About Closedness by Convolution of the Tsallis Maximizers

C. Vignat*, A. O. Hero III and J. A. Costa

E.E.C.S., University of Michigan, Ann Arbor, USA

(*) L.T.C.I., Université de Marne la Vallée, France

E-mail: `cvignat@umich.edu`

Abstract

In this paper, we study the stability under convolution of the maximizing distributions of the Tsallis entropy under energy constraint (called hereafter Tsallis distributions). These distributions are shown to obey three important properties: a stochastic representation property, an orthogonal invariance property and a duality property. As a consequence of these properties, the behaviour of Tsallis distributions under convolution is characterized. At last, a special random convolution, called Kingman convolution, is shown to ensure the stability of Tsallis distributions.

Key words: Tsallis entropy, convolution

Introduction

It is well-known that the Boltzmann distributions are the maximizers of the Shannon-Boltzmann entropy under energy (covariance) constraint. These Boltzmann distributions enjoy the stability property for addition: if $\mathbf{X}$ and $\mathbf{Y}$ are independent vectors, each distributed according to a Boltzmann law, then their sum $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ is again of the Boltzmann type. This property is important in physics since it is involved in some very general results like the central limit theorem [6].

The Tsallis entropy was introduced by Tsallis [10], and proved as a very efficient tool to describe the behavior of complex systems like the interior solar plasma [1] or self-gravitating systems [2]. A particular case of the Tsallis entropy family is the Shannon-Boltzmann entropy.

The maximizers of the Tsallis entropy under covariance constraint - called Tsallis distributions in this paper - were studied in detail by several authors, in the scalar case in [3], in the multivariate case in [4] and more recently in [5]. A natural question is thus to explore the stability under convolution of these Tsallis distributions. This is the problem addressed in this paper.

1 Tsallis distributions

The order-q Tsallis entropy $H_q(f)$ of a continuous probability density f writes

$$H_q(f) = \frac{1}{q-1}\left(1 - \int_\Omega f^q\right).$$

It can be easily checked that $\lim_{q\to 1} H_q(f)$ exists and coincides with the celebrated Boltzmann-Shannon entropy

$$H_1(f) = -\int f \log f.$$

In this paper, $\mathbf{x}$ denotes an $n-$dimensional real-valued random vector with covariance matrix $K{=}\mathbf{E}(\mathbf{x}-\boldsymbol{\mu}_X)(\mathbf{x}-\boldsymbol{\mu}_X)^T$. Without loss of generality, we consider only the centered case $\boldsymbol{\mu}_X = 0$. We define next the following $n-$variate probability density f_q as follows:

$$\text{if } \frac{n}{n+2} < q, \quad f_q(\mathbf{x}) = A_q\left(1-(q-1)\beta\mathbf{x}^T K^{-1}\mathbf{x}\right)_+^{\frac{1}{q-1}} \quad \forall \mathbf{x}\in\mathbf{R}^n \tag{1}$$

with $x_+ = \max(0,x)$, $\beta = \frac{1}{2q-n(1-q)}$. The problem of maximization of the Tsallis entropy under energy constraint can be easily solved using a Bregman information divergence, as described in the following theorem.

Theorem 1 *f_q defined by (1) is the only probability density that verifies*

$$f_q = \arg\max_{f\,:\,E\mathbf{x}\mathbf{x}^T=\mathbf{K}} H_q(f).$$

PROOF. *Consider the following non-symmetric Bregman divergence*

$$D_q(f||g) = sign(q-1)\int \frac{f^q}{q} + \frac{q-1}{q}g^q - fg^{q-1}. \tag{2}$$

The positivity of $D_q(f||g)$ (with nullity if and only if $f=g$ pointwise) is a consequence of the convexity of function $x \mapsto sign(q-1)\,x^q/q$. Suppose for example $q>1$: the fact that distribution f has the same covariance K as f_q defined by (1) can be expressed by

$$\int f_q^q = \int f_q^{q-1} f$$

so that

$$0 \leq D_q\left(f||f_q\right) = \int \frac{f^q}{q} + \frac{q-1}{q} f_q^q - f_q^q = \frac{1}{q}\int f^q - f_q^q = \frac{q-1}{q}\left(H_q\left(f_q\right) - H_q\left(f\right)\right).$$

The proof in the case $q < 1$ follows accordingly.

