Sample Solutions Exercise 8

Exercise 8.1: Maximal entropy (3P)

The Shannon entropy of a probability distribution is given by

$$H = -\sum_c p_c \ln p_c \,.$$

(a) Use Jensens inequality for convex functions to show that the Shannon entropy attains its global maximum for a uniform distribution. (1P)

(b) Prove the same statement in the framework of variational calculus. Use the method of Langrange multipliers to take the normalization constraint of the probability distribution into account. (2P)

Sample Solution

(a) The claim of Jensen's inequality is the following: If a random variable X is mapped by a convex function f, the corresponding expectation values $\langle ... \rangle$ will obey the inequality

$$f\big(\langle X\rangle\big) \leq \langle f(X)\rangle \,.$$

Applied to a probability distribution $\{p_c\}$ this implies the inequality

$$f\Big(\sum_c p_c x_c\Big) \leq \sum_c p_c f(x_c)$$

Let us now choose $f(x_c) = \ln(x_c)$ and $x_c = 1/p_c$. Here one has to take into account that the logarithm is not a convex function, it is rather concave so that we have to flip the inequality. This leads to the following expression for the entropy

$$H = -\sum_c p_c \ln p_c = +\langle \ln(1/p_c)\rangle \leq \ln\Big(\langle 1/p_c\rangle\Big) = \ln \sum_c p_c \frac{1}{p_c} = \ln|\Omega| \,.$$

Therefor, H is bounded from above by $\ln|\Omega|$. In the case of equally probable configurations (microcanonical equilibrium) this value is actually reached.

(b) The aim is to maximize the Shannon entropy $H = -\sum_c p_c \ln p_c$ under the constraint of the normalization $\sum_c p_c = 1$. To this end we apply the method of Lagrange multiplier, minimizing the functional

$$F = -\sum_c p_c \ln p_c + \lambda\Big(\sum_c p_c - 1\Big) \,,$$

where we assume the p_c to be independent. This results into $|\Omega|+1$ equations, namely:

$$0 = -\ln p_c - 1 + \lambda \qquad \forall c \in \Omega$$
$$0 = \sum_c p_c - 1$$

The first equation gives

$$p_c = e^{-1+\lambda}\,.$$

Therefore, one obtains one solution which is proportional to a constant distribution. The second multiplier has to be chosen in such a way that the second equation is obeyed. This results into a normalized constant distribution of maximal entropy.

Exercise 8.2: Master equation of embedded subsystems (3P)

A laboratory system with the configuration space Ω^{sys} is embedded into the environment by means of a projection $\pi : \Omega^{\text{tot}} \mapsto \Omega^{\text{sys}} : c \mapsto s$ (see lecture notes). The total system is ergodic and and has a unique stationary state.

(a) Suppose that the probability distribution $P_c(t)$ of the total system evolves ergodically according to the master equation $\frac{\mathrm{d}}{\mathrm{d}t}P_c(t) = \sum_{c'} P_{c'}(t) w_{c'\to c} - P_c(t) \sum_{c'} w_{c\to c'}$ with symmetric time-independent rates $w_{c'\to c} = w_{c\to c'} \geq 0$. Show that the probability distribution of the laboratory system is exactly given by a master equation with certain effective rates that may be non-symmetric and time-dependent. (1P)

(b) Prove that for a laboratory system in a genuine non-equilibrium steady state (stationary and violating detailed balance), the corresponding total system must be non-stationary and infinitely large. (2P)

Sample Solution

(a) The marginal probabilities in the laboratory system are defined by

$$P_s(t) = \sum_{c\in s} P_c(t)\,.$$

Likewise, the marginal probability currents in the laboratory system read

$$J_{s\to s'}(t) = \sum_{c\in s} \sum_{c'\in s'} J_{c\to c'}(t) = \sum_{c\in s} \sum_{c'\in s'} P_c(t) w_{c\to c'}\,.$$

Therefore, the effective rates have to be defined by

$$w_{s\to s'}(t) = \frac{J_{s\to s'}(t)}{P_s(t)} = \frac{\sum_{c\in s}\sum_{c'\in s'} P_c(t) w_{c\to c'}}{\sum_{c\in s} P_c(t)}$$

Note that these rates are generally non-symmetric and time-dependent. So far this is all written in the lecture notes.

