

Maxwell-Boltzmann Distribution

by Tony Shing

Overview:

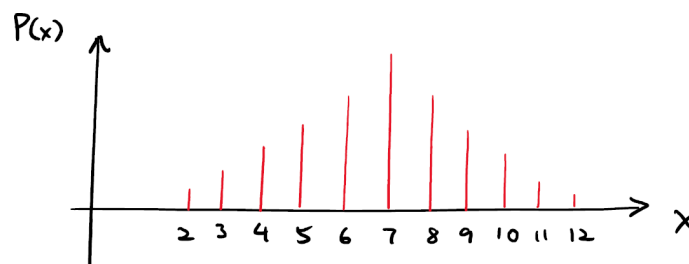
- Prerequisite: Probability, extending to continuous random variables
- Boltzmann factor
- Deriving Maxwell-Boltzmann distribution

1 Probability Distribution: Discrete & Continuous

1.1 Probability Mass Function (PMF)

For random variables that can be picked from a finite number cases, we describe them using the **probability mass function** (PMF), $P(x)$, or most of the time we just call them probabilities.

E.g. By throwing 2 dices and sum their values, we obtain a list of different probabilities for getting any integer between 2 to 12.



- **Normalization requirement:** The probability of all possible outcomes add up to 1.

$$\sum_{\text{all possible } x_i} P(x_i) = 1$$

- **Expected value:** i.e. The weighted average of all outcomes according to the probabilities, which is the average value you would get after repeating the process ∞ times.

$$E[x] \stackrel{\text{def}}{=} \sum_{\text{all possible } x_i} x_i P(x_i)$$

- **Variance:** It is the average of (distance)² of all outcomes relative to the expected value. Taking square root gives the **standard derivation** (S.D.) = $\sqrt{\text{Variance}}$.

$$\text{Var}[x] \stackrel{\text{def}}{=} \underbrace{E[(x - E[x])^2]}_{\substack{\text{Then take average} \rightarrow \\ \text{P.1}}} \quad \underbrace{(x - E[x])^2}_{\substack{\text{Square for taking magnitude only} \\ (x - E[x]) = x\text{'s distance from the expected value}}}$$

Side note:

In a special case where there are N different outcomes and every outcome has equal probability to occur,

$$P(x_i) = \frac{1}{N}$$

Then the expected value, variance and S.D. will reduce to the formula that we can find in high school textbook.

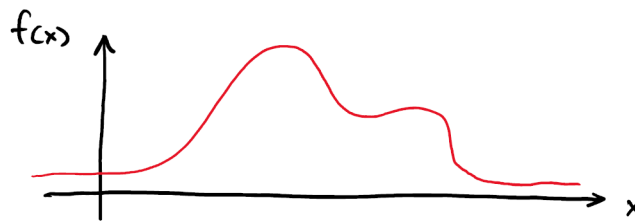
$$E[x] = \sum_{\text{all possible } x_i} x_i P(x_i) = \frac{\sum_i x_i}{N} \stackrel{\text{def}}{=} \bar{x}$$

$$\text{Var}[x] = E[(x - E[x])^2] = \frac{\sum_i (x_i - E[x])^2}{N} = \frac{\sum_i (x_i - \bar{x})^2}{N}$$

$$\text{S.D.} = \sqrt{\text{Var}[x]} = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{N}}$$

1.2 Probability Density Function (PDF)

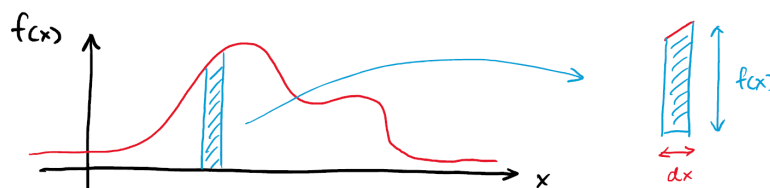
For random variables that can appear as any value within an interval, we describe them using the **probability density function** (PDF).



Pay attention here: the output value of the PDF, $f(x)$, is NOT the probability of getting x . This is because there are always infinitely many real number along any interval of real numbers. **The probability of getting an exact number is limited to 0.**

Instead, PDF can only describe the probability of getting a value within an interval $[x_0, x_0 + dx]$.

$$P(\underbrace{x_0 < x < x_0 + dx}_{\text{Prob. inside the interval}}) = \underbrace{f(x)}_{\text{height}} \underbrace{dx}_{\text{width}}$$



As interpreted, the probability is represented by the area under curve $f(x) dx$, not the height $f(x)$. This is also why it is named "density" function - the probability per "length" of x .

