Exercise 1.1: Deformed entropy measures (4P)

(a) Prove the the Tsallis entropy $S_q^T = (q-1)^{-1}(1-\sum_i p_i^q)$ (see lecture notes) reduces to the usual Shannon entropy S in the limit $q \to 1$. (1P)

(b) Let A and B two uncorrelated systems, i.e. $p_{i,j}^{AB} = p_i^A p_j^B$ (cf. lecture notes). Show that Tsallis entropy is non-extensive, i.e., (2P)

$$S_q^T(AB) = S_q^T(A) + S_q^T(B) + (1-q)S_q^T(A)S_q^T(B)$$

(c) Show that the Rényi entropy, in contrast to Tsallis entropy, is always extensive on uncorrelated subsystems, i.e. $S_q^R(AB) = S_q^R(A) + S_q^R(B)$ for all $\alpha > 0$. (1P)

Sample Solution

(a) First we rewrite the Tsallis entropy as

$$S_1^T = \lim_{q\to 1} \frac{1-\sum_i p_i^q}{q-1} = \lim_{q\to 1} \frac{1-\sum_i e^{q\ln p_i}}{q-1}$$

Both nominator and denominator vanish in for $q=1$. We therefore apply the l'Hoŷital's rule

$$S_1^T = \frac{-\sum_i (\ln p_i) e^{q\ln p_i}}{1}\Big|_{q=1} = -\sum_i p_i \ln p_i = S\,. \qquad \square$$

Similarly we can prove the Shannon-limit of the Rńyi entropy:

$$S_1^R = \lim_{\alpha\to 1} \frac{\ln \sum_i p_i^\alpha}{1-\alpha} = -\underbrace{\frac{1}{\sum_i p_i^\alpha}}_{\to 1} \sum_j (\ln p_j) p_j^\alpha \Big|_{\alpha=1} = -\sum_j p_j \ln p_j = S.$$

(2P)

(b) The proof of non-extensivity is straight-forward:

$$\begin{aligned}
S_q^T(AB) &= \frac{1-\sum_{ij}(p_{ij}^{AB})^q}{q-1} = \frac{1-\sum_{ij}(p_i^A p_j^B)^q}{q-1} \\
&= \frac{1-\left(\sum_i (p_i^A)^q\right)\left(\sum_j (p_j^B)^q\right)}{q-1} \\
&= \frac{1-\sum_i (p_i^A)^q}{q-1} + \frac{1-\sum_j (p_j^B)^q}{q-1} - \frac{\left(1-\sum_i (p_i^A)^q\right)\left(1-\sum_j (p_j^B)^q\right)}{q-1} \\
&= S_q^T(A) + S_q^T(B) + (1-q)S_q^T(A)S_q^T(B) \qquad \square
\end{aligned}$$

(2P)

(c) The Rényi entropy is indeed extensive:

$$\begin{aligned} S_\alpha^R(AB) &= \frac{1}{1-\alpha}\ln\Big[\sum_{ij}(p_{ij}^{AB})^\alpha\Big] \;=\; \frac{1}{1-\alpha}\ln\Big[\sum_{ij}(p_i^A p_j^B)^\alpha\Big] \\ &= \frac{1}{1-\alpha}\ln\Big[\Big(\sum_i (p_i^A)^\alpha\Big)\Big(\sum_j (p_j^B)^\alpha\Big)\Big] \\ &= \frac{1}{1-\alpha}\ln\Big[\sum_i (p_i^A)^\alpha\Big] + \frac{1}{1-\alpha}\ln\Big[\sum_j (p_j^B)^\alpha\Big] \;=\; S_\alpha^E(A) + S_\alpha^R(B). \qquad \square \end{aligned}$$

Alternative proof: S_α^R is basically a cumulant-generating function. Cumulant generating functions are additive on uncorrelated subsystems. Finished. (1P)

Exercise 1.2: Relative entropy **(2P)**

Let$\{p_1, \ldots, p_N\}$ and $\{q_1, \ldots, q_N\}$ be two discrete probability distributions. The Kullback-Leibler divergence (also called relative entropy or Kullback-Leibler distance) between the two distributions is defined as

$$D(p||q) = \sum_i p_i \ln \frac{p_i}{q_i}\,.$$

Note that the Kullback-Leibler divergence is generally non-symmetric. Show that

(a) $D(p||q) \geq 0$. Hint: Use Jensen's inequality. (1P)

(b) $D(p||q) = 0$ if and only if the two distributions coincide. (1P)

Sample Solution

(a) Jensens inequality states that for every convex function f and non-negative λ_i which add up to 1 we have

$$f\Big(\sum_{i=1}^n \lambda_i x_i\Big) \leq \sum_{i=1}^n \lambda_i f\left(x_i\right).$$

With Jensens inequality it is straight forward to show (1P)

$$D(p||q) = \sum_i p_i \ln\frac{p_i}{q_i} = -\sum_i p_i \ln\frac{q_i}{p_i} \geq -\ln\sum_i p_i\frac{q_i}{p_i} = 0$$

(b) If the two distributions coincide ($p_i = q_i$), then $D(p||q)$ vanishes trivially. Conversially, if $D(p||q) = 0$, we can again use Jensens inequality which is known to be sharp if and only if the x_i are constant or f is linear. Since f is non-linear in our case, we can conclude that $p_i/q_i = const$. Since both distributions are normalized, this implies that the distributions have to be identical. (1P)

($\Sigma = 6$P)