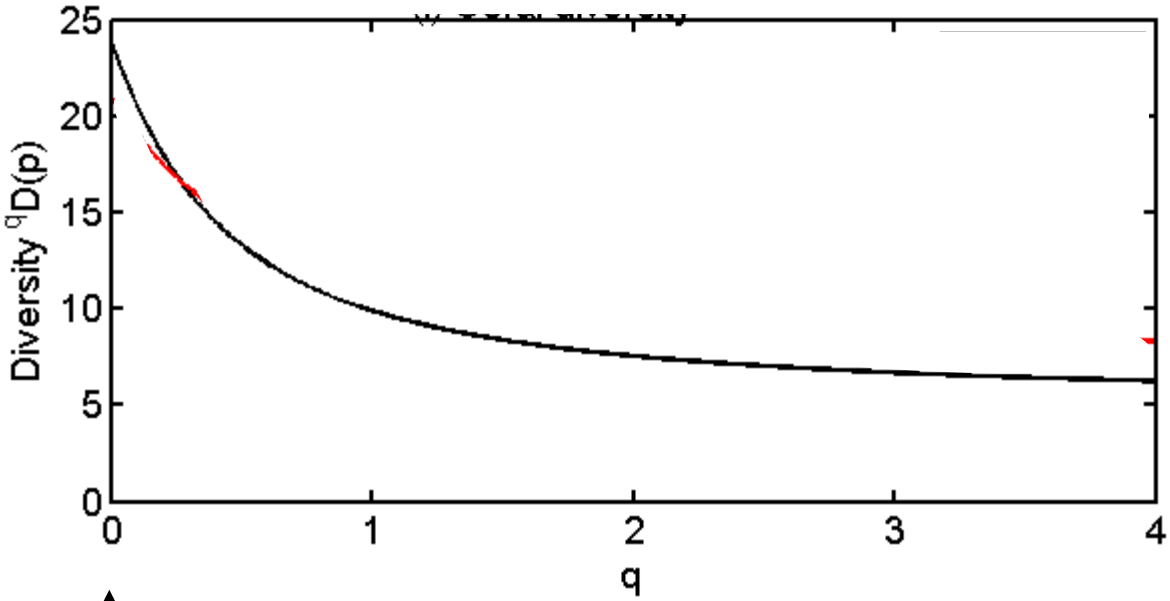
What is the role of q? Diversity profile



‘view point parameter’

University of Glasgow

What is the role of q? Diversity profile



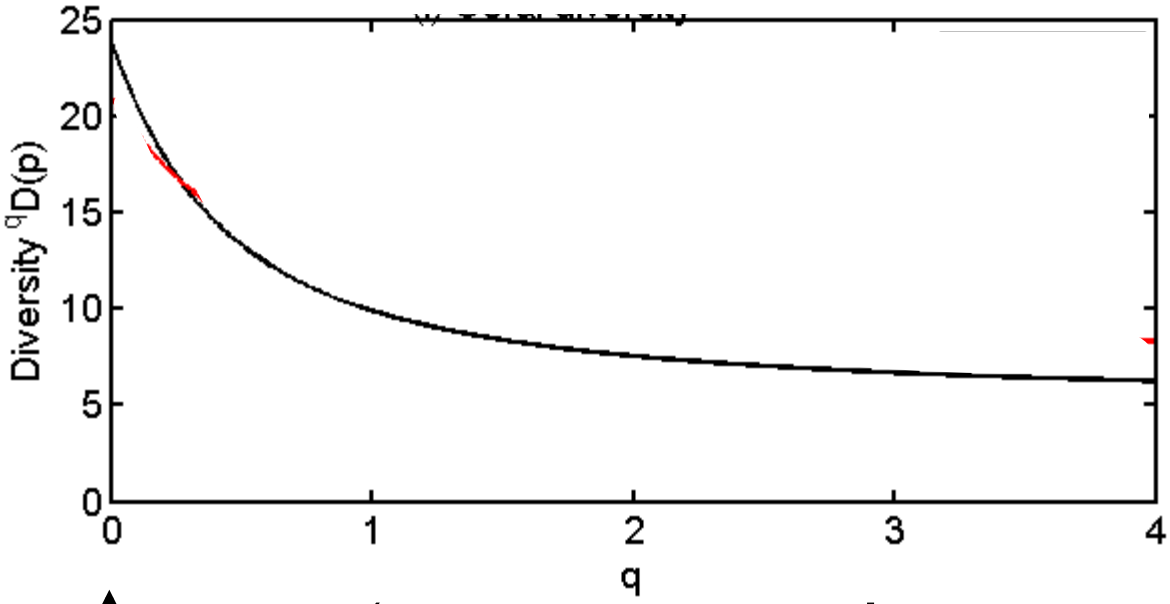
‘view point parameter’

