Elementary large deviation

origin of entropy — Boltzmann

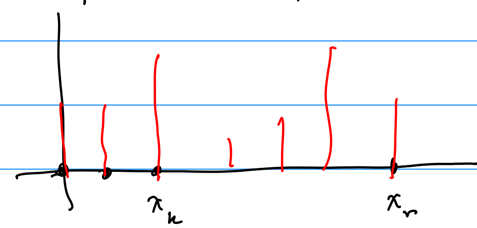
$X = \{x_1, \dots, x_r\} \qquad p_1, \dots, p_r$

long sequence ω of independent trials

e $\quad n = (n_1, \dots, n_r), \quad n_k = n_k(\omega) = \#$ of occurrences of x_k in ω

$=$ the occupation numbers of ω

$|n| = \sum n_k = N$



$(1+|n|)^r$ occupation numbers

$$p(\omega) = p_1^{n_1} p_2^{n_2} \cdots p_r^{n_r}$$

lost details of ω only know occupation nos.

$$p(n) = (\#\text{ of } \omega \text{ with } n)\, p(\omega) = \sum_{\omega \sim n} p(\omega)$$

$$= \binom{|n|}{n} p(\omega)$$

$$p(n) = e^{-|n| A} = e^{-|n| \cdots\cdots - \log\binom{|n|}{n}}$$

$$\log m! = m \log m - m + O(\log m)$$

$$A = -\left(\frac{n_1}{|n|}\log p_1 + \frac{n_2}{|n|}\log p_2 + \cdots + \frac{n_r}{|n|}\log p_r\right) - \frac{1}{|n|}\log\binom{|n|}{n}$$

$$= \sum \rho_k \log \frac{\rho_k}{p_k} + \text{small error}$$

$\Phi(\rho \| p) =$ Kullback-Leibler relative entropy of ρ w.r.to p

$$= \sum \rho_k \log \frac{\rho_k}{p_k} \geq 0 \qquad \text{Jensen}$$

$p =$ uniform $\Longrightarrow$ $\Phi(\rho \| p) =$ ordinary entropy

$$p(n) = e^{-|n|\left(\Phi(q \| p) + \text{small error}\right)}$$

$$\longrightarrow 0 \quad \text{unless} \quad q \longrightarrow p$$

make more precise:

$$A_\varepsilon = \left\{ q : \max \left| \frac{n_k}{|n|} - p_k \right| \geq \varepsilon \right\} \qquad p(A_\varepsilon) \to 0 \text{ as } |n| \to \infty$$

LLN

$$p(A_\varepsilon) \leq (1+|n|)^r \, e^{-|n|\,\alpha(\varepsilon)}$$

$$\alpha(\varepsilon) = \min_{A_\varepsilon} \Phi(q|p) + \text{small error}$$

With probability = 1, all trials cluster around state of minimum entropy

Return to energy

ω : set of $\overset{U}{|n|+1} \longrightarrow X$

length of trial

Let ψ be a function on X, $\psi \geq 0$

$$E(\omega) = \sum_{U} \psi(\omega(u)) = \sum_{k=1}^{r} \psi(x_k)\, n_k \qquad \text{energy of } \omega$$

$$p(\omega) = \frac{1}{Z} e^{-E(\omega)}$$

$$\Phi(q \| p) = F(q) + \log Z$$

$$F(q) = \sum \psi(x_k)\, q_k + \sum q_k \log q_k$$

free energy of q

- explained entropy as result of changing the ensemble
- explained distribution corresponding to minimum energy as the most probable one

Dalton's law $\sim$ independent trials

Information theory example

prefix codes