Entropy & Probability

(i) Derive $S = f(p)$

(ii) Why nature prefers Maxwell-Boltzmann distributions at equilibrium?

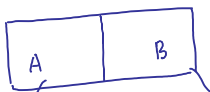
$p(v) \sim N e^{-mv^2/k_B T}$ why not $\cancel{f(v) = e^{-\alpha v} ? \; = v^n \; \vdots}$

Boltzmann's def^n of Entropy :

$$S = k_B \ln \Omega$$

Why not $\cancel{S \sim e^{\Omega}}$? or $\cancel{S \sim \Omega^N}$?

S must be additive (Extensive !)



has Ω_A states

$S_A \sim \ln \Omega_A$

has Ω_B states

$S_B \sim \ln \Omega_B$

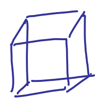
$$S = \text{Total Entropy of } A+B = \ln \Omega_A + \ln \Omega_B$$
$$= \ln (\Omega_A \Omega_B)$$
$$= \ln (\Omega_{total})$$

$$\Omega_{total} = \Omega_A \Omega_B$$

Hence $S \sim \ln \Omega$ becomes extensive (additive)

Chasing connection $S = f(p)$?

Example: Dice with 6 faces



Take N such dice

Roll them!

6 5 6 . . . 1 $\rightarrow$ out comes "One possible state of N-dice system"

Classify this one state: $n_1 = 7$

$n_2 = 11$

$\vdots$

$n_6 = 19$

$\sum_{i=1}^{6} n_i = N$ "Total no. of dice"

4 3 6 . . . 5 $\underline{\{n_1 = 7, n_2 = 11, \ldots n_6 = 19\}}$

$\vdots$

Four coins: $N = 4$

$n_H = 2$

$n_T = 2$

$n_H + n_T = 4 = N$

H H T T

H T T H

$\vdots$ $\Big\}$ $n_H = 2, n_T = 2$

$\Omega = \frac{4!}{2!\,2!} = 6$ micro-states

$n_1, n_2, n_3, n_4, n_5, n_6$

$\sum_{i=1}^{6} n_i = N$ = Total no. of dice

$$\Omega = \frac{N!}{\prod_{i=1}^{6} (n_i!)}$$

Now if $N >> 1$, $N! = (N/e)^N$ "Stirling Approximation"

$= (N/e)^N \sqrt{2\pi N}$ "better"

$$\Omega = \frac{(N/e)^N}{\prod_{i=1}^{6} (n_i/e)^{n_i}}$$

$$\ln \Omega = S/k_B = \ln (N/e)^N - \sum_{i=1}^{6} \ln \left(\frac{n_i}{e}\right)^{n_i}$$

$$S/k_B = N\ln N - \cancel{N} - \sum_{i=1}^{6}\left(n_i \ln n_i - \cancel{n_i}\right) \qquad \because \sum_{i=1}^{6} n_i = N$$

$$= N\ln N - \sum_{i=1}^{6} n_i \ln\left(\frac{n_i}{N}\cdot N\right)$$

$$= N\ln N - \sum_{i=1}^{6}\left[n_i \ln\left(\frac{n_i}{N}\right) + n_i \ln N\right]$$

we know $n_i/N = p_i$

$$= \cancel{N\ln N} - \sum_{i=1}^{6} n_i \ln p_i \cancel{- \underbrace{\sum_{i=1}^{6} n_i}_{N} \ln N}$$

$$= -N \sum_{i=1}^{6} \frac{n_i}{N} \ln p_i$$

$$= -N \sum_{i=1}^{6} p_i \ln p_i \qquad \because p_i = n_i/N$$

$$\boxed{S/k_B = -N \sum_{i=1}^{t} p_i \ln p_i}$$

$t = 6$ in this case!

Shannon Entropy.

(ii) Entropy maximation or "Maximum entropy principle".

Optimization under constraint:

$f(x,y)$ is some function.