q=0, species richness

rare species are important

University of Glasgow

What is the role of q? Diversity profile



'view point parameter'

q=0, species richness
rare species are important

q=infinity, 1/Berger-Parker
rare species are unimportant

'Importance of evenness'

University of Glasgow

What are some of the limitations of Hill numbers

- "…associated with the idea of diversity is the concept of 'distance' i.e. some measure of the dissimilarity of the resources in question"

OECD Handbook on biodiversity valuation: A guide for policy makers

- "A community of ten species of barnacle should clearly be less diverse than a community of ten very different species"

The essence of a point made by E.C. Pielou, Ecological Diversity

- There is not always a good notion of what constitutes a species e.g. microbial communities

University of Glasgow

Leinster and Cobbold (2012) Ecology

Including species similarity

model of community

Similarity data, **Z** Abundance data, **p**

S × S matrix (S = number of species)
Z_{ij} = similarity between ith and jth species

$$0 \leq Z_{i,j} \leq 1$$

totally dissimilar ← 0 ≤ $Z_{i,j}$ ≤ 1 → *identical*

University of Glasgow

Leinster and Cobbold (2012) Ecology

Including species similarity

model of community

Similarity data, **Z** Abundance data, **p**

S × S matrix (S = number of species)
Z_{ij} = similarity between ith and jth species

$0 \leq Z_{i,j} \leq 1$

totally dissimilar ← 0 ≤ $Z_{i,j}$ ≤ 1 → *identical*

Example

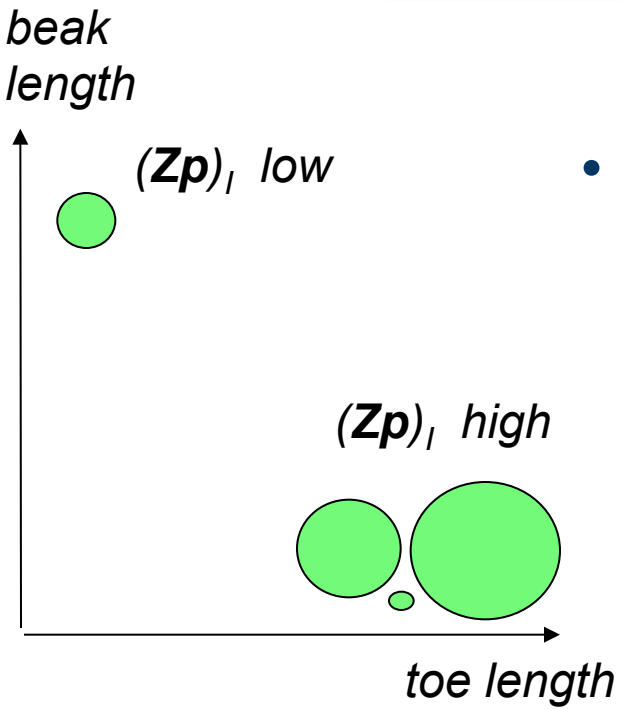
This model assumes that distinct species are totally dissimilar

