Rescuing the MaxEnt treatment for q-generalized entropies

A. Plastino[1,2,4], M.C.Rocca[1,2,3],

[1] Departamento de Física, Universidad Nacional de La Plata,

[2] Departamento de Matemática, Universidad Nacional de La Plata,

[3] Consejo Nacional de Investigaciones Científicas y Tecnológicas

(IFLP-CCT-CONICET)-C. C. 727, 1900 La Plata - Argentina

[4] SThAR - EPFL, Lausanne, Switzerland

July 5, 2018

Abstract

It has been recently argued, via very clever arguments, that the MaxEnt variational problem would not adequately work for Renyi's and Tsallis' entropies. We constructively show here that this is not so if one formulates the associated variational problem in a more orthodox functional fashion.

Keywords: Tsallis entropy, Renyi entropy, MaxEnt

1 Introduction

Renyi's and Tsallis' information measures (S_R and S_T), respectively) are generalizations of the Shannon one, quantifying a system's diversity, uncertainty, or randomness. The degree of generalization is parameterized by a real parameter. Both are very important quantities in variegated areas of scientific research. One can mention, for instance, ecology, quantum information, the Heisenberg XY spin chain model, theoretical computer science, diffusion processes, several biological processes, high energy physics, etc. As a small sample, see for example, [1, 2, 3, 4, 5, 6, 7, 8, 9, 10], [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22].

Information Theory (IT) yields a quite powerful inference methodology, abbreviated as MaxEnt [23], that is able to describe general properties of arbitrary systems, in most areas the of scientific endeavor, on the basis of scarce information. MaxEnt yields the least-biased description that can be gotten according to specific data-amounts, in any possible circumstances [23]. In statistical mechanics (SM), Jaynes pioneered the use of MaxEnt so as to both i) reformulate SM and ii) generalize its foundations [23]. Here we focus on the MaxEnt technique as applied to Renyi-Tsallis information measures.

A very interesting result was recently reported in [24]. Let a and β stand for the two Lagrange multipliers appearing in a canonical ensemble MaxEnt treatment. Then, one confidently expects MaxEnt to yield, for the canonical probability distribution (PD) P (Z being the partition function, f an appropriate function, and ϵ the level energy),

$$P_i = f(a + \beta \epsilon_i)/Z, \tag{1.1}$$

$$Z = \sum_i f(a + \beta \epsilon_i). \tag{1.2}$$

It is claimed in [24], by recourse to an ingenious procedure, that for this to be possible f must be endowed with certain properties. These properties, it is shown in [24], are NOT possessed by the PDs that maximize either Tsallis' or Renyi's entropies. This constitutes an intriguing result that deserves further consideration.

We show here by construction, appealing to conventional techniques of Functional Analysis, not used in [24], that the *functional forms* (1.1), (1.2) are

indeed recovered à la Tsallis (or Renyi), thus surmounting the formidable obstacle posed by [24] to the use of MaxEnt for these two entropies. We are not claiming that the arguments of [24] are wrong, but that they are Calculus-based, while the proper MaxEnt scenario requires functional Analysis. The central notion required here is that of increment h of a functional. The general theory of Variational Calculus is concerned with Banach Space (BS) [25], examples of which are Hilbert's space and classical phase space. The MaxEnt technique in Banach space demands a first variation that has to vanish and a second one that ascertains the characteristics of the ensuing extremum. The main need is then to evaluate the increment h of a functional F at the point x of the operative BS. *One must remember that functional calculus is not identical to ordinary calculus (involving ordinary functions), particularly when one is looking for extremes [26, 25].*

2   Tsallis MaxEnt Treatment

The concomitant Tsallis' procedure has been developed in [26]. The Tsallis information functional is [11]

$$F_S(P) = -\sum_{i=1}^{n} P_i^q \ln_q(P_i) + \lambda_1 \left(\sum_{i=1}^{n} P_i U_i - <U> \right) + \lambda_2 \left(\sum_{i=1}^{n} P_i - 1 \right) \quad (2.1)$$

