Maximizing biological diversity

Tom Leinster

School of Mathematics
University of Edinburgh

Boyd Orr Centre
for Population and Ecosystem Health
University of Glasgow

Where we're going

I'll show you a maximum entropy theorem...

... with entropy interpreted as biodiversity...

... but maybe you can interpret the theorem in other interesting ways.

Plan:

1. Measuring biological diversity

2. The theorem

3. Examples and consequences

4. Unanswered questions

1. Measuring biological diversity

joint with Christina Cobbold

A spectrum of viewpoints on biodiversity

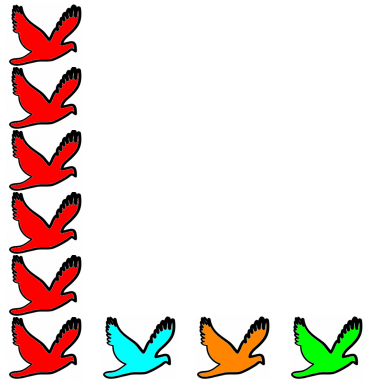
Conserving *species*
is what matters

Rare species
count for as much
as common ones
—every species is precious

Conserving *communities*
is what matters

Common species
are the really
important ones
—they shape the community

This ⟶

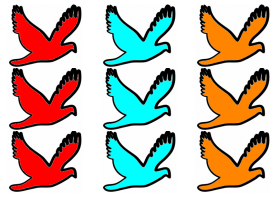


⟵ This

is more diverse than

is less diverse than

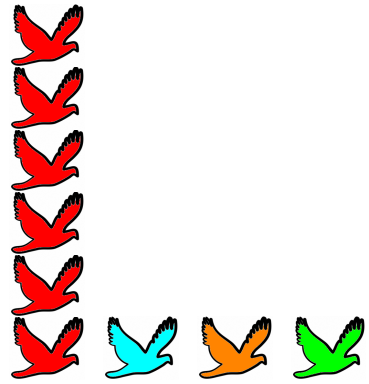
that ⟶



⟵ that

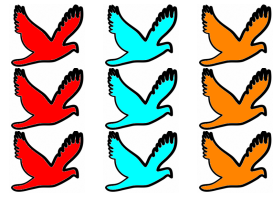
A spectrum of viewpoints on biodiversity

