Lecture-05: The Boltzmann Distribution

1 The Boltzmann Distribution

The fundamental purpose of statistical physics is to understand how microscopic interactions of particles (atoms, molecules, etc.) can lead to macroscopic phenomena. It is unreasonable to try to calculate how each and every particle is behaving. Instead, we use probability and statistics to model the behaviour of a large group of particles as a whole.

A physical system can be described probabilistically as:

- A **space of configurations** $\mathcal{X}$: The state/configuration of the ith particle is represented by the random variable $x_i \in \mathcal{X}$. If there are N particles, then the configuration of the system is represented by $x = (x_1, x_2, \ldots, x_N)$, where each $x_i \in \mathcal{X}$. The configuration space for a N particle system is the product space $\underbrace{\mathcal{X} \times \mathcal{X} \times \ldots \times \mathcal{X}}_{N} = \mathcal{X}^N$. We will limit ourselves to configuration spaces which are

  - (i) finite sets, or
  - (ii) smooth, compact, finite dimensional manifolds.

- A **set of obervables** which are real-valued functions from the configuration space to $\mathbb{R}$. That is, any observable is $\mathcal{O} : \mathcal{X}^N \to \mathbb{R}$ such that for any configuration $x \in \mathcal{X}^N$, we have the observable $\mathcal{O}(x)$. A key point to note is that observables can, at least in principle, be measured through an experiment. In contrast, the configuration of a system usually cannot be measured.

- One special observable is the **energy function** $E(x)$. The form of the energy function depends on the level of interaction of the particles.

Example 1.1 (Interacting Particle System Energy). We consider three different examples of energy function for an N-particle system.

1. An energy function of the form $E(x) = \sum_{i=1}^{N} E_i(x_i)$ is called a **non-interacting system**, as the total energy of the system depends only on the energies of the individual particles.

2. An energy function of the form

$$E(x) = \sum_{i=1}^{N} E_i(x_i) + \alpha \sum_{\substack{i,j=1 \\ i<j}}^{N} E_{ij}(x_i, x_j)$$

   is called a **two-body interaction**, as interaction between pairs of particles is taken into account.

3. In general, if the energy function has a term of the form

$$\sum_{i_1, \ldots, i_k} E_{i_1, \ldots, i_k}(x_{i_1}, \ldots, x_{i_k})$$

   the system has a **k-body interaction**. However, in real physical systems, interactions above the two or three body level are rarely considered.

Definition 1.2. Given a finite configuration space $\mathcal{X}$ and energy function $E : \mathcal{X} \to \mathbb{R}$, the **Boltzmann distribution** is the probability distribution for the system to be in state $x \in \mathcal{X}$, given by

$$P\{X = x\} = \mu_\beta(x) = \frac{e^{-\beta E(x)}}{Z(\beta)},$$

where $Z(\beta)$ is the normalizing constant such that $\sum_{x \in \mathcal{X}} \mu_\beta(x) = 1$.

Definition 1.3. Given a finite configuration space $\mathcal{X}$ and the energy function $E : \mathcal{X} \to \mathbb{R}$, the **partition function** is defined as

$$Z(\beta) = \sum_{x \in \mathcal{X}} e^{-\beta E(x)}.$$

Here, the quantity β is the inverse temperature, i.e. $\beta = 1/T$. (In physics, β is actually defined as $1/(k_B T)$, where k_B is the Boltzmann constant. Here, we simply define $k_B = 1$.)

The partition function is so called because it describes how particles are distributed or "partitioned" into different energy levels of a system.

Example 1.4 (Independent coin tosses). Let us consider the case of n independent coin tosses with coin bias being p. We can show that for large n, the number of ways of getting qn heads in n coin tosses is

$$\binom{n}{qn} p^{qn} (1 - p)^{qn} \approx 2^{-n\mathcal{D}(q\|p)} = e^{-n\mathcal{D}(q\|p)\ln 2}.$$

We take n coin tosses as n particles, and their states $x_i \in \mathcal{X} = \{H, T\}$. Hence, the configuration space is $\mathcal{X}^n = \{H, T\}^n$. The energy of the state $x \in \mathcal{X}^n$ is given

$$E(x) = \frac{n}{\beta} \mathcal{D}\left(\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{x_i = H\}} \,\Big\|\, p\right) \ln 2.$$

The most probable fraction of heads is $q = p$, as we proved in class. We can see that the "energy" is minimum for this configuration. The extremes of $q = 0$ or $q = 1$, corresponding to no heads or all heads, have the maximum "energy".