For the increment one has

$$F_S(P+h) - F_S(P) = -\sum_{i=1}^{n} (P_i + h_i)^q \ln_q(P_i + h_i) + \lambda_1 \left[\sum_{i=1}^{n} (P_i + h_i) U_i - <U> \right] +$$

$$\lambda_2 \left[\sum_{i=1}^{n} (P_i + h_i) - 1 \right] + \sum_{i=1}^{n} P_i^q \ln_q(P_i) - \lambda_1 \left(\sum_{i=1}^{n} P_i U_i - <U> \right) -$$

$$\lambda_2 \left(\sum_{i=1}^{n} P_i - 1 \right) \quad (2.2)$$

Eq. (2.2) can be rewritten as

$$F_S(P + h) - F_S(P) = \sum_{i=1}^{n} \left[\left(\frac{q}{1-q} \right) P_i^{q-1} + \lambda_1 U_i + \lambda_2 \right] h_i$$

3

$$\sum_{i=1}^{n} qP_i^{q-2}\frac{h_i^2}{2} + O(h^3), \tag{2.3}$$

Eq. (2.3) leads to

$$\left(\frac{q}{1-q}\right)P_i^{q-1} + \lambda_1 U_i + \lambda_2 = 0, \tag{2.4}$$

$$-\sum_{i=1}^{n} qP_i^{q-2}h_i^2 \le C\|h\|^2 \tag{2.5}$$

Eq. (2.4) is the Euler-Lagrange one while (2.5) gives bounds originating from the second variation. Thus, (2.4) entails

$$\lambda_1 = \beta q Z^{1-q} \tag{2.6}$$

$$\lambda_2 = \frac{q}{q-1}Z^{1-q} \tag{2.7}$$

The following two basic relations for our task follow now, namely,

$$P_i = \frac{[1 + \beta(1-q)U_i]^{\frac{1}{q-1}}}{Z} \tag{2.8}$$

$$Z = \sum_{i=1}^{n} [1 + \beta(1-q)U_i]^{\frac{1}{q-1}}. \tag{2.9}$$

These two relations should not exist according to [24]. That we can construct them follows from Functional Analysis, that was not appealed to in that reference.

3 Renyi's MaxEnt Treatment

Renyi's S_R is defined as [9]:

$$S_R = \frac{1}{1-\alpha}\ln\left(\sum_{i=1}^{n} P_i^\alpha\right), \tag{3.1}$$

and the accompanying (canonical ensemble) MaxEnt probability distribution P arises from the maximization of the functional $F_{S_R}(P)$ [where U denotes the energy and $<U>$ its mean value]

$$F_{S_R}(P) = \frac{1}{1-\alpha}\ln\left(\sum_{i=1}^{n} P_i^{\alpha}\right) + \lambda_1\left(\sum_{i=1}^{n} P_i U_i - <U>\right) + \lambda_2\left(\sum_{i=1}^{n} P_i - 1\right),$$
(3.2)

Following standard *Functional Analysis* procedures (which are not employed in Ref- [24], we consider the functional-h increment [26, 25]

$$F_{S_R}(P+h) = \frac{1}{1-\alpha}\ln\left[\sum_{i=1}^{n}(P_i+h_i)^{\alpha}\right] + \lambda_1\left[\sum_{i=1}^{n}(P_i+h_i)U_i - <U>\right] +$$

$$\lambda_2\left[\sum_{i=1}^{n}(P_i+h_i) - 1\right],$$
(3.3)

so that

$$F_{S_R}(P+h) - F_{S_R}(P) = \frac{1}{1-\alpha}\ln\left[\sum_{i=1}^{n}(P_i+h_i)^{\alpha}\right] - \frac{1}{1-\alpha}\ln\left(\sum_{i=1}^{n} P_i^{\alpha}\right) +$$

$$\lambda_1\sum_{i=1}^{n} h_i U_i + \lambda_2\sum_{i=1}^{n} h_i.$$
(3.4)