Rare species are important — *Rare species are unimportant*

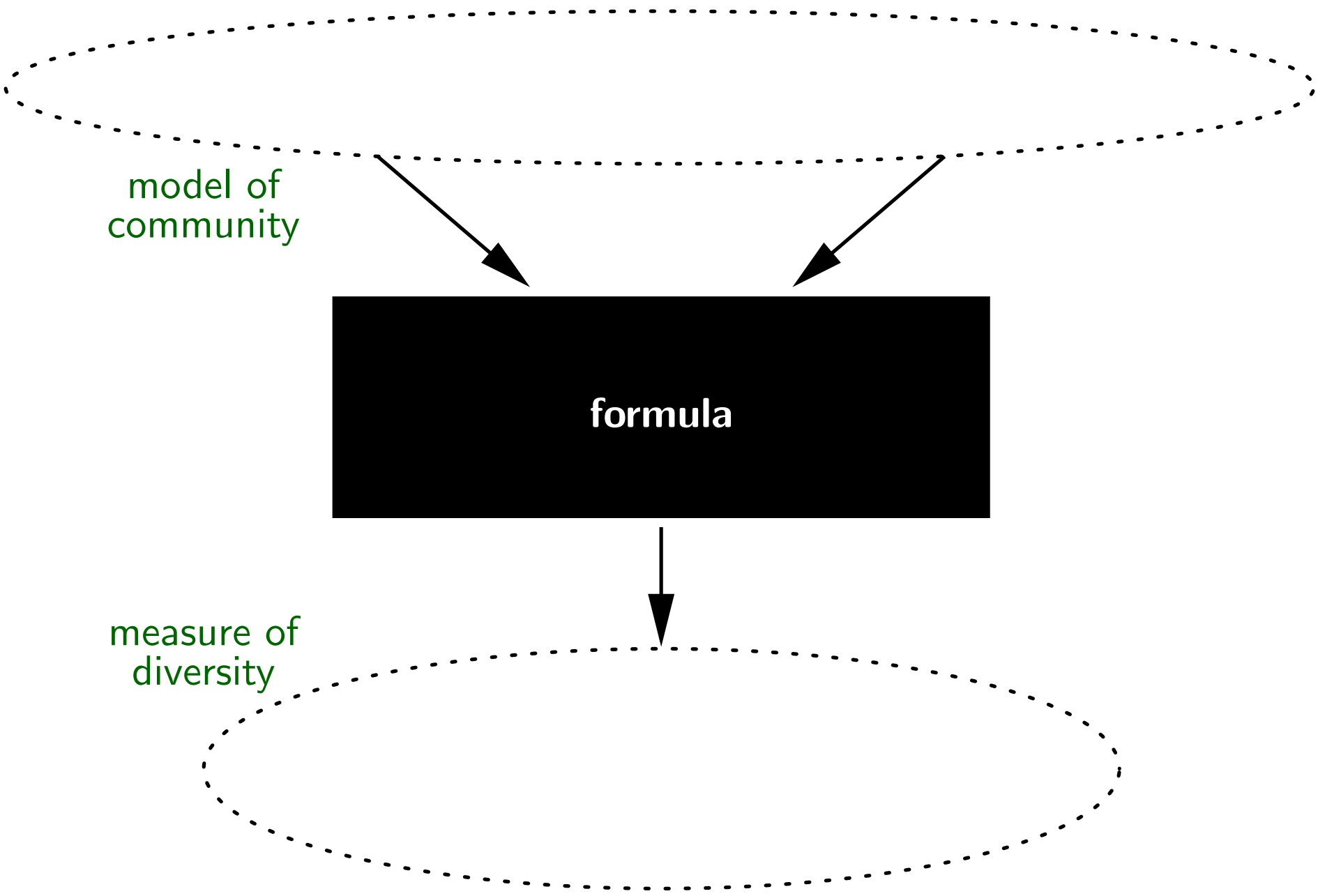


This ⟶ is more diverse than ⟶ that

⟵ This is less diverse than ⟵ that



Quantifying diversity



Quantifying diversity

Quantifying diversity

similarity matrix Z

$n \times n$ matrix (n = number of species)

Z_{ij} = similarity between ith and jth species = Z_{ji}

$0 \leqslant Z_{ij} \leqslant 1$ and $Z_{ii} = 1$

totally dissimilar

identical

E.g.: Naive model: $Z = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ (species have *nothing* in common)

E.g.: Genetic similarity.

E.g.: Taxonomic: e.g. $Z_{ij} = \begin{cases} 1 & \text{if same species} \\ 0.7 & \text{if different species but same genus} \\ 0 & \text{otherwise.} \end{cases}$

