Source coding with Tsallis entropy

François CHAPEAU-BLONDEAU, Agnès DELAHAIES, David ROUSSEAU

Laboratoire d'Ingénierie des Systèmes Automatisés (LISA), Université d'Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France.

An extension is presented to the source coding theorem traditionally based on the Shannon entropy and latter generalized to the Rényi entropy. Another possible generalization is demonstrated, with a lower bound realized by the Tsallis entropy, when the performance is measured by a generalized average coding length which is exhibited, and with the optimal codelengths expressed from the escort probability distribution also known in nonextensive thermodynamics.

Introduction: As a quantitative measure of information, the Rényi entropy is defined, for a probability distribution p_i, $i = 1$ to N, as [1]

$$H_\alpha(p_i) = \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{N} p_i^\alpha\right) , \tag{1}$$

and represents an extension to the traditional Shannon entropy $H_1(p_i) = -\sum_{i=1}^{N} p_i \log(p_i)$ obtained at the limit $\alpha = 1$. The Rényi entropy of Eq. (1) can be motivated by its satisfying a set of axiomatic requirements, and, as an important complement, it has been shown to convey operational significance for practical problems, a very essential one being source coding [2], that we review below and complement.

Rényi source coding: The N symbols of a source alphabet are encoded into D-ary codewords, with length ℓ_i for symbol i having probability p_i, for $i = 1$ to N. For a uniquely decipherable code, the lengths ℓ_i must satisfy [2] the Kraft inequality

$$\sum_{i=1}^{N} D^{-\ell_i} \leq 1 . \tag{2}$$

The traditional approach to optimal source coding measures the elementary cost of encoding symbol i directly by its codelength ℓ_i, and then seeks those lengths ℓ_i that minimize the average coding length $\sum_{i=1}^{N} p_i\ell_i = \overline{\ell}$ while satisfying (2). The optimal lengths come out as $\ell_i^* = -\log_D(p_i)$, for $i = 1$ to N, and these achieve the minimum average coding length $\overline{\ell}^* = -\sum_{i=1}^{N} p_i \log_D(p_i)$, i.e. the Shannon entropy, which also forms a lower bound to $\overline{\ell}$ for any other codelengths ℓ_i.

The generalized approach to source coding of [2] measures the elementary cost of encoding symbol i as $D^{\beta\ell_i}$, introducing a cost which is an exponential function of the codelength ℓ_i, with a parameter $\beta > 0$ to have the cost an increasing function of the length ℓ_i. The global cost of encoding the source is expressed by the exponential average

$$C_\beta = \left(\sum_{i=1}^{N} p_i D^{\beta\ell_i}\right)^{1/\beta} . \tag{3}$$

Minimizing the cost C_β of Eq. (3) is equivalent to minimizing the monotonic increasing function of C_β as

$$L_\beta = \log_D(C_\beta) . \tag{4}$$

By measuring the coding performance with L_β of Eq. (4), the traditional approach is recovered at the limit $0 \leftarrow \beta$, when $L_0 = \overline{\ell}$ in Eq. (4). As $0 \leftarrow \beta$, the length L_β tends to distribute the weights among the codewords in proportion of their lengths, while as $\beta \to \infty$ more weight is put on long codewords [2, 3].

In this generalized approach to source coding, it can be proved [2] that the Rényi entropy forms a lower bound to the average coding length L_β, as

$$L_\beta = \log_D(C_\beta) \geq H_\alpha(p_i) , \tag{5}$$

with $\alpha = 1/(\beta + 1)$, or equivalently, for the exponential mean of Eq. (3),

$$C_\beta \geq D^{H_\alpha(p_i)} . \tag{6}$$

And the optimal lengths

$$\ell_i^* = -\log_D \frac{p_i^\alpha}{\displaystyle\sum_{j=1}^{N} p_j^\alpha} , \tag{7}$$

achieve equality in Eqs. (5) and (6) at the minimum average length $L_\beta^* = H_\alpha(p_i)$ while satisfying Eq. (2).

We now show that a similar approach can be developed, so as to provide in the same way an operational significance with a source coding problem, to a more recent measure of information consisting in the Tsallis entropy.

Tsallis source coding: The Tsallis entropy is defined, for a probability distribution p_i, $i = 1$ to N, as [4]

$$S_q(p_i) = \frac{1}{\ln(D)} \frac{1}{q-1} \left(1 - \sum_{i=1}^{N} p_i^q\right) , \tag{8}$$

and represents another extension to the traditional Shannon entropy $S_1(p_i) = -\sum_{i=1}^{N} p_i \log_D(p_i)$ obtainable at the limit $q = 1$. The Tsallis entropy has been postulated to form the ground of a nonextensive generalization to statistical mechanics [4].