We now tackle h^2 contributions so as to assess second variations of F_{S_R} [26, 25]

$$F_{S_R}(P+h) - F_{S_R}(P) = \frac{1}{1-\alpha}\ln\left\{\sum_{i=1}^{n}\left[P_i^{\alpha} + \alpha h_i P_i^{\alpha-1} + \frac{\alpha(\alpha-1)}{2}h_i^2 P_i^{\alpha-2}\right]\right\} -$$

$$\frac{1}{1-\alpha}\ln\left(\sum_{i=1}^{n} P_i^{\alpha}\right) + \lambda_1\sum_{i=1}^{n} h_i U_i + \lambda_2\sum_{i=1}^{n} h_i.$$
(3.5)

or, equivalently,

$$F_{S_R}(P+h) - F_{S_R}(P) = \frac{1}{1-\alpha}\ln\left\{1 + \frac{\sum_{i=1}^{n}\left[\alpha h_i P_i^{\alpha-1} + \frac{\alpha(\alpha-1)}{2}h_i^2 P_i^{\alpha-2}\right]}{\sum_{i=1}^{n} P_i^{\alpha}}\right\} +$$

5

$$\lambda_1 \sum_{i=1}^{n} h_i U_i + \lambda_2 \sum_{i=1}^{n} h_i, \tag{3.6}$$

so that one finally arrives at

$$F_{S_R}(P + h) - F_{S_R}(P) = \frac{1}{1 - \alpha} \frac{\sum_{i=1}^{n} \left[\alpha h_i P_i^{\alpha-1} + \frac{\alpha(\alpha-1)}{2} h_i^2 P_i^{\alpha-2} \right]}{\sum_{i=1}^{n} P_i^{\alpha}} +$$

$$- \frac{1}{2(1 - \alpha)} \left(\frac{\sum_{i=1}^{n} \alpha h_i P_i^{\alpha-1}}{\sum_{i=1}^{n} P_i^{\alpha}} \right)^2 + \lambda_1 \sum_{i=1}^{n} h_i U_i + \lambda_2 \sum_{i=1}^{n} h_i. \tag{3.7}$$

Summing up, we have for the first variation

$$\frac{\alpha}{1 - \alpha} \frac{P_i^{\alpha-1}}{\sum_{i=1}^{n} P_i^{\alpha}} + \lambda_1 U_i + \lambda_2 = 0. \tag{3.8}$$

Multiplying (3.8) by P_i we find

$$\frac{\alpha}{1 - \alpha} \frac{P_i^{\alpha}}{\sum_{i=1}^{n} P_i^{\alpha}} + \lambda_1 P_i U_i + \lambda_2 P_i = 0. \tag{3.9}$$

Integrating now we are led to

$$\frac{\alpha}{1 - \alpha} + \lambda_1 < U > + \lambda_2 = 0. \tag{3.10}$$

This is an important result, showing that λ_1 and λ_2 are not independent Lagrange multipliers. We are authorized to write

$$\lambda_2 = \frac{\alpha}{\alpha - 1} - \lambda_1 < U >, \tag{3.11}$$

and replacing this value of λ_2 in (3.8) we get

$$\frac{\alpha}{1 - \alpha} \frac{P_i^{\alpha-1}}{\sum_{i=1}^{n} P_i^{\alpha}} + \lambda_1 (U_i - < U >) + \frac{\alpha}{\alpha - 1} = 0, \tag{3.12}$$

6

whose solution is given by

$$\lambda_1 = -\beta\alpha \tag{3.13}$$

$$P_i = \frac{[1 + \beta(1-\alpha)(U_i - <U>)]^{\frac{1}{\alpha-1}}}{Z} \tag{3.14}$$

$$Z = \sum_{i=1}^{n} [1 + \beta(1-\alpha)(U_i - <U>)]^{\frac{1}{\alpha-1}}. \tag{3.15}$$

Ref. [24] showed, via an alternative, calculus-based procedure, that Eqs. (3.14) and (3.15) are impossible to obtain in this form for Renyi's entropy. We are able here to contradict such assertion because we employed Functional Analysis and not just ordinary calculus. Finally, note that Eqs. (8) and (11) of [24] *define* both the exponential and the logarithm functions. Thus, their co-implication is obvious.

4   Conclusions

The MaxEnt variational treatment in the canonical ensemble is believed to yield, for the probability distribution (PD) P that maximizes the entropy the forms

$$P_i = f(\alpha + \beta\epsilon_i)/Z, \tag{4.1}$$

$$Z = \sum_i f(\alpha + \beta\epsilon_i), \tag{4.2}$$

where α and β are Lagrange multipliers, Z is the partition function, f is an appropriate function, and ϵ is a level-energy.