Quantifying diversity

Quantifying diversity

similarity matrix Z

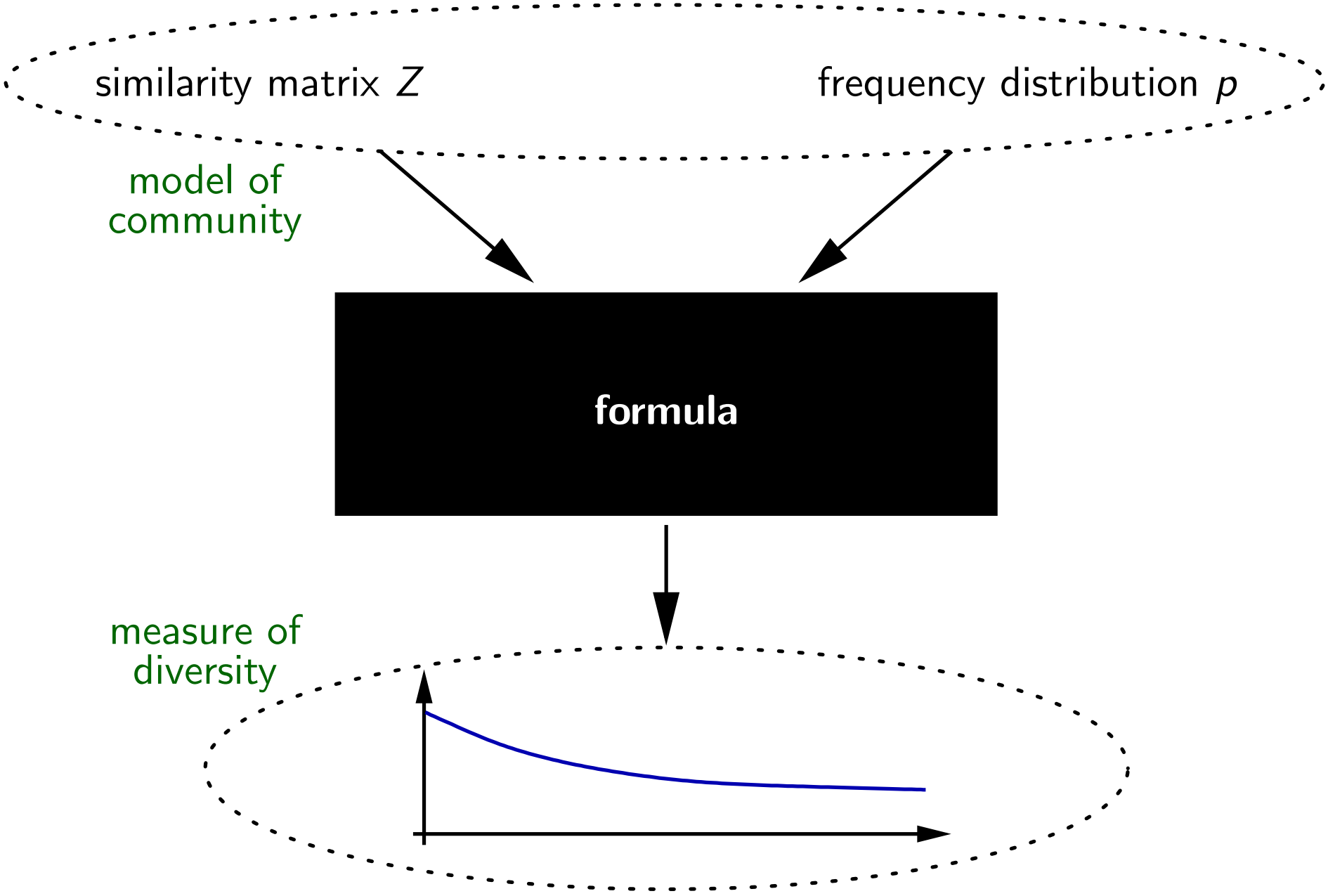
frequency distribution p

$$p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$$

p_i = relative frequency, or relative abundance, of the ith species

$p_i \geqslant 0$ and $\sum p_i = 1$

Quantifying diversity



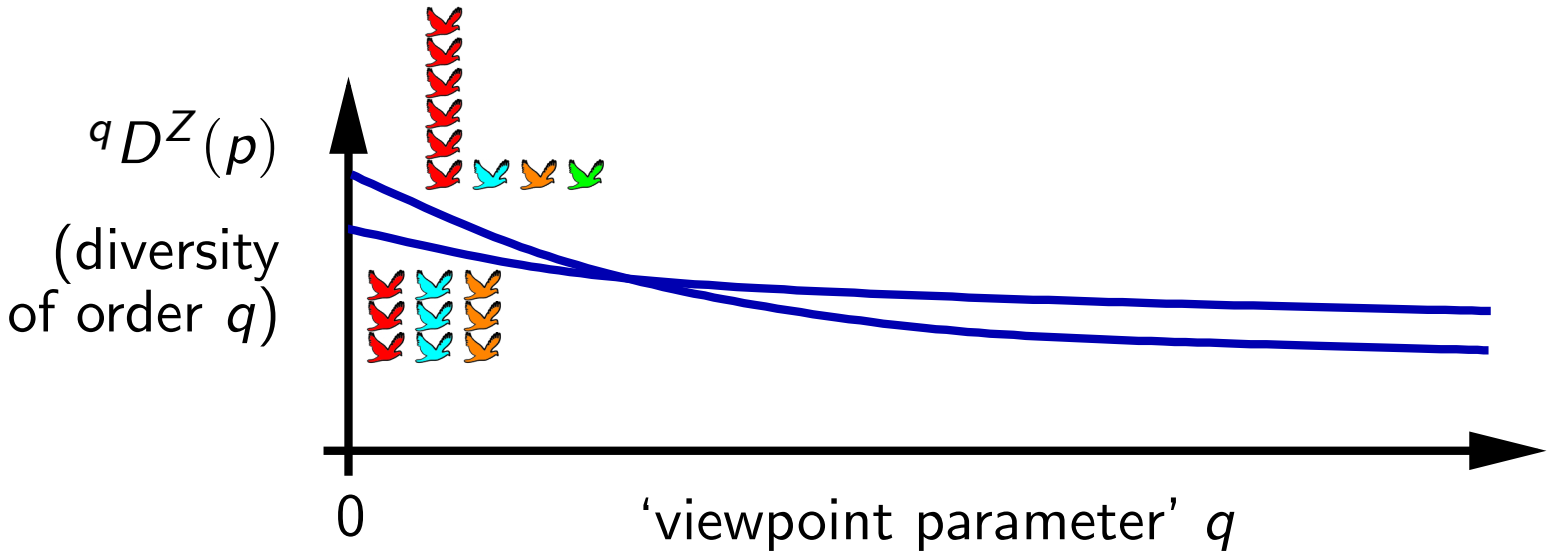
Quantifying diversity

similarity matrix Z