It is possible to express Eq. (8) through a generalization [5, 4] of the traditional logarithm of base D, under the form of the q-logarithm of base D defined as

$$\log_D^q(x) = \frac{1}{\ln(D)} \frac{1 - x^{1-q}}{q-1} . \tag{9}$$

Inversion of $y = \log_D^q(x)$ defines the q-exponential function

$$\exp_D^q(y) = \Big[1 + \ln(D)(1-q)y\Big]^{1/(1-q)} . \tag{10}$$

At $q = 1$, one recovers the traditional logarithm $\log_D^1(x) = \log_D(x)$ and exponential $\exp_D^1(y) = D^y$. The Tsallis entropy of Eq. (8) is an expectation of the q-logarithm as

$$S_q(p_i) = \mathrm{E}\big[\log_D^q(1/p_i)\big] = \sum_{i=1}^{N} p_i \log_D^q(1/p_i) . \tag{11}$$

ELECTRONICS LETTERS 3rd February 2011 Vol. 47 No. 3 pages 187-188

By extracting the common factor $\sum_i p_i^\alpha$ from Eq. (1) and replacing in Eq. (8), a relation between Rényi H_α and Tsallis S_q entropies is obtained as $S_q = \mathcal{R}_q(H_q)$ with the function

$$\mathcal{R}_q(x) = \frac{1}{\ln(D)} \frac{1}{q-1} \left[1 - D^{(1-q)x}\right] , \qquad (12)$$

where D is the logarithm base used in Eq. (1) for H_q. From Eq. (9), another form is accessible for the relation based on Eq. (12), as

$$S_q = \log_D^q \left(D^{H_q}\right) . \qquad (13)$$

The transformation by the monotonic increasing function of Eq. (13) can be applied to Eq. (5) to yield

$$K_\beta = \log_D^\alpha(C_\beta) \geq S_\alpha(p_i) , \qquad (14)$$

or equivalently, analogous to Eq. (6),

$$C_\beta \geq \exp_D^\alpha[S_\alpha(p_i)] . \qquad (15)$$

A new generalized length K_β emerges with Eq. (14), which parallels L_β of Eq. (5), and which implements another monotonic increasing transformation of the exponential average C_β of Eq. (3) as a measure of the coding performance. Both L_β and K_β reduce to the traditional length $\overline{\ell}$ in the configuration $(\beta = 0, \alpha = 1/(\beta + 1) = 1)$, and otherwise stand as two distinct measures of the coding performance. Eq. (14) expresses a new coding theorem, where the Tsallis entropy forms a lower bound to the new length K_β, much like the Rényi entropy to the length L_β, and the Shannon entropy to the length $\overline{\ell}$. The new length K_β, contrary to L_β, is not a generalized average of the type $\varphi^{-1}[\sum_i p_i \varphi(\ell_i)]$ as considered in [3]. Instead, K_β emerges as a simple expression of the q-deformed logarithm, a basic tool of nonextensive thermodynamics [5, 4], where also the Tsallis entropy plays a central role. Furthermore, consistently with the derivation of Eq. (14), the optimal codelengths ℓ_i^* that minimize K_β down to the Tsallis entropy bound, are again those of Eq. (7). These optimal codelengths are also expressible as $\ell_i^* = -\log_D P_i^{(\alpha)}$, with $\{P_i^{(\alpha)}\}$ the escort probability distribution also playing an important role in nonextensive thermodynamics [4]. A comparable source coding study appeared in [6] yet with an entropy different from the plain Tsallis of Eq. (8), a coding length different from K_β of Eq. (14), and with optimal codelengths not given by the escort distribution.

Conclusion: Eq. (14) confers an operational role to the Tsallis entropy through a new coding theorem on a generalized length. Next is to investigate practical situations where the new generalized length K_β could be specifically useful. Complex conditions, like long-range correlations, heavy-tailed distributions, as usually associated with nonextensive thermodynamics, could offer a relevant context.

References

[1] A. Rényi, "On measures of entropy and information," in *Proceedings 4th Berkeley Symposium on Mathematical Statistics and Probability*, (Berkeley, USA), pp. 547–561, University of California Press, 1961.

[2] L. L. Campbell, "A coding theorem and Rényi's entropy," *Information and Control*, vol. 8, pp. 423–429, 1965.

[3] M. B. Baer, "Source coding for quasiarithmetic penalties," *IEEE Transactions on Information Theory*, vol. 52, pp. 4380–4393, 2006.

[4] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics*. New York: Springer, 2009.

[5] E. P. Borges, "A possible deformed algebra and calculus inspired in nonextensive thermostatistics," *Physica A*, vol. 340, pp. 95–101, 2004.

[6] T. Yamano, "Source coding theorem based on a nonadditive information content," *Physica A*, vol. 305, pp. 190–195, 2002.

7 October 2010 doi: 10.1049/el.2010.2792