It is claimed in [24], by recourse to an admirable, Calculus-based procedure, that for this to be possible f must be endowed with special properties. These properties, one finds in [24], are NOT possessed by the PDs that maximize either Tsallis' or Renyi's entropies.

We have shown here (by construction), appealing to conventional techniques of Functional Analysis, not used in [24], that the *functional forms* (1.1), (1.2) are indeed encountered in dealing with Tsallis' (or Renyi's) entropies. This

removes the rather formidable obstacle posed by [24] to the use of MaxEnt for these two entropies. Note that we are not claiming that the arguments of [24] are wrong, but that they are Calculus-based, while the proper MaxEnt scenario requires functional Analysis.

References

[1] C. M. Herdman, Stephen Inglis, P.-N. Roy, R. G. Melko, and A. Del Maestro, Phys. Rev. E **90** (2014) 013308.

[2] Mohammad H. Ansari and Yuli V. Nazarov, Phys. Rev. B **91** (2015) 174307.

[3] Lei Wang and Matthias Troyer, Phys. Rev. Lett. **113** (2014) 110401.

[4] Matthew B. Hastings, Ivn Gonzlez, Ann B. Kallin, and Roger G. Melko, Phys. Rev. Lett **104** (2010) 157201.

[5] Richard Berkovits, Phys. Rev. Lett. **115** (2015) 206401.

[6] Nima Lashkari, Phys. Rev. Lett. **113** (2014) 051602 .

[7] Gabor B. Halasz and Alioscia Hamma, Phys. Rev. Lett. **110** (2013) 170605 .

[8] MB Hastings, I Gonzlez, AB Kallin, RG Melko, Phys. Rev. Lett. **104**, 157201 (2010); A. De Gregorio, S.M. lacus, **179** (2009) 279 .

[9] Leila Golshani, Einollah Pasha, Gholamhossein Yari, Information Sciences, **179** (2009) 2426; J.F. Bercher, Information Sciences **178** (2008) 2489.

[10] EK Lenzi, RS Mendes, LR da Silva, Physica A **280** (2000) 337.

[11] M. Gell-Mann and C. Tsallis, Eds. *Nonextensive Entropy: Interdisciplinary applications*, Oxford University Press, Oxford, 2004; C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*, Springer, New York, 2009.

[12] See http://tsallis.cat.cbpf.br/biblio.htm for a regularly updated bibliography on the Tsallis-subject.

[13] P-Jizba, J. Korbel, V. Zatloukal, Phys. Rev. E **95** (2017) 022103.

[14] I. S. Oliveira: Eur. Phys. J. B **14** (2000) 43.

[15] E. K. Lenzi , R. S. Mendes: Eur. Phys. J. B **21** (2001) 401.

[16] C. Tsallis: Eur. Phys. J. A **40** (2009) 257.

[17] P. H. Chavanis: Eur. Phys. J. B **53** (2003) 487.

[18] G. Ruiz , C. Tsallis: Eur. Phys. J. B **67** (2009) 577.

[19] P. H. Chavanis , A. Campa: Eur. Phys. J. B **76** (2010) 581.

[20] N. Kalogeropoulos: Eur. Phys. J. B **87** (2014) 56.

[21] N. Kalogeropoulos: Eur. Phys. J. B **87** (2014) 138.

[22] A. Kononovicius, J. Ruseckas: Eur. Phys. J. B **87** (2014) 169.

[23] ET Jaynes, Phys. Rev. **106** (1957) 620; *Papers on probability, statistics and statistical physics*, edited by R. D. Rosenkrantz, Reidel, Dordrecht, Holland, 1983; L. Brillouin, *Science and Information Theory*, Academic Press, New York (1956); WT Grandy, Jr., and PW Milonni, *Physics and probability: Essays in honor of E. T. Jaynes*, Cambridge University Press, Cambridge, England, 1993.

[24] T. Oikonomoua, G. B. Bagci, Phys. Lett. A **381** (2017) 207.

[25] G.Y. Shilov, *Mathematical Analysis*, Pergamon Press, NY, 1965.

[26] A. Plastino, M. C. Rocca, Physica A **436**, 572 (2015).