CHAPTER 5

On Bounds of R-Norm Mean Code Lengths

5.1 Introduction

Let us consider the set of positive real numbers, not equal to 1 and denote this
by $\mathbb{R}^*$ defined as $\mathbb{R}^* = \{ R : R > 0, R \neq 1 \}$. Let $\Delta_n$ with $n \geq 2$ is the set of all probability
distributions $P = \{(p_1, p_2, \ldots, p_n), p_i \geq 0, \text{ for each } i \text{ and } \sum_{i=1}^{n} p_i = 1 \}$.

Boekee and Lubbe (1965) studied R-norm information of the distribution $P$
defined for $R \in \mathbb{R}^*$ by

$$H_R(P) = \frac{R}{R-1} \left[ 1 - \left( \sum_{i=1}^{n} p_i^{R} \right)^{\frac{1}{R}} \right] \quad \text{(5.1.1)}$$

The measure (5.1.1) is different from entropies of Shannon (1948), Renyi
(1976), Havrda and Charvat (1967) and Daroczy (1970). The main property of this
measure is that when $R \to 1$, (5.1.1) approaches to Shannon’s entropy and when
$R \to \infty$, $H_R(P) \to (1 - \max_i p_i)$, for each $i = 1, 2, \ldots, n$.

The measure (5.1.1) can be generalized in so many ways. Hooda and Ram (1998)
proposed and characterized the following parametric generalization of (5.1.1):

$$H_R^\beta(P) = \frac{R}{R+\beta-2} \left[ 1 - \left( \sum_{i=1}^{n} p_i^{R} \right)^{\frac{2-\beta}{R}} \right], 0 < \beta \leq 1, R(>0) \neq 1 \quad \text{(5.1.2)}$$

The above measure (5.1.2) was known as generalized R-norm entropy of degree $\beta$
which reduces to (5.1.1) when $\beta = 1$. Further when $R=1$, (5.1.2) reduces to
\[ H_i^\beta(P) = \frac{1}{\beta - 2} \left[ 1 - \left( \sum_{i=1}^{n} p_i^{2-\beta} \right)^{2-\beta} \right], \quad 0 < \beta \leq 1, \] (5.1.3)

In case \( \gamma = \frac{1}{2 - \beta} \), (5.1.3) reduces to

\[ H^\gamma(P) = \frac{\gamma}{\gamma - 1} \left[ 1 - \left( \sum_{i=1}^{n} p_i^{\gamma} \right)^{\frac{1}{\gamma}} \right], \quad \frac{1}{2} < \gamma \leq 1 \] (5.1.4)

This is an information measure which has been given by Arimoto (1971). It can be seen that (5.1.4) also reduces to Shannon’s entropy when \( \gamma \to 1 \).

Hooda and Sharma (2008) proposed and studied the following parametric generalization of (5.1.1):

\[ H_R^{\alpha,\beta}(P) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left( \sum_{i=1}^{n} p_i^{\frac{R}{2\alpha-\beta}} \right)^{\frac{2\alpha-\beta}{R}} \right], \] (5.1.5)

\( \alpha \geq 1, \ 0 < \beta \leq 1, \ R(>0) \neq 1, \ 0 < R + \beta \neq 2\alpha. \)

They called (5.1.5) as the generalized R-norm information measure of type \( \alpha \) and degree \( \beta \). (5.1.5) reduces to (5.1.2) when \( \alpha = 1 \), which further reduces to (5.1.1) when \( \beta = 1 \).

In this chapter we discuss the characterization of the generalized R-norm information measure (5.1.5) by infimum method in section 5.2. In section 5.3 we discuss properties of the generalized measure (5.1.5). In section 5.4 we study its monotone behaviour with regard to parameters \( \alpha, \beta \) and \( R \). In section 5.5 we give a brief account of some existing mean codeword lengths. In section 5.6 we define a new generalized mean code length and study its bounds in terms of generalized R-norm information measure given by (5.1.5), we also discuss particular cases.
5.2 Characterization of Generalized R-Norm Information Measure of Type $\alpha$ and Degree $\beta$

The measure (5.1.5) can be characterize in many ways but in this section we shall discuss characterization by infimum method (refer to Hooda and Ram(1998)).

We can consider the generalized R-norm entropy (5.1.5) as weighted arithmetic mean representation of elementary R-norm entropies of type $\alpha$ and degree $\beta$.

**Theorem 5.2.1** Let

$$f_R^{\alpha,\beta}(\hat{p}_i) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \hat{p}_i \frac{R + \beta - 2\alpha}{R} \right], \quad \alpha \geq 1, 0 < \beta \leq 1, R(> 0) \neq 1, 0 < R + \beta \neq 2\alpha.$$  

(5.2.1)

then

$$H_R^{\alpha,\beta}(P) = \inf_{\hat{p}_i} \sum_{i=1}^{n} p_i f_R^{\alpha,\beta}(\hat{p}_i).$$  

(5.2.2)

Where the infimum operation is taken over the probability distribution $(\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n) \in \Delta_n$.

