Quantum Information Theory

— Lecture Notes —

Preliminary Version

October 16, 2017

Prof. Dr. Haye Hinrichsen

Lehrstuhl für Theoretische Physik III
Fakultät für Physik und Astronomie
Universität Würzburg, Germany

Winter term 2017/18

Contents

1 Classical Information Theory

1.1 Probability

1.1.1 Probability distribtions and probability densities

1.1.2 Generating functions

1.1.3 Moments

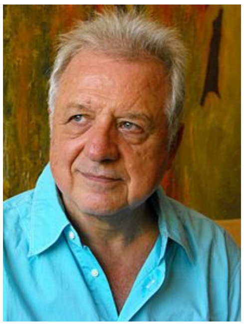
1.1.4 Cumulants

1.2 Information / Entropy

The notion of entropy

1.2.1 Individual and average information

1.2.2 Deformed entropy measures



C. Tsallis [Wikimedia]

Tsallis entropy: In the literature you will find a variety of alternative entropy measures. A well-known example is the so-called *Tsallis entropy* introduced in 1988 by Constantino Tsallis. For a given probability distribution $\{p_i\}$ with $\sum_i p_i = 1$ the Tsallis entropy is defined by

$$S_q^T = \frac{1}{1-q}\Big(1 - \sum_i p_i^q\Big)\,, \tag{1.1}$$

where $q \in \mathbb{R}$ is a free parameter. As can be shown easily, when taking $q \to 1$ the Tsallis entropy reduces to the ordinary Shannon-Boltzmann entropy:

$$\lim_{q\to 1} S_q^T = -\sum_i p_i \ln p_i \tag{1.2}$$

meaning that the standard entropy is included as a special case $S_1^T \equiv S$. Moving away from $q = 1$ the expression S_q^T changes continuously. For this reason the parameter q is often referred to as a *deformation parameter*. Likewise one can define this measure for continuous probability densities by

$$S_q^R = \frac{1}{1-q}\Big(1 - \int [p(x)]^q \, \mathrm{d}x\Big)\,, \tag{1.3}$$

where $p(x)\,dx$ is the probability to find the random variable x in the infinitesimal interval $[x, x + \mathrm{d}x]$.

In the literature the physical relevance of Tsallis entropy has been debated controversially, mainly because there is no immediate information-theoretic interpretation. However, the concept could be applied successfully to a large variety of physical systems and empirical data, ranging from spin glasses to high-energy experiments.

It is important to note that the deformation destroys the extensivity of the measure. As we have seen above, for two *uncorrelated* systems A and B with the probability distributions $p_i^{(A)}$ and $p_j^{(B)}$ the joint probability distribution is given by the product $p_{ij}^{(AB)} = p_i^{(A)} p_j^{(B)}$ so that the standard and P is additive, i.e., $S(AB) = S(A) + S(B)$. In the case of Tsallis entropy, this relationship is replaced by

$$S_q^T(AB) = S_q(A) + S_q(B) + (1-q) S_q(A) S_q(B). \tag{1.4}$$

Therefore, unlike the standard entropy, Tsallis entropy does not work for uncorrelated systems. However, the hope is that it might work for correlated systems, for example, particles interacting by long-range forces such as gravity or electromagnetic interactions.

A. Rényi [alchetron.com]

Rényi entropy: Another important deformed entropy measure is the so-called *Rényi entropy*, named after the Hungarian mathematician Alfred Rényi (1921-1970). For a given probability distribution $\{p_i\}$ the Rényi entropy is defined by

$$S_\alpha^R = \frac{1}{1-\alpha} \ln\Big(\sum_i p_i^\alpha\Big)\,, \tag{1.5}$$

where α is again a deformation parameter. The Rényi entropy includes four important special cases:

α	expression	expression
0	$S_0^R = \ln \lvert\Omega\rvert$	Hartley entropy
1	$S_1^R = -\sum_i p_i \ln p_i$	Shannon / information entropy
2	$S_2^R = -\ln \sum_i p_i^2$	Correlation entropy
∞	$S_\infty^R = -\ln \sup_i p_i$	Minimal individual entropy

In quantum information theory the Rényi entropy is frequently studied while Tsallis entropy is almost unknown. In particular, as we will see below, the Rényi entropy appears in a remarkable generalized "Second Law" of quantum thermodynamics[1]. Moreover,

[1] While the usual Second Law tells us that the ordinary entropy $S = S_1^R$ cannot decrease, the generalized second law states that the Rényi entropy cannot decrease for every value of α.

opposed to Tsallis entropy, the Rényi entropy has a very clear and significant mathematical interpretation. To see this let us first recall that for a given probability distribution $\{p_i\}$ each state or event labeled by i carries an individual entropy

$$S_i = -\ln p_i \tag{1.6}$$

and that the ordinary Shannon entropy is nothing but its expectation value (arithmetic mean)

$$S = \langle S_i \rangle_i = \sum_i p_i S_i = -\sum_i p_i \ln p_i. \tag{1.7}$$

Thus, if we know the Shannon entropy of the system, we actually know very little about the entire probability distribution, namely, only a single number which is the first moment of the individual entropies (the average information). However, it is of course also possible to define higher moments of the information content, i.e.,

$$M_n \;:=\; \sum_i p_i S_i^n \;=\; \sum_i p_i \Big(-\ln p_i\Big)^n. \tag{1.8}$$

These moments characterize the distribution of the individual information content, in particular $M_1 = S$ is just the standard entropy. As outlined above, this allows us to define a corresponding moment-generating function

$$\begin{aligned} M(t) &= \sum_{n=0}^{\infty} \frac{M_n}{n!} t^n \;=\; \sum_{n=0}^{\infty} \frac{1}{n!} \sum_i p_i \Big(-\ln p_i\Big)^n t^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_i p_i \Big((-\ln p_i)t\Big)^n \;=\; \sum_i p_i \sum_{n=0}^{\infty} \frac{1}{n!} \Big((-\ln p_i)t\Big)^n \\ &= \sum_i p_i \exp\Big((-\ln p_i)t\Big) \;=\; \sum_i p_i\, p_i^{-t} \;=\; \langle p_i^{-t} \rangle_i \end{aligned} \tag{1.9}$$

In fact, we can retrieve the standard entropy by computing the first derivative:

$$\frac{\mathrm{d}}{\mathrm{d}t} M(t)\Big|_{t=0} \;=\; \sum_i \frac{\mathrm{d}}{\mathrm{d}t} p_i^{1-t}\Big|_{t=0} \;=\; \sum_i (-p_i^{1-t} \ln p_i)\Big|_{t=0} \;=\; -\sum_i p_i \ln p_i = S. \tag{1.10}$$

As before, we would argue that moments are good but cumulants are better. For this reason let us define cumulant-generating function

$$K(t) = \ln M(t) = \ln \sum_i p_i^{1-t}. \tag{1.11}$$

Comparing the cumulant-generating function with the original expression for the Rńyi entropy given above, we observed that are closely related by

$$S_\alpha^R = \frac{K(1-\alpha)}{1-\alpha}. \tag{1.12}$$

This means that, apart from the renamed argument and an additional factor, the **Rényi entropy is just the cumulant-generating function** of the information content. Thus, knowing the Rényi entropy for all values of the deformation parameter α, we could in principle reconstruct all individual entropies S_i and therewith the whole probability distribution.

Index