frequency distribution p

$q = 0$:
rare species are
important
↓

$q = \infty$:
rare species are
unimportant
↓



Quantifying diversity

similarity matrix Z

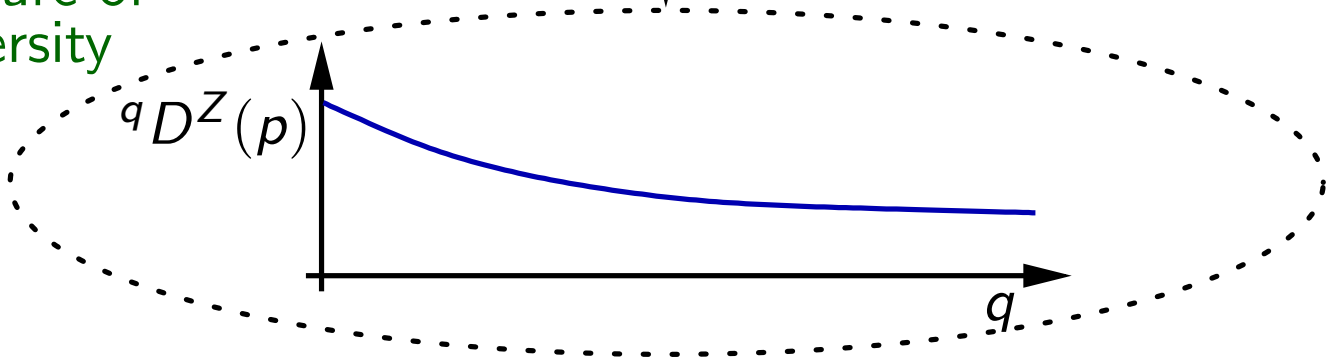
frequency distribution p

model of community

The diversity of order q is

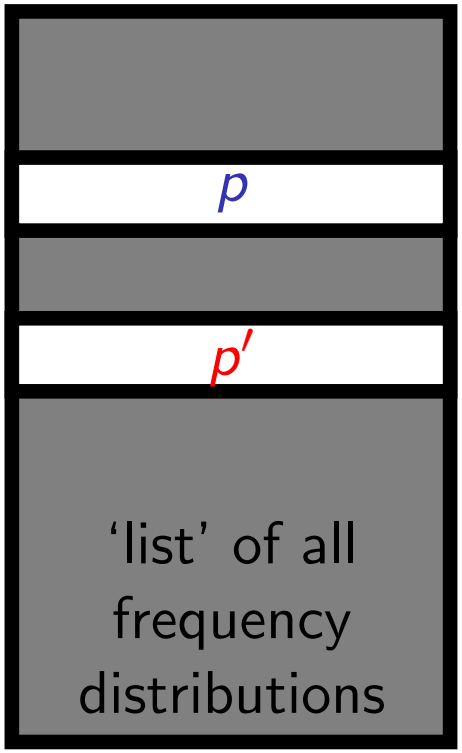
$$ {}^qD^Z(p) = \left(\sum_{i:\, p_i > 0} p_i (Zp)_i^{q-1} \right)^{\frac{1}{1-q}} $$

