Diversity, Entropy and Thermodynamics

John Baez
http://math.ucr.edu/home/baez/biodiversity/

July 5, 2012
The Mathematics of Biodiversity
CRM

The Shannon entropy

$$S(p) = -\sum_{i=1}^{n} p_i \ln(p_i)$$

appears in thermodynamics and information theory, but it can also be used to measure biodiversity. Is this a coincidence?

The Shannon entropy

$$S(p) = -\sum_{i=1}^{n} p_i \ln(p_i)$$

appears in thermodynamics and information theory, but it can also be used to measure biodiversity. Is this a coincidence?

No:

- In thermodynamics, the entropy of system is the expected amount of information we gain by learning its precise state.
- In biodiversity studies, the entropy of an ecosystem is the expected amount of information we gain about an organism by learning its species.

The Shannon entropy

$$S(p) = -\sum_{i=1}^{n} p_i \ln(p_i)$$

appears in thermodynamics and information theory, but it can also be used to measure biodiversity. Is this a coincidence?

No:

- In thermodynamics, the entropy of system is the expected amount of information we gain by learning its precise state.
- In biodiversity studies, the entropy of an ecosystem is the expected amount of information we gain about an organism by learning its species.

Can we connect biodiversity more deeply to thermodynamics?

Starting from any probability distribution, we can quickly reach all the main ideas of thermodynamics!

Starting from any probability distribution, we can quickly reach all the main ideas of thermodynamics!

These include:

- entropy, S
- temperature, T
- energy, E
- the partition function, Z
- the free energy, $F = E - TS$

Starting from any probability distribution, we can quickly reach all the main ideas of thermodynamics!

These include:

- entropy, S
- temperature, T
- energy, E
- the partition function, Z
- the free energy, $F = E - TS$

So, all these and much more are available to biodiversity studies.

Suppose we have a finite list of probabilities p_i summing to 1.

Suppose we have a finite list of probabilities p_i summing to 1. Fixing a number $T_0 > 0$, write

$$p_i = e^{-E_i/T_0}$$

for some **energies** E_i.

Suppose we have a finite list of probabilities p_i summing to 1. Fixing a number $T_0 > 0$, write

$$p_i = e^{-E_i/T_0}$$

for some **energies** E_i.

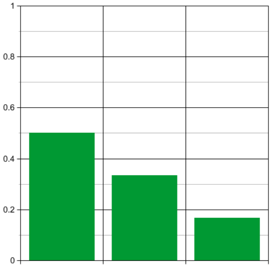
This lets us define probabilities depending on the **temperature** T:

$$p_i(T) = \frac{1}{Z(T)} e^{-E_i/T}$$

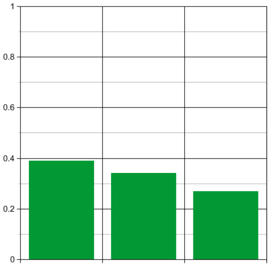
where $Z(T)$ is called the **partition function**:

$$Z(T) = \sum_i e^{-E_i/T}$$

As we raise the temperature, the probabilities $p_i(T)$ become more evenly distributed:

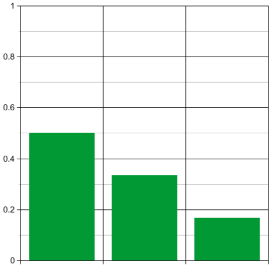


$T = 1$

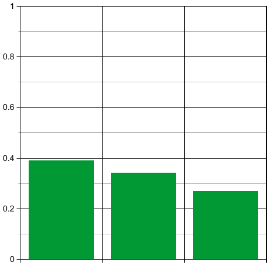


$T = 3$

As we raise the temperature, the probabilities $p_i(T)$ become more evenly distributed:



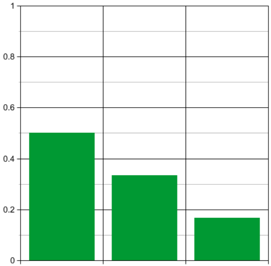
$T = 1$



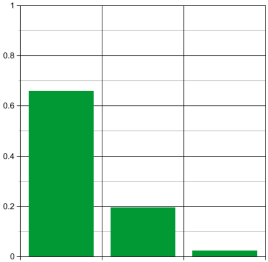
$T = 3$

When something gets hotter, all possible situations become closer to being equally probable.