Definition 1.5. The **Boltzmann average** of an obervable $\mathcal{O}$ is defined as the expactation value of that obervable under the Boltzmann distribution. It is denoted by $\langle \mathcal{O} \rangle$.

$$\langle \mathcal{O} \rangle = \sum_{x \in \mathcal{X}} \mathcal{O}(x) \mu_\beta(x) = \frac{1}{Z(\beta)} \sum_{x \in \mathcal{X}} e^{-\beta E(x)} \mathcal{O}(x).$$

Example 1.6 (Spin-1/2 system). One of the intrinsic properties of elementary particles is spin, denoted by σ. An **Ising spin** takes values in $\sigma \in \mathcal{X} = \{+1, -1\}$. The energy of the particle in spin state σ in a magnetic field B is given by $E(\sigma) = -B\sigma$. Since there are only two states, the partition function is

$$Z(\beta) = \sum_{\sigma = \{+1, -1\}} e^{-\beta E(\sigma)} = e^{-\beta B} + e^{\beta B}.$$

The probability of the particle being in spin state σ is given by

$$\mu_\beta(\sigma) = \frac{e^{-\beta E(\sigma)}}{Z(\beta)} = \frac{e^{\beta B \sigma}}{e^{-\beta B} + e^{\beta B}}.$$

The average value of the spin, called the **magnetization** is

$$\langle \sigma \rangle = \sum_{\sigma = \{+1, -1\}} \sigma \mu_\beta(\sigma) = \frac{e^{\beta B} - e^{-\beta B}}{e^{-\beta B} + e^{\beta B}} = \tanh(\beta B).$$

- At high temperatures ($T >> |B|$ or $\beta \to 0$), we have $e^{\beta B} \approx e^{-\beta B}$, so $\langle \sigma \rangle \approx 0$.

- At low temperatures ($\beta \to 0$), we have two cases.

  If $B > 0$, then $e^{\beta B} >> e^{-\beta B}$, so $\langle \sigma \rangle \approx 1$. If $B < 0$, then $e^{\beta B} << e^{-\beta B}$, so $\langle \sigma \rangle \approx -1$. That is, we have $\langle \sigma \rangle \approx \text{sign}(B)$.

Example 1.7 (Pott's spin). Another spin variable is the **Potts spin** with q states, which takes values in $\mathcal{X} = [q] = \{1, 2, \ldots, q\}$. In this case, the energy of a particle with Potts spin σ, when it is placed in a magnetic field of intensity B pointing in direction r, is given by

$$E(\sigma) = -B\mathbb{1}_{\{\sigma = r\}}.$$

Now that we have the configuration space and the distribution function, we can calculate the relevant quantities. The partition function is given by

$$Z(\beta) = e^{\beta B} + \underbrace{e^0 + e^0 + \ldots + e^0}_{q-1 \text{ times}} = e^{\beta B} + (q - 1).$$

The Boltzmann distribution is computed to be

$$\mu_\beta(\sigma) = \frac{e^{-\beta E(\sigma)}}{Z(\beta)} = \frac{e^{\beta B \mathbb{1}_{\{\sigma = r\}}}}{e^{\beta B} + (q - 1)}.$$

The Boltzmann average of energy is given by

$$\langle E \rangle = -B\frac{e^{\beta B}}{e^{\beta B} + (q - 1)} + (\text{all other terms are } 0) = \frac{-Be^{\beta B}}{e^{\beta B} + (q - 1)}.$$

Example 1.8 (Particles in a bottle). Let us consider the case of a particle in a closed container (bottle), which is placed in a gravitational field. The space of configurations is the three-dimensional volume of the bottle, which can be represented as $\mathcal{X} = \text{BOTTLE} \subset \mathbb{R}^3$.

The (potential) energy of a particle in a gravitational field is $E(x) = mgh(x) = wh(x)$, where $h(x)$ corresponds to the height of the configuration x. Here, x is not to be confused with the x-coordinate; it is simply a representation of the state of the particle.