measure of diversity



Visualizations

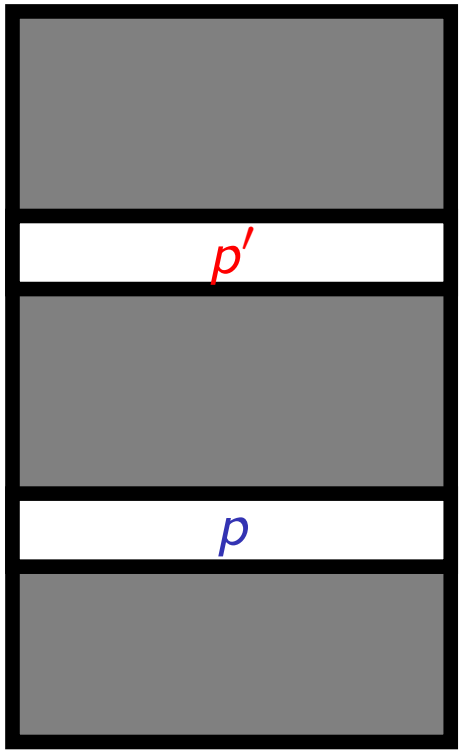
$q = 0$
(sensitive to rare species)

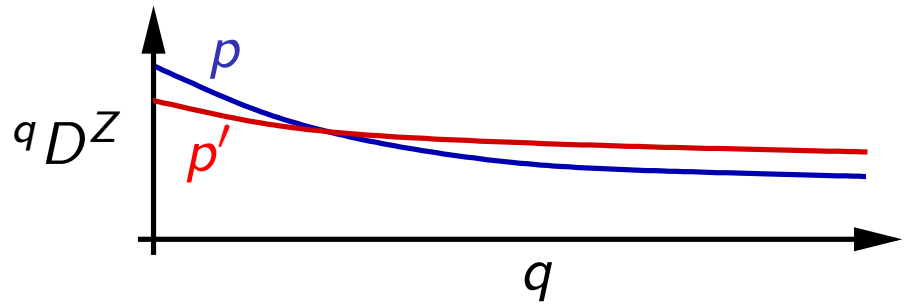


more diverse

less diverse

$q = 10$
(insensitive to rare species)





2. The theorem

The central questions

Take a list of species, with known similarity matrix Z.

Questions

- Which frequency distribution(s) maximize the diversity?
- What is the value of the maximum diversity?

Remember the birds!

In principle, the answers depend on the 'viewpoint parameter' q.