The Tsallis distributions verify three important properties:

- the stochastic representation property. If $\mathbf{X}$ is Tsallis distributed with parameter $q < 1$ and covariance matrix K then

$$\mathbf{X} \stackrel{d}{=} \frac{C\mathbf{N}}{A} \tag{3}$$

  where A is a chi random variable with $m = -n + \frac{2}{1-q}$ degrees of freedom, independent of the Gaussian vector $\mathbf{N}$ ($E\mathbf{N}\mathbf{N}^T = I$) and with $C = (m-2)\,K$. If $\mathbf{Y}$ is Tsallis distributed with parameter $q > 1$, then

$$\mathbf{Y} \stackrel{d}{=} \frac{C\mathbf{N}}{\sqrt{A^2 + \|\mathbf{N}\|_2^2}} \tag{4}$$

  where A is a chi random variable with $\frac{2}{q-1} + 2$ degrees of freedom. Note that the denominator is again a chi random variable, but that contrary to the case $q < 1$, it is now dependent on the numerator.
- the orthogonal invariance property. The Tsallis distributions write as

$$f_q\left(\mathbf{x}\right) = \phi_q\left(\mathbf{x}^T K^{-1}\mathbf{x}\right).$$

  This property is a direct consequence of the invariance under orthogonal transformation of the covariance constraint.
- the duality property. There is a natural bijection between the cases $q < 1$ and $q > 1$: if $\mathbf{X}$ is Tsallis with parameter $q < 1$, $m = -n + \frac{2}{1-q}$ and $C = (m-2)\,K$ then

$$\mathbf{Y} = \frac{\mathbf{X}}{\sqrt{1 - \mathbf{X}^T C^{-1}\mathbf{X}}}$$

  is Tsallis with covariance matrix $\frac{m-2}{m+2}K$ and parameter $q' > 1$ such that

$$\frac{1}{q'-1} = \frac{1}{1-q} - \frac{n}{2} - 1.$$

2 The stability issue

The stability problem can be solved using the properties described above, and the main theorem is the following.

Theorem 2 *(1) if* $\mathbf{X}$ *is a Tsallis* $n-$*vector with parameter* q *and if* H *is a full-rank* $\tilde{n} \times n$ *matrix with* $\tilde{n} \leq n$ *then* $\tilde{\mathbf{X}} = H\mathbf{X}$ *is a Tsallis* $\tilde{n}-$*vector with parameter* $\tilde{q}$ *such that*

$$\frac{2}{1-\tilde{q}} - \tilde{n} = \frac{2}{1-q} - n$$

(2) as a particular case, if $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are mutually Tsallis distributed (as components of a Tsallis vector* $\mathbf{X}^T = \left[\mathbf{X}_1^T, \mathbf{X}_2^T\right]$*)*[1]*, then any linear combination*

$$\mathbf{Y} = H_1\mathbf{X}_1 + H_2\mathbf{X}_2$$

where H_1 *and* H_2 *are full-rank matrices, is a Tsallis vector*

(3) if $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are both Tsallis but* ***independent*** *(and then* $\mathbf{X}^T = \left[\mathbf{X}_1^T, \mathbf{X}_2^T\right]$ *is not Tsallis), then a linear combination* $Y = H_1\mathbf{X}_1 + H_2\mathbf{X}_2$ *is not Tsallis*

(4) if X_1 *and* X_2 *are scalar, each with parameter* $q < 1$*, then there exists a convolution of the random type (called Kingman convolution)*

$$Y = X_1 \oplus X_2$$

such that Y *is Tsallis* ***with the same parameter*** q *as* X *and* Y*.*

The three first results, that hold in both cases $q < 1$ and $q > 1$, are a direct consequence of the orthogonal invariance property, and their detailed proofs can be found in [5]. The last statement is developed and proved in the next section. We stress the point that, contrary to the Boltzmann ($q = 1$) case, stability under addition holds only if $\mathbf{X}_1$ and $\mathbf{X}_2$ are dependent, in the sense that they share the same mixing random variable A as introduced in (3) and (4). This result sheds a new light on the relationship between independence and stability in the set of Tsallis distributions.

3 Kingman convolution (case $q < 1$)

The results about Kingman convolution are described in the following theorem.

Theorem 3 *If* X *and* Y *are independent and scalar Tsallis random variables with parameter* $\frac{1}{2} < q \leq 1$ *and if* λ *is a Tsallis random variable independent of* X *and* Y *and with parameter* $q' = \frac{q}{2q-1} \geq 1$*, then the random variable*

$$Z = \frac{XY}{\sqrt{X^2 + Y^2 + 2\lambda XY}} \tag{5}$$

[1] note that this assumption implies that $\mathbf{X}_1$ and $\mathbf{X}_2$ are dependent

is Tsallis with parameter q and variance σ_Z^2 such that

$$\sigma_Z^{-1} = \sigma_X^{-1} + \sigma_Y^{-1}.$$

PROOF. *A classical result about Hankel transform writes (see [9, formula 6.565.4])*

$$\int_{-\infty}^{+\infty} \Omega_m \left(u \left|x\right|\right) f_q \left(x\right) dx = e^{-\frac{1}{\sigma_X \sqrt{m-2u}}} \quad \forall u > 0 \tag{6}$$

with $m = \frac{2}{1-q} - 1$, where the Bessel convolution kernel Ω_m writes

$$\Omega_m \left(u\right) = j_{\frac{m}{2}-1} \left(\frac{1}{u}\right) \quad u > 0 \tag{7}$$

and where j_ν denotes the normalized Bessel function of the first kind

$$j_\nu \left(x\right) = 2^\nu \Gamma \left(\nu + 1\right) x^{-\nu} J_\nu \left(x\right).$$