In this case, the partition function is difficult to calculate, because $\mathcal{X}$ is an infinite set. However, we can still compute some related quantities in a straightforward fashion. Letting the partition function be $Z(\beta)$, we see that the Boltzmann density is

$$\mu_\beta(x) = \frac{e^{-\beta wh(x)}}{Z(\beta)}, \qquad \mu_\beta(y) = \frac{e^{-\beta wh(y)}}{Z(\beta)}, \qquad \frac{\mu_\beta(x)}{\mu_\beta(y)} = e^{\beta w(h(y) - h(x))}.$$

Let us fix $y = 0$. We get

$$\frac{\mu_\beta(x)}{\mu_\beta(0)} = e^{-\beta wh(x)}$$

The above equation tells us the greater the height of the particle, the smaller the probability of the particle being in that state. This makes physical sense, because we expect most of the particles to be at low heights due to gravity. Note that this result was obtained without any explicit knowledge of the partition function.

2 Temperature Limits

Let us now see that happens to the Boltzmann distribution at high and low temperatures. The probability of the particle being in state x according to the Boltzmann distribution is

$$\mu_\beta(x) = \frac{e^{-\beta E(x)}}{\sum_{x \in \mathcal{X}} e^{-\beta E(x)}}.$$

At high temperatures $\beta \to 0$, and therefore we have $e^{-\beta E(x)} \to 1$ for all values of energy. The partition function is simply $Z(\beta) = \sum_{x \in \mathcal{X}} e^{-\beta E(x)} = |\mathcal{X}|$. The probability of the particle being in any state x is

$$\lim_{\beta \to 0} \mu_\beta(x) = \frac{1}{|\mathcal{X}|}.$$

which is a uniform distribution, i.e. in the high temperature limit, we expect every state to become equally probable. Note that this is in accordance with the Ising spin model. At high temperatures, the spin of the particles become randomly orientated, so the average spin (magnetization) will be 0.

The case of low temperature is more interesting. We need the following definitions to understand the low temperature limit.

Definition 2.1. The **ground state energy** of the system is the lowest energy attainable by it, denoted by $E_0 = \inf\{E(x) : x \in \mathcal{X}\}$. A state corresponding to the ground state energy is called a **ground state**, and the set of ground states is given by $\mathcal{X}_0 = \{x \in \mathcal{X} : E(x) = E_0\}$.

In terms of the ground state $\mathcal{X}_0$, we can write the Boltzmann distribution as

$$\mu_\beta(x) = \frac{e^{-\beta(E(x) - E_0)}}{|\mathcal{X}_0| + \sum_{x \in \mathcal{X} \setminus \mathcal{X}_0} e^{-\beta(E(x) - E_0)}}$$

Note that since $E(x) > E_0$ for all $x \in \mathcal{X} \setminus \mathcal{X}_0$, the second term in the denominator goes to 0 as $\beta \to \infty$. The term in the numerator is 1 if $E(x) = E_0$ and it is 0 otherwise. Therefore, the probability distribution can be written as

$$\mu_\beta(x) = \frac{1}{|\mathcal{X}_0|} \mathbb{1}_{\{x \in \mathcal{X}_0\}}.$$

So, at low temperatures, all the particles will settle to the ground state, and all the other energy levels are left unpopulated. We will be primarily interested in the low temperature limit of systems, as it provides insight into the "low energy structure" of the system. At high temperatures, the states are just random, which does not give us much information.

3 Thermodynamic Potentials

In the study of the Boltzmann distribution, it is often useful to study a few functions depending on the system which summarize the system and provide useful information. Such functions are called **thermodynamic potentials** and are written as functions of the inverse temperature β. The following are the commonly used thermodynamic potentials.

Definition 3.1. The **free energy** of a system is defined as

$$F(\beta) \triangleq -\frac{1}{\beta} \ln Z(\beta).$$

The **free entropy** of a system is defined as

$$\Phi(\beta) \triangleq -\beta F(\beta) = \ln Z(\beta).$$

The **internal energy** of a system is defined as

$$U(\beta) \triangleq \frac{\partial}{\partial \beta} (-\Phi(\beta)) = \frac{\partial}{\partial \beta} (\beta F(\beta)).$$

The **cannonical entropy** of a system is defined as

$$S(\beta) \triangleq \beta^2 \frac{\partial F(\beta)}{\partial \beta} = \beta[\frac{\partial}{\partial \beta}(\beta F(\beta)) - F(\beta)].$$

Using these definitions, we derive some relations between the potentials which will be useful later.

