Math 343, Lecture 12

① The Boltzmann model

Now

By allowing ourselves this extra freedom

We still require

That is

Let $\mathcal{C}$ be a combinatorial class and $C(x)$ its generating function.

Rather than

Suppose $C(x)$ has radius of convergence $\rho > 0$

how plausible is this assumption $\rightarrow$ lets check some examples we've done before

Recall from MACM 201

Def A finite probability space with sample space $\mathcal{S}$ and probability function P, is a finite set $\mathcal{S}$ and a function P from the set of subsets of $\mathcal{S}$ to the interval $[0,1]$ satisfying

1. $P(A) \geq 0 \quad \forall A \subseteq \mathcal{S}$
2. $P(\mathcal{S}) = 1$
3. If $A, B \subseteq \mathcal{S}$ with $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$

take $A = \{a_1, \ldots, a_n\} \subseteq \mathcal{S}$. By repeated application of 3

$$P(A) =$$

Then we interpret

Combinatorial classes aren't finite, but they're still discrete so we can define

Def A discrete probability space with sample space $\mathcal{S}$ and probability function P, is a countable set $\mathcal{S}$ and a function P from the set of subsets of $\mathcal{S}$ to the interval $[0,1]$ satisfying

1.
2.
3.

Note

Evaluating the generating function gives a useful discrete probability space with $\mathcal{C}$ as the sample space

Def Let $\mathcal{C} \neq \emptyset$ be a combinatorial class and let $\rho > 0$ be the radius of convergence of $C(x)$.

The **Boltzmann model** at $0 < x < \rho$ is a discrete probability space with sample space $\mathcal{C}$ and probability function given by

$$P_x(c) =$$

Lets check that this is a discrete probability space

The other property we wanted was that our probabilities were uniform by size. That is

Proposition Let $\mathcal{C} \neq \emptyset$ be a combinatorial class and $0 < x < \rho$ where ρ is the radius of convergence of $C(x)$. Let P_x be the probability function for the Boltzmann model.

Suppose

Then

proof

So we know the Boltzmann model is uniform by size but which sizes are preferred?

Prop Let $\mathcal{C}$, ρ, P_x be as above. The expected size is

$$E_x = x \frac{C'(x)}{C(x)}$$

proof The formula for expected value from probability is

$$E_x =$$

If you didn't know that, you can take the proposition as a definition of the expected size

Suppose

Then

So our goal is

② Boltzmann samplers

So how do we actually build a Boltzmann sampler?

Lets first start with the class $\mathcal{E}$

So

Algorithm Boltzmann $\mathcal{E}$

input x

$\mathcal{Z}$ is similarly silly:

Algorithm Boltzmann $\mathcal{Z}$

input x

Suppose $\mathcal{A} = \mathcal{B} + \mathcal{C}$ and we have a Boltzmann sampler for $\mathcal{B}$ and $\mathcal{C}$

What is the probability that a generated $a \in \mathcal{A}$ came from $\mathcal{B}$?

Algorithm Boltzmann $\mathcal{A} = \mathcal{B} + \mathcal{C}$

input x

Suppose $\mathcal{A} = \mathcal{B} \times \mathcal{C}$ and suppose we have Boltzmann samplers for $\mathcal{A}$ and $\mathcal{B}$

take $a = (b, c) \in \mathcal{A}$ so $b \in \mathcal{B}$ $c \in \mathcal{C}$

then a should be generated with probability

$\longrightarrow$

Algorithm Boltzmann $\mathcal{A} = \mathcal{B} \times \mathcal{C}$

input x

eg What about our binary trees

$$\mathcal{T} = \mathcal{E} + \mathcal{Z} \times \mathcal{T}^2$$

We know

Thus we have

lets try it (see notes for the other kind of binary rooted tree)

That kind of worked but

(3) What about Sequence?

There are two approaches to sequence

(1) note

$\mathcal{A} = \mathrm{Seq}(\mathcal{B})$ is equivalent to

The cost of this translation is

But there's also a direct way for Boltzmann generators

(2) Say to generate an element of Seq $\mathcal{B}$ we first pick a k and then make k calls to Boltzmann $\mathcal{B}$

The question is with what probability should each k occur?

What do we need?

$$P_x((b_1, \dots, b_k)) =$$

This is a geometric distribution

Def The random variable X is **geometrically distributed** if

$$P(X=k) = (1-\lambda)\lambda^k$$

for some $0<\lambda\le 1$

How do we know $B(x)\le 1$?

So assuming we have a geometric random number generator:

Algorithm Boltzmann $A = Seq(\mathcal{B})$

input x

④ Next time

Theorems and practicalities

→ how to build geometric_rand

→ which x to pick

→ how fast is Boltzman generation

→ which specifications work well and how to tweak the ones which don't.