Measuring diversity: the axiomatic approach

Tom Leinster

School of Mathematics
University of Edinburgh

Boyd Orr Centre
for Population and Ecosystem Health
University of Glasgow

These slides: on my web page

Plan

1. Overview
2. Three popular diversity measures (and their problems)
3. The Hill numbers (the solution)
4. The axiomatic approach

1. Overview

Who cares about diversity?

- Ecologists

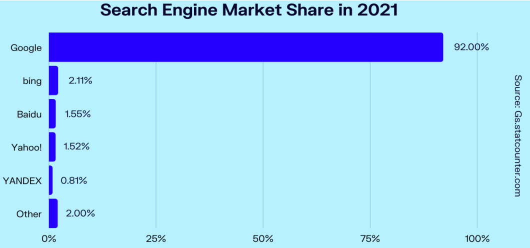
  There is a biodiversity crisis. . . but what exactly does it mean to say “biodiversity has decreased by 20%”?

- Social scientists

  E.g. ethnic diversity and gender diversity.

- Economists (using different words)

  E.g. in an industry, is the market distributed among many small companies, or is there a monopoly?



Who cares about diversity?

- Ecologists

  There is a biodiversity crisis. . . but what exactly does it mean to say "biodiversity has decreased by 20%"?

- Social scientists

  E.g. ethnic diversity and gender diversity.

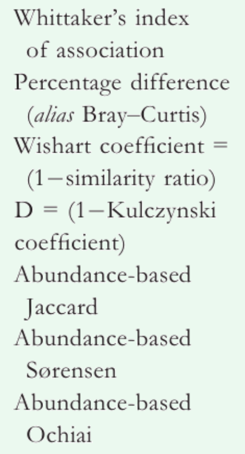
- Economists (using different words)

  E.g. in an industry, is the market distributed among many small companies, or is there a monopoly?

- Information theorists (nearly)

  Diversity is very closely related to entropy.

Ecologists have proposed and used many, many diversity measures. . .



Whittaker's index of association

Percentage difference (*alias* Bray–Curtis)

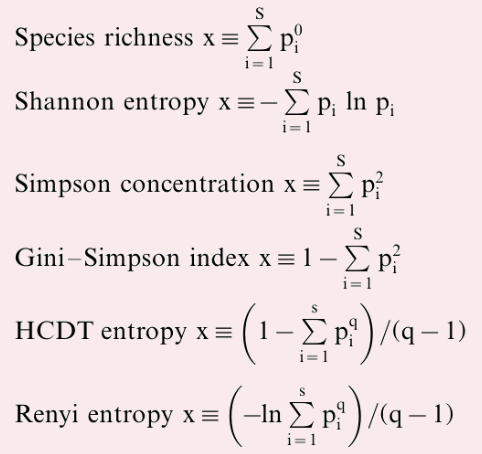
Wishart coefficient = (1−similarity ratio)

D = (1−Kulczynski coefficient)

Abundance-based Jaccard

Abundance-based Sørensen

Abundance-based Ochiai



Species richness $x \equiv \sum_{i=1}^{S} p_i^0$

Shannon entropy $x \equiv -\sum_{i=1}^{S} p_i \ln p_i$

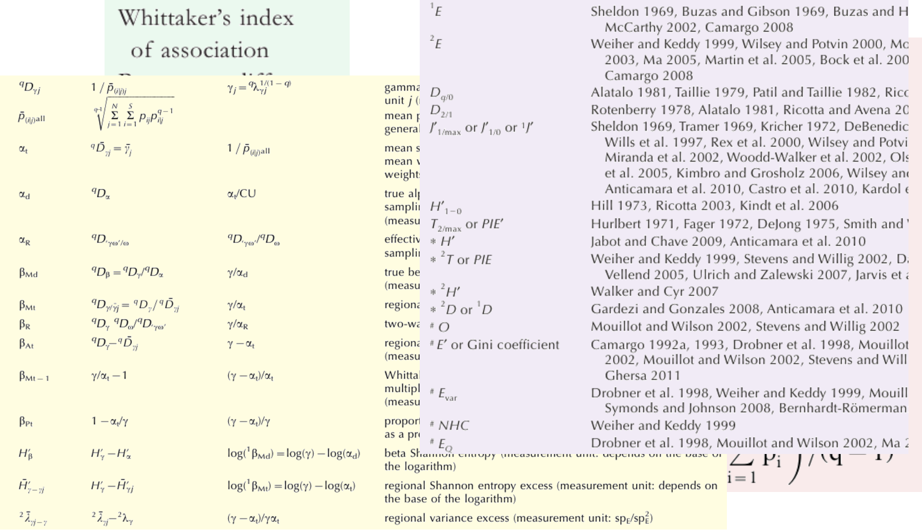
Simpson concentration $x \equiv \sum_{i=1}^{S} p_i^2$