As we lower the temperature, the biggest probabilities increase, while the rest go to zero:

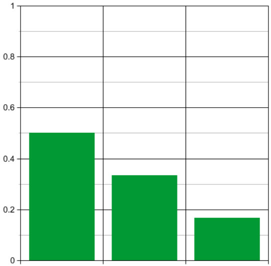


$T = 1$

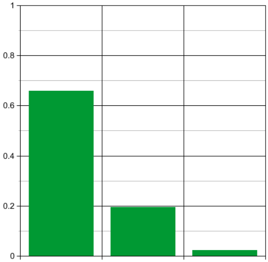


$T = 1/3$

As we lower the temperature, the biggest probabilities increase, while the rest go to zero:



$T = 1$



$T = 1/3$

When something gets colder, the chance that it's in a low-energy state goes up.

What is special about the probability distribution $p_i(T)$?

What is special about the probability distribution $p_i(T)$?

It minimizes the **free energy** F, which we can define for any probability distribution r_i:

$$\begin{aligned} \text{free energy} \quad &= \quad \text{expected energy} \quad - \quad \text{temperature} \quad \text{times} \quad \text{entropy} \\ &= \quad \sum_i r_i E_i \quad + \quad T \sum_i r_i \ln(r_i) \end{aligned}$$

What is special about the probability distribution $p_i(T)$?

It minimizes the **free energy** F, which we can define for any probability distribution r_i:

$$\begin{aligned} \text{free energy} &= \text{expected energy} - \text{temperature times entropy} \\ &= \sum_i r_i E_i + T \sum_i r_i \ln(r_i) \end{aligned}$$

So:

- When it gets hotter, $p_i(T)$ tries harder to maximize entropy.
- When it gets colder, $p_i(T)$ tries harder to minimize energy.

Using the probabilities $p_i(T)$, entropy becomes temperature-dependent:

$$S(T) = -\sum_i p_i(T) \ln(p_i(T))$$

Using the probabilities $p_i(T)$, entropy becomes temperature-dependent:

$$S(T) = -\sum_i p_i(T) \ln(p_i(T))$$

So does the expected value of the energy:

$$E(T) = \sum_i p_i(T) E_i$$

Using the probabilities $p_i(T)$, entropy becomes temperature-dependent:

$$S(T) = -\sum_i p_i(T) \ln(p_i(T))$$

So does the expected value of the energy:

$$E(T) = \sum_i p_i(T) E_i$$

Thus, so does the free energy:

$$F(T) = E(T) - TS(T)$$

Given any probability distribution p_i, we get a 1-parameter family of entropies $S(T)$. Are these the Rényi or Tsallis entropies?

Given any probability distribution p_i, we get a 1-parameter family of entropies $S(T)$. Are these the Rényi or Tsallis entropies?

No. The **Rényi entropy** is:

$$S_q(p) = \frac{1}{1-q} \ln \sum_i p_i^q = \frac{1}{1-q} \ln \sum_i e^{-E_i q / T_0}$$

Given any probability distribution p_i, we get a 1-parameter family of entropies $S(T)$. Are these the Rényi or Tsallis entropies?

No. The **Rényi entropy** is:

$$S_q(p) = \frac{1}{1-q} \ln \sum_i p_i^q = \frac{1}{1-q} \ln \sum_i e^{-E_i q / T_0}$$

If we let q be the ‘cooling factor’:

$$T = T_0 / q$$

this gives

$$S_q(p) = \frac{1}{1-q} \ln \sum_i e^{-E_i / T} = \frac{1}{1-q} \ln Z(T)$$

What is the meaning of this equation:

$$S_q(p) = \frac{1}{1-q} \ln Z(T) \text{ ?}$$

What is the meaning of this equation:

$$S_q(p) = \frac{1}{1-q} \ln Z(T) \text{ ?}$$

We can exponentiate both sides to get the **Hill numbers**:

$$D_q(p) = Z(T)^{\frac{1}{1-q}}$$

which Lou Jost argues are a better measure of biodiversity.

What is the meaning of this equation:

$$S_q(p) = \frac{1}{1-q} \ln Z(T) \ ?$$

We can exponentiate both sides to get the **Hill numbers**:

$$D_q(p) = Z(T)^{\frac{1}{1-q}}$$

which Lou Jost argues are a better measure of biodiversity.

Challenge: If Hill numbers are fundamental to biodiversity, while the partition function is fundamental to thermodynamics, why this funny relationship?

