CHAPTER-II

NEW MEASURES OF ENTROPY BASED UPON
STATISTICAL CONSTANTS, SAMPLING AND PROBABILITY DISTRIBUTIONS

2.1. Introduction:

The most successful measure of information is certainly Shannon’s (1948) entropy, which is interpreted as the uncertainty content of a random experiment ruled by the probability distribution \( P \). The “universe of discourse”, that is, the set of all possible outcomes of the experiment, called the sample space, or the alphabet, is assumed to be finite. In fact, one can approach Shannon’s (1948) measure of entropy with two different approaches. The first approach, which is preferred by coding theorists, is the approach under which, Shannon entropy is the solution to a coding-theoretic problem, and represents the ideal rate of compression. The second one is the axiomatic approach which is more relevant to the point of view of information measures in the strict sense. In this case, one lists properties, or axioms, which an adequate information measure should possess. It was an American Mathematician C.E. Shannon who in 1948, remarked that since uncertainty is always associated with every probability distribution \( P = (p_1, p_2, \ldots, p_n) \), the uncertainty measure must be a function of probabilities and with certain desirable postulates, investigated and found that the only function which satisfies these postulates is given by

\[
H(P) = -\sum_{i=1}^{n} p_i \ln p_i
\]  

(2.1.1)

He called the measure (2.1.1) as entropy and studied its many interesting properties which, later on proved to be extremely useful in many disciplines of Mathematics, Statistics, Operations Research and Computer Science.
After the invention of Shannon’s (1948) measure of entropy, many scientists became interested in the field. Renyi (1961) introduced entropy of order $\alpha$, given by the following expression:

$$H_{\alpha}(P) = \frac{1}{1-\alpha} \ln \left( \sum_{i=1}^{n} p_i^\alpha \right) \left/ \sum_{i=1}^{n} p_i \right., \quad \alpha \neq 1, \alpha > 0$$  \hspace{1cm} (2.1.2)

Havrada and Charvat (1967) introduced another measure called the non-additive measure of entropy, given by:

$$H^\alpha(P) = \frac{1}{1-\alpha} \left( \sum_{i=1}^{n} p_i^\alpha - 1 \right), \quad \alpha \neq 1, \alpha > 0$$  \hspace{1cm} (2.1.3)

Burg (1972) developed his non-parametric measure of entropy, given by

$$H^1(P) = \sum_{i=1}^{n} \log p_i$$  \hspace{1cm} (2.1.4)

Rao et al. (2004) used the cumulative distribution of a random variable to define its information content and thereby developed an alternative measure of uncertainty that extends Shannon’s (1948) entropy to random variables with continuous distributions. They called this measure of entropy as cumulative residual entropy (CRE) and explained the salient features of CRE are:

1. it possesses more general mathematical properties than the Shannon entropy, and
2. it can be easily computed from sample data and these computations asymptotically converge to the true values.

The properties of CRE and a precise formula relating CRE and Shannon entropy have been provided by the authors and finally, they have given the applications of CRE to reliability engineering and computer vision.

Chakrabarti (2005) presented a new axiomatic derivation of Shannon’s (1948) entropy for discrete probability distribution on the basis of postulates of additivity and concavity of the entropy function. Then they modified Shannon’s entropy to account for observational uncertainty. The modified entropy reduced, in the
limiting case, to the form of Shannon differential entropy. As an application, they derived the expression for classical entropy of statistical mechanics from the quantized form of the entropy. It was Schroeder (2004), who pointed out that Shannon entropy, while useful in communications theory, is conceptually inadequate as a measure of information.

Herremoes and Vignat (2003) extended the classical entropy power inequality for continuous random variables to the family of binomial random variables with parameter half. Asadi et al. (2004) have given the applications of dynamic entropy and consequently provided maximum dynamic entropy models for several well-known distributions. Ang and Jowitt (2005) considered the entropy used in the analysis of water distribution networks and proposed a new approach to calculate its value. The author first recalled the definition of this entropy and then switched to the new method for its calculations, based on a concept of path entropy. The method is described for branching-trees networks and then it is shown that it provides a direct way to calculate the maximum entropy value of any given single-source water distribution network topology, nodal demands and flow directions for every pipe.


There exist many well-known measures of information which are frequently used by the researchers working in Biological Sciences for measuring diversity and equitability of different communities. Some of these measures are due to Shannon (1948), Renyi (1961), Simpson (1949), Weiner (1949) etc. Of course Shannon’s (1948) measure is most widely applicable and possesses many interesting and desirable properties. But, it has a limitation that it deals with exponential families only whereas in actual practice, there are many distributions which are non-exponential, there are many growth curves which do not follow
exponential law. Thus there is a need for developing new measures to extend the scope of their applications. Such measures have been developed in this chapter.

