Journal of Statistical Physics, Vol. 27, No. 2, 1982

Instabilities of Rényi Entropies

Bernhard Lesche[1]

Received March 23, 1981

We show that for systems with a large number of microstates Rényi entropies do not represent experimentally observable quantities except the Rényi entropy that coincides with the Shannon entropy.

KEY WORDS: Entropies; mixing character; convex functions.

Rényi entropies[1] are mixing homomorphic functions[2] which are additive with respect to the composition of statistically independent systems. They are characterized by a real parameter $\alpha > 0$. Their definition reads as follows:

$$I_\alpha(p) = \frac{1}{1-\alpha} \ln \sum_{i=1}^{n} (p_i)^\alpha \qquad \text{if } \alpha \neq 1$$

and (1)

$$I_1(p) = \lim_{\alpha \to 1} I_\alpha(p) = -\sum_{i=1}^{n} p_i \ln p_i$$

where p_i is the probability of the microstate i according to the probability assignment p defined on the set of n microstates. The Rényi entropy with $\alpha = 1$ coincides with the Shannon entropy. For probability assignments that take the constant value $1/m$ on a subset of m elements all Rényi entropies have the same value

$$I_\alpha(p) = \ln m$$

which coincides with the Boltzmann entropy (up to the Boltzmann factor),

Work supported by the DFG (1978); author is recipient of a Feodor-Lynen grant from the Alexander von Humboldt Stiftung.

[1] Instituto de Física, U.N.A.M., Apdo. Postal 20-364, 01000 México, D.F.

0022-4715/82/0200-0419$03.00/0

if one considers the subset of m microstates to represent a macrostate. In the special case $m = n$ one gets the maximal entropy of the space of n states:

$$I_{\max} = \ln n \tag{2}$$

In general, however, Rényi entropies differ from each other and the question arises which of these functions might possibly be related to some experimentally observable quantity.

A necessary condition for a quantity G to be observable is that their values $G(x)$ do not change dramatically if the state x of the system in consideration is changed by an unobservably small amount δx. The states x of our problem are the probability assignments p on the set of n microstates. Using mathematical statistics one can show[3] that the experimental effort necessary to distinguish between two probability assignments p and p' with an appreciable reliability is related to the l_1 distance

$$\|p - p'\|_1 = \sum_{i=1}^{n} |p_i - p_i'|$$

in a way that is independent of n. Therefore we ask the following question:

For which values of α can one find for every $\epsilon > 0$ a $\delta_\epsilon > 0$ such that for all n and for all p, p' one has

$$\|p' - p\|_1 \leqslant \delta_\epsilon \Rightarrow \frac{|I_\alpha(p') - I_\alpha(p)|}{I_{\max}} < \epsilon$$

We will show that this is possible for $\alpha = 1$ and impossible for any other value of α.

1. The Case $\alpha = 1$. Consider the functions

$$A(S, p) = \sum_{i=1}^{n} (p_i - e^{-s})^+$$

where $x^+ = \max\{x, 0\}$. We remark in passing that mixing character[2] can be defined with the aid of these functions:

$$m[p'] \succ m[p] \Leftrightarrow \forall S \geqslant 0: \qquad A(S, p') \leqslant A(S, p)$$

These functions have the following properties:

$$|A(S, p) - A(S, p')| \leqslant \|p - p'\|_1 \qquad \text{for all } S \geqslant 0 \tag{3}$$

$$(1 - \exp[-S + \ln n])^+ \leqslant A(S, p) < 1 \tag{4}$$

$$I_1(p) = -1 + \int_0^\infty [1 - A(S, p)]\, dS \tag{5}$$

Equation (5) yields

$$|I_1(p) - I_1(p')| = \left| \int_0^\infty [A(S,p) - A(S,p')]\, dS \right|$$
$$\leqslant \int_0^{a+\ln n} |A(S,p) - A(S,p')|\, dS$$
$$+ \int_{a+\ln n}^\infty |A(S,p) - A(S,p')|\, dS$$

where $a \geqslant -\ln n$ is arbitrary. Supposing $a \geqslant 0$ we apply inequality (3) to the first integral and inequality (4) to the second one:

$$|I_1(p) - I_1(p')| \leqslant \|p - p'\|_1 (a + \ln n) + e^{-a}$$

Now we choose $a \geqslant 0$ so that the right-hand side becomes a minimum. For $\|p - p'\|_1 < 1$ the minimum shows up at $a = -\ln \|p - p'\|$. So, for $\|p - p'\|_1 < 1$ one has

$$|I_1(p) - I_1(p')| \leqslant \|p - p'\|_1 (1 + \ln n) - \|p - p'\|_1 \ln \|p - p'\|_1$$