Gini–Simpson index $x \equiv 1 - \sum_{i=1}^{S} p_i^2$

HCDT entropy $x \equiv \left(1 - \sum_{i=1}^{s} p_i^q\right)/(q-1)$

Renyi entropy $x \equiv \left(-\ln \sum_{i=1}^{s} p_i^q\right)/(q-1)$

Ecologists have proposed and used many, many diversity measures. . .



Whittaker's index
of association

$^{q}D_{\gamma j}$	$1/\bar{p}_{(i\|j)j}$	$\gamma_j = {}^{q}\lambda_{\gamma j}^{1/(1-q)}$	gamma unit j (
$\bar{p}_{(i\|j)\text{all}}$	$\sqrt[q-1]{\sum_{j=1}^{N}\sum_{i=1}^{S} p_{ij}p_{i\|j}^{q-1}}$		mean p general
α_t	${}^{q}\bar{D}_{\gamma j} = \bar{\gamma}_j$	$1/\bar{p}_{(i\|j)\text{all}}$	mean s mean v weights
α_d	${}^{q}D_{\alpha}$	α_t/CU	true alp samplin (measu
α_R	${}^{q}D_{\cdot\gamma\omega'/\omega}$	${}^{q}D_{\cdot\gamma\omega'}/{}^{q}D_{\omega}$	effectiv samplin
β_{Md}	${}^{q}D_{\beta} = {}^{q}D_{\gamma}/{}^{q}D_{\alpha}$	γ/α_d	true be (measu
β_{Mt}	${}^{q}D_{\gamma/\bar{\gamma}j} = {}^{q}D_{\gamma}/{}^{q}\bar{D}_{\gamma j}$	γ/α_t	regiona
β_R	${}^{q}D_{\gamma}\ {}^{q}D_{\omega}/{}^{q}D_{\cdot\gamma\omega'}$	γ/α_R	two-wa
β_{At}	${}^{q}D_{\gamma} - {}^{q}\bar{D}_{\gamma j}$	$\gamma - \alpha_t$	regiona (measu
β_{Mt-1}	$\gamma/\alpha_t - 1$	$(\gamma - \alpha_t)/\alpha_t$	Whittal multipl (measu
β_{Pt}	$1 - \alpha_t/\gamma$	$(\gamma - \alpha_t)/\gamma$	proport as a pr
H'_{β}	$H'_{\gamma} - H'_{\alpha}$	$\log({}^{1}\beta_{Md}) = \log(\gamma) - \log(\alpha_d)$	beta Shannon entropy (measurement unit: depends on the base of the logarithm)
$\bar{H}'_{\gamma-\gamma j}$	$H'_{\gamma} - \bar{H}'_{\gamma j}$	$\log({}^{1}\beta_{Mt}) = \log(\gamma) - \log(\alpha_t)$	regional Shannon entropy excess (measurement unit: depends on the base of the logarithm)
${}^{2}\bar{\lambda}_{\gamma j-\gamma}$	${}^{2}\bar{\lambda}_{\gamma j} - {}^{2}\lambda_{\gamma}$	$(\gamma - \alpha_t)/\gamma\alpha_t$	regional variance excess (measurement unit: sp_E/sp_E^2)