$$\mathbf{Z} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

University of Glasgow

Leinster and Cobbold (2012) Ecology

Similarity-sensitive diversity measures



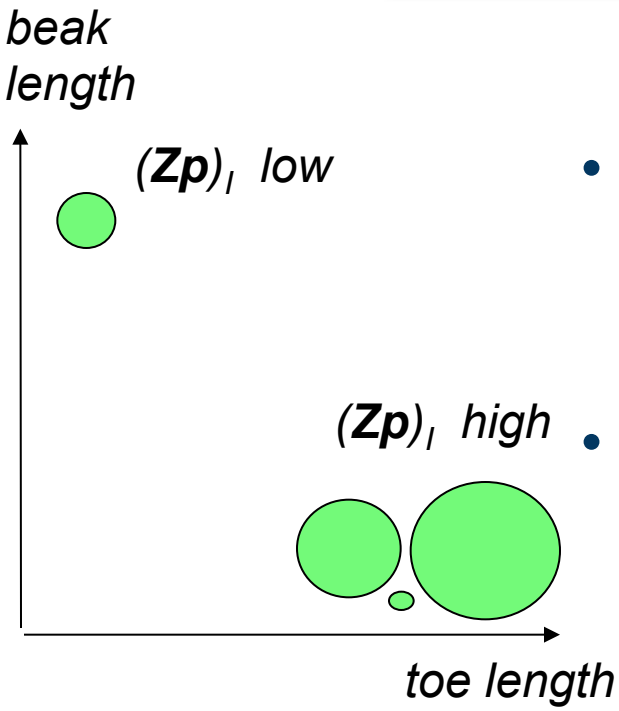
- The ordinariness of the *i*-th species is

$$(Zp)_i = \sum_{j=1}^{S} Z_{ij} p_j$$

University of Glasgow

Leinster and Cobboold (2012) Ecology

Similarity-sensitive diversity measures



- The ordinariness of the *i*-th species is

$$(Zp)_i = \sum_{j=1}^{S} Z_{ij} p_j$$

- Average ordinariness of an individual within the community

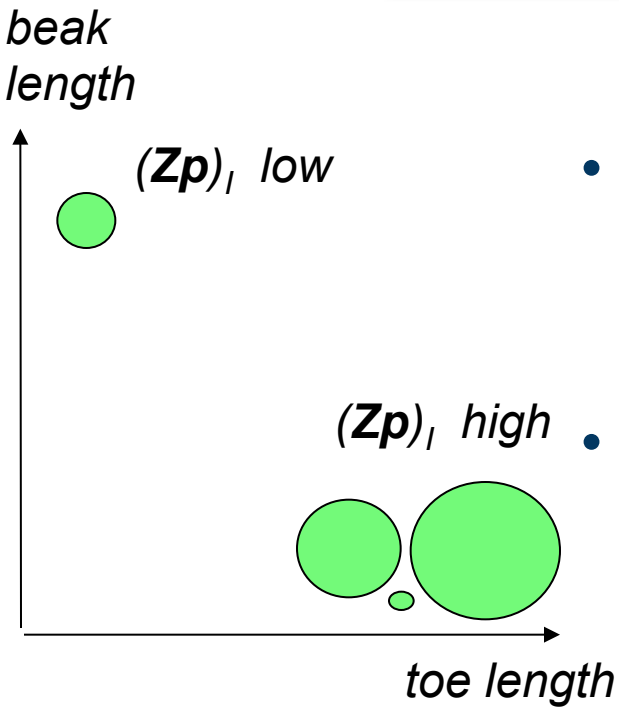
$$\sum_{i=1}^{S} p_i (Zp)_i$$

which measures lack of diversity.

University of Glasgow

Leinster and Cobbold (2012) Ecology

Similarity-sensitive diversity measures



- The ordinariness of the *i*-th species is

$$(Zp)_i = \sum_{j=1}^{S} Z_{ij} p_j$$

- Average ordinariness of an individual within the community

$$\sum_{i=1}^{S} p_i (Zp)_i$$

which measures lack of diversity.

- So one measure of diversity is

$$1 \Big/ \sum_{i=1}^{S} p_i (Zp)_i$$

University of Glasgow

Leinster and Cobboold (2012) Ecology

Spectrum of view points: general q

- Generalised weighted power mean of $x=(x_1 \ldots x_S)$ has the form

$$\left(\sum_{i=1}^{S} p_i x_i^{\,q-1} \right)^{1/(q-1)}$$

- So the generalised power mean of ordinariness is

$$\left(\sum_{i=1}^{S} p_i (Zp)_i^{q-1} \right)^{1/(q-1)}$$

University of Glasgow

Leinster and Cobbold (2012) Ecology

Spectrum of view points: general q

- Generalised weighted power mean of $x=(x_1 \ldots x_S)$ has the form

$$\left(\sum_{i=1}^{S} p_i x_i^{q-1} \right)^{1/(q-1)}$$

- So the generalised power mean of ordinariness is

$$\left(\sum_{i=1}^{S} p_i (Zp)_i^{q-1} \right)^{1/(q-1)}$$

- Diversity of order q is 1/ generalised power mean

$$^{q}D^{\mathbf{Z}}(\mathbf{p}) = \left(\sum_{i:p_i>0} p_i (Z\mathbf{p})_i^{q-1} \right)^{\frac{1}{1-q}}$$