$f(x) = -x \ln x$ is an increasing nonnegative function in the interval $[0, 1/e]$, so one has

$$|I_1(p) - I_1(p')| \leqslant \delta(1 + \ln n) - \delta \ln \delta$$

for $\|p - p'\|_1 < \delta \leqslant 1/e$. Equation (2) gives with $n \geqslant 2$

$$\frac{|I_1(p) - I_1(p')|}{I_{\max}} \leqslant \delta\left(\frac{1}{\ln n} + 1\right) - \frac{\delta \ln \delta}{\ln n} \leqslant \delta\left(\frac{1}{\ln 2} + 1\right) - \frac{\delta \ln \delta}{\ln 2}$$

if $\|p - p'\| < \delta \leqslant 1/e$. Thus, it is clear that one can find an appropriate δ_ϵ for every ϵ, because the right-hand side is a continuous function of δ approaching 0 for $\delta \to 0$.

2. The Case $\alpha > 1$. Let p and p' be defined as

$$p_i = \frac{1}{(n-1)}(1 - \delta_{1i})$$

$$p_i' = \frac{\delta}{2}\delta_{1i} + \left(1 - \frac{\delta}{2}\right)\frac{1}{(n-1)}(1 - \delta_{1i})$$

One has $\|p - p'\|_1 = \delta$ and

$$I_\alpha(p) - I_\alpha(p') = \frac{1}{1-\alpha} \ln \frac{\sum_i p_i^\alpha}{\sum_j p_j'^\alpha}$$
$$= \frac{1}{1-\alpha} \ln \left[\frac{(n-1)^{1-\alpha}}{(\delta/2)^\alpha + (n-1)^{1-\alpha}(1-\delta/2)^\alpha} \right]$$

For large n the asymptotic behavior of this difference is

$$I_\alpha(p) - I_\alpha(p') \curvearrowright \frac{1}{1-\alpha} \ln \frac{(n-1)^{1-\alpha}}{(\delta/2)^\alpha}$$

and thus

$$\lim_{n\to\infty} \frac{|\Delta I_\alpha|}{I_{\max}} = 1$$

no matter how small δ might be.

3. The Case $\alpha < 1$. We choose

$$p_i = \delta_{1i}, \qquad p_i' = \left(1 - \frac{\delta}{2}\right)\delta_{1i} + \frac{1}{n-1}\,\frac{\delta}{2}(1 - \delta_{1i})$$

which gives $\|p_i - p_i'\| = \delta$ and

$$I_\alpha(p') - I_\alpha(p) = I_\alpha(p') = \frac{1}{1-\alpha} \ln\left[\left(1 - \frac{\delta}{2}\right)^\alpha + (n-1)^{1-\alpha}\left(\frac{\delta}{2}\right)^\alpha\right]$$

The asymptotic behavior of this difference for large n is

$$\Delta I_\alpha \curvearrowright \frac{1}{1-\alpha} \ln\left[(n-1)^{1-\alpha}\left(\frac{\delta}{2}\right)^\alpha\right]$$

and thus

$$\lim_{n\to\infty} \frac{|\Delta I_\alpha|}{I_{\max}} = 1$$

no matter how small δ might be. The first counterexample illustrates that Rényi entropies with $\alpha > 1$ overestimate a high peak of probability. Therefore it can occur that the whole rest is completely ignored despite the fact that its overall probability is practically 1 and that it contains all relevant information. The second counterexample illustrates that Rényi entropies with $\alpha < 1$ overestimate a large number of occupied states even if their overall probability is so small that they are of no physical relevance.

ACKNOWLEDGMENTS

This paper is dedicated to Ernst Ruch on the occasion of his 60th birthday.

REFERENCES

1. Rényi, A., *Wahrscheinlichkeitstheorie* (De Gruyter, Berlin, 1974).
2. Ruch, E., *Theor. Chim. Acta* **38**:167 (1975).
3. Lesche, B., Hat das Prinzip vom zunehmenden Mischungscharakter Konsequenzen im Bereich makroskopischer Phänomene, die über das Entropieprinzip messbar hinausgehen? Dissertation, TU-Berlin, Berlin (1978) D83.