Lemma 3.2. *Basic identities of thermodynamic potentials*

1. *Relation between free energy, internal energy, and canonical entropy:* $F(\beta) = U(\beta) - \frac{1}{\beta}S(\beta) = -\frac{1}{\beta}\Phi(\beta)$.

   Proof. From the definition, we can write

   $$U(\beta) = \frac{\partial}{\partial \beta}(\beta F(\beta)) = F(\beta) + \beta \frac{\partial F(\beta)}{\partial \beta} = F(\beta) + \frac{1}{\beta}S(\beta).$$

   Rearranging, we get the required relation. □

2. *Internal energy is the expected energy under the Boltzmann distribution. That is, $U(\beta) = \langle E \rangle$.*

   Proof. From the definition of free entropy, we can write

   $$U(\beta) = \frac{\partial}{\partial \beta}(-\Phi(\beta)) = \frac{\partial}{\partial \beta}(-\ln Z(\beta)) = -\frac{1}{Z(\beta)}\frac{\partial Z(\beta)}{\partial \beta}.$$

   For a finite configuration space $\mathcal{X}$, we can exchange the partial derivative and the summation, to write

   $$U(\beta) = -\frac{1}{Z(\beta)}\frac{\partial}{\partial \beta}\sum_{x \in \mathcal{X}} e^{-\beta E(x)} = -\frac{1}{Z(\beta)}\sum_{x \in \mathcal{X}} e^{-\beta E(x)}(-E(x)) = \sum_{x \in \mathcal{X}} \mu_\beta(x)E(x) = \langle E \rangle.$$

   □

3. *Canonical entropy is the Shannon entropy of the Boltzmann distribution. That is, $(\beta) = H(\mu_\beta)\ln 2$.*

   Proof. From the result in first two parts and the definition of free entropy, we can write the canonical entropy as

   $$S(\beta) = \Phi(\beta) + \beta U(\beta) = \ln Z(\beta) + \sum_{x \in \mathcal{X}} \beta E(x)\mu_\beta(x) = \sum_{x \in \mathcal{X}} \mu_\beta(x)(\beta E(x) + \ln Z(\beta)).$$

   Now note that $\mu_\beta(x) = \frac{e^{-\beta E(x)}}{Z(\beta)}$. Taking natural log on both sides, we have $\ln \mu_\beta(x) = -\beta E(x) - \ln Z(\beta)$. Therefore, we recognize that $S(\beta) = -\sum_{x \in \mathcal{X}} \mu_\beta \ln \mu_\beta(x) = H(\mu_\beta)\ln 2$. □

4. *The free entropy function $\Phi(\beta)$ is convex in inverse temperature β.*

   Proof. We can write the second partial derivative of free entropy as first partial derivative of the negative free energy, and hence

   $$\frac{\partial^2}{\partial \beta^2}(\Phi(\beta)) = -\frac{\partial}{\partial \beta}(U(\beta)) = -\frac{\partial}{\partial \beta}\left(\frac{\sum_{x \in \mathcal{X}} e^{-\beta E(x)}E(x)}{Z(\beta)}\right) = \sum_{x \in \mathcal{X}} \mu_\beta(x)E(x)^2 - \sum_{x \in \mathcal{X}} \mu_\beta(x)E(x)\frac{1}{Z(\beta)}\frac{\partial Z(\beta)}{\partial \beta}.$$

   We recall that $\frac{\partial \ln Z(\beta)}{\partial \beta} = \langle E \rangle$, to obtain

   $$\frac{\partial^2}{\partial \beta^2}(\Phi(\beta)) = -\frac{\partial^2}{\partial \beta^2}(\beta F(\beta)) = \langle E^2 \rangle - \langle E \rangle^2.$$

   This is auto-covariance of energy E under the Boltzmann distribution μ_β, and is always positive. This implies the convexity of free entropy as a function of inverse temperature. □