Considering a scalar Tsallis random variable X with parameter $q < 1$, equality (6) writes equivalently

$$E_{X_m} \Omega_m \left(u \left|X\right|\right) = e^{-\frac{1}{\sigma_X \sqrt{m-2u}}} \quad \forall u > 0.$$

Now Sonine-Gegenbauer's integral writes ([7] cited by [8]):

$$\Omega_m \left(x\right) \Omega_m \left(y\right) = \frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)} \int_{-1}^{+1} \Omega_m \left(\frac{1}{\sqrt{x^{-2} + y^{-2} + 2\lambda x^{-1} y^{-1}}}\right) \left(1 - \lambda^2\right)^{\frac{m-3}{2}} d\lambda$$

and we remark that the function $\frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)} \left(1 - \lambda^2\right)^{\frac{m-3}{2}}$ is the distribution of a Tsallis random variable λ with parameter q_λ such as

$$q_\lambda = \frac{m-1}{m-3} > 1.$$

Now choose X and Y independent and Tsallis with parameter $q < 1$ and respective variances σ_X^2 and σ_Y^2 so that

$$\Omega_m \left(u \left|X\right|\right) \Omega_m \left(u \left|Y\right|\right) = E_\lambda \Omega_m \left(\frac{\left|XY\right|}{\sqrt{X^2 + Y^2 + 2\lambda XY}} u\right), \quad \forall u > 0.$$

But as

$$E_X \Omega_m \left(u \left|X\right|\right) E_Y \Omega_m \left(u \left|Y\right|\right) = e^{-\frac{1}{u\sqrt{(m-2)}}\left(\frac{1}{\sigma_X} + \frac{1}{\sigma_Y}\right)},$$

we deduce, by defining

$$Z = \frac{XY}{\sqrt{X^2 + Y^2 + 2\lambda XY}}$$

that

$$E_{Z,\lambda}\Omega_m\left(u\left|Z\right|\right) = e^{-\frac{1}{u\sqrt{(m-2)}}\left(\frac{1}{\sigma_X}+\frac{1}{\sigma_Y}\right)}. \tag{8}$$

From equality (8), we conclude, using [8, lemma 4], that Z is Tsallis with parameter q and variance σ_Z^2 such that

$$\sigma_Z^{-1} = \sigma_X^{-1} + \sigma_Y^{-1}.$$

We note that the Kingman convolution expressed by (5) is associative and commutative (see [8]); moreover, it is a convolution of random type, since it combines both random variables X and Y with a third one, λ, that should be chosen independent of X and Y, and acts as a "mixing variable". The physical interpretation of these results is currently under study.

Acknowledgements The first author would like to thank Pr. P.H. Chavanis and Pr. J. Naudts for fruitful discussions during Next 2003 Conference.

References

[1] G. Kaniadakis, A. Lavagno and P. Quarati, *Generalized statistics and solar neutrinos*, Physics Letters B, 369, 3-4, (1996), 308–312

[2] A. R. Plastino, A. Plastino, *Stellar polytropes and Tsallis' entropy*, Phys. Lett. A 174 (1993), 5-6, 384–386

[3] S. Moriguti, *A lower bound for a probability moment of any absolutely continuous distribution with finite variance,* Ann. Math. Stat. 23, 1952, 286–289

[4] J. N. Kapur, *Generalized Cauchy and Student's distributions as maximum-entropy distributions*, Proc. Nat. Acad. Sci. India Sect. A 58 (1988), 2, 235–246

[5] C. Vignat, J. Costa and A. Hero, *On Solutions to Multivariate Maximum alpha-entropy Problems*, Lecture Notes in Computer Science, Vol. 2683, 2003, 211–228

[6] A.R. Barron, *Entropy and the Central limit Theorem*, The Annals of Probability, 14-1, (1986), 336–342

[7] G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge, 1944

[8] J.F.C. Kingman, *Random Walks with Spherical Symmetry*, Act. Math. 109, 1963, 11–53

[9] I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series and Products*, Sixth edition, Alan Jeffrey and Daniel Zwillinger (eds.)

[10] C. Tsallis, *Possible Generalization of Boltzmann-Gibbs Statistics*, J. Stat. Phys. 52, 1988, 479–487