${}^{1}E$	Sheldon 1969, Buzas and Gibson 1969, Buzas and H McCarthy 2002, Camargo 2008
${}^{2}E$	Weiher and Keddy 1999, Wilsey and Potvin 2000, Mo 2003, Ma 2005, Martin et al. 2005, Bock et al. 200 Camargo 2008
$D_{q/0}$	Alatalo 1981, Taillie 1979, Patil and Taillie 1982, Ricc
$D_{2/1}$	Rotenberry 1978, Alatalo 1981, Ricotta and Avena 20
$J'_{1/\max}$ or $J'_{1/0}$ or ${}^{1}J'$	Sheldon 1969, Tramer 1969, Kricher 1972, DeBenedic Wills et al. 1997, Rex et al. 2000, Wilsey and Potvi Miranda et al. 2002, Woodd-Walker et al. 2002, Ols et al. 2005, Kimbro and Grosholz 2006, Wilsey and Anticamara et al. 2010, Castro et al. 2010, Kardol e
H'_{1-0}	Hill 1973, Ricotta 2003, Kindt et al. 2006
$T_{2/\max}$ or PIE'	Hurlbert 1971, Fager 1972, DeJong 1975, Smith and
$*\ H'$	Jabot and Chave 2009, Anticamara et al. 2010
$*\ {}^{2}T$ or PIE	Weiher and Keddy 1999, Stevens and Willig 2002, Da Vellend 2005, Ulrich and Zalewski 2007, Jarvis et a
$*\ {}^{2}H'$	Walker and Cyr 2007
$*\ {}^{2}D$ or ${}^{1}D$	Gardezi and Gonzales 2008, Anticamara et al. 2010
# O	Mouillot and Wilson 2002, Stevens and Willig 2002
# E' or Gini coefficient	Camargo 1992a, 1993, Drobner et al. 1998, Mouillot 2002, Mouillot and Wilson 2002, Stevens and Will Ghersa 2011
# E_{var}	Drobner et al. 1998, Weiher and Keddy 1999, Mouill Symonds and Johnson 2008, Bernhardt-Römerman
# NHC	Weiher and Keddy 1999
# E_{Q}	Drobner et al. 1998, Mouillot and Wilson 2002, Ma 2

A very simple model of an ecological community

Take an ecological community with n species.

Let p_i be the relative abundance (frequency) of the ith species.
So $p_1 + \cdots + p_n = 1$.

Write $\boldsymbol{p} = (p_1, \ldots, p_n)$.

You can measure abundance however you like: for example,

- number of individuals in each species
- total mass of each species (so ant $<$ elephant)
- area covered by each species (for plants).

"Species" does not have to mean species: could be some other grouping.

Why is measuring diversity difficult?

There are serious practical and statistical difficulties.

I'll ignore these.

There are also conceptual difficulties.

I won't ignore these!

For example, which of these bird communities is more diverse?

more species,
but dominated by
one species

fewer species,
but better balanced

It depends on your viewpoint.

2. Three popular diversity measures (and their problems)

Three popular diversity measures

Three of the measures most often used in ecology are:

- Species richness: the number of species present in the community.
- Shannon entropy. The motivation comes from information theory.
- Gini–Simpson index: the probability that two randomly chosen individuals are from different species.

I'll explain what these are.

Recall that we model a community by the relative abundances (frequencies) $p_1, \ldots, p_n$ of its species.

Species richness

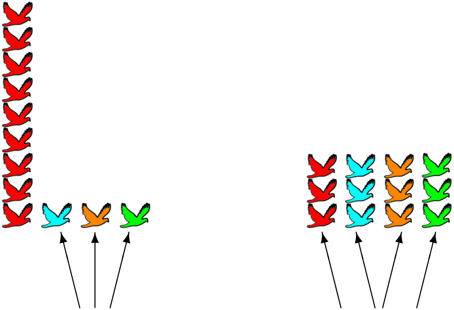
The species richness of a community is the number of species present:

$$\left|\{i \ : \ p_i \neq 0\}\right|.$$

This is the most popular measure of diversity, in both the scientific literature and in newspapers, on television, etc.

But it's very crude.

It says nothing about whether species are common or rare.



very few of these

all doing well

These two communities
have the same
species richness, 4

Shannon entropy and its problems

The Shannon entropy of a community with species relative abundances $\boldsymbol{p} = (p_1, \ldots, p_n)$ is

$$H(\boldsymbol{p}) = -\sum_{i : p_i \neq 0} p_i \log p_i.$$

When all species are equally abundant—

$$\boldsymbol{p} = \left(\tfrac{1}{n}, \ldots, \tfrac{1}{n}\right),$$

—the Shannon entropy is $\log n$.

Shannon entropy and its problems

Example (Lou Jost) A continent has 10^6 equally common species.

An asteroid strike kills 90% of the species, leaving the other 10% unchanged.