**Proof:** Let us consider

$$\sum_{i=1}^{n} p_i f_R^{\alpha,\beta}(\hat{p}_i) = \frac{R}{R + \beta - 2\alpha} \sum_{i=1}^{n} p_i \left[ 1 - \hat{p}_i \frac{R + \beta - 2\alpha}{R} \right].$$  

(5.2.3)

We minimize (5.2.3) subject to natural constraint

$$\sum_{i=1}^{n} \hat{p}_i = 1$$  

(5.2.4)

For this we consider Lagrangian method of multiplier.
\[ L = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \sum_{i=1}^{n} p_i (\hat{p}_i)^{\frac{R+\beta-2\alpha}{R}} \right] + \lambda \left( \sum_{i=1}^{n} \hat{p}_i - 1 \right) \]

Differentiating with respect to \( \hat{p}_i \), we have

\[ \frac{\partial L}{\partial \hat{p}_i} = -p_i (\hat{p}_i)^{\frac{\beta-2\alpha}{R}} + \lambda. \quad (5.2.5) \]

For extreme value we put \( \frac{\partial L}{\partial (\hat{p}_i)} = 0 \), which gives

\[ \hat{p}_i = \frac{R}{p_i^{\frac{2\alpha-\beta}{\beta-2\alpha}}} + \frac{R}{\alpha}. \quad (5.2.6) \]

We see that \( \frac{\partial^2 L}{\partial (\hat{p}_i)^2} > 0 \), when \( \hat{p}_i = \frac{R}{p_i^{\frac{2\alpha-\beta}{\beta-2\alpha}}} \). Hence the value of \( \hat{p}_i \) given by \( (5.2.4) \), we can find the value of \( \lambda \) and consequently, we have

\[ \hat{p}_i = \frac{R}{p_i^{\frac{2\alpha-\beta}{\beta-2\alpha}}}, \quad \alpha \geq 1, 0 < \beta \leq 1, R(> 0) \neq 1. \quad (5.2.7) \]

Now we consider R.H.S. of \( (5.2.2) \)

\[
\inf \sum_{i=1}^{n} p_i f_{R}^{q, \beta}(\hat{p}_i) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \sum_{i=1}^{n} \frac{R + \beta - 2\alpha}{R} \left( \frac{R}{\sum_{i=1}^{n} \hat{p}_i^{\frac{2\alpha-\beta}{\beta-2\alpha}}} \right) \right] \]

\[ = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \sum_{i=1}^{n} \frac{R}{\hat{p}_i^{\frac{2\alpha-\beta}{\beta-2\alpha}}} \right] \]
Further without any loss of generality, we may assume that corresponding to the observed probability distribution \( P \in \Delta_n \), there is a prior probability distribution \( Q \in \Delta \), and replacing \( f^{\alpha, \beta}_R(\hat{p}_i) \) by \( f^{\alpha, \beta}_R(q_i) \) in (5.2.2) we have

\[
H^{\alpha, \beta}_R(P \mid Q) = \inf_{q_i} \sum_{i=1}^{n} p_i f^{\alpha, \beta}_R(q_i)
\]  

(5.2.9)

In case we do not apply the operation of infimum to (5.2.9), then also it depends on two probability distributions \( P \) and \( Q \). For \( R=1 \) and \( \alpha=1 \), \( f^{\alpha, \beta}_R(q_i) \) is analogous to

\[
\frac{1}{\beta-1}(1-q_i^{-\beta-1})
\]

which reduces to \( \log \frac{1}{q_i} \) in case \( \beta \to 1 \), (5.2.9) becomes

\[
H^{\alpha, \beta}_R(P \mid Q) = \frac{1}{1-\beta} \sum_{i=1}^{n} p_i (q_i^{\beta-1} - 1),
\]

(5.2.10)

which is a generalized inaccuracy measure of degree \( \beta \) characterized by Sharma and Taneja(1975). Therefore, we can represent (5.2.2) via \( f^{\alpha, \beta}_R(q_i) \). Hence

\[
H^{\alpha, \beta}_R(P \mid Q) = \sum_{i=1}^{n} p_i f^{\alpha, \beta}_R(q_i)
\]

(5.2.11)

Actually, (5.2.11) can also be described as the average of elementary \( R \)-norm inaccuracies \( f^{\alpha, \beta}_R(q_i) \), \( i=1,2,\ldots,n \) and so can be called as \( R \)-norm inaccuracy.
measure of type $\alpha$ and degree $\beta$. Thus it seems possible that (5.2.11) may be characterized and then by taking its infimum we can arrive at (5.2.11).

In the next theorem we characterized the elementary information function $f^\alpha_\beta (q_i)$ by assuming only two axioms and applying infimum operation.

**Theorem 5.2.2** Let $f$ be a real valued continuous self-information function defined on (0,1] satisfying the following axioms:

**Axiom A$_1$.** $f(xy) = f(x) + f(y) - \frac{R + \beta - 2\alpha}{R} f(x)f(y)$

**Axiom A$_2$.**

$$f\left(\frac{1}{n}\right) = \frac{R}{R + \beta - 2\alpha} \left( 1 - n \frac{2\alpha - R - \beta}{R} \right), \quad \alpha \geq 1, \; 0 < \beta \leq 1, \; R(> 0) \neq 1, \; 0 < R + \beta \neq 2\alpha.$$ and $n = 2,3, \ldots, \ldots, \ldots$, is a maximally constant. If $f^\alpha_\beta (p_i)$ is replaced by $f^\alpha_\beta (q_i)$ in (5.2.8), the result holds.

**Proof:** By taking $f(x) = \frac{R}{R + \beta - 2\alpha} (1 - \phi(x))$ in axiom A$_1$, we get

$$\frac{R}{R + \beta - 2\alpha} (1 - \phi(xy)) = \frac{R}{R + \beta - 2\alpha} (1 - \phi(x)) + \frac{R}{R + \beta - 2\alpha} (1 - \phi(y)) - \frac{R}{R + \beta - 2\alpha} (1 - \phi(x))(1 - \phi(xy))$$

or

$$\phi(xy) = \phi(x) + \phi(y) \quad (5.2.12)$$

The relation (5.2.12) is well known Cauchy’s functional equation (refer Aczel (1996)). The continuous solution of (5.2.12) is $\phi(x) = x^\alpha$, where $\alpha \neq 0$ is an arbitrary constant.