University of Glasgow

Example: Butterfly data in the Ecuadorian rainforest

Abundance data of Charaxinae Butterflies at a rainforest site in Ecuador

Species	Canopy	Understorey
Prepona laertes	15	0
Archaeoprepona demophon	14	37
Zaretis itys	25	11
Memphis arachne	89	23
Memphis offa	21	3
Memphis xenocles	32	8

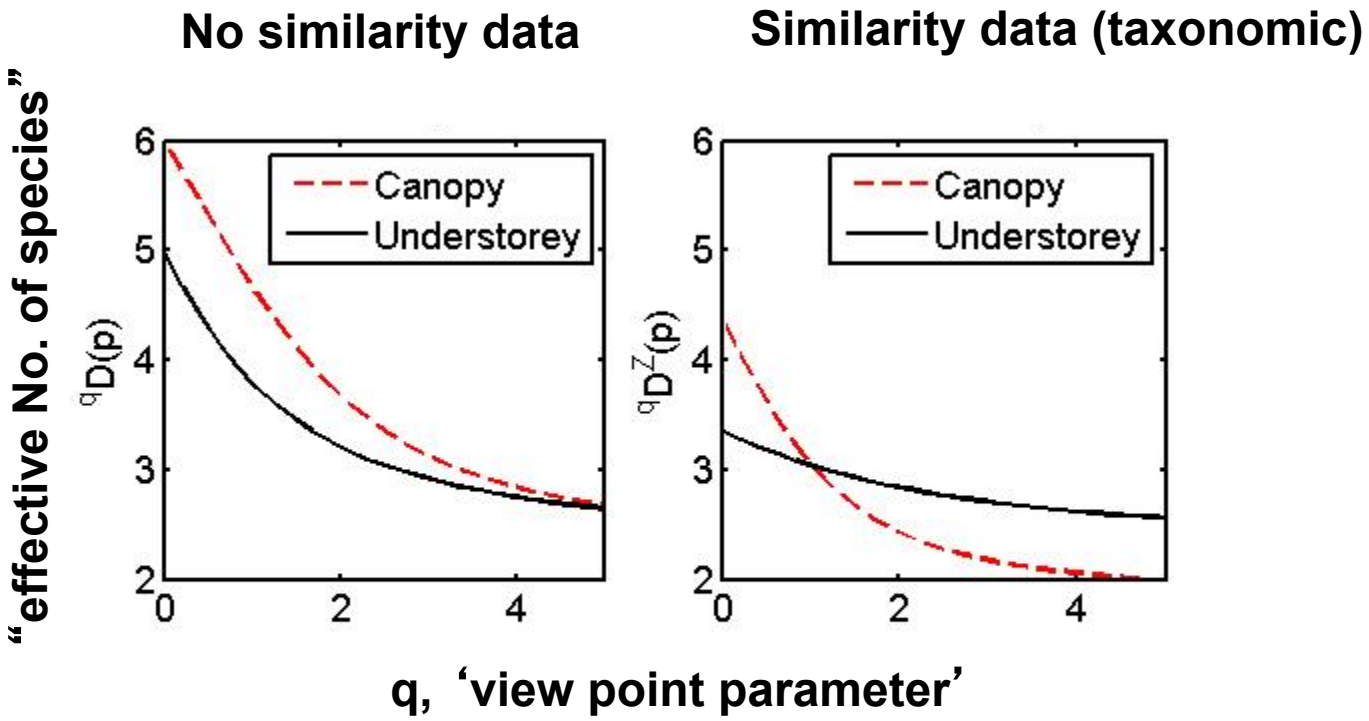
Taxonomic similarity matrix

$$Z_{ij} = \begin{cases} 0 & \text{if of different genera} \\ 0.5 & \text{if different but congeneric} \\ 1 & \text{if i = j,} \end{cases}$$