The solution

Theorem (2009) Neither answer depends on q. That is,

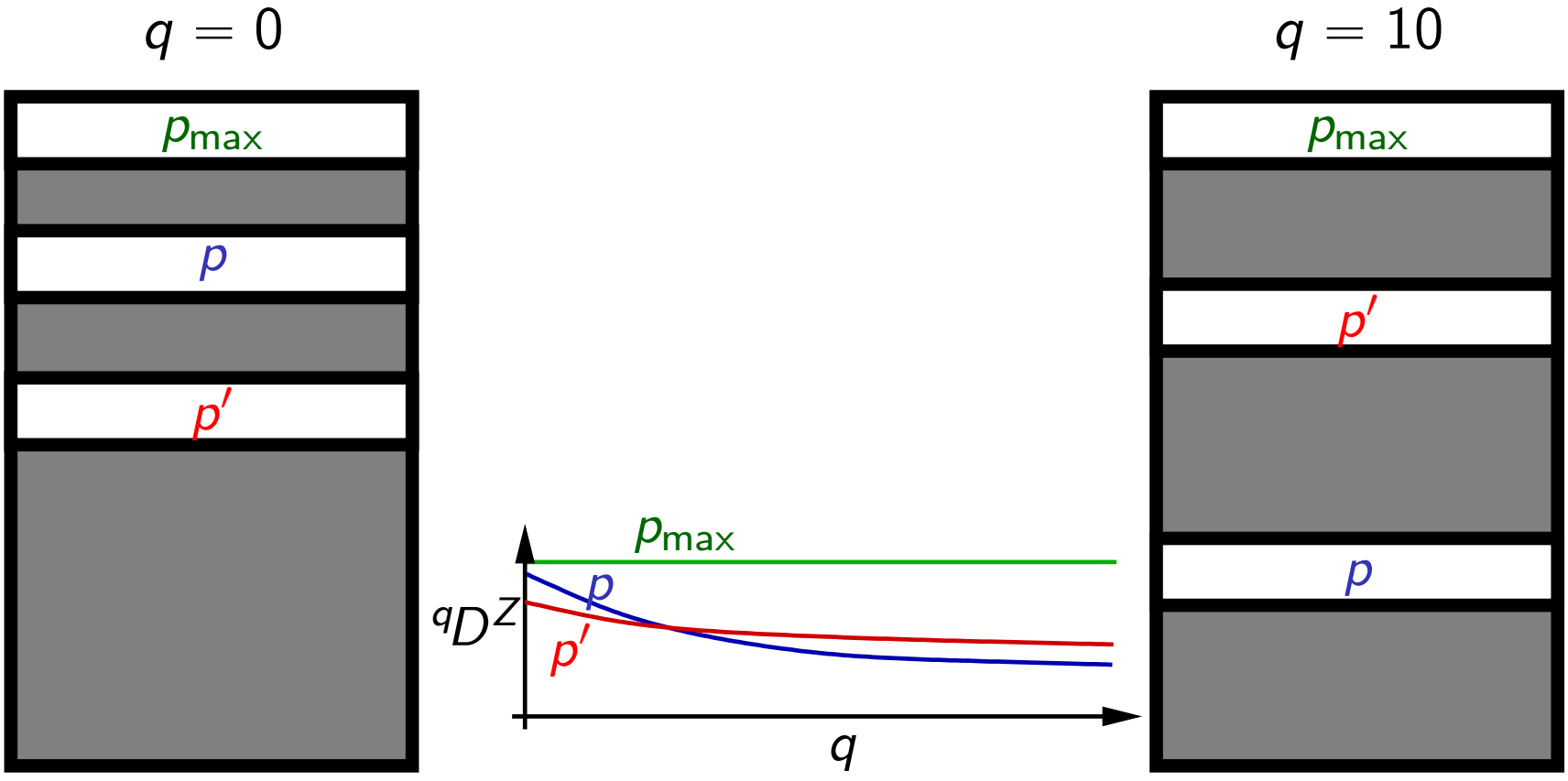
- There is a single frequency distribution $p_{\max}$ that maximizes diversity of all orders q simultaneously $(0 \leqslant q \leqslant \infty)$

  —a 'best of all possible worlds'.

- The maximum diversity, ${}^{q}D^{Z}(p_{\max})$, is the same for all q.

The proof of the theorem gives a construction of $p_{\max}$.

How is that possible?



Different values of the viewpoint parameter q produce different judgements on which distributions are more diverse than which others.

But there is a single distribution p_{max} that is optimal for all q.

3. Examples and consequences

The naive model

Put

$$Z_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

(Then ${}^qD^Z(p)$ is exponential of Rényi entropy.)

The maximizing distribution is uniform: every species has equal abundance.

Why? Can prove it directly, or reason as follows:

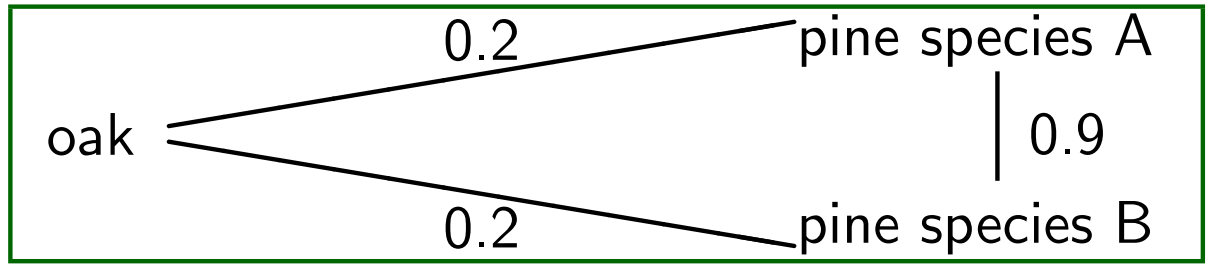
One corollary of the theorem is that if a distribution is maximizing for *some* $q > 0$, then it is maximizing for *all* q.