On using axiom A$_2$, we get $a = \frac{R + \beta - 2\alpha}{R}$ and hence

$$f(x) = \frac{R}{R + \beta - 2\alpha} \left( 1 - x \frac{R + \beta - 2\alpha}{R} \right).$$
which is exactly of the form of (5.2.1). Next the measure (5.2.1) can be easily obtained by applying the operation infimum on the equation (5.2.11) on the lines of theorem 5.2.1

**Remarks:** For an incomplete probability distribution scheme

\[ P = (p_1, p_2, \ldots, p_n) \quad p_i \geq 0, \sum_{i=1}^{n} p_i \leq 1, i = 1, 2, \ldots, n, \]

associated with individual events may be worked out. Then as in case of (5.2.11) we may define

\[ H^{\alpha, \beta}_{R}(P / Q) = \left( p_1, p_2, \ldots, p_n; q_1, q_2, \ldots \right) \]

By using the operation infimum with respect to \( q_i \)'s the equation (5.2.13) gives

\[ H^{\alpha, \beta}_{R}(P / Q) = \frac{1}{\sum_{i=1}^{n} p_i} \sum_{i=1}^{n} p_i f^{\alpha, \beta}_{R}(q_i) \]

\[ (5.2.13) \]

which is the R-norm entropy of type \( \alpha \) and degree \( \beta \) of incomplete probability distribution. It is also worth mentioning that if we take arithmetic average with weights as continuous function \( w(.) \), then we get the general expression

\[ H^{\alpha, \beta}_{R}(P / Q) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left( \sum_{i=1}^{n} p_i \frac{R}{2\alpha - \beta} \right)^{\frac{2\alpha - \beta}{R}} \right] \]

\[ (5.2.14) \]

On considering different weight \( w(.) \) satisfying the condition \( w(pq) = w(p)w(q), \) where \( w(.) \neq 0 \), we can obtain various generalized R-norm entropies by using the operation infimum with respect to \( q_i \)'s.
5.3 Important Properties of the New Generalized Measure

The generalized information measure of type $\alpha$ and degree $\beta$ given by (5.1.5) satisfy the algebraic and differential properties as given below:

5.3.1 Algebraic Properties

(i) $H_R^{\alpha,\beta}(p_1, p_2, \ldots, p_n)$ is a symmetric function of $(p_1, p_2, \ldots, p_n)$.

(ii) $H_R^{\alpha,\beta}(P)$ is expansible i.e. $H_{n-1,R}(p_1, p_2, \ldots, p_n, 0) = H_R^{\alpha,\beta}(p_1, p_2, \ldots, p_n)$.

(iii) $H_R^{\alpha,\beta}(P)$ is decisive : $H_R^{\alpha,\beta}(1,0) = H_R^{\alpha,\beta}(0,1) = 0$.

(iv) $H_R^{\alpha,\beta}(P)$ is non-recursive.

(v) $H_R^{\alpha,\beta}(PQ) = H_R^{\alpha,\beta}(P) + H_R^{\alpha,\beta}(Q) - \frac{R + \beta - 2\alpha}{R} H_R^{\alpha}(P)H_R^{\alpha}(Q)$ i.e. $H_R^{\alpha}(P)$ is non-additive.

Proof: Properties (i) to (iii) can be verified easily. However, for (iv) we consider

\[
H_R^{\alpha,\beta}\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left(\frac{\frac{R}{p_1^{2\alpha-\beta}} + \frac{R}{p_2^{2\alpha-\beta}}}{p_1 + p_2}\right)^{\frac{2\alpha-\beta}{R}}\right]
\]

(5.3.1)

and

\[
H_R^{\alpha,\beta}(p_1 + p_2, p_3, \ldots, p_n) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left(\frac{R}{(p_1 + p_2)^{2\alpha-\beta}} + \frac{R}{p_3^{2\alpha-\beta}} + \ldots + \frac{R}{p_n^{2\alpha-\beta}}\right)^{\frac{2\alpha-\beta}{R}}\right]
\]

(5.3.2)

By combining (5.3.1) and (5.3.2), we have
Thus $H_R^{\alpha, \beta}(P)$ is non-recursive. Next we prove the property (v).

Let $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots, B_m$ be two sets of events associated with probability distribution $P \in \Delta$, and $Q \in \Delta_m$. We denote the probability of the joint occurrence of events $A_i = (i = 1, 2, 3, \ldots, n)$ and $B_j = (j = 1, 2, 3, \ldots, m)$ as $p(A_i \cap B_j)$. Then the generalized R-norm entropy of type $\alpha$ and degree $\beta$ is given by

$$H_R^{\alpha, \beta}(P \ast Q) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left( \sum_{i=1}^{n} p^{2\alpha - \beta} (A_i) \right) \left( \sum_{j=1}^{m} p^{2\alpha - \beta} (B_j) \right) \right].$$