Using these rates one can prove the validity of the master equation in the laboratory system simply by reorganizing the sums: (1P)

$$\begin{aligned}\dot{P}_s(t) &= \sum_{c\in s}\dot{P}_c(t) = \sum_{c\in s}\sum_{c'}(J_{c'\to c}(t) - J_{c\to c'}(t))\\ &= \sum_{c\in s}\Big[\sum_{s'}\sum_{c'\in s'}\Big](J_{c'\to c}(t) - J_{c\to c'}(t))\\ &= \sum_{s'}\sum_{c\in s}\sum_{c'\in s'}(J_{c'\to c}(t) - J_{c\to c'}(t)) = \sum_{s'}(J_{s'\to s}(t) - J_{s\to s'}(t))\\ &= \sum_{s'}(P_{s'}(t)w_{s'\to s}(t) - P_s(t)w_{s\to s'}(t))\end{aligned}$$

Note that this master equation is exact.

(b) Proof by contradiction. We have to show: If the corresponding total system is either finite or stationary, then the laboratory system (the subsystem) cannot be in a non-equilibrium steady state (NESS).

- If the total system is finite, we know that the stationary state is unique, so the largest relaxation time of the total system is finite, and this means that we reach the stationary state of the total system (the equally distributed state) within finite time. This state, as we have shown, obeys detailed balance. (1P)
- If the total system is (infinite but) stationary, then – knowing that the stationary state is unique – the we are already in the equally distributed state. This state obeys detailed balance. (1P)

Exercise 8.3: Mutual information (6P)

With this exercise we want to recall some basic notions of statistics in bipartite systems, namely, marginal and conditional probabilities, entropies, as well as mutual information (see Wikipedia or any textbook): Consider two random variables $X, Y \in \{0,1\}$ with the joint probability distribution $P_{XY}(x,y)$ defined by

$$P_{XY}(0,0) = 1/2, \quad P_{XY}(0,1) = 1/4, \quad P_{XY}(1,0) = 0, \quad P_{XY}(1,1) = 1/4\,.$$

Let $P_X(x)$ and $P_Y(y)$ be the corresponding marginal probabilities. Please compute

(a) the joint (total,full) entropy H_{XY}. (1P)

(b) the marginal entropies H_X and H_Y. (1P)

(c) the conditional entropies $H_{X|Y}$ and $H_{Y|X}$. (2P)

(d) the mutual information $I_{X:Y} = H_X + H_Y - H_{XY}$. Verify that the mutual information is also given by $I_{X:Y} = H_X - H_{X|Y} = H_Y - H_{Y|X}$. (2P)

Sample Solution

The marginal probabilities read

$$P_X(1) = 1 - P_X(0) = \frac{1}{4}\,, \qquad P_Y(1) = 1 - P_Y(0) = \frac{1}{2}.$$

(a)

$$H_{XY} = -\sum_{x,y=0}^{1} P_{XY}(x,y)\log_2 P_{XY}(x,y) = \ldots = \frac{3}{2}$$

(b)

$$H_X = -\sum_{x=0}^{1} P_X(x)\log_2 P_X(x) = \ldots = 2 - \frac{3}{4}\log_2 3 = 0.811278$$

$$H_Y = -\sum_{y=0}^{1} P_Y(y)\log_2 P_y(y) = \ldots = 1$$

(c) The conditional probabilities read

$$P_{X|0}(0) = 1 - P_{X|0}(1) = 1\,, \qquad P_{X|1}(0) = 1 - P_{X|1}(1) = \frac{1}{2}\,,$$

$$\Rightarrow \quad H_{X|0} = 0\,, \qquad H_{X|1} = 1$$

Likewise:

$$P_{Y|0}(0) = 1 - P_{Y|0}(1) = \frac{2}{3}\,, \qquad P_{X|1}(0) = 1 - P_{X|1}(1) = 0.$$

$$\Rightarrow \quad H_{Y|0} = \log_2(2) - \frac{2}{3} = 0.918296\,, \qquad H_{Y|1} = 0$$

This gives:

$$H_{X|Y} = H_{X|0}P_Y(0) + H_{X|1}P_Y(1) = \frac{1}{2}$$

$$H_{Y|X} = H_{Y|0}P_X(0) + H_{Y|1}P_X(1) = \frac{1}{4}\log_2\frac{27}{4} = 0.688722$$

(d) The mutual information quantifies how much one part knows about the other. It is symmetric. Here we get:

$$I_{X:Y} = \frac{3}{4}\log_2\frac{4}{3} = \frac{3}{2} - \frac{3}{4}\log_2(3) = 0.311278\,.$$

(Σ = 12P)