Applied Mathematics, 2015, 6, 1968-1976
Published Online November 2015 in SciRes. http://www.scirp.org/journal/am
http://dx.doi.org/10.4236/am.2015.612174

An Algorithm to Generate Probabilities with Specified Entropy

Om Parkash, Priyanka Kakkar

Department of Mathematics, Guru Nanak Dev University, Amritsar, India
Email: omparkash777@yahoo.co.in, priyanka_kakkar85@yahoo.com

Received 8 September 2015; accepted 8 November 2015; published 11 November 2015

Abstract

The present communication offers a method to determine an unknown discrete probability distribution with specified Tsallis entropy close to uniform distribution. The relative error of the distribution obtained has been compared with the distribution obtained with the help of mathematica software. The applications of the proposed algorithm with respect to Tsallis source coding, Huffman coding and cross entropy optimization principles have been provided.

Keywords

Entropy, Source Coding, Kraft's Inequality, Huffman Code, Mean Codeword Length, Uniquely Decipherable Code

1. Introduction

After the publication of his first paper "A mathematical theory of communication", Shannon [1] made a remarkable discovery of entropy theory which immediately caught the interest of engineers, mathematicians and other scientists from various disciplines. Naturally one had speculated before Shannon about the nature of information but at the qualitative level but it was Shannon who for the first time introduced the following quantitative measure of information in a statistical framework:

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \tag{1}$$

with the convention $0\log 0 := 0$.

Shannon's main focus was related with the type of communication problems related with engineering sciences but as the field of information theory progressed, it became clear that Shannon's entropy was not the only feasible information measure. Indeed, many modern communication processes, including signals, images and coding systems, often operate in complex environments dominated by conditions that do not match the basic

How to cite this paper: Parkash, O. and Kakkar, P. (2015) An Algorithm to Generate Probabilities with Specified Entropy. *Applied Mathematics*, **6**, 1968-1976. http://dx.doi.org/10.4236/am.2015.612174

tenets of Shannon's communication theory. For instance, coding can have a non-trivial cost functions, codes might have variable lengths, sources and channels may exhibit memory or losses, etc. Post-Shannon developments of non-parametric entropy, it was realized that generalized parametric measures of entropy can play a significant role to deal with the prevailing situations, since these measures introduce flexibility in the system and also helpful towards maximization problems.

An extension to the Shannon entropy proposed by Renyi [2], is given by

$$H_\alpha(P) = \frac{1}{1-\alpha}\log\left(\sum_{i=1}^{n} p_i^\alpha\right), \quad \alpha \neq 1, \alpha > 0 \tag{2}$$

The Renyi entropy offers a parametric family of measures, from which the Shannon entropy is accessible as a special case when $\alpha \to 1$.

Another information theorist Tsallis [3] introduced his measure of entropy, given by

$$S^q(P) = \frac{1}{1-q}\left(\sum_{j=1}^{n} p_j^q - 1\right), \quad q \neq 1, q > 0 \tag{3}$$

When $q \to 1$, the Tsallis entropy recovers the Shannon entropy for any probability distribution. The Tsallis [4] entropy has been postulated to form the ground of a nonextensive generalization to statistical mechanics. Tsallis pioneering work has stimulated the exploration of the properties of other generalized or alternative information measures [5] [6]. Oikonomou and Bagci [7] maximized the Tsallis entropy based on complete deformed functions to show that the escort distributions are redundant. Bercher [8] showed that Tsallis distributions can be derived from the standard (Shannon) maximum entropy setting, by incorporating a constraint on the divergence between the distribution and another distribution imagined as its tail.

Information theory provides a fundamental performance limits pertaining to certain tasks of information processing, such as data compression, error-correction coding, encryption, data hiding, prediction, and estimation of signals or parameters from noisy observations. Shannon [1] provided an operational meaning to his entropy through a source coding theorem by establishing the limits to possible data compression. Bercher [9] discussed the interest of escort distributions and Rényi entropy in the context of source coding whereas Parkash and Kakkar [10] developed new mean codeword lengths and proved source coding theorems. Huffman [11] introduced a procedure for designing a variable length source code which achieves performance close to Shannon's entropy bound. Baer [12] provided new lower and upper bounds for the compression rate of binary prefix codes optimized over memoryless sources. Mohajer *et al*. [13] studied the redundancy of Huffman codes whereas Walder *et al*. [14] provided algorithms for fast decoding of variable length codes.