Since the events are stochastically independent, therefore, we have

$$H_R^{\alpha, \beta}(P \ast Q) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left( \sum_{i=1}^{n} p^{2\alpha - \beta} (A_i) \right) \left( \sum_{j=1}^{m} p^{2\alpha - \beta} (B_j) \right) \right].$$

$$H_R^{\alpha, \beta}(P \ast Q) = \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left( \sum_{i=1}^{n} p^{2\alpha - \beta} (A_i) \right) \left( \sum_{j=1}^{m} p^{2\alpha - \beta} (B_j) \right) \right].$$

$$= \frac{R}{R + \beta - 2\alpha} - \frac{R}{R + \beta - 2\alpha} \left[ 1 - \frac{R + \beta - 2\alpha}{R} H_R^{\alpha, \beta}(P) \right] \left[ 1 - \frac{R + \beta - 2\alpha}{R} H_R^{\alpha, \beta}(Q) \right].$$

$$= H_R^{\alpha, \beta}(P) + H_R^{\alpha, \beta}(Q) - \frac{R + \beta - 2\alpha}{R} H_R^{\alpha, \beta}(P) H_R^{\alpha, \beta}(Q)$$

(5.3.3)

This completes the proof of property (v).
5.3.2 Differential Properties

(i) $H_{R}^{\alpha,\beta}(P)$ is non-negative.

(ii) $H_{R}^{\alpha,\beta}(P) \geq H_{R}^{\alpha,\beta}(1,0,\ldots,0)$.

(iii) $H_{R}^{\alpha,\beta}(P) \leq H_{R}^{\alpha,\beta}\left(\frac{1}{n},\ldots,\frac{1}{n}\right) = \frac{R}{R + \beta - 2\alpha} \left[1 - n \frac{2\alpha - \beta - R}{R}\right]$. 

(iv) $H_{R}^{\alpha,\beta}(P)$ is continuous at $R \in \mathbb{R}^+$. 

(v) $H_{R}^{\alpha,\beta}(P)$ is stable in $p_i$, $i=1,2,\ldots,n$.

(vi) $H_{R}^{\alpha,\beta}(P)$ is small for small probabilities.

(vii) $H_{R}^{\alpha,\beta}(P)$ is concave function for all $P \in \Delta$.

(viii) $\lim_{R \to \infty} H_{R}^{\alpha,\beta}(P) = 1 - \max_{i} p_i$.

Proof: Properties (i) to (vi) can be verified easily. However we prove that the generalized entropy is concave function of $P$.

Definition 5.3.1 A function $f$ over a set $S$ is said to be concave, if for all choice of $x_1, x_2, \ldots, x_m \in S$ and for all scalars $\lambda_1, \lambda_2, \ldots, \lambda_m$, such that $\sum_{i=1}^{m} \lambda_i = 1$, the following holds:

$$f \left( \sum_{i=1}^{m} \lambda_i x_i \right) \geq \sum_{i=1}^{m} \lambda_i f(x_i). \quad (5.3.4)$$

We consider random variable $x$ taking its values in the set $S$ such that $x_1, x_2, \ldots, x_m \in S$ and probability distributions over $S$ as follows:

$$P_k(x) = \{ p_k(x_1), \ldots, p_k(x_n) \} , \quad p_k(x_i) \geq 0 \text{ and } \sum_{i=1}^{n} p_k(x_i) = 1, \quad k = 1, 2, \ldots, r.$$ 

Let us define another probability distribution over $S$: 

99
\[ P_0(x) = \{ p_0(x_1), \ldots, p_0(x_n) \} , \text{ such that for all } i \text{'s } p_0(x_i) = \sum_{k=1}^{r} \lambda_k p_k(x_i) \text{ where} \]

\[ \lambda_k 's \text{ are non-negative scalars satisfying } p_0(x) = \sum_{k=1}^{r} \lambda_k = 1. \]

Then \( H^{\alpha, \beta}_R(P) \) is concave if \( D < 0 \), where

\[ D = \sum_{k=1}^{r} \lambda_k H^{\alpha, \beta}_R(P_k) - H^{\alpha, \beta}_R(P_0), R(>0) \neq 1 \text{ and } 0 < \beta \leq 1. \]

So let us consider

\[ D = \sum_{k=1}^{r} \lambda_k \left[ H^{\alpha, \beta}_R(P_k) - H^{\alpha, \beta}_R(P_0) \right] \]

\[ = \sum_{k=1}^{r} \lambda_k \left[ 1 - \left( \sum_{i=1}^{n} p_k^{2\alpha-\beta}(x_i) \right) \frac{2\alpha-\beta}{x} \right] \frac{R}{R + \beta - 2\alpha} - \frac{R}{R + \beta - 2\alpha} \left[ 1 - \left( \sum_{i=1}^{n} p_0^{2\alpha-\beta}(x_i) \right) \frac{2\alpha-\beta}{R} \right] \]

\[ = \frac{R}{R + \beta - 2\alpha} \left[ \sum_{i=1}^{n} \left( \sum_{k=1}^{r} \lambda_k p_k(x_i) \right) \frac{R}{2\alpha-\beta} \right] + \sum_{i=1}^{n} \left( \sum_{k=1}^{r} \lambda_k p_k(x_i) \right) \frac{R}{2\alpha-\beta} \]

\[ \frac{R}{R + \beta - 2\alpha} \]

Now using the inequality \( \left( \sum_{i=1}^{r} a_i x_k^i \right) \left( \sum_{i=1}^{n} a_i x_k^i \right) \) according as \( \left( \sum_{i=1}^{r} a_i x_k^i \right) < 1 \), we have

\[ \left( \sum_{k=1}^{r} \lambda_k p_k(x_i) \right) \frac{R}{2\alpha-\beta} \left( \sum_{i=1}^{n} \lambda_k p_k(x_i) \right) \left( \sum_{i=1}^{n} \lambda_k p_k(x_i) \right) \text{ according as } \frac{R}{2\alpha-\beta} \left[ \sum_{i=1}^{n} \lambda_k p_k(x_i) \right] < 1 \]