How does the Shannon entropy change?

$$\begin{aligned} H(\boldsymbol{p}_{\text{before}}) &= \log(10^6) = 6 \log 10, \\ H(\boldsymbol{p}_{\text{after}}) &= \log(10^5) = 5 \log 10, \end{aligned}$$

so the percentage decrease in entropy is $\frac{1}{6} \approx$ 17%.

Example A small patch of grass has 4 equally abundant insect species.

A bird eats one of them.

How does the Shannon entropy change?

$$\begin{aligned} H(\boldsymbol{p}_{\text{before}}) &= \log 4, \\ H(\boldsymbol{p}_{\text{after}}) &= \log 3, \end{aligned}$$

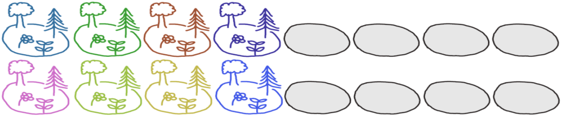
so the percentage decrease in entropy is $\approx$ 21%.

This is a bigger percentage decrease than the asteroid disaster!

Shannon entropy and its problems

Imagine 16 equal-sized islands, each with 4 unique, equally abundant species:

An oil company plans work which will destroy all life on half of the islands:



Oil company's lawyer: "The diversity now is $\log 64$, the diversity afterwards will be $\log 32$, and $\frac{\log 32}{\log 64} \approx 0.83$, so we will preserve 83% of the diversity."

Environmentalists' lawyer: "The diversity now is $\log 64$, the diversity of the islands destroyed is $\log 32$, and $\frac{\log 32}{\log 64} \approx 0.83$, so they will destroy 83% of the diversity."

The Gini–Simpson index (and its problems)

For a community with species relative abundances $p_1, \ldots, p_n$, the Gini–Simpson index is $1 - \sum_i p_i^2$.

This is the probability that two random individuals are of different species.

It is an intuitively appealing measure of diversity.

But it suffers from the same problems as Shannon entropy.

3. The Hill numbers (the solution)

Mark Hill

Effective numbers

The fundamental problem with Shannon entropy as a diversity measure is that $H\big(\frac{1}{n}, \ldots, \frac{1}{n}\big)$ is $\log n$, not n.

We saw this in the asteroid and oil company examples.

Definition A diversity measure D is an effective number if $D\big(\frac{1}{n}, \ldots, \frac{1}{n}\big) = n$ for all $n \geq 1$: the diversity of a community of n equally common species is n.

Effective numbers let us speak naturally.

Example If D is an effective number and $D(\boldsymbol{p}) = 18.2$, our community is slightly more diverse than a community of 18 equally abundant species.

- Species richness is an effective number.
- Shannon entropy is not, since $\log n \neq n$.
- The Gini–Simpson index is not either.

But there is hope!

Converting diversity measures to effective numbers

If a diversity measure is not an effective number, it can (and should!) be transformed into one.

Example Shannon entropy $H(\boldsymbol{p}) = -\sum p_i \log p_i$ is not an effective number. But

$$e^{H(\boldsymbol{p})} = \frac{1}{p_1^{p_1} p_2^{p_2} \cdots p_n^{p_n}}$$

is an effective number, and carries the same information.

Example The Gini–Simpson index $1 - \sum p_i^2$ is not an effective number. But

$$1 \Big/ \sum p_i^2$$

is an effective number, and carries the same information.

With effective numbers, the asteroid and oil company problems are solved: the percentage changes in diversity match our intuition.

The Hill numbers

Species richness, $e^{H(\boldsymbol{p})}$ and $1/\sum p_i^2$ all belong to a single family of diversity measures.

Definition Let $-\infty \leq q \leq \infty$.
The Hill number of order q of $\boldsymbol{p} = (p_1, \ldots, p_n)$ is

$$D_q(\boldsymbol{p}) = \left(\sum_{i : p_i \neq 0} p_i^q \right)^{\frac{1}{1-q}}$$

$(q \neq 1, \pm\infty)$. In the exceptional cases, take limits:

$$D_1(\boldsymbol{p}) = \prod_{i : p_i \neq 0} p_i^{-p_i} = e^{H(\boldsymbol{p})},$$

$$D_\infty(\boldsymbol{p}) = 1 \Big/ \max_{i : p_i \neq 0} p_i,$$

$$D_{-\infty}(\boldsymbol{p}) = 1 \Big/ \min_{i : p_i \neq 0} p_i.$$

(We usually only consider $q \geq 0$.)