In Section 2, we have provided an algorithm to find a discrete distribution closer to uniform distribution with specified Tsallis [3] entropy. We have proved the acceptability of the algorithm by comparing the relative error of the probability distribution generated through algorithm with the probability distribution generated through Mathematica software through an example. Section 3 provides the brief introduction to source coding and the study of source coding with the Tsallis entropy. Also, we have extended the applications of the algorithm with respect to Tsallis source coding, Huffman coding and cross entropy optimization principles.

2. Generating Probability Distribution Closer to Uniform Distribution with Known Entropy

Tsallis introduced the generalized q-logarithm function is defined as

$$\log^q x = \frac{x^{1-q} - 1}{1-q} \tag{4}$$

which for $q = 1$, becomes the common natural logarithm. Its inverse is the generalized q-exponential function, given by

$$e_q^x = \left[1 + (1-q)x\right]^{\frac{1}{1-q}} \tag{5}$$

which becomes the exponential function for $q = 1$.

The q-logarithm satisfies the following pseudo additive law

$$\log^q xy = \log^q x + \log^q y + (1-q)\log^q x \log^q y \tag{6}$$

It is to be noted that the classical power and the additive laws for the logarithm and exponential do no longer hold for (4) and (6). Except for $q=1$, in general $\log^q x^\alpha \neq \alpha \log^q x$

The Tsallis entropy (3) can be written as an expectation of the generalized q-logarithm as

$$S_n^q(p) = E\left[\log^q\left(\frac{1}{p_j}\right)\right] = \sum_{j=1}^{n} p_j \log^q\left(\frac{1}{p_j}\right)$$

Let us suppose that there are $n \geq 3$ probabilities to be found. Separating the n^{th} probability and renaming it q_n, that is, $p_n = q_n$, we have

$$S_n^q = \sum_{j=1}^{n-1} p_j \log^q\left(\frac{1}{p_j}\right) + q_n \log^q\left(\frac{1}{q_n}\right)$$

Multiplying and dividing by $(1-q_n)$ (assuming $q_n \neq 1$, that is, $S_n^q \neq 0$) yields

$$S_n^q = (1-q_n)\sum_{j=1}^{n-1}\frac{p_j}{1-q_n}\log^q\left(\frac{(1-q_n)}{p_j(1-q_n)}\right) + q_n\log^q\left(\frac{1}{q_n}\right)$$

$$= (1-q_n)\sum_{j=1}^{n-1}\frac{p_j}{1-q_n}\left(\log^q\left(\frac{(1-q_n)}{p_j}\right) + \log^q\left(\frac{1}{(1-q_n)}\right) + (1-q)\log^q\left(\frac{(1-q_n)}{p_j}\right)\log^q\left(\frac{1}{(1-q_n)}\right)\right) + q_n\log^q\left(\frac{1}{q_n}\right)$$

Defining $r_j = \frac{p_j}{1-q_n}$, $j = 1, 2, \cdots, n-1$, the above expression can be written as

$$\begin{aligned} S_n^q = {} & (1-q_n)\sum_{j=1}^{n-1} r_j \log^q\left(\frac{1}{r_j}\right) + (1-q_n)\log^q\left(\frac{1}{(1-q_n)}\right)\sum_{j=1}^{n-1} r_j \\ & + (1-q_n)(1-q)\log^q\left(\frac{1}{(1-q_n)}\right)\sum_{j=1}^{n-1} r_j \log^q\frac{1}{r_j} + q_n\log^q\left(\frac{1}{q_n}\right) \end{aligned} \tag{7}$$

