CHAPTER 5

SOME CODING THEOREMS FOR NONADDITIVE GENERALIZED MEAN-VALUE ENTROPIES

5.1 Introduction

In this chapter we give an optimality characterization of non-additive generalized mean-value entropies from suitable non-additive and generalized mean-value properties of the measure of average length. The results obtained by us covers many results obtained by other authors as particular cases, as well as the ordinary length due to Shannon [98]. The main instrument $l(n_i)$ is the function of the word lengths in obtaining the average length of the code.

Given a discrete random variable X taking a finite number of values $(x_1, x_2, ..., x_n)$ with probabilities $P = (p_1, p_2, ..., p_n),\ p_i \geq 0,\ \sum_{i=1}^{n} p_i = 1,$ the Shannon's entropy of the probability distribution is given by

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i \ , \qquad (5.1.1)$$

where the base of the logarithm is in general arbitrary.

Shannon entropy is a very useful and powerful measure having rich meanings. Entropy has an important connection with noiseless coding. If $X = (x_1, x_2, ..., x_n)$ represents an information source with n messages and input probabilities $p_1, p_2, ..., p_n$, as given above, that is encoded into words of lengths $N = \{n_1, n_2, ..., n_n\}$ forming an instantaneous code, then

$$\sum_{i=1}^{n} D^{-n_i} \leq 1, \qquad (5.1.2)$$

where D is the size of the code alphabet.

The average length L, for the instantaneous code is such that

$$L = p_i n_i \geq -\sum_{i=1}^{n} p_i \log_D p_i \ , \qquad (5.1.3)$$

with equality iff for each i

$$p_i = D^{-n_i}. \qquad (5.1.4)$$

Eq. (5.1.4), refer Shannon [98], characterizes Shannon's entropy as a measure of optimality of a linear function, viz. L, under the relation defined by Eq. (5.1.2).

Several generalizations of Shannon's entropy have been studied. All these become special cases of those studied by Sharma and Taneja [107] or Sharma and Mittal [101].

We define

$$H^*(p_1, p_2, ...p_n; \alpha, \beta) = (2^{\frac{\beta-1}{\alpha}} - 1)^{-1}\left[\exp_2(\frac{\beta-1}{\alpha})\sum_{i=1}^{n} p_i \log_2 p_i - 1\right]$$

$$\beta > 0, \beta \neq 1 \qquad (5.1.5)$$

and

$$H^*(p_1, p_2, ..., p_n; \alpha, \beta) = (2^{\frac{1-\beta}{\alpha}} - 1)^{-1}\left[\left(\sum_{i=1}^{n} p_i^{\alpha}\right)^{\frac{1}{\alpha}\left(\frac{\beta-1}{\alpha-1}\right)}\right]$$

$$\beta > 0, \beta \neq 1, \ \alpha > 0, \alpha \neq 1. \qquad (5.1.6)$$

These quantities satisfy the non-additivity

$$H^*(P * Q) = H^*(P) + H^*(Q) + (2^{\frac{1-\beta}{\alpha}} - 1)H^*(P)H^*(Q). \qquad (5.1.7)$$

In this chapter, we give an optimality characterization of entropies given by Eq. (5.1.5) and Eq. (5.1.6) from suitable non-additive and generalized mean-value properties of the measure of average length. The results obtained cover many results obtained by other authors as particular cases, as well as the Eq. (5.1.3). The main instrument is the $l(.)$ function of the word lengths in obtaining the average length of the code.

5.2 Non-additive Measure of Code Length

Let us consider two independent sources $X = \{x_1, x_2, ..., x_n\}$ and $Y = \{y_1, y_2, ..., y_n\}$ with associated probability distribution $P = (p_1, p_2, ..., p_n)$

and $Q=(q_1,q_2,...,q_m)$. Then the probability distribution of the product $XY=\{(x_i,y_i)\}$ is $P*Q=(p_1q_1,p_1q_2,...,p_1q_m,...,p_nq_m)$. Let the source Y be encoded with a code of length $M=\{m_1,m_2,...,m_m\}$ and the pair (x_i,y_j) be represented by a sequence for x_i and y_j put side by side, so that the product source has code length sequence

$$N+M=\{n_1+m_1,...,n_1+m_n,...,n_n+m_m\}.$$

The additive measure of mean length is required to satisfy the requirement, refer Campbell [24],

$$L(P*Q,N+M,\phi)=L(P,N,\phi)+L(Q,M,\phi), \quad (5.2.1)$$

where

$$L(P,N,\phi)=\phi^{-1}\left(\sum_{i=1}^{n}p_i\phi(n_i)\right), \quad (5.2.2)$$

ϕ being a continuous strictly monotonic increasing function.