These are all effective numbers!

Examples of Hill numbers

- $D_0(\boldsymbol{p}) = \left(\sum_{i: p_i \neq 0} p_i^0 \right)^{\frac{1}{1-0}} = |\{ i \ : \ p_i \neq 0 \}|$: species richness.
- $D_1(\boldsymbol{p}) = e^{H(\boldsymbol{p})}$ is the effective number version of Shannon entropy.
- $D_2(\boldsymbol{p}) = 1 \Big/ \sum p_i^2$ is the effective number version of the Gini–Simpson index.
- $D_\infty(\boldsymbol{p}) = 1 / \max_i p_i$ is the Berger–Parker index, which is high when no species is too dominant.

Generally, $D_q(\boldsymbol{p})$ is the exponential of the Rényi entropy of order q (which arose in information theory).

What do the different Hill numbers do?

Changing the parameter q changes how much D_q is sensitive to rare species.

When q is low, D_q is very sensitive to rare species.

Example $D_0(\boldsymbol{p})$, the number of species present, makes rare species as important as common species.

When q is high, D_q mostly ignores rare species.

Example $D_\infty(\boldsymbol{p}) = 1/\max_i p_i$ depends only on the most common species.

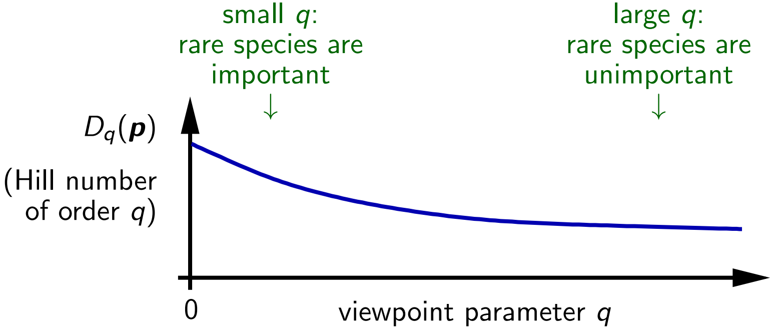
We call q the viewpoint parameter.

So which Hill numbers should we use?

All of them!

Otherwise, we might be tempted to use the one that gives the result we like best.

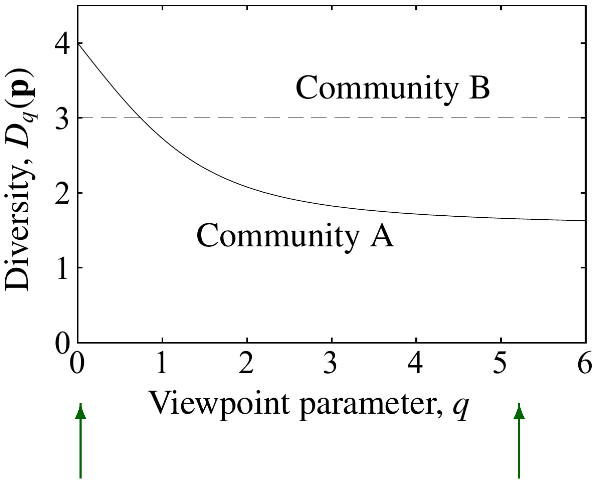
We can draw the diversity profile of the community:



Example: the bird communities

Community A

Community B



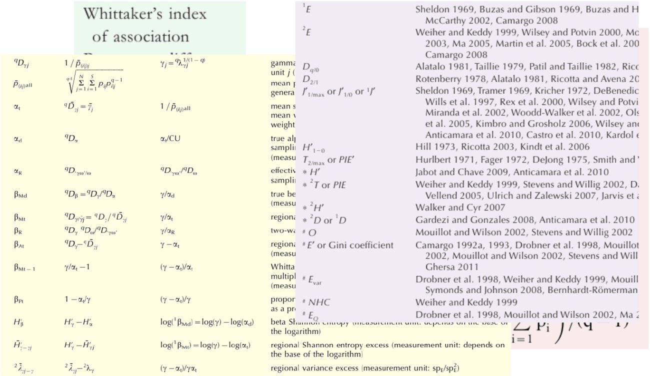
low q:
community A
more diverse

high q:
community B
more diverse

4. The axiomatic approach

Stories versus properties

It's easy to make up a new diversity measure, tell a good story about where it comes from, find some examples where it gives reasonable numbers. . .