Therefore,

\[ \sum_{i=1}^{n} \left( \sum_{k=1}^{r} \lambda_k p_k(x_i) \right) \frac{R}{2\alpha-\beta} \left( \sum_{i=1}^{n} \lambda_k p_k(x_i) \right) \left( \sum_{i=1}^{n} \lambda_k p_k(x_i) \right) \text{ according as } \frac{R}{2\alpha-\beta} \left[ \sum_{i=1}^{n} \lambda_k p_k(x_i) \right] < 1 \]
\[ D_1 = \left( \sum_{i=1}^{n} \left( \sum_{k=1}^{r} \lambda_k p_k(x_i) \right) \right)^{\frac{2\alpha - \beta}{R}} \leq \left( \sum_{k=1}^{n} p_k \right)^{\frac{2\alpha - \beta}{R}} \]  
\[ \text{according as } \frac{R}{2\alpha - \beta} \begin{cases} < 1 & \text{for } < \frac{R}{2\alpha - \beta} \end{cases} \]  

Moreover
\[ \left( \sum_{k=1}^{r} \sum_{i=1}^{n} p_k^{\frac{2\alpha - \beta}{R}} (x_i) \right)^{\frac{2\alpha - \beta}{R}} \leq \left( \sum_{i=1}^{n} \sum_{k=1}^{r} \lambda_k p_k(x_i) \right)^{\frac{2\alpha - \beta}{R}} = D_2 \]  
\[ \text{according as } \frac{R}{2\alpha - \beta} \begin{cases} < 1 & \text{for } < \frac{R}{2\alpha - \beta} \end{cases} \]  

Thus, \( D_1 \geq D_2 \) according as \( \frac{R}{2\alpha - \beta} < 1 \), which implies
\[ D < 0 \text{ in the view of the sign of } \frac{R}{R + \beta - 2\alpha} \]  
according as \( \frac{R}{2\alpha - \beta} > 1 \).

Thus we have proved that \( H_{\alpha,\beta}^R(P) \) is a concave function of \( P \).

In particular we set \( \max p_i = p_k \). Assuming \( n_0 = 1, 2, \ldots; R + \beta > 2\alpha \) and \( 0 < \beta \leq 1 \),

Then we find
\[ \left( \sum_{i=1}^{n} \frac{R}{2\alpha - \beta} \right)^{\frac{2\alpha - \beta}{R}} \leq \left( n_0 \frac{R}{2\alpha - \beta} \right)^{\frac{2\alpha - \beta}{R}} = n_0 \frac{R}{2\alpha - \beta} p_k. \]  
\[ \text{It may be also noted that for } R + \beta > 2\alpha \]
\[ \left( \sum_{i=1}^{n} \frac{R}{2\alpha - \beta} \right)^{\frac{2\alpha - \beta}{R}} \geq p_k \]  
\[ \text{Combining (5.3.8) and (5.3.9), we get} \]
\[ p_k \leq \left( \sum_{i=1}^{n} \frac{R}{2\alpha - \beta} \right)^{\frac{2\alpha - \beta}{R}} \leq n_0 \frac{R}{2\alpha - \beta} p_k \]
Taking limit for \( R \to \infty \) in (5.3.10), we have

\[
\lim_{R \to \infty} \left( \sum_{i=1}^{n_i} p_i^\frac{R}{2^{\alpha-\beta}} \right)^\frac{2^{\alpha-\beta}}{R} = p_k = \max_i p_i
\]

and finally

\[
\lim_{R \to \infty} H^{\alpha,\beta}_R(P) = \lim_{R \to \infty} \left[ 1 - \sum_{i=1}^{n_i} p_i^\frac{R}{2^{\alpha-\beta}} \right] = 1 - \max_i p_i.
\]

This proves the property (viii).

5.4 Monotone Behaviour of Generalized \( R \)-Norm Measure

In this section we study the monotone behaviour of generalized information measure given by (5.1.5) with regards to parameters.

Let \( P=\{0.2,0.2, 0.2,0.1 ,0.3\} \) be a set of probabilities.

Assuming \( R=0.1 \) and \( \alpha = 2 \). We tabulate the values of \( H^{\alpha,\beta}_R(P) \) for different values of \( \beta \) as given in the following table:

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H^{1.5,\beta}_{0.6}(P) )</td>
<td>57.3007</td>
<td>46.9200</td>
<td>8.4468</td>
<td>31.5267</td>
<td>25.8721</td>
<td>21.2488</td>
<td>17.4663</td>
<td>14.3699</td>
<td>11.8332</td>
</tr>
</tbody>
</table>

Next we draw the graph of the table (5.4.1) and illustrate from figure (5.4.1) that \( H^{\alpha,\beta}_R(P) \) is monotonic decreasing with increasing values of \( \beta \).

Let \( P=\{0.2,0.2, 0.2,0.1 ,0.3\} \) be a set of probabilities.