Since $r_j, j = 1, 2, \cdots, n-1$ form a full set of probabilities, we have $\sum_{j=1}^{n-1} r_j = 1$.Thus, the above expression (7) becomes

$$\begin{aligned} S_n^q = {} & (1-q_n)\sum_{j=1}^{n-1} r_j \log^q\left(\frac{1}{r_j}\right) + (1-q_n)\log^q\left(\frac{1}{(1-q_n)}\right) \\ & + (1-q_n)(1-q)\log^q\left(\frac{1}{(1-q_n)}\right)\sum_{j=1}^{n-1} r_j \log^q\frac{1}{r_j} + q_n\log^q\left(\frac{1}{q_n}\right) \\ = {} & (1-q_n)\sum_{j=1}^{n-1} r_j \log^q\left(\frac{1}{r_j}\right)\left(1 + (1-q)\log^q\left(\frac{1}{(1-q_n)}\right)\right) + (1-q_n)\log^q\left(\frac{1}{(1-q_n)}\right) + q_n\log^q\left(\frac{1}{q_n}\right) \\ = {} & (1-q_n)\sum_{j=1}^{n-1} r_j \log^q\left(\frac{1}{r_j}\right)\left(1 + \frac{(1-q)\left\{\left(\frac{1}{(1-q_n)}\right)^{1-q} - 1\right\}}{1-q}\right) + (1-q_n)\log^q\left(\frac{1}{(1-q_n)}\right) + q_n\log^q\left(\frac{1}{q_n}\right) \\ = {} & (1-q_n)^q \sum_{j=1}^{n-1} r_j \log^q\left(\frac{1}{r_j}\right) + (1-q_n)\log^q\left(\frac{1}{(1-q_n)}\right) + q_n\log^q\left(\frac{1}{q_n}\right) \\ = {} & (1-q_n)^q S_{n-1}^q(r) + S_b^q(q_n) \end{aligned} \tag{8}$$

where $S_{n-1}^q(r)$ is the expression for Tsallis entropy of the $(n-1)$-ary vector r and $S_b^q(.)$ is the binary entropy function

$$S_b^q(x) = x\log^q\left(\frac{1}{x}\right) + (1-x)\log^q\left(\frac{1}{(1-x)}\right)$$

Rearranging the terms in Equation (8) gives

$$S_{n-1}^q(r) = \frac{S_n^q - S_b^q(q_n)}{(1-q_n)^q} \tag{9}$$

Thus,

$$S_{n-1}^q(r) = L\left(S_n^q, q_n\right) \tag{10}$$

where $L\left(S_n^q, q_n\right) = \frac{S_n^q - S_b^q(q_n)}{(1-q_n)^q}$

The maximum value of Tsallis entropy subject to natural constraint, that is, $\sum_{i=1}^{n} p_i = 1$ is given by

$$S_{n(\max)}^q = \frac{n^{1-q}-1}{1-q} = \log^q n$$

which is obtained at uniform distribution $\left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$.

So, we have

$$0 \le S_n^q \le \log^q n$$

In a similar way, $S_{n-1}^q(r)$, being the entropy of $(n-1)$ variables with the probability vector *r* must satisfy

$$0 \le S_{n-1}^q \le \log^q (n-1)$$

Hence, q_n must satisfy the requirements

$$0 < q_n < 1, \tag{11}$$

$$0 \le L\left(S_n^q, q_n\right) \le \log^q (n-1) \tag{12}$$

The objective of present paper is to find q_n subject to conditions (11) and (12) so as to obtain next stage entropy $L\left(S_n^q, q_n\right)$, that is, $S_{n-1}^q(r)$. It is to be noted that the next iteration's entropy $S_{n-1}^q(r)$, may be larger or smaller than $S_n^q(r)$. The procedure may be iterated until only two variables remain, r_1 and r_2. The remaining entropy for these normalized variables, $S_2^q = L\left(S_3^q, q_3\right)$ satisfies the usual binary entropy function

$$S_2^q = r_1\log^q\left(\frac{1}{r_1}\right) + r_2\log^q\left(\frac{1}{r_2}\right) \tag{13}$$

To finish the selection of the q_j, take $q_1 = r_1$ and $q_2 = r_2 = 1 - q_1$ as one of the two solutions of (13). The set of scaled values q_1 through q_n is obtained at the end. Swaszek and Wali [15] made use of Shannon's entropy and to find the probability distribution for the same, provided the following relation between p_k's and q_k's after recursing through the sequential definitions of *r* vectors.

$$p_k = \begin{cases} q_n, & k = n, \\ q_k\left(1 - \sum_{j=k+1}^{n} p_j\right), & k = n-1, n-2, \cdots, 3, 2, \\ q_1\left(1 - \sum_{j=3}^{n} p_j\right), & k = 1. \end{cases} \tag{14}$$

We also use relation (14) to find probability distribution $(p_1, p_2, \cdots, p_n)$ for Tsallis entropy which is close to uniform distribution.

