APMA 2610: Recent Applications of Probability and Statistics

Milan Capoor

Spring 2025

1 Jan 22

1.1 Maximum Entropy Principle

A strange though experiment of Gibbs: Imagine a physical system S (say a gas) in an "infinite bath". Let x be the state of every particle (positions, velocities, ...) in S.

For simplicity, let S be be 3 particles in $\mathbb{Z}^2$ with $x \in \mathbb{Z}^6$ being the positions. Let s be the number of states of particles in S.

What is $p(x)$, the probability that S has state x?

In the simplest case (each particle is independent and the state distribution is uniform), we trivially have $P(x) = \frac{1}{s}$. But in general, these are incredibly strong assumptions.

We can create some constraints to do better.

1. Assume that the average kinetic energy $\mathcal{E}$ of the infinite heat bath is some constant θ.

   In this case, we expect the average kinetic energy of S is approximately θ:

$$\sum_x p(x)\mathcal{E}(x) = \theta$$

2. Trivially, p is a probability distribution, so

$$\sum_x p(x) = 1$$

But still this is far from enough: this gives us only 2 constraints for s many unknowns!

However, we can approximate with the LLN. Sample $n \gg s \gg 1$ iid copies of S, $S_1, S_2, \dots, S_n$ with positions $x_1, x_2, \dots, x_n$.

Define the **empirical distribution**

$$\widehat{p}_x = \frac{\#\{i : X_i = x\}}{n}$$

So with large n, $\widehat{p} = p$, and

$$\sum_x \widehat{p}(x)\mathcal{E}(x) \approx \theta$$

Claim: The vast majority of assignments of states to $X_1, \dots, X_n$ yield a single empirical distribution $\widehat{p}$.

Consider $C(\widehat{p})$, the number of ways to assign a state to each of n systems that would yield $\widehat{p}$. Then, with $\widehat{n}_x = \widehat{p}_x \cdot n = \#\{i : X_i = x\}$,

$$C(\widehat{p}) = \binom{n}{\prod_{i=1}^s n_i}$$