Whittaker's index
of association

${}^qD_{\gamma j}$	$1/\bar{p}_{(i\|j)j}$	$\gamma_j = {}^q\lambda_{\gamma j}^{1/(1-q)}$	gamma unit j (
$\bar{p}_{(ij)\text{all}}$	$\sqrt[q-1]{\sum_{j=1}^{N}\sum_{i=1}^{S} p_{ij}p_{i\|j}^{q-1}}$		mean p genera
α_t	${}^q\bar{D}_{\gamma j} = \bar{\gamma}_j$	$1/\bar{p}_{(i\|j)\text{all}}$	mean s mean v weight
α_d	${}^qD_\alpha$	α_t/CU	true alp samplin (measu
α_R	${}^qD_{\cdot\gamma\omega'/\omega}$	${}^qD_{\cdot\gamma\omega'}/{}^qD_\omega$	effectiv samplin
β_{Md}	${}^qD_\beta = {}^qD_\gamma/{}^qD_\alpha$	γ/α_d	true be (measu
β_{Mt}	${}^qD_{\gamma/\bar{\gamma}j} = {}^qD_\gamma/{}^q\bar{D}_{\gamma j}$	γ/α_t	regiona
β_R	${}^qD_\gamma\ {}^qD_\omega/{}^qD_{\cdot\gamma\omega'}$	γ/α_R	two-wa
β_{At}	${}^qD_\gamma - {}^q\bar{D}_{\gamma j}$	$\gamma - \alpha_t$	regiona (measu
β_{Mt-1}	$\gamma/\alpha_t - 1$	$(\gamma - \alpha_t)/\alpha_t$	Whitta multipl (measu
β_{Pt}	$1 - \alpha_t/\gamma$	$(\gamma - \alpha_t)/\gamma$	propor as a pr
H'_β	$H'_\gamma - H'_\alpha$	$\log({}^1\beta_{Md}) = \log(\gamma) - \log(\alpha_d)$	beta Sh the logarithm)
$\bar{H}'_{\gamma-\gamma j}$	$H'_\gamma - \bar{H}'_{\gamma j}$	$\log({}^1\beta_{Mt}) = \log(\gamma) - \log(\alpha_t)$	regional Shannon entropy excess (measurement unit: depends on the base of the logarithm)
${}^2\bar{\lambda}_{\gamma j-\gamma}$	${}^2\bar{\lambda}_{\gamma j} - {}^2\lambda_\gamma$	$(\gamma - \alpha_t)/\gamma\alpha_t$	regional variance excess (measurement unit: $\text{sp}_E/\text{sp}_E^2$)

${}^1E$	Sheldon 1969, Buzas and Gibson 1969, Buzas and H McCarthy 2002, Camargo 2008
${}^2E$	Weiher and Keddy 1999, Wilsey and Potvin 2000, Mo 2003, Ma 2005, Martin et al. 2005, Bock et al. 200 Camargo 2008
$D_{q/0}$	Alatalo 1981, Taillie 1979, Patil and Taillie 1982, Ricc
$D_{2/1}$	Rotenberry 1978, Alatalo 1981, Ricotta and Avena 20
$J'_{1/\max}$ or $J'_{1/0}$ or ${}^1J'$	Sheldon 1969, Tramer 1969, Kricher 1972, DeBenedic Wills et al. 1997, Rex et al. 2000, Wilsey and Potvi Miranda et al. 2002, Woodd-Walker et al. 2002, Ols et al. 2005, Kimbro and Grosholz 2006, Wilsey and Anticamara et al. 2010, Castro et al. 2010, Kardol e
H'_{1-0}	Hill 1973, Ricotta 2003, Kindt et al. 2006
$T_{2/\max}$ or PIE'	Hurlbert 1971, Fager 1972, DeJong 1975, Smith and
$*\ H'$	Jabot and Chave 2009, Anticamara et al. 2010
$*\ {}^2T$ or PIE	Weiher and Keddy 1999, Stevens and Willig 2002, D Vellend 2005, Ulrich and Zalewski 2007, Jarvis et a
$*\ {}^2H'$	Walker and Cyr 2007
$*\ {}^2D$ or ${}^1D$	Gardezi and Gonzales 2008, Anticamara et al. 2010
$\#\ O$	Mouillot and Wilson 2002, Stevens and Willig 2002
$\#\ E'$ or Gini coefficient	Camargo 1992a, 1993, Drobner et al. 1998, Mouillot 2002, Mouillot and Wilson 2002, Stevens and Will Ghersa 2011
$\#\ E_{\text{var}}$	Drobner et al. 1998, Weiher and Keddy 1999, Mouill Symonds and Johnson 2008, Bernhardt-Römerman
$\#\ NHC$	Weiher and Keddy 1999
$\#\ E_Q$	Drobner et al. 1998, Mouillot and Wilson 2002, Ma 2