2.1. Method A

1) For given q, k, S_k^q, find the solutions of the following equation $\frac{x^q - x}{1-q} = \frac{S_k^q}{k}$.

2) Pick the solution z_k that lies in $\left(\frac{1}{k}, 1\right]$.

3) Generate the random number q_k in the interval $[0, z_k]$ for which $0 \le L\left(S_k^q, q_k\right) \le \log^q(k-1)$.

4) Repeat the above three steps for $k = n-1, n-2, \cdots, 3$.

5) For $k = 2$, solve Equation (13) for getting r_1 and r_2.

6) Take $q_1 = r_1$ and $q_2 = r_2$.

7) Use equation (14) to get probability distribution $p = (p_1, p_2, \cdots, p_n)$ closer to uniform distriburtion.

Note: Before specifying the value of parameter q and entropy S_n^q, make sure to choose that value of q for which value given entropy is less than or equal to its maximum value which is obtained at uniform distribution, that is,

$$S_n^q \le \log^q n.$$

2.2. Numerical

Let $n = 8,\ q = 0.3,\ S_8^q = 4.5$ bits. Applying above method gives $q_8 = 0.109755$ and $L\left(S_8^q, q_8\right) = S_7^q = 3.94803$. Proceeding on similar lines, we obtain remaining probabilities as depicted in the following **Table 1**.

Note that the values q_1 and q_2 do not lie in the neighbourhood of $\frac{1}{n}$ because they are exact solutions of equation of (13) and hence is the case for the values p_1 and p_2. With this exception, the other values, that is, $p_3, p_4, p_5, p_6, p_7, p_8$ are very close to the uniform distribution and the associated entropy is 4.5 bits.

2.3. Use of Mathematica Software

The above mentioned problem is also solved using the Mathematica software by using the same input. NMinimize command is used for this purpose which has several inbuilt optimization methods available. Since the problem is to find the discrete distribution $P = (p_1, p_2, \cdots, p_n)$ having a specific Tsallis entropy closer to the discrete uniform distribution $\left(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}\right)$, therefore the optimization problem becomes:

Minimize $Z(P) = \sum_{i=1}^{n}\left(p_i - \frac{1}{n}\right)^2$ subject to constraints

$$\sum_{i=1}^{n} p_i \log^q\left(\frac{1}{q}\right) = S_n^q,$$

Table 1. Probabilities obtained through method A.

n	q	S_n^q	q_n	p_n
8	0.3	4.5	$q_8 = 0.109755$	$p_8 = 0.109755$
7	0.3	3.94803	$q_7 = 0.139772$	$p_7 = 0.124431$
6	0.3	3.36823	$q_6 = 0.14697$	$p_6 = 0.112552$
5	0.3	2.75961	$q_5 = 0.172774$	$p_5 = 0.112867$
4	0.3	2.11183	$q_4 = 0.201015$	$p_4 = 0.108628$
3	0.3	1.41409	$q_3 = 0.281574$	$p_3 = 0.121574$
2	0.3	0.631956	$q_2 = 0.921147$ $q_1 = 0.0788534$	$p_2 = 0.285733$ $p_1 = 0.0244598$

$$\sum_{i=1}^{n} p_i = 1,$$

$$0 \le p_i \le 1, \quad i = 1, 2, \cdots, n.$$

The solution obtained is $P = (0.141287, 0.141287, 0.141287, 0.141287, 0.141287, 0.141287, 0.141287, 0.0109883)$ for $n = 8$.

2.4. Relative Errors

The relative error is calculated using the formula $\varepsilon = \frac{\sum_{i=1}^{n} \left| p_i - \frac{1}{n} \right|}{\sqrt{Z(P)}}$. It is found that relative error in case of probability distribution found by mathematica software is $\varepsilon_M = 1.87081$ whereas it is $\varepsilon_A = 1.67671$ in case of method A. This implies that method A provides an acceptable discrete distribution and hence the method itself is acceptable.

