mgr Marek Śmieja
Uniwersytet Jagielloński
Instytut Informatyki

Entropy of the mixture of sources

Suppose that we are given two sources of information S_1, S_2 on measurable space (X, Σ), represented by the probability measures μ_1, μ_2. Let $\mathcal{Q}$ be an "error-control". We assume that we lossy-code information from S_1 with $\mathcal{Q}$-acceptable alphabet $\mathcal{P}_1$ and from S_2 with $\mathcal{Q}$-acceptable alphabet $\mathcal{P}_2$.

Consider a new source S which sends a signal produced by source S_1 with probability a_1 and by source S_2 with probability $a_2 = 1 - a_1$. We provide a simple greedy algorithm which constructs a $\mathcal{Q}$-acceptable coding alphabet $\mathcal{P}$ of S such that the entropy $h(\mathcal{P})$ satisfies:

$$h(S;\mathcal{P}) \leq a_1\, h(S_1;\mathcal{P}_1) + a_2\, h(S_2;\mathcal{P}_2) + 1.$$

In the proof of the above formula the basic role is played by a new equivalent definition of entropy based on measures instead of partitions which we call weighted entropy.

Weighted entropy describes the statistical amount of information needed in the random lossy-coding. Moreover, it provides the computation and interpretation of the entropy with respect to "formal" convex combination $a_1\mathcal{P}_1 + a_2\mathcal{P}_2$, where $\mathcal{P}_1, \mathcal{P}_2$ are partitions (which clearly does not make sense in the classical approach).

As a consequence we obtain an estimation of the entropy and Rényi entropy dimension of the convex combination of measures. In particular if probability measures μ_1, μ_2 have entropy dimension then

$$\dim_E(a_1\mu_1 + a_2\mu_2) = a_1 \dim_E(\mu_1) + a_2 \dim_E(\mu_2).$$