Stories versus properties

It's easy to make up a new diversity measure, tell a good story about where it comes from, find some examples where it gives reasonable numbers. . .

But measures of diversity should behave logically!

The axiomatic approach:
Always analyse properties

- What axioms (or properties) does a measure satisfy? Are they desirable?
- Is it the *only* measure with these properties?
- If not, why not use one of the other ones?
- Ideally, prove a theorem saying that your measure *is* the only one satisfying a list of properties.

A characterization theorem for the Hill numbers

In the next six slides, I'll state six properties a diversity measure should have.

Then I'll state a theorem:

The Hill numbers are the only diversity measures with these properties.

Notation For $n \geq 1$, write

$$\Delta_n = \{(p_1, \ldots, p_n) \in \mathbb{R}^n \ : \ p_i \geq 0, p_1 + \cdots + p_n = 1\}.$$

In the following slides, let D be a sequence of functions

$$\big(D \colon \Delta_n \to \mathbb{R}^+\big)_{n \geq 1}.$$

I'll call D a diversity measure.

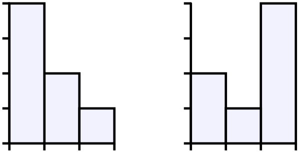
Property 1: symmetry

A diversity measure should not depend on the order in which the species are listed.

Mathematically:

$$D(p_{\sigma(1)}, \ldots, p_{\sigma(n)}) = D(p_1, \ldots, p_n)$$

for all $(p_1, \ldots, p_n) \in \Delta_n$ and permutations σ of $\{1, \ldots, n\}$.



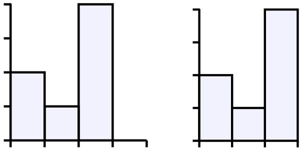
Property 2: absent species

If a species has relative abundance 0, the diversity should be the same as if it was not mentioned at all.

Mathematically:

$$D(p_1, \ldots, p_n, 0) = D(p_1, \ldots, p_n)$$

for all $(p_1, \ldots, p_n) \in \Delta_n$.

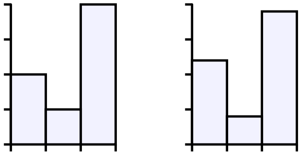


Property 3: continuity in nonzero abundances

Small changes in the abundances of the species present should only cause small changes in the diversity.

Mathematically: D is continuous on

$$\{(p_1, \ldots, p_n) \in \Delta_n \ : \ p_1, \ldots, p_n \neq 0\}.$$



Note We don't ask that D is continuous on the *whole* of Δ_n.

Species richness D_0 is continuous in *nonzero* abundances but not continuous everywhere: e.g. $(0.01, 0.99)$ has richness 2 but $(0, 1)$ has richness 1.

We don't want to exclude D_0.

Property 4: normalization

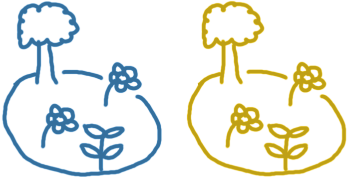
The diversity of a community consisting of just one species is 1.

Mathematically:
$D(1) = 1$, where (1) is the unique element of Δ_1.

Note This is a case of the effective number property, $D\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) = n$.

Property 5: replication

Take a group of k islands of equal size.
Suppose they have the same species *distributions*, but no shared species (no species lives on two different islands).