Campbell [23] proved a noiseless coding theorem for Renyi's [95] entropy of order α in terms of mean length given by Eq. (5.2.2) defined for $\phi(n_i)=D^{n_i}$.

The mean length concerning Shannon's entropy and order α entropy of Renyi are both additives as they satisfy additivity of type given by Eq. (5.2.1).

Here we deal with non-additive measures of length denoted by L^* which satisfy 'non-additivity relation'

$$L^*(P+Q,N+M,\phi)=L^*(P,N,\phi)+L^*(Q,M,\phi)+\lambda L^*(P,N,\phi)L^*(Q,M,\phi) \quad (5.2.3)$$

and the mean value property

$$L^*(P,N,\phi)=\phi^{-1}\left(\frac{\sum_{i=1}^{n}p_i\phi(l(n_i))}{\sum_{i=1}^{n}p_i}\right), \quad (5.2.4)$$

where $l(n_i)$ is the function of length of a single element with code word length n_i, which is non-additive.

5.3 Characterization of Non-additive Measures of Code Length

We take the mean value non-additive measures of length given by Eq. (5.2.4) to satisfy the Eq. (5.2.3), where the expression and notations used there have their meanings explained earlier.

First of all we shall determine the non-additive length function l of the code word length in satisfying the non-additivity relation

$$l(n+m) = l(n) + l(m) + \lambda l(n)l(m), \quad \lambda \neq 0. \tag{5.3.1}$$

This by taking $f(n) = 1 + \lambda l(n)$ gives

$$f(n+m) = f(n)f(m). \tag{5.3.2}$$

The most general non-zero solution of Eq. (5.3.2), is

$$f(n) = D^{\left(\frac{1-\beta}{\alpha}\right)n},$$

where $\alpha, \beta \neq 1$ are arbitrary constants(we have taken base D with a purpose here).

So that

$$l(n) = \frac{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}{\lambda}, \quad \lambda \neq 0. \tag{5.3.3}$$

At this stage we make a proper choice of the constant λ. By analogy its value is dictated as

$$\lambda = (D^{\left(\frac{1-\beta}{\alpha}\right)} - 1), \ \beta \neq 1 \text{ as } \lambda \neq 0.$$

Another purpose served with this value of λ is that when it tends to zero, i.e. when $\beta \to 1$, the function of length $l(n)$ should reduce to additive one which is $l(n) = n$. This value of λ can also be obtained by imposing a boundary condition $l(n) = 1$.

So that finally, we have

$$l\{n) = \frac{D^{\left(\frac{1-\beta}{\alpha}\right)n} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}, \quad \beta \neq 1. \tag{5.3.4}$$

Next we proceed to determine $L^*(P,N,\phi)$ by first evaluating the values of ϕ. To achieve this we put the value of $l(n)$ from Eq. (5.3.4) in Eq. (5.2.4) and then use the Eq. (5.2.3) with $\lambda = (D^{\left(\frac{1-\beta}{\alpha}\right)} - 1),\ \beta \neq 1$, to get

$$\phi^{-1}\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m} p_i q_j \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)(n_i+m_j)} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{m} p_i q_j}\right)$$

$$= \phi^{-1}\left(\frac{\sum_{i=1}^{n} p_i \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n_i} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}\right)}{\sum_{i=1}^{n} p_i}\right) + \phi^{-1}\left(\frac{\sum_{j=1}^{m} q_j \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)m_j} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}\right)}{\sum_{j=1}^{m} q_j}\right)$$

$$+ \left(D^{\left(\frac{1-\beta}{\alpha}\right)} - 1\right)\phi^{-1}\left(\frac{\sum_{i=1}^{n} p_i \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n_i} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}\right)}{\sum_{i=1}^{n} p_i}\right)\phi^{-1}\left(\frac{\sum_{j=1}^{m} q_j \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)m_j} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}\right)}{\sum_{j=1}^{m} q_j}\right).$$

(5.3.5)

Now let us take $Q = \{q\}$ and $M = \{m\}$, so that for

$P = (p_1, p_2, ..., p_n),\ p_i \geq 0,\ \sum_{i=1}^{n} p_i = 1,$ (5.3.5) gives after some simplification

$$\phi^{-1}\left(\sum_{i=1}^{n} p_i \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)(n_i+m)} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}\right)\right) = \phi^{-1}\left(\sum_{i=1}^{n} p_i \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n_i} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}\right)\right) D^{\left(\frac{1-\beta}{\alpha}\right)m} + \frac{D^{\left(\frac{1-\beta}{\alpha}\right)m} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1}$$

or

$$\psi_m^{-1}\left(\sum_{i=1}^{n} p_i \psi_m\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n_i}-1}{D^{\left(\frac{1-\beta}{\alpha}\right)}-1}\right)\right)=\phi^{-1}\left(\sum_{i=1}^{n} p_i \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n_i}-1}{D^{\left(\frac{1-\beta}{\alpha}\right)}-1}\right)\right)D^{\left(\frac{1-\beta}{\alpha}\right)m}+\frac{D^{\left(\frac{1-\beta}{\alpha}\right)m}-1}{D^{\left(\frac{1-\beta}{\alpha}\right)}-1}$$