The formula for PMF (discrete cases) can be extended to PDF (continuous cases).

- **Normalization requirement:** The probability of all possible cases add up to 1. From the condition for PMF, we can extend the sum to integral for PDF.

$$\sum_{\text{all possible } x_i} P(x_i) = 1 \quad \Rightarrow \quad \int_{-\infty}^{\infty} f(x) dx = 1$$

- **Expected value:** The weighted average of all outcomes according to the probabilities. Can be understood by the weighted sum interpretation of integration.

$$E[x] \stackrel{\text{def}}{=} \sum_{\text{all possible } x_i} x_i P(x_i) \quad \Rightarrow \quad E[x] \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} x f(x) dx$$

- **Variance:** Same as in the discrete case, it is the average of (distance)² of all outcomes relative to the expected value.

$$\text{Var}[x] \stackrel{\text{def}}{=} E[(x - E[x])^2] \quad \Rightarrow \quad \text{Var}[x] \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} (x - E[x])^2 f(x) dx$$

Taking square root gives the standard deviation (S.D.) = $\sqrt{\text{Var}[x]}$.

2 Thermal Fluctuation

In thermodynamics, every state parameter is a statistical property - a state being labeled with parameter $x = x_0$ only means the expected value of observing x is x_0 . Its actual value can random fluctuate, different in value from time to time.

For example, in the 2-boxes system with 1000 balls, the macrostates are described by one state parameter N_R = the number of balls in the right box. Because the balls are always free to move between the boxes, N_R must fluctuate.

- Starting from 500L-500R. One of the ball may temporarily move from the left box to the right box. The system's state change to 499L-501R.
- After some time, another ball moves from the right box to the left box. The system's state returns to 500L-500R.

(add figure here: moving ball in and out)

Because of this randomness, we expect that measurement to any parameter x will form some probability distribution centered around its expected value $x = x_0$.

What is the probability for x to deviate by Δx ? i.e. Observing x to be $x_0 + \Delta x$?

2.1 Entropy Fluctuation

Recall that the probability of observing a macrostate with parameter x is proportional to the number of microstates (i.e. multiplicity) it corresponds to,

$$P(x = x_0) = \frac{\left(\begin{array}{c} \# \text{ of microstates that} \\ \text{leads to observing } x = x_0 \end{array} \right)}{(\text{Total } \# \text{ of microstates)}}$$

Side note:

Recall the statement of Clausius theorem - entropy of a closed system must either be a constant or increase along any process. However this is only true for the expected value.

$$\Delta S \geq 0 \quad \xrightarrow{\text{More accurately}} \quad \Delta S_{\text{avg.}} \geq 0$$

Irreversibility exists only because systems tend to evolve from macrostates of low multiplicity to high multiplicity. The reverse is allowed to happen, although at a much lower chance.

The fluctuation in multiplicity is most likely observed after the system has reached its maximum entropy state - because any fluctuation will cause the system evolves into a lower multiplicity state. After some time multiplicity increases again due to probability.

And multiplicity is related to the entropy of the macrostate. So if we know the function form of multiplicity or entropy related to the parameter, i.e. $W(x)$ or $S(x)$, we can tell the probability of observing a particular value of x .

$$P(x = x_0) = \frac{e^{\frac{S(x_0)}{k}}}{Z} \quad \text{where} \quad Z = \sum_{\substack{\text{all possible} \\ x_0}} e^{\frac{S(x_0)}{k}}$$

In particular, if the system is at its maximum entropy state when $x = x^*$, we can relate the fluctuation in entropy ΔS can be related to the fluctuation of the state parameter Δx by Taylor series expansion:

$$\begin{aligned} S(x^* + \Delta x) &= S(x^*) + \Delta S \\ &= S(x^*) + \left. \frac{\partial S}{\partial x} \right|_{x=x^*} (\Delta x) + \frac{1}{2} \left. \frac{\partial^2 S}{\partial x^2} \right|_{x=x^*} (\Delta x)^2 + \dots \end{aligned}$$

Suppose the system has a state parameter x (E.g. In the 2-boxes model, we have x = number of balls in the right box). What is the probability to observe it fluctuates by a magnitude Δx ?

In general, we can express the entropy S of a system as a function of another state parameter x

2.2 Boltzmann Distribution

The fluctuation for internal energy is a bit more interesting - because energy of a closed system must be a constant, there shall not be fluctuation for the total internal energy. However, we can consider the fluctuation of internal energy on individual objects in the system.