Maximise $f(x,y)$ under constraint $g(x,y) = 0$

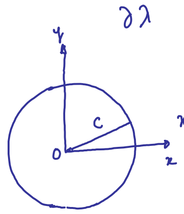
Lagrange multipliers can be used.

Lagrangian $\mathcal{L} = f(x,y) - \underbrace{\lambda}_{\text{Lagrange multiplier}} g(x,y)$

$$\left.\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} &= \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0\end{aligned}\right\} \text{Maximization!}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = g = 0 \quad \{\text{constraint}\}$$

Eg.



$x^2 + y^2 = c^2$

What is pt. of circle where $x+y$ becomes maximum?

Maximize $f(x,y) = x+y$

subject to constraint : $g(x,y) = x^2+y^2-c^2 = 0$

$$\mathcal{L} = f - \lambda g$$

$$= (x+y) - \lambda (x^2+y^2-c^2)$$

$$\frac{\partial \mathcal{L}}{\partial x} = 1 - 2x\lambda = 0 \implies x = 1/2\lambda$$

$$\frac{\partial \mathcal{L}}{\partial y} = 1 - 2y\lambda = 0 \implies y = 1/2\lambda$$

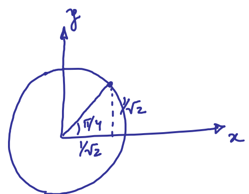
$$\frac{\partial \mathcal{L}}{\partial \lambda} = x^2+y^2-c^2 = 0 \implies 1/4\lambda^2 + 1/4\lambda^2 - c^2 = 0$$

$$\implies \lambda = +\sqrt{\frac{1}{2}} \cdot \frac{1}{c}$$

$c^2 = 1$ (Unit circle)

$\lambda = \sqrt{\frac{1}{2}}$.

$(x^*, y^*) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$



$x+y\big|_{x=1/\sqrt{2},\, y=1/2} = \sqrt{2}$

Maximum Entropy distributions:

(i) **Discrete distribution:** Dice, Coins ....

Take t-faced dice. $\sum_{i=1}^{t} p_i = 1$

Constraint: $g(\{p_i\}) = \sum_{i=1}^{t} p_i - 1 = 0$

Function: $f(\{p_i\}) = \frac{S}{Nk_B} = -\sum_{i=1}^{t} p_i \ln p_i$

Maximise $f(\{p_i\}) = S/Nk_B$ subject to constraint $g(\{p_i\}) = 0$

$$\mathcal{L} = f(\{p_i\}) - \lambda g(\{p_i\})$$

$\{p_i\} = \{p_1, p_2 \dots p_t\}$

$$\frac{\partial \mathcal{L}}{\partial p_j} = \frac{\partial}{\partial p_j}\left[f(\{p_i\}) - \lambda g(\{p_i\})\right]$$

$$= \frac{\partial}{\partial p_j}\left[-\sum_{i=1}^{t} p_i \ln p_i - \lambda\left\{\sum_{i=1}^{t} p_i - 1\right\}\right]$$

$$= -(1 + \ln p_j) - \lambda = 0$$

.... p_j is the maximum of $f = \frac{S}{Nk_B}$

$$\Rightarrow 1 + \ln p_j = -\lambda$$

$$\Rightarrow \underbrace{p_j}_{\text{depends on } j^{th} \text{ face}} = \underbrace{e^{-\lambda-1}}_{\text{Const.}}$$

Const. → implies unbiased dice.

Normalize p_j: $\mathcal{N} = \sum_{j=1}^{t} p_j$

$$= t\, e^{-\lambda-1}$$

$$p_j = \frac{e^{-\lambda-1}}{t\, e^{-\lambda-1}} = \frac{1}{t}$$

$t = 6$ face dice.

$$p_j = \frac{1}{6}$$ "Equal-apriori – Unbiased"

Probability distribution is normalized (Only constraint)