3. Source Coding

In source coding, one considers a set of symbols $X = (x_1, x_2, \cdots, x_n)$ and a source that produces symbols x_i from *X* with probabilities p_i where $\sum_{i=1}^{n} p_i = 1$. The aim of source coding is to encode the source using an alphabet of size *D*, that is to map each symbol x_i to a codeword c_i of length l_i expressed using the *D* letters of the alphabet. It is known that if the set of lengths l_i satisfies the Kraft's [16] inequality

$$\sum_{i=1}^{n} D^{-l_i} \le 1, \tag{15}$$

then there exists a uniquely decodable code with these lengths, which means that any sequence $c_{i1} c_{i2} \cdots c_{in}$ can be decoded unambiguously into a sequence of symbols $x_{i1} x_{i2} \cdots x_{in}$ Furthermore, any uniquely decodable code satisfies the Kraft's inequality (15).

The Shannon [1] source coding theorem indicates that the mean codeword length

$$L = \sum_{i=1}^{n} p_i l_i \tag{16}$$

is bounded below by the entropy of the source, that is, Shannon's entropy $H(P)$, and that the best uniquely decodable code satisfies

$$H(P) \le L < H(P) + 1$$

where the logarithm in the definition of the Shannon entropy is taken in base *D*. This result indicates that the Shannon entropy $H(P)$ is the fundamental limit on the minimum average length for any code constructed for the source. The lengths of the individual codewords, are given by

$$l_i = -\log_D p_i. \tag{17}$$

The characteristic of these optimum codes is that they assign the shorter codewords to the most likely symbols and the longer codewords to unlikely symbols.

Source Coding with Campbell Measure of Length

Implicit in the use of average codeword length (16) as a criteria of performance is the assumption that cost varies linearly with code length. But this is not always the case. Campbell [17] introduced the mean codeword length which implies that cost is an exponential function of code length. The cost of encoding the source is expressed by the exponential average

$$C_t = \left(\sum_{i=1}^{n} p_i D^{l_i t} \right)^{\frac{1}{t}} \tag{18}$$

where $t>0$ is some parameter related to cost.

Minimizing the cost is equivalent to minimizing the monotonic increasing function of C_t defined as

$$L_t = \log_D\left(C_t\right)$$

So, L_t is the exponentiated mean codeword length given by Campbell [12] which approaches to L as $t \to 0$.

Campbell proved that Renyi [2] entropy forms a lower bound to the exponentiated codeword length L_t as

$$L_t = \log_D\left(C_t\right) \geq H_\alpha\left(P\right) \tag{19}$$

where $\alpha = \frac{1}{1+t}$ or equivalently

$$C_t \geq D^{H_\alpha(P)} \tag{20}$$

subject to Kraft's inequality (15) with optimal lengths given by

$$l_i = -\log_D \frac{p_i^\alpha}{\sum_{j=1}^{n} p_j^\alpha} \tag{21}$$

By choosing a smaller value of α, the individual lengths can be made smaller than the Shannon lengths $l_i = -\log_D p_i$, specially for small p_i.

Similar approach is applied to provide an operational significane to Tsallis [3] entropy through a source coding problem.

From Renyi's entropy of order α where logarithm is taken to the base D, we have

$$\sum_{i=1}^{n} p_i^\alpha = D^{(1-\alpha)H_\alpha} \tag{22}$$

Substituting (22) in (3) where parameter q is replaced by α gives

$$S^\alpha\left(P\right) = \frac{1}{1-\alpha}\left(D^{(1-\alpha)H_\alpha} - 1\right)$$

or equivalently

$$S^\alpha = \log^\alpha\left(D^{H_\alpha}\right) \tag{23}$$

Equation (23) establishes a relation between Renyi's entropy and Tsallis entropy.

From (20), we have

$C_t \geq D^{H_\alpha}$ where $\alpha = \frac{1}{1+t}$

Case-I Now, when $0 < \alpha < 1$

$$\begin{aligned} & C_t^{(1-\alpha)} \geq D^{(1-\alpha)H_\alpha} \\ & \Rightarrow \frac{C_t^{(1-\alpha)} - 1}{1-\alpha} \geq \frac{D^{(1-\alpha)H_\alpha} - 1}{1-\alpha} \\ & \Rightarrow G_\alpha = \log^\alpha\left(C_t\right) \geq \log^\alpha\left(D^{H_\alpha}\right) = S^\alpha \end{aligned} \tag{24}$$