A direct consequence of entropy fluctuation is that objects in a closed system are allowed to exchange energy (as long as total energy conserves).

Now we can apply the fluctuation. However one thing very different between energy and other state parameters is that - energy must conserve in any scenario.

In the simplest 2-boxes model, without energy consideration, we used $x = n_L$ = number of balls in the left box as the state parameter.

when two such model is put together, total multiplicity is just a

When $x=nL$, nL can take whatever value,

energy is a conserved quantity if one object has high energy then the other must have low energy

Multiplicity under Energy Constraint

Recall that in the 2-boxes system with energy difference, the multiplicity of the system is a function to internal energy. If the system is closed, i.e. its internal energy is constant because it is not in thermal contact with anything, then neither its multiplicity or entropy can vary.

(add figure here: bar chart S)

But in reality, the only closed system is the whole universe because 100% heat insulation from heat radiation does not exist. Every object in the universe is constantly under thermal contact with other objects - sometimes you may observe its internal energy to be some value U_1 , and sometimes you may observe it to be another value U_2 .

What is the probability to observe that the object carry internal energy U_i ?

Previously, we have only dealt with closed system - any objects that involve in heat exchange are considered as part of the system.

2.3 The Boltzmann Factor

We model the object as thermally contacting with a HUGE heat bath (i.e. the universe), where the only form of energy exchange is by heat transfer.

(add figure here: heat bath contact)

We require the following assumptions:

1. The heat bath is so HUGE that any energy exchange has no effect on it. Its state parameters (e.g. temperature, entropy,...) remains constant all the time.
2. The total energy of the object + heat bath is conserved.

$$U_{\text{total}} = U_{\text{bath}} + U_{\text{obj}} = \text{const.}$$

3. Energy carried by the heat bath is much greater than the energy carried by the object. i.e.

$$U_{\text{bath}} \gg U_{\text{obj}}$$

Let the entropy of the heat bath be S_{bath} , which must be some function to its internal energy U_{bath} (although we do not know what this function is). We can make a Taylor expansion by

$$\begin{aligned}
 S_{\text{bath}} &= S_{\text{bath}}(U_{\text{bath}}) \\
 &= S_{\text{bath}}(U_{\text{total}} - U_{\text{obj}}) \\
 &\approx S_{\text{bath}}(U_{\text{total}}) - \left. \frac{dS_{\text{bath}}}{dU} \right|_{U=U_{\text{total}}} \cdot U_{\text{obj}} \\
 &= S_{\text{bath}}(U_{\text{total}}) - \frac{U_{\text{obj}}}{T_{\text{bath}}}
 \end{aligned}$$

The total multiplicity of the {heat bath + object} is thus

$$\begin{aligned}
 S_{\text{total}} &= S_{\text{bath}} + S_{\text{obj}} \\
 W_{\text{total}} &= e^{\frac{S_{\text{bath}} + S_{\text{obj}}}{k}} \\
 &\approx e^{\frac{S_{\text{bath}}(U_{\text{total}}) - \frac{U_{\text{obj}}}{T_{\text{bath}}}}{k}} \cdot e^{\frac{S_{\text{obj}}}{k}} \\
 &= e^{\frac{S_{\text{bath}}(U_{\text{total}})}{k}} \cdot e^{-\frac{U_{\text{obj}}}{kT_{\text{bath}}}} \cdot e^{\frac{S_{\text{obj}}}{k}} \\
 &= \underbrace{e^{\frac{S_{\text{bath}}(U_{\text{total}})}{k}}}_{\substack{\text{U}_{\text{total}} \text{ is a constant} \\ \text{so this term is just a constant}}} \cdot e^{-\frac{U_{\text{obj}}}{kT_{\text{bath}}}} \cdot \underbrace{e^{\frac{S_{\text{obj}}}{k}}}_{\substack{\text{Multiplicity of the object} \\ \text{is some function of } U_{\text{obj}}}} = W_{\text{obj}}
 \end{aligned}$$

Now the total multiplicity is expressed purely a function of U_{obj} . This is the total multiplicity (i.e. number of microstates) when the object has a particular value of internal energy U_{obj} with the heat bath maintained at temperature T_{bath} :

$$W_{\text{total}}(U_{\text{obj}}) = (\text{const.}) \cdot e^{-\frac{U_{\text{obj}}}{kT_{\text{bath}}}} \cdot W_{\text{obj}}(U_{\text{obj}})$$