Assuming \( R=0.1 \) and \( \beta = 0.1 \). We tabulate the values of \( H^{\alpha,\beta}_R(P) \) for different values of \( \alpha \) as given in the following table:

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H^{0.5,\beta}_{0.6}(P) )</td>
<td>57.3007</td>
<td>46.9200</td>
<td>8.4468</td>
<td>31.5267</td>
<td>25.8721</td>
<td>21.2488</td>
<td>17.4663</td>
<td>14.3699</td>
<td>11.8332</td>
</tr>
</tbody>
</table>

102
Table 5.4.2: Monotone behaviour of \( H_{R}^{\alpha,\beta}(P) \) with respect to \( \alpha \)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>1.0</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
<th>1.7</th>
<th>1.8</th>
<th>1.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_{0.6}^{\alpha,0.1}(P) )</td>
<td>1.7986</td>
<td>2.5982</td>
<td>3.7702</td>
<td>5.4961</td>
<td>8.0478</td>
<td>11.8332</td>
<td>17.4663</td>
<td>25.8721</td>
<td>38.4468</td>
</tr>
</tbody>
</table>

Next we draw the graph of the table (5.4.2) and illustrate from figure (5.4.2) that \( H_{R}^{\alpha,\beta}(P) \) is monotonic increasing with increasing values of \( \alpha \).

Let \( P=\{0.2,0.2,0.2,0.1,0.3\} \) be a set of probabilities.

Assuming \( \alpha=2 \) and \( \beta=0.1 \). We tabulate the values of \( H_{R}^{\alpha,\beta}(P) \) for different values of \( R \) as given in the following table:

Table 5.4.3: Monotone behaviour of \( H_{R}^{\alpha,\beta}(P) \) with respect to \( R \)

<table>
<thead>
<tr>
<th>( R )</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_{R}^{2,0.1}(P) )</td>
<td>436.6944</td>
<td>3.4336</td>
<td>0.7724</td>
<td>0.5634</td>
<td>0.2178</td>
<td>0.1411</td>
<td>0.0902</td>
<td>0.0510</td>
<td>0.0177</td>
</tr>
</tbody>
</table>

Next we draw the graph of the table (5.4.3) and illustrate from figure (5.4.3) that \( H_{R}^{\alpha,\beta}(P) \) is monotonic decreasing with increasing values of \( R \).
Figure (5.4.1): Monotone behaviour of $H_R^{\alpha,\beta}(P)$ with respect to $\beta$

Figure (5.4.2): Monotone behaviour of $H_R^{\alpha,\beta}(P)$ with respect to $\alpha$
5.5 A New Generalized Mean Codeword Length

Let a finite set of n source symbols $X = \{x_1, x_2, \ldots, x_n\}$ with probabilities $P = \{p_1, p_2, p_3, \ldots, p_n\}$ be encoded using D code alphabets, then there is a uniquely decipherable instantaneous code with lengths $l_1, l_2, \ldots, l_n$ if and only if

$$\sum_{i=1}^{n} D^{-l_i} \leq 1$$

(5.5.1)

(5.5.1) is known as Kraft’s inequality.

Let $L = \sum_{i=1}^{n} p_i / l_i \log D$

(5.5.2)

be the average codeword length associated with input symbols $(x_1, x_2, \ldots, x_n)$.  

Figure (5.4.3): Monotone behaviour of $H_{R}^{\alpha, \beta}(P)$ with respect to $R$
Under the Kraft’s inequality Shannon’s (1948) proved the following result for a noiseless channel:

\[ H(P) \leq L < H(P) + \log D, \, D \geq 2 \tag{5.5.3} \]

with equality if and only if \( l_i = -\log D p_i \).

Further, it was proved that mean codeword length can be arbitrary close to \( H(P) \) by suitably encoding the message in blocks. Campbell (1965) introduced the exponential mean codeword length which is given by

\[ L(t) = \frac{1}{t} \log \left\{ \sum_{i=1}^{n} p_i D_i^{l_i} \right\}, \quad -1 < t < \infty, \tag{5.5.4} \]

and proved the following theorem:

\[ H^\alpha(P) \leq L(t) \leq H^\alpha(P) + 1 \tag{5.5.5} \]

under Kraft inequality \( \sum_{i=1}^{n} D_i^{-l_i} \leq 1 \),

where \( H_\alpha(P) \) is Renyi’s entropy of of order \( \alpha = \frac{1}{1+t} \) and \( l_i \) is the codeword length corresponding to source symbol \( x_i \).

Kiffer (1979) defined a class of decision rules and showed \( H_\alpha(P) \) is the best decision rule for deciding which of the two sources can be coded with expected cost of sequences of length \( l \) when \( l \to \infty \), where the cost of the encoding a sequence is assumed to be a function of codeword length only. Further Jelinek (1980) showed that coding with respect to \( L(t) \) is useful in minimizing the problem of buffer overflow which occurs when the source symbols are being produced at a fixed rate and the code words are stored temporarily in a finite buffer.

Hooda and Bhaker (1997) considered the following generalization of (5.5.4):

\[ L^\beta(t) = \frac{1}{t} \log_D \left\{ \frac{\sum_{i=1}^{n} p_i^\beta D_i^{l_i\beta}}{\sum_{i=1}^{n} p_i^\beta} \right\}, \quad \beta \geq 1 \quad \text{and} \quad -1 < t < \infty \tag{5.5.6} \]

and proved
under the condition
\[
\sum_{i=1}^{n} p_i^{\beta - 1} D^{-i} \leq \sum_{i=1}^{n} p_i^\beta , \tag{5.5.8}
\]
where \( H_\alpha^\beta (P) \) is generalized entropy of order \( \alpha = \frac{1}{1+t} \) and type \( \beta \) given by Aczel and Daroczy (1963) and Kapur (1967). It may be seen that the mean codeword length (5.5.2) had been generalized parametrically and their bounds have been studied in terms of generalized measures of entropies.

Since (5.5.8) reduces to Kraft’s inequality when \( \beta = 1 \), therefore, it is called generalized Kraft inequality and codes obtained under this generalized inequality are called personal codes.