Case-II when $\alpha > 1$

$$\begin{aligned} & C_t^{(1-\alpha)} \leq D^{(1-\alpha)H_\alpha} \\ & \Rightarrow \frac{C_t^{(1-\alpha)} - 1}{1-\alpha} \geq \frac{D^{(1-\alpha)H_\alpha} - 1}{1-\alpha} \\ & \Rightarrow G_\alpha = \log^\alpha\left(C_t\right) \geq \log^\alpha\left(D^{H_\alpha}\right) = S^\alpha \end{aligned} \tag{25}$$

From (24) and (25), it is observed that Tsallis entropy S^α forms a lower bound to G_α which is nothing but the new generalized length and is a monotonic increasing function of C_t. It reduces to mean codeword length L

when $\alpha \to 1$. The optimal codeword lengths are given by $l_i = -\log_D \frac{p_i^\alpha}{\sum_{j=1}^{n} p_j^\alpha}$ which is similar as in case of Campbell's mean codeword length. G_α is not an average of the type $\phi^{-1}\left(\sum_i p_i \phi(l_i)\right)$ as introduced by Kolmogorov [18] and Nagumo [19] but is a simple expression of the q-deformed logarithm.

4. Application of Method A

1) Huffman [11] introduced a method for designing variable length source code in which he showed that the average length of a Huffman code is always within one unit of source entropy, that is, $H(P) \le L < H(P)+1$ where L is defined by (16) and $H(P)$ is defined by (1). By using method A as mentioned in Section 2, different sets of probability distributions can be generated which are closer to uniform distribution and which has same Tsallis entropy. The probability distributions thus generated can be used to develop Huffman code and to see that whether the lengths of Huffman codewords satisfies relation (24) or (25) where the Tsallis entropy forms a lower bound to generalized length G_α. Thus, performance of Huffman algorithm can be judged in case of source coding with Tsallis entropy.

In the following example, Huffman code is constructed using the probability distribution obtained in **Table 1**. Let $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$ be an array of symbols with probabilities in decreasing order as shown below where Huffman method is employed to construct an optimal code.

x_1	0.285733	x_1	0.285733	x_1	0.285733	x_1	0.285733	x_1	0.285733
x_2	0.124431	$x_{7,8}$	0.133088	$x_{5,6}$	0.222307	$x_{3,4}$	0.234441	$x_{2,78}$	0.257519
x_3	0.121574	x_2	0.124431	$x_{7,8}$	0.133088	$x_{5,6}$	0.222307	$x_{3,4}$	0.234441
x_4	0.112867	x_3	0.121574	x_2	0.124431	$x_{7,8}$	0.133088	$x_{5,6}$	0.222307
x_5	0.112552	x_4	0.112867	x_3	0.121574	x_2	0.124431		
x_6	0.109755	x_5	0.112552	x_4	0.112867				
x_7	0.108628	x_6	0.109755						
x_8	0.0244598								

$x_{3,456}$	0.456748	$x_{1,278}$	0.543252
x_1	0.285733	$x_{3,456}$	0.456748
$x_{2,78}$	0.234441		

Optimal code is obtained as follows

$x_{1,278}$	0	x_1	00	x_1	00	x_1	00
$x_{3,456}$	1	$x_{2,78}$	01	x_2	010	x_2	010
		x_3	10	$x_{7,8}$	011	x_7	0110
		$x_{4,56}$	11	x_3	10	x_8	0111
				x_4	110	x_3	10
				$x_{5,6}$	111	x_4	110
						x_5	1110
						x_6	1111

Hence the optimal code is $(00, 010, 10, 110, 1110, 1111, 0110, 0111)$ and the lengths of the codewords are $l_1 = 2,\ l_2 = 3,\ l_3 = 2,\ l_4 = 3,\ l_5 = 4, l_6 = 4, l_7 = 4, l_8 = 4$.

So, using the probability distribution generated in **Table 1** along with known value of $\alpha(q = 0.3)$ and above mentioned codeword lengths, value of G_α is obtained as 6.2327 which is greater that Tsallis entropy (=4.5 bits), thus satisfies relation (24).

2) The problem of determining an unknown discrete distribution closer to uniform distribution with known Tsallis entropy as discussed in Section 2 can be looked upon as minimum cross entropy principle which states that given any priori distribution, we should choose that distribution which satisfies the given constraints and which is closest to priori distribution. So, cross entropy optimization principles offer a relevant context for the application of method A.