To find the probability which the object has a particular value of internal energy U_{obj} , we simply divide it by the total multiplicity (total number of microstates), which is summed from all possible U_{obj} .

$$P(U_{\text{obj}}) = \frac{W_{\text{total}}(U_{\text{obj}})}{\sum_{\substack{\text{all possible} \\ U_{\text{obj}}}} W_{\text{total}}(U_{\text{obj}})}$$

But the thing we are actually interested in is the probability to find the object having certain value of internal energy. Recall the relation between multiplicity and probability:

$$W(U) =$$

Let the entropy-energy relation of the heat bath and the object be some functions (which we may or may not know its formula),

$$, \quad S_{\text{obj}} = S_{\text{obj}}(U_{\text{obj}})$$

2.4 Example: System with Discrete Energy States

3 Density of State of Ideal Gas

The most important application of the above ideas is to find the energy-probability relation for states in ideal gas, when heat exchange is allowed.

i.e. States with which energy are most probable to be found, given that it is at temperature T ?

Since energy of ideal gas can be any real number, we first need to adapt the formula from last section for continuous distribution.

$$\sum_{\substack{\text{all possible} \\ U}} e^{-\frac{U}{kT}} W(U) \quad \rightarrow \quad \int_0^{\infty} e^{-\frac{U}{kT}} W(U) dU$$

W

What is the $W(U)$ for ideal gas?

3.1 Entropy of Ideal Gas

3.2 Average Speed of Ideal Gas

In textbooks you can find 2 different formulas related to the average speed about ideal gas. How are they derived?

1. Average Speed = Expected value to $|v|$, i.e. $E(|v|)$.

$$\begin{aligned} E(|v|) &= \int_0^{\infty} |v| \cdot P(|v|) d|v| \\ &= \int_0^{\infty} |v| \cdot \sqrt{\frac{2m^3}{\pi k^2 T^3}} |v|^2 e^{-\frac{m|v|^2}{2kT}} d|v| \\ &= \sqrt{\frac{2m^3}{\pi k^2 T^3}} \left(\frac{2kT}{m}\right)^2 \int_0^{\infty} \left(\sqrt{\frac{m}{2kT}} v\right)^3 e^{-(\sqrt{\frac{m}{2kT}} v)^2} d\left(\sqrt{\frac{m}{2kT}} v\right) \\ &= \sqrt{\frac{32kT}{\pi m}} \int_0^{\infty} x^3 e^{-x^2} dx \\ &= \sqrt{\frac{32kT}{\pi m}} \cdot \frac{1}{2} \\ &= \sqrt{\frac{8kT}{\pi m}} \stackrel{\text{def}}{=} v_{\text{avg}} \end{aligned}$$

2. Root Mean Square Speed = $\sqrt{\text{Expected value to } |v|^2}$, i.e. $\sqrt{E(|v|^2)}$.

$$\begin{aligned}
 E(|v|^2) &= \int_0^\infty |v|^2 \cdot P(|v|) \, d|v| \\
 &= \int_0^\infty |v|^2 \cdot \sqrt{\frac{2m^3}{\pi k^2 T^3}} |v|^2 e^{-\frac{m|v|^2}{2kT}} \, d|v| \\
 &= \sqrt{\frac{2m^3}{\pi k^2 T^3}} \left(\frac{2kT}{m}\right)^3 \int_0^\infty \left(\sqrt{\frac{m}{2kT}} v\right)^3 e^{-(\sqrt{\frac{m}{2kT}} v)^2} \, d\left(\sqrt{\frac{m}{2kT}} v\right) \\
 &= \sqrt{\frac{64k^2 T^2}{\pi m^2}} \int_0^\infty x^4 e^{-x^2} \, dx \\
 &= \sqrt{\frac{64k^2 T^2}{\pi m^2}} \cdot \frac{3\sqrt{\pi}}{8} \\
 &= \frac{3kT}{m} \\
 \sqrt{E(|v|^2)} &= \sqrt{\frac{3kT}{m}} \stackrel{\text{def}}{=} v_{\text{rms}}
 \end{aligned}$$

Note: Because the average KE (which is something we can measure) is calculated as

$$\begin{aligned}
 \text{Avg. KE} &= E\left(\frac{1}{2}m|v|^2\right) \\
 &= \frac{1}{2}mE(|v|^2) \\
 &= \frac{1}{2}mv_{\text{rms}}^2
 \end{aligned}$$

The root mean square speed is way more useful than the average speed.

— The End —