Further Hooda (2001) generalized (5.5.2) by
\[
L \left( \frac{R}{2 - \beta} \right) = \frac{R}{R + \beta - 2} \left( 1 - \sum_{i=1}^{n} p_i D^{-p_i^{(R+\beta-2)}} \right), 0 < \beta \leq 1. \tag{5.5.9}
\]
which reduces to
\[
L(R) = \frac{R}{R-1} \left( 1 - \sum_{i=1}^{n} p_i D^{-p_i^{(R-1)}} \right), \text{ when } \beta = 1. \tag{5.5.10}
\]
(5.5.10) is average codeword length due to Boekee and Lubbe (1965) and it reduces to (5.5.2) when \( R \rightarrow 1 \). Hooda (2001) also studied the bounds of (5.5.9) in terms of the generalized \( R \)-norm entropy of type \( \beta \) given by (5.1.2).

We see that (5.5.9) can be generalized parametrically in many ways. However, we present one generalization given as below and study its bounds:
\[
L \left( \frac{R}{2\alpha - \beta} \right) = \frac{R}{R + \beta - 2\alpha} \left( 1 - \sum_{i=1}^{n} p_i D^{-p_i^{(R+\beta-2\alpha)}/R} \right) \tag{5.5.11}
\]
It may be noted that if \( \alpha = 1 \), (5.5.11) reduces to (5.5.9) and further (5.5.9) reduces to (5.5.2) when \( \beta = 1 \) and \( R \rightarrow 1 \).
5.6 A Coding Theorem on Bounds

In this section we prove coding theorem on the bounds of \( L\left(\frac{R}{2\alpha - \beta}\right) \) in terms of generalized R-norm information measure given by (5.1.5).

**Theorem 5.6.1.** If \( l_i, i = 1, 2, \ldots, n \) is the length of codeword \( x_i \) satisfying (5.5.11), then

\[
H_{R}^{\alpha, \beta}(P) \leq L\left(\frac{R}{2\alpha - \beta}\right) < D \frac{(2\alpha - \beta - R)}{R} \left[ H_{R}^{\alpha, \beta}(P) + \frac{R}{R + \beta - 2\alpha} \right] \left[ 1 - D \frac{(2\alpha - \beta - R)}{R} \right]
\]

(5.6.1)

and the sign of equality holds iff

\[
D^{-1} = \left\{ \frac{R(2\alpha - \beta)}{n} \right\}_{i=1}^{n} \sum_{i=1}^{n} p_i \]

(5.6.2)

where \( H_{R}^{\alpha, \beta}(P) \) is given by (5.1.5)

**Proof:** First of all we shall prove the lower bound of \( L\left(\frac{R}{2\alpha - \beta}\right) \).

From Holder’s inequality we know that

\[
\left( \sum_{i=1}^{n} x_i^{p} \right)^{\frac{1}{p}} \left( \sum_{i=1}^{n} y_i^{q} \right)^{\frac{1}{q}} \leq \sum_{i=1}^{n} x_i y_i
\]

(5.6.3)

where \( x_i, y_i > 0 \) for each \( i \) and \( \frac{1}{p} + \frac{1}{q} = 1 \).

Putting \( x_i = p_i^{\frac{R}{R + \beta - 2\alpha}} D^{-1}, y_i = p_i^{\frac{R}{2\alpha - R - \beta}} \) \(, p = \frac{(R + \beta - 2\alpha)}{R} \) and \( q = \frac{(2\alpha - R - \beta)}{(2\alpha - \beta)} \) in (5.6.3), we have
\[
\sum_{i=1}^{n} \left( p_i \frac{R^{(R+\beta-2\alpha)}}{R^{R+\beta-2\alpha}} \right) \leq \sum_{i=1}^{n} p_i \frac{R^{2\alpha-\beta}}{R^{2\alpha-R-\beta}} = \sum_{i=1}^{n} q_i = 1.
\]

It implies that
\[
\sum_{i=1}^{n} p_i D^{-\frac{R^{(R+\beta-2\alpha)}}{R^{R+\beta-2\alpha}}} \leq \left[ \sum_{i=1}^{n} p_i \frac{R^{2\alpha-\beta}}{R^{2\alpha-R-\beta}} \right]^{\beta-2\alpha} \tag{5.6.4}
\]

Now two cases arises

**Case (i)** Raising power \( \frac{R+\beta-2\alpha}{R} > 0 \) to both sides of (5.6.4), we have
\[
\sum_{i=1}^{n} p_i D^{-\frac{R^{(R+\beta-2\alpha)}}{R^{R+\beta-2\alpha}}} \geq \left[ \sum_{i=1}^{n} p_i \frac{R^{2\alpha-\beta}}{R^{2\alpha-R-\beta}} \right]^{\beta-2\alpha}
\]

or
\[
1 - \left[ \sum_{i=1}^{n} p_i \frac{R^{2\alpha-\beta}}{R^{2\alpha-R-\beta}} \right]^{R^{\frac{R+\beta-2\alpha}{R}}} \geq 1 - \sum_{i=1}^{n} p_i D^{-\frac{R^{(R+\beta-2\alpha)}}{R^{R+\beta-2\alpha}}} \tag{5.6.5}
\]

Multiplying (5.6.5) by \( \frac{R}{R+\beta-2\alpha} < 0 \) throughout, we get
\[
H_{R}^{\alpha,\beta}(P) \leq L\left( \frac{R}{2\alpha-\beta} \right) \tag{5.6.6}
\]