Acknowledgements

The authors are thankful to University Grants Commission and Council of Scientific and Industrial Research, New Delhi, for providing the financial assistance for the preparation of the manuscript.

References

[1] Shannon, C.E. (1948) A Mathematical Theory of Communication. *Bell System Technical Journal*, **27**, 379-423. http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

[2] Renyi, A. (1961) On Measures of Entropy and Information. *Proceedings of the* 4*th Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 547-561.

[3] Tsallis, C. (1988) Possible Generalization of Boltzmann-Gibbs Statistics. *Journal of Statistical Physics*, **52**, 479-487. http://dx.doi.org/10.1007/BF01016429

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

[5] Papa, A.R.R. (1998) On One-Parameter-Dependent Generalizations of Boltzmann Gibbs Statistical Mechanics. *Journal of Physics A*: *Mathematical and General*, **31**, 5271-5276. http://dx.doi.org/10.1088/0305-4470/31/23/009

[6] Abe, S. (1997) A Note on the q-Deformation-Theoretic Aspect of the Generalized Entropies in Nonextensive Physics. *Physics Letters A*, **224**, 326-330. http://dx.doi.org/10.1016/S0375-9601(96)00832-8

[7] Oikonomou, T. and Bagci, G.B. (2010) The Maximization of Tsallis Entropy with Complete Deformed Functions and the Problem of Constraints. *Physics Letters A*, **374**, 2225-2229. http://dx.doi.org/10.1016/j.physleta.2010.03.038

[8] Bercher, J.F. (2008) Tsallis Distribution as a Standard Maximum Entropy Solution with “Tail” Constraint. *Physics Letters A*, **372**, 5657-5659. http://dx.doi.org/10.1016/j.physleta.2008.06.088

[9] Bercher, J.F. (2009) Source Coding with Escort Distributions and Renyi Entropy Bounds. *Physics Letters A*, **373**, 3235-3238. http://dx.doi.org/10.1016/j.physleta.2009.07.015

[10] Parkash, O. and Kakkar, P. (2012) Development of Two New Mean Codeword Lengths. *Information Sciences*, **207**, 90-97. http://dx.doi.org/10.1016/j.ins.2012.04.020

[11] Huffman, D.A. (1952) A Method for the Construction of Minimum Redundancy Codes. *Proceedings of the IRE*, **40**, 1098-1101. http://dx.doi.org/10.1109/JRPROC.1952.273898

[12] Baer, M.B. (2011) Redundancy-Related Bounds for Generalized Huffman Codes. *IEEE Transactions on Information Theory*, **57**, 2278-2290. http://dx.doi.org/10.1109/TIT.2011.2110670

[13] Mohajer, S., Pakzad, P. and Kakhbod, A. (2012) Tight Bounds on the Redundancy of Huffman Codes. *IEEE Transactions on Information Theory*, **58**, 6737-6746. http://dx.doi.org/10.1109/TIT.2012.2208580

[14] Walder, J., Kratky, M., Baca, R., Platos, J. and Snasel, V. (2012) Fast Decoding Algorithms for Variable-Lengths Codes. *Information Sciences*, **183**, 66-91. http://dx.doi.org/10.1016/j.ins.2011.06.019

[15] Swaszek, P.F. and Wali, S. (2000) Generating Probabilities with Specified Entropy. *Journal of the Franklin Institute*, **337**, 97-103. http://dx.doi.org/10.1016/S0016-0032(00)00010-7

[16] Kraft, L.G. (1949) A Device for Quantizing Grouping and Coding Amplitude Modulated Pulses. M.S. Thesis, Electrical Engineering Department, MIT.

[17] Campbell, L.L. (1965) A Coding Theorem and Renyi’s Entropy. *Information and Control*, **8**, 423-429. http://dx.doi.org/10.1016/S0019-9958(65)90332-3

[18] Kolmogorov, A. (1930) Sur la notion de la moyenne. *Atti R. Accad. Naz. Lincei*, **12**, 388.

[19] Nagumo, M. (1930) Uber eine Klasse der Mittelwerte. *Japanese Journal of Mathematics*, **7**, 71-79.