We know that the uniform distribution maximizes the Shannon entropy ($q = 1$).

It follows that the uniform distribution maximizes all the Rényi entropies.

A three-species example

$$Z = \begin{pmatrix} 1 & 0.2 & 0.2 \\ 0.2 & 1 & 0.9 \\ 0.2 & 0.9 & 1 \end{pmatrix}$$



Which frequency distribution maximizes the diversity?

Not this:

$$\begin{matrix} & & 0.333\ldots \\ 0.333\ldots & & \\ & & 0.333\ldots \end{matrix}$$

Or this:

$$\begin{matrix} & & 0.25 \\ 0.5 & & \\ & & 0.25 \end{matrix}$$

In fact, it's this:

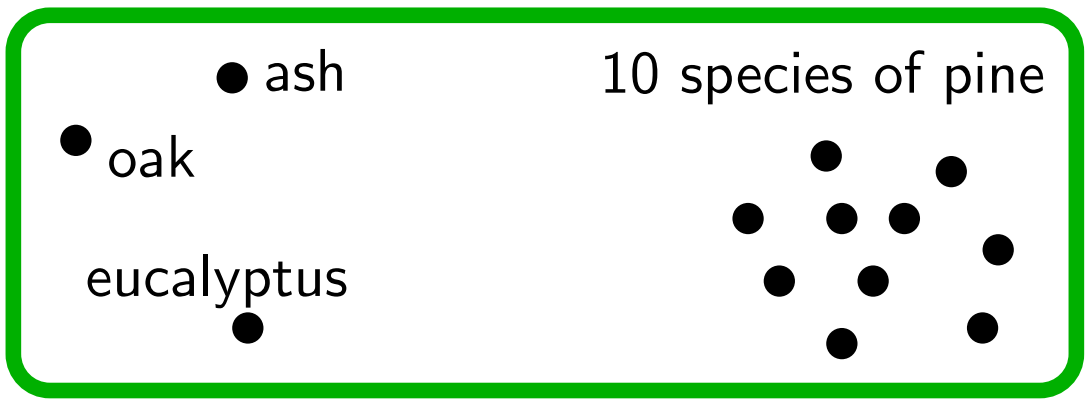
$$\begin{matrix} & & 0.26 \\ 0.48 & & \\ & & 0.26 \end{matrix}$$

This distribution maximizes diversity for *all* q, with constant value $1.703\ldots$.

Properties of maximizing distributions

- There can be more than one maximizing distribution.

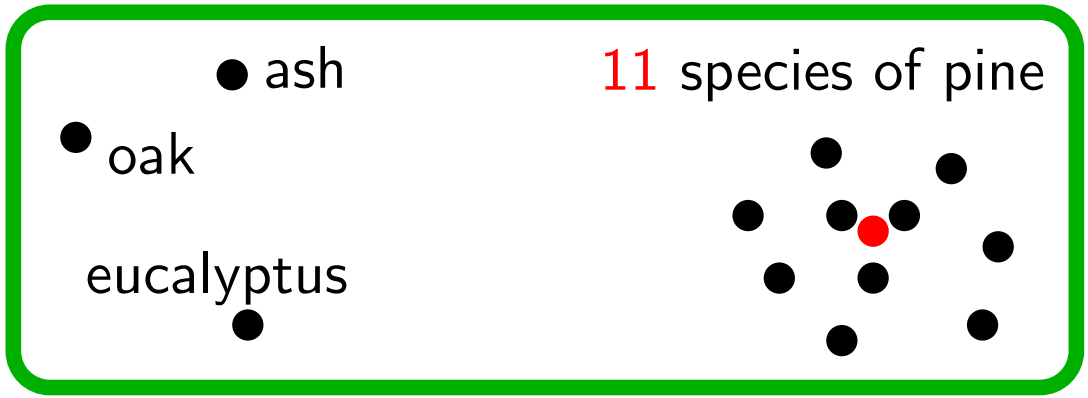
- A maximizing distribution can eliminate some species entirely.
  How can that be?
  To understand, run time backwards. . .