**Case (ii)** Again raising power \( \frac{R+\beta-2\alpha}{R} < 0 \) to both sides of (5.6.4), we get
\[
\sum_{i=1}^{n} p_i D^{-\frac{R^{(R+\beta-2\alpha)}}{R^{R+\beta-2\alpha}}} \leq \left[ \sum_{i=1}^{n} p_i \frac{R^{2\alpha-\beta}}{R^{2\alpha-R-\beta}} \right]^{\beta-2\alpha}
\]

or
\[
1 - \left[ \sum_{i=1}^{n} p_i \frac{R^{2\alpha-\beta}}{R^{2\alpha-R-\beta}} \right]^{\frac{R}{R^{\frac{R+\beta-2\alpha}{R}}}} \leq 1 - \sum_{i=1}^{n} p_i D^{-\frac{R^{(R+\beta-2\alpha)}}{R^{R+\beta-2\alpha}}} \tag{5.6.7}
\]
Multiplying (5.6.7) by \( \frac{R}{R + \beta - 2\alpha} > 0 \) throughout, we have

\[
H_R^{\alpha, \beta}(P) \leq L\left( \frac{R}{2\alpha - \beta} \right) \tag{5.6.8}
\]

If \( D^{-i} = \frac{p_i^{2\alpha - \beta}}{R} \), \( i = 1, 2, \ldots, n \) \( \sum_{i=1}^{n} p_i^{2\alpha - \beta} \)

Then equality sign holds in (5.6.8).

Next we prove the inequality (5.6.1) for upper bound of \( L\left( \frac{R}{2\alpha - \beta} \right) \). Equation (5.6.9) can be written as

\[
l_i = -\log_D p_i^{\frac{R}{2\alpha - \beta}} + \log_D \left( \sum_{i=1}^{n} p_i^{\frac{R}{2\alpha - \beta}} \right), \quad i = 1, 2, \ldots, n.
\]

By choosing \( l_i \) such that

\[
-\log_D p_i^{\frac{R}{2\alpha - \beta}} + \log_D \left( \sum_{i=1}^{n} p_i^{\frac{R}{2\alpha - \beta}} \right) \leq l_i \leq -\log_D p_i^{\frac{R}{2\alpha - \beta}} + \log_D \left( \sum_{i=1}^{n} p_i^{\frac{R}{2\alpha - \beta}} \right) + 1.
\]

We obtain the following inequality

\[
D^{-i} > \frac{p_i^{2\alpha - \beta}}{R} D \sum_{i=1}^{n} p_i^{\frac{R}{2\alpha - \beta}}. \tag{5.6.10}
\]

Here also two cases arise

(i) When \( R + \beta > 2\alpha \) then \( \frac{R + \beta - 2\alpha}{R} > 0 \). Now raising the power \( \frac{R + \beta - 2\alpha}{R} \) to both sides of (5.6.10) and summing over \( i \) we have

\[
\sum_{i=1}^{n} p_i D^{-i} \left( \frac{R + \beta - 2\alpha}{R} \right) > \left( \sum_{i=1}^{n} p_i^{\frac{R}{2\alpha - \beta}} \right) \frac{2\alpha - \beta}{R} \frac{D}{\left( \frac{2\alpha - \beta - R}{R} \right)}
\]

or

110
\[ 1 - \sum_{i=1}^{n} p_i D^{\frac{(R+\beta-2\alpha)}{R}} < 1 - \left( \sum_{i=1}^{n} p_i \frac{R}{2\alpha-\beta} \right) D^{\frac{(2\alpha-\beta-R)}{R}} \]  

(5.6.11)

On multiplying (5.6.11) by \( \frac{R}{R + \beta - 2\alpha} \) and simplifying, we get

\[ L \left( \frac{R}{2\alpha-\beta} \right) < H_R^{\alpha,\beta}(P)D^{\frac{2\alpha-R-\beta}{R}} \frac{R}{R + \beta - 2\alpha} \left( 1 - D^{\frac{(2\alpha-R-\beta)}{R}} \right), \]  

(5.6.12)

(ii) When \( R + \beta < 2\alpha \), we can show on the same lines that right hand inequality of (5.6.1) holds.

Hence the proof of Theorem 5.6.1 completes.

**Particular Cases**

(a) If \( \alpha = 1, \beta = 1 \), (5.6.1) reduces to

\[ H_R(P) \leq L(R) < D^{\frac{(1-R)}{R}} H_R(P) + \frac{R}{R-1} \left[ 1 - D^{\frac{(1-R)}{R}} \right], \]  

(5.6.13)

which is the result due to Boekee and Lubee (1965).

(b) When \( R \to 1 \) in (5.6.1), we get

\[ \frac{H(P)}{\log D} \leq \overline{L} < \frac{H(p)}{\log D} + 1, \text{ where } \overline{L} = \sum_{i=1}^{n} p_i l_i, \]  

(5.6.14)

which is the well known result given by Shannon (1948).

**5.7 Conclusion**

We know that optimal code is that code for which the value \( L \left( \frac{R}{2\alpha-\beta} \right) \) is equal to its lower bound. From the result of the theorem 5.4.1, it can be seen that the mean codeword length of the optimal code is dependent on three parameters \( R, \alpha \) and \( \beta \), while in the case of Shannon’s theorem it does not depend on any parameter. So it can be reduced significantly by taking suitable values of parameters. We have also
discussed the properties of generalized measure and studied its monotonic behaviour. A new mean code word length has been defined and its bounds in term of the generalized R-norm information measure of type $\alpha$ and degree $\beta$ have been studied.