Data from DeVries et al (1997)

University of Glasgow

Butterfly data in the Ecuadorian rainforest



Data from DeVries et al (1997)

University of Glasgow

Comparing sites:

	Site 1	…	Site k	….	Ecosystem
Species 1			p_1		g_1
Species 2			p_2		g_2
Species 3			p_3		g_3
Species 4			p_4		g_4
…			…		…
Species S			p_S		g_S

- *How does diversity of site k contribute to the ecosystem diversity?*
- *Is site k typical or distinct in some way?*

University of Glasgow

Reeve et al http://arxiv.org/abs/1404.6520

Average ordinariness of site j from the ecosystem point of view

- The ordinariness of the *i*-th species in the *ecosystem*

$$(Zg)_i = \sum_{j=1}^{S} Z_{ij} g_j$$

University of Glasgow

Reeve et al http://arxiv.org/abs/1404.6520

Average ordinariness of site j from the ecosystem point of view

- The ordinariness of the *i*-th species in the *ecosystem*

$$(Zg)_i = \sum_{j=1}^{S} Z_{ij} g_j$$

- Generalised power mean of ordinariness of species in site k when viewed from the ecosystem perspective:

$$\left(\sum_{i=1}^{S} p_i (Zg)_i^{q-1} \right)^{1/(q-1)}$$

University of Glasgow

Reeve et al http://arxiv.org/abs/1404.6520

Average ordinariness of site j from the ecosystem point of view

- The ordinariness of the *i*-th species in the *ecosystem*

$$(Zg)_i = \sum_{j=1}^{S} Z_{ij} g_j$$

- Generalised power mean of ordinariness of species in site k when viewed from the ecosystem perspective:

$$\left(\sum_{i=1}^{S} p_i (Zg)_i^{q-1} \right)^{1/(q-1)}$$

- Renyi's cross entropy

$$(1-q)^{-1} \ln \left(\sum_{i=1}^{S} p_i (Zg)_i^{q-1} \right)$$

University of Glasgow

"Cross diversity" – exp (cross entropy)

- Which site/s has highest cross diversity?

Species	Site 1	Site 2	Site 3	Ecosystem (my back garden)
	1	1	10000	10002
	1	10	1	12
	1	10	1	12
	10	1	1	12

University of Glasgow

"Cross diversity" – exp (cross entropy)

- Which site/s has highest cross diversity?

Species	Site 1	Site 2	Site 3	Ecosystem (my back garden)
	1	1	10000	10002
	1	10	1	12
	1	10	1	12
	10	1	1	12

University of Glasgow

Reeve et al http://arxiv.org/abs/1404.6520

Relationship between entropy and cross entropy

- For Z=I and q=1

Relative entropy=Cross entropy – Entropy

$$(1-q)^{-1}\ln\left(\sum_{i=1}^{S} p_i (Zp/Zg)_i^{q-1}\right) = (1-q)^{-1}\ln\left(\sum_{i=1}^{S} p_i (Zg)_i^{q-1}\right) - (1-q)^{-1}\ln\left(\sum_{i=1}^{S} p_i (Zp)_i^{q-1}\right)$$

University of Glasgow

Reeve et al http://arxiv.org/abs/1404.6520

Relationship between entropy and cross entropy

- For Z=I and q=1

Relative entropy=Cross entropy – Entropy

$$(1-q)^{-1}\ln\left(\sum_{i=1}^{S} p_i (Zp/Zg)_i^{q-1}\right) = (1-q)^{-1}\ln\left(\sum_{i=1}^{S} p_i (Zg)_i^{q-1}\right) - (1-q)^{-1}\ln\left(\sum_{i=1}^{S} p_i (Zp)_i^{q-1}\right)$$

- This does not hold for general q and Z, but does relative entropy tell us something about diversity?