(5.3.6)

where

$$\psi_m\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n_i}-1}{D^{\left(\frac{1-\beta}{\alpha}\right)}-1}\right)=\phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)(n_i+m)}-1}{D^{\left(\frac{1-\beta}{\alpha}\right)}-1}\right). \quad (5.3.7)$$

Now, refer Hardy Littlewood and Polya [37], there must be a linear relation between ψ_m and ϕ, i.e.

$$\psi_m(n) = A(m)\phi(n) + B(m), \quad (5.3.8)$$

where $A(m)$ and $B(m)$ are independent of n.

Using Eq. (5.3.7) and Eq. (5.3.8), we have

$$g(n+m) = A(m)g(n) + B(m), \quad (5.3.9)$$

where

$$g\{n) = \phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n}-1}{D^{\left(\frac{1-\beta}{\alpha}\right)}-1}\right). \quad (5.3.10)$$

or

$$G(n+m) = A(m)G(n) + G(m), \quad (5.3.11)$$

where

$$G(n) = g(n) - g(0) = g(n) - a,\ \ a = g(0). \quad (5.3.12)$$

From the symmetry of Eq. (5.3.11), we get

$$A(m)G(n) + G(m) = A(n)G(m) + G(n)$$

It implies $\frac{G(n)}{A(n)-1} = \frac{G(m)}{A(m)-1} = \frac{1}{K}$ (say).

Thus

$$A(n) - 1 = K(G(n)) \text{ for all values of } n. \quad (5.3.13)$$

There are two cases, viz., $K = 0$ and $K \neq 0$.

If $K = 0$, $A(n) = 1$ and Eq. (5.3.11) gives

$$G(n+m) = G(n) + G(m) \ ,$$

the most general continuous solution of which is given by

$$G(n) = cn \, , \qquad (5.3.14)$$

where c is an arbitrary constant.

This by Eq. (5.3.12) and Eq. (5.3.10) gives

$$\phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1} \right) = a + cn \, ,$$

which gives

$$\phi(n) = a + \frac{c\alpha}{1-\beta} \log_D [1 + (D^{\frac{1-\beta}{\alpha}} - 1)n], \qquad \beta \neq 1 \qquad (5.3.15)$$

Again if $K \neq 0$, we have the relation obtained from Eq. (5.3.13) and Eq. (5.3.11)

$$A(n+m) = A(n)A(m), \qquad (5.3.16)$$

the general continuous solution of which are $A(n) = 0$ (which we neglect) for all n,

and

$$A(n) = D^{tn} \, ,$$

where t is an arbitrary nonzero constant.

This by using Eq. (5.3.13), Eq. (5.3.12) in Eq. (5.3.10), gives

$$\phi\left(\frac{D^{\left(\frac{1-\beta}{\alpha}\right)n} - 1}{D^{\left(\frac{1-\beta}{\alpha}\right)} - 1} \right) = a + \frac{D^{tn} - 1}{K} \, ,$$

which gives

$$\phi(n) = a + \frac{((D^{\frac{1-\beta}{\alpha}} - 1)n + 1)^{t\alpha/(1-\beta)} - 1}{K} \qquad (5.3.17)$$

$$\beta \neq 1, \; t \neq 0.$$

The value of ϕ given by Eq. (5.3.15) and Eq. (5.3.17) determine the following two non-additive measures of length defined by Eq. (5.2.4), i.e.

$$L^*(P,N;1,\alpha,\beta) = (D^{\frac{1-\alpha}{\beta}} - 1)^{-1}\left[\left(\sum_{i=1}^{n} p_i^{\beta}\right)^{\frac{1}{\beta}\left(\frac{\alpha-1}{\beta-1}\right)} - 1\right], \quad (5.3.18)$$

$$L^*(P,N;t,\alpha,\beta) = (D^{\frac{1-\alpha}{\beta t}\log_D \sum_{i=1}^{n} p_i D^{n_i t}} - 1). \quad (5.3.19)$$

These code length denoted by $L^*(P,N;1,\alpha,\beta)$ and $L^*(P,N;t,\alpha,\beta)$ may be named as non-additive type (α,β) lengths of order 1 and t respectively.

These results are contained in the following theorem:

Theorem 5.3.1. The mean length given by Eq. (5.2.4) of a sequence of lengths n_i, $i = 1,2,...,n$ formed of the code alphabet of size D of a probability distribution $P = (p_1, p_2, ..., p_n)$, $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$ satisfying $\sum_{i=1}^{n} D^{-n_i} \leq 1$ and non-additivity relation can be only of one of the two forms given in Eq. (5.3.18) and Eq. (5.3.19).