Alternatively, we can write

$$S_q(p) = \frac{1}{1-q} \ln Z(T) = \frac{1}{1 - T_0/T} \ln Z(T) = \frac{T \ln Z(T)}{T - T_0}$$

Alternatively, we can write

$$S_q(p) = \frac{1}{1-q} \ln Z(T) = \frac{1}{1 - T_0/T} \ln Z(T) = \frac{T \ln Z(T)}{T - T_0}$$

Then, use a wonderful identity relating free energy to the partition function:

$$F(T) = -T \ln Z(T)$$

Alternatively, we can write

$$S_q(p) = \frac{1}{1-q} \ln Z(T) = \frac{1}{1 - T_0/T} \ln Z(T) = \frac{T \ln Z(T)}{T - T_0}$$

Then, use a wonderful identity relating free energy to the partition function:

$$F(T) = -T \ln Z(T)$$

to get

$$S_{T_0/T}(p) = -\frac{F(T)}{T - T_0}$$

Alternatively, we can write

$$S_q(p) = \frac{1}{1-q}\ln Z(T) = \frac{1}{1 - T_0/T}\ln Z(T) = \frac{T\ln Z(T)}{T - T_0}$$

Then, use a wonderful identity relating free energy to the partition function:

$$F(T) = -T\ln Z(T)$$

to get

$$S_{T_0/T}(p) = -\frac{F(T)}{T - T_0}$$

But $Z(T_0) = 1$, so $F(0) = 0$.

Alternatively, we can write

$$S_q(p) = \frac{1}{1-q} \ln Z(T) = \frac{1}{1 - T_0/T} \ln Z(T) = \frac{T \ln Z(T)}{T - T_0}$$

Then, use a wonderful identity relating free energy to the partition function:

$$F(T) = -T \ln Z(T)$$

to get

$$S_{T_0/T}(p) = -\frac{F(T)}{T - T_0}$$

But $Z(T_0) = 1$, so $F(0) = 0$. Thus

$$S_{T_0/T}(p) = -\frac{F(T) - F(T_0)}{T - T_0}$$

$$S_{T_0/T}(p) = -\frac{F(T) - F(T_0)}{T - T_0}$$

Moral: Rényi entropy is minus the change in free energy divided by the change in temperature.

$$S_{T_0/T}(p) = -\frac{F(T) - F(T_0)}{T - T_0}$$

Moral: Rényi entropy is minus the change in free energy divided by the change in temperature.

Taking $T \to T_0$ we recover a famous formula for the Shannon entropy:

$$S(p) = -\left.\frac{dF(T)}{dT}\right|_{T=T_0}$$

$$S_{T_0/T}(p) = -\frac{F(T) - F(T_0)}{T - T_0}$$

Moral: Rényi entropy is minus the change in free energy divided by the change in temperature.

Taking $T \to T_0$ we recover a famous formula for the Shannon entropy:

$$S(p) = -\left.\frac{dF(T)}{dT}\right|_{T=T_0}$$

Challenge: What do these facts mean for biodiversity?

The Second Law of Thermodynamics says that Shannon entropy must increase under certain conditions. Biodiversity does not always increase.

The Second Law of Thermodynamics says that Shannon entropy must increase under certain conditions. Biodiversity does not always increase.

However, Marc Harper has shown that something *similar* to the Second Law holds when a population approaches an ‘evolutionary optimum’!

Suppose the vector of populations $P = (P_1, \ldots, P_n)$ evolve with time according to a generalized Lotka–Volterra equation:

$$\frac{dP_i}{dt} = f_i(P_1, \ldots, P_n)P_i$$

Suppose the vector of populations $P = (P_1, \ldots, P_n)$ evolve with time according to a generalized Lotka–Volterra equation:

$$\frac{dP_i}{dt} = f_i(P_1, \ldots, P_n) P_i$$

Let p be the corresponding probability distribution:

$$p_i = \frac{P_i}{\sum_j P_j}$$

Suppose the vector of populations $P = (P_1, \ldots, P_n)$ evolve with time according to a generalized Lotka–Volterra equation:

$$\frac{dP_i}{dt} = f_i(P_1, \ldots, P_n) P_i$$

Let p be the corresponding probability distribution:

$$p_i = \frac{P_i}{\sum_j P_j}$$

Let q be a *fixed* probability distribution, and let

$$I(q,p) = \sum_i \ln\left(\frac{q_i}{p_i}\right) q_i$$

be the **relative Shannon information.**

Then

$$\frac{d}{dt} I(q,p) \le 0$$

if q is an **evolutionary optimum**:

$$p \cdot f(P) \le q \cdot f(P)$$

for all P: i.e., the mean fitness of a small sample of ‘invaders’ distributed according to the distribution q exceeds or equals the mean fitness of any population P.

So: the information ’left to learn’ never increases as the population’s distribution evolves toward an evolutionary optimum. Not biodiversity, but *relative biodiversity*, matters here!