University of Glasgow

Relative entropy and ordinariness

- The ordinariness of the *i*-th species at *site j* compared to the ordinariness of the *i*-th species in the *ecosystem*

$$(Zp / Zg)_i = \sum_{j=1}^{S} Z_{ij} p_j \Big/ \sum_{j=1}^{S} Z_{ij} g_j$$

University of Glasgow

Relative entropy and ordinariness

- The ordinariness of the *i*-th species at *site j* compared to the ordinariness of the *i*-th species in the *ecosystem*

$$(Zp / Zg)_i = \sum_{j=1}^{S} Z_{ij} p_j \Big/ \sum_{j=1}^{S} Z_{ij} g_j$$

- *Averaging relative ordinariness* weighted by relative abundance of species i at site j

$$\left(\sum_{i=1}^{S} p_i (Zp/Zg)_i^{q-1} \right)^{1/(q-1)}$$

- Relative diversity = 1/ Av relative ordinariness

University of Glasgow

"Relative diversity" – exp (relative entropy)

- Which site/s has highest relative diversity?

Species	Site 1	Site 2	Site 3	Ecosystem (my back garden)
	1	1	10000	10002
	1	10	1	12
	1	10	1	12
	10	1	1	12

University of Glasgow

"Relative diversity" – exp (relative entropy)

- Which site/s has highest relative diversity?

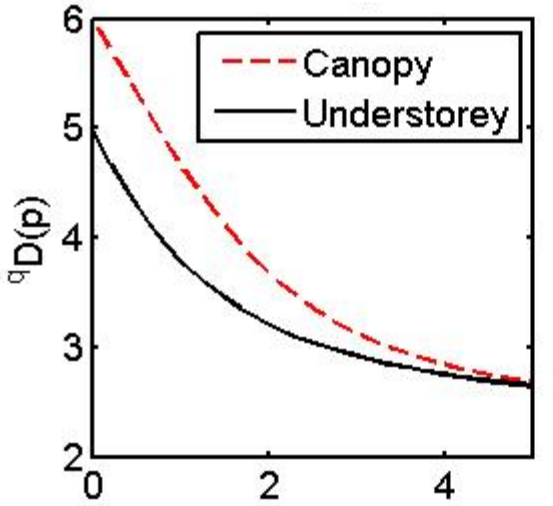
Species	Site 1	Site 2	Site 3	Ecosystem (my back garden)
	1	1	10000	10002
	1	10	1	12
	1	10	1	12
	10	1	1	12

University of Glasgow

Summary

- Generalised entropy allows us to consider evenness and dominance

Species	Site 1	Site 2	Site 3
	1	1	10000
	1	10	1
	1	10	1
	10	1	1



- Similarity sensitive diversity – functional diversity, genetic diversity etc
- Cross entropy allows us to consider sites containing “rarity”
- Relative entropy allows to consider “distinctness” of sites

University of Glasgow

Acknowledgements to….

Tom Leinster
School of Mathematics
University of Edinburgh

Louise Matthews
Boyd Orr Centre
University of Glasgow

Richard Reeve
Boyd Orr Centre
University of Glasgow

University of Glasgow

Microbe communities: notion of a species is problematic

- Expected similarity between a randomly chosen pair of individuals:

$$1/{}^{2}D^{Z}(p) = \mu_2$$

- Now consider q>2 (whole number). Given q individuals of species $i_1, i_2 \ldots, i_q$ *then a measure of group similarity is:*

$$Z_{i_1 i_2} Z_{i_1 i_3} \cdots Z_{i_1 i_q}$$

- *Let μ_q be the expected similarity of a randomly chosen group of q individuals (sampled with replacement). Then*

$${}^{q}D^{Z}(p) = \mu_q^{1/(1-q)}$$

University of Glasgow

Properties of diversity of order q

Elementary properties:

1. If two species are **nearly identical**, then merging them into one leaves the diversity nearly unchanged.
2. Diversity is unchanged by adding a new species of abundance 0
3. The order we list species in does not change diversity.

University of Glasgow

Properties of diversity of order q

Partitioning properties:

4. It is an **effective number**: diversity of a community of S equally abundant, totally dissimilar species is S.

5. **Replication**: If *m* islands of equal size each have diversity *d*, and species on different islands are totally dissimilar, the the diversity of the whole is *md*

6. **Modularity:** If the *m* islands have different sizes and diversities (species on different islands totally dissimilar), the diversity of the whole is determined by the sizes and diversities of the islands

University of Glasgow

Properties of diversity of order q

Similarity properties:

7. *Monotonicity*: increasing species similarity decreases diversity
8. *Naïve model*: ignoring species similarity increases diversity.
9. *Range:* The diversity of a community of S species is between 1 and S