Properties of maximizing distributions

- There can be more than one maximizing distribution.

- A maximizing distribution can eliminate some species entirely.
  How can that be?
  To understand, run time backwards. . .



Introducing an 11th species of pine *decreases* diversity.

So if we start with a forest containing all 11 pine species, eliminating the 11th will *increase* diversity.

Computing the maximizing distribution

Start with an $n \times n$ similarity matrix Z.

- Computing the maximizing distribution(s) takes 2^n steps.
- But each step is fast. E.g. can do 25 species in a few seconds.
- And that's for *arbitrary* similarity matrices Z.
  For some *special* types of Z, the computation is near-instant.

Tree-based similarity matrices

Suppose we define similarity via a taxonomic or phylogenetic tree.

Example

Put

$$Z_{ij} = \begin{cases} 1 & \text{if same species } (i = j) \\ 0.7 & \text{if different species but same genus} \\ 0.3 & \text{if different genera but same family} \\ 0 & \text{otherwise.} \end{cases}$$

Then:

- There's a unique maximizing distribution $p_{\max}$.
- It eliminates no species; that is, $(p_{\max})_i > 0$ for all i.

4. Unanswered questions

Why is the theorem true?

Why is there a *single frequency distribution* that maximizes diversity from all viewpoints simultaneously?

- We can prove it. . .
- And it's easy in the naive case where Z is the identity (Rényi entropies). Then the maximizing distribution is uniform.

But we lack intuition as to why it's true in general.

What about maximization under constraints?

The theorem concerns maximization of diversity without constraints.

For some constraints, the theorem fails: there is no distribution that maximizes diversity for all q simultaneously.

Under which types of constraint *can* we maximize diversity for all q simultaneously?

What is the significance of the maximum diversity itself?

The theorem also says that the maximum diversity is independent of the choice of viewpoint parameter q.

So, any similarity matrix Z ('list of species') gives rise to a number, $D_{\max}(Z) = {}^{q}D^{Z}(p_{\max})$, which does not depend on the choice of q.

What does $D_{\max}(Z)$ *mean*?

It has geometric significance. . .

ways better behaved. For a compact (not necessarily positive definite) metric space A, the **maximum diversity** of A is

$$|A|_+ = \sup_{\mu \in P(A)} \left(\int \int e^{-d(a,b)} \, d\mu(a) \, d\mu(b) \right)^{-1}, \tag{4.3}$$

where $P(A)$ denotes the space of Borel probability measures on A. By renormalization, this

(Mark Meckes)

equivalence (7.1) of packing and covering numbers. When $\overline{\dim}_{\mathrm{Mink}} A = \underline{\dim}_{\mathrm{Mink}} A$, or equivalently $\lim_{\varepsilon \to 0^+} \frac{\log N(A,\varepsilon)}{\log(1/\varepsilon)}$ exists, the **Minkowski dimension** $\dim_{\mathrm{Mink}} A$ is equal to this limit.

The **upper**, **lower**, and ordinary **diversity dimension** of an arbitrary compact metric space A are defined analogously to magnitude dimensions, using the maximum diversity

How else can it be understood?

What are the biological implications of the theorem?

The theorem is fundamentally a maximum entropy theorem.

Interpreting entropy as diversity, it says there's a 'best of all possible worlds' (if we think diversity is good).

But there are other biological interpretations of entropy!

Under those interpretations, what does the theorem tell us?

Summary

Summary

- There is a one-parameter family of diversity measures, $({}^qD^Z(p))_{0\leqslant q\leqslant\infty}$, generalizing Rényi (and Shannon) entropy.
- They take into account not only the frequencies p_i of the species, but also the inter-species similarities (Z_{ij}).

- Different values of the parameter q correspond to different viewpoints on the relative importance of rare/common species.
- Different values of q produce different judgements on which communities are more diverse.

- Nevertheless:
  - Given a similarity matrix Z, there's a single frequency distribution $p_{\max}$ that maximizes diversity from all viewpoints q simultaneously.
  - Moreover, the maximum diversity value ${}^qD^Z(p_{\max})$ is the same for all q.

- Still, we don't fully understand:
  - why the theorem is true
  - what its biological implications are.