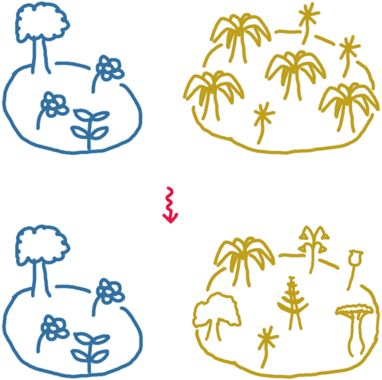
Then the diversity of the whole island group is k times the diversity of each island.

Mathematically:

$$D\big(\underbrace{\tfrac{1}{k}p_1,\ldots,\tfrac{1}{k}p_n,\quad \ldots,\quad \tfrac{1}{k}p_1,\ldots,\tfrac{1}{k}p_n}_{k\text{ times}}\big) = k\cdot D(p_1,\ldots,p_n)$$

for all $(p_1,\ldots,p_n)\in\Delta_n$ and $k\geq 1$.

Property 6: monotonicity

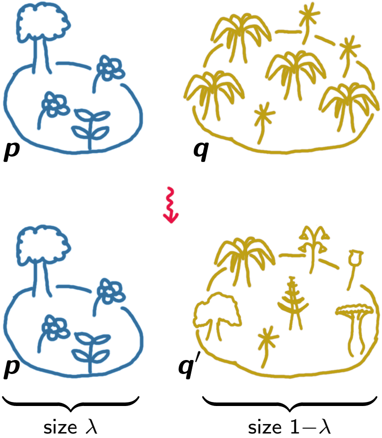


Take a pair of islands with no shared species.

Increase the diversity of one of the islands, without changing its size.

Then the diversity of the group increases (or stays the same).

Property 6: monotonicity



Take a pair of islands with no shared species.

Increase the diversity of one of the islands, without changing its size.

Then the diversity of the group increases (or stays the same).

Mathematically: if $D(q_1, \ldots, q_m) \leq D(q'_1, \ldots, q'_m)$ then

$$\begin{aligned} & D\big(\lambda p_1, \ldots, \lambda p_n, (1-\lambda)q_1, \ldots, (1-\lambda)q_m\big) \\ \leq\ & D\big(\lambda p_1, \ldots, \lambda p_n, (1-\lambda)q'_1, \ldots, (1-\lambda)q'_m\big) \end{aligned}$$

for all $\boldsymbol{p} \in \Delta_n$, $\boldsymbol{q}, \boldsymbol{q}' \in \Delta_m$ and $0 \leq \lambda \leq 1$.

The theorem

Theorem (*Entropy and Diversity*, Theorem 7.4.3)
Let $(D\colon \Delta_n \to \mathbb{R}^+)_{n\geq 1}$ *be a sequence of functions.*
The following are equivalent:

- *D has the six properties above: symmetry, absent species, continuity, normalization, replication, monotonicity;*
- *$D = D_q$ for some $q \in [-\infty, \infty]$.*

Interpretation:

The Hill numbers are the only sensible diversity measures

. . . at least, for *this* simple model of a community.

A criticism of everything I have said today

The model of the community is too simple!

Intuitively, a field made up of ten similar species of grass is less diverse than a field made up of ten very different plant species.

Diversity measures should reflect how *different* the species are.

Mathematically, we should consider not probability distributions on a *set*, but probability distributions on something like a *metric space*.

Next week, I will explain:

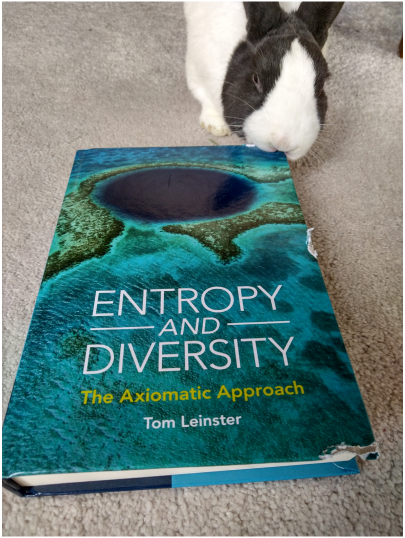
- how to do this
- how it leads to more realistic ecological judgements
- how it is mathematically connected to invariants in geometry and analysis.

References

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

Lou Jost.
Entropy and diversity.
Oikos 113 (2006), 363–375.

Lou Jost.
Partitioning diversity into independent alpha and beta components.
Ecology 88 (2007), 2427–2439.



Thank you for listening