5.3.1 Limiting and Particular Cases

It is immediate to see that

1. $\lim_{t \to 0} L^*(P,N;t,\alpha,\beta) = L^*(P,N;1,\alpha,\beta)$. (5.3.20)

2. $\lim_{\beta \to 1} L^*(P,N;1,\alpha,\beta) = \sum_{i=1}^{n} p_i n_i = L$. (5.3.21)

3. $\lim_{\beta \to 1} L^*(P,N;t,\alpha,\beta) = \frac{1}{t}\log \sum_{i=1}^{n} p_i D^{t n_i}$, (5.3.22)

length of order t defined by Campbell [23].

For $n_1 = n_2 = ... = n_n = n$ (say), both the expressions for length given by Eq. (5.3.18) and Eq. (5.3.19) reduce to

$$\left(D^{\left(\frac{1-\beta}{\alpha}\right)} - 1\right)^{-1}\left[D^{\left(\frac{1-\beta}{\alpha}\right)n} - 1\right],$$

which in the limiting case when β approaches unity, reduces to n.

Thus we have shown that $L^*(P,N;1,\alpha,\beta)$ and $L^*(P,N;t,\alpha,\beta)$ are type β generalizations of

$$L=\sum_{i=1}^{n} p_i n_i \text{ and } L(t)=\frac{1}{t}\log_D \sum_{i=1}^{n} p_i D^{tn_i}$$

respectively.

We now prove the following:

Theorem 5.3.2 If $n_1 = n_2 = ... = n_n = n$ denote the lengths of an instantaneous/ uniquely decipherable code formed of code alphabet of size D

(i) $L^*(P,N;1,\alpha,\beta) \geq H^*(P;\alpha,\beta),$ (5.3.23)

with equality iff $n_i = -\log_D p_i$ for all i.

(ii) $L^*(P,N;t,\alpha,\beta) \geq H^*(P;\alpha,\beta),$ (5.3.24)

with equality iff

$$n_i = -\alpha \log_D p_i + \log_D\left(\sum_{i=1}^{n} p_i^{\alpha}\right),$$

where $\alpha=(1+t)^{-1}$.

Proof . As in Shannon's case refer Feinstien [31]

$$\sum_{i=1}^{n} p_i n_i \geq -\sum_{i=1}^{n} p_i \log_D p_i.$$

Now $(D^{\left(\frac{1-\beta}{\alpha}\right)}-1)>0$ or <0 according as $\beta<1$ or >1.

Therefore from the above after suitable manipulation, we get the inequality

$$\frac{D^{\frac{1-\beta}{\alpha}\sum_{i=1}^{n} p_i n_i}-1}{D^{\frac{1-\beta}{\alpha}}-1} \geq \frac{D^{\frac{\beta-1}{\alpha}\sum_{i=1}^{n} p_i \log_D p_i}-1}{D^{\frac{1-\beta}{\alpha}}-1}, \quad \beta \neq 1$$

which is the Eq. (5.3.23).

The case of equality can be discussed, as for Shannon's, which holds only when $n_i = -\log_D p_i$, for each i.

We now proceed to part (ii) . If $t=0$, the result is the same proved in part (i). For other values, we use Holder's inquality

$$\left(\sum_{i=1}^{n} x_i^p\right)^{1/p} . \left(\sum_{i=1}^{n} y_i^q\right) \le x_i y_i, \qquad (5.3.25)$$

where $p^{-1} + q^{-1} = 1$ and $p < 1$.

Making the substitutions $p = -t$, $q = (1-\alpha)$, $x_i = p_i^{-1/t} D^{-n_i}$ and $y_i = p_i^{1/t}$ in Eq. (5.3.25), we get after suitable manipulations

$$\left(\sum_{i=1}^{n} p_i D^{tn_i}\right)^{1/t} \ge \left(\sum_{i=1}^{n} p_i^{\alpha}\right)^{1/1-\alpha}, \qquad (5.3.26)$$

with $\alpha = (1+t)^{-1}$

Raising the power to $(1-\beta)/\alpha$ of both sides of Eq. (5.3.26) and using $(D^{\left(\frac{1-\beta}{\alpha}\right)} - 1) > 0$ or <0 according as $\beta < 1$ or >1, respectively then after simple manipulation we get Eq. (5.3.24).

Hence the theorem is proved.

5.4 Conclusions

Coding theorems for non-additive generalized mean-value entropies have been proved. Our main results cover many results obtained by other authors as particular cases, as well as the ordinary length due to Shannon [98]. The main instrument $l(n_i)$ is the function of the word lengths in obtaining the average length of the code. The results obtained are used as application in chater 6.