The theory of fuzzy sets introduced by Zadeh (1965) received a good response and after its introduction, many researchers started working around this field. Thus, keeping in view the idea of fuzzy sets, De Luca and Termini (1972) introduced a fuzzy entropy corresponding to Shannon’s (1948) entropy, given by

\[
H(A) = -\sum_{i=1}^{n} \left[ \mu_A(x_i) \log \mu_A(x_i) + (1 - \mu_A(x_i)) \log (1 - \mu_A(x_i)) \right]
\] (2.1.5)

After this development, a large number of measures of fuzzy entropy were discussed, characterized and generalized by various authors. Kapur (1997) introduced the following measure of fuzzy entropy:

\[
H_{\alpha, \beta}(A) = \frac{1}{\beta - \alpha} \log \left( \sum_{i=1}^{n} \left\{ \mu_A^{\alpha}(x_i) + (1 - \mu_A(x_i))^\alpha \right\} \right), \alpha \geq 1, \beta \leq 1
\] (2.1.6)

Parkash and Sharma (2004) introduced two measures of fuzzy entropy, given by

\[
K_a(A) = \sum_{i=1}^{n} \left[ \log (1 + a \mu_A(x_i)) + \log (1 + a(1 - \mu_A(x_i))) - \log (1 + a) \right], a \geq 0
\] (2.1.7)

\[
H_a(A) = -\frac{1}{a} \sum_{i=1}^{n} \left[ \left( 1 + a \mu_A(x_i) \right) \log (1 + a \mu_A(x_i)) + \left( 1 + a(1 - \mu_A(x_i)) \right) \log \left( 1 + a(1 - \mu_A(x_i)) \right) - (1 + a) \log (1 + a) \right]
\] (2.1.8)


In section 2.2, we have introduced some new probabilistic measures of information based upon measures of central tendency, dispersion and other means...
like power mean, exponential mean and quadratic mean etc. Section 2.3 deals with the development of some new probabilistic measures of entropy based upon sampling distributions along with the study of their properties whereas some new generalized information theoretic measures based upon probability distributions have been investigated and introduced in section 2.4.

2.2. New Measures Of Information Based Upon Measures Of Central tendency And Dispersion

In this section, we introduce different measures of information which can find their applications to different fields of biological sciences.

2.2.1. Information Measure In Terms of Measures of Central Tendency:
Let a random variable \( X \) takes the values \( x_1, x_2, \ldots, x_n \). Then, geometric mean \( G \), arithmetic mean \( M \) and harmonic mean \( H \) of these \( n \) observations are given by:

\[
G = \left( x_1 x_2 x_3 \ldots x_n \right)^{\frac{1}{n}}, \quad x_i \geq 0
\]  

\[
M = \frac{1}{n} \sum_{i=1}^{n} x_i
\]  

\[
H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots + \frac{1}{x_n}}
\]

Equations (2.2.1), (2.2.2) and (2.2.3) can be rewritten as

\[
G = \left( \frac{x_1 x_2 x_3 \ldots x_n}{\left( \sum_{i=1}^{n} x_i \right)^n} \right)^{\frac{1}{n}} \sum_{i=1}^{n} x_i
\]

or

\[
G = \left( \frac{p_1 p_2 p_3 \ldots p_n}{\left( \sum_{i=1}^{n} x_i \right)^n} \right)^{\frac{1}{n}} \sum_{i=1}^{n} x_i
\]

where \( p_i = \frac{x_i}{\sum_{i=1}^{n} x_i} \), Also

\[
nM = \sum_{i=1}^{n} x_i
\]
\begin{equation}
H = \frac{n \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i + \sum_{i=2}^{n} x_i + \ldots + \sum_{i=n}^{n} x_i}
\end{equation}

or
\begin{equation}
\frac{H}{M} = \frac{n^2}{\sum_{i=1}^{n} \frac{1}{p_i}}
\tag{2.2.6}
\end{equation}

Again equation (2.2.4) can be written as
\begin{equation}
\frac{G}{M} = n \left( p_1 p_2 p_3 \ldots p_n \right)^{\frac{1}{n}}
\end{equation}

or
\begin{equation}
\frac{1}{2^n} \sum_{i=1}^{n} \log p_i = \frac{1}{2^n} n \log \left( \frac{G}{nM} \right)
\tag{2.2.7}
\end{equation}

Adding equations (2.2.6) and (2.2.7), we get
\begin{equation}
\frac{1}{2^n} n \log \left( \frac{G}{nM} \right) + \frac{H}{M} = \frac{1}{2^n} \sum_{i=1}^{n} \log p_i + \frac{n^2}{\sum_{i=1}^{n} \frac{1}{p_i}}
\tag{2.2.8}
\end{equation}

Now, we introduce an information theoretic measure depending upon geometric mean G, arithmetic mean M and harmonic mean H. This measure is given by
\begin{equation}
\phi_n(P) = \frac{1}{2^n} \sum_{i=1}^{n} \log p_i + \frac{n^2}{\sum_{i=1}^{n} \frac{1}{p_i}}, \quad n \geq 2, \quad 0 < p_i < 1
\tag{2.2.9}
\end{equation}

We shall prove that the R.H.S. of equation (2.2.9) represents an information measure. To prove this, we study the essential properties of an entropy function as follows.

(i) $\phi_n(P)$ is permutationally symmetric as it doesn’t change if $p_1, p_2, \ldots, p_n$ are rearranged among themselves. This property is desirable since the labeling of the outcome shouldn’t affect the entropy.
(ii) Since $\frac{1}{p_i}$ is continuous function for $0 < p_i < 1$, $\phi_n(P)$ is also continuous everywhere in the same interval. Thus $\phi_n(P)$ is a continuous function of $p_1, p_2, ..., p_n$ for all $p_i$ lying between 0 and 1. It is highly desirable that $\phi_n(P)$ is a continuous function for all permissible probability distributions, since when the values of some probabilities are changed by small amounts, the entropy should also change by only a small amount.

(iii) Since $0 < p_i < 1$, $\phi_n(P)$ is never negative. This property is again desirable, since entropy should never be negative.

(iv) **Maximum Value:**

To find the maximum value, we consider the following function:

$$L = \frac{1}{2^n} \sum_{i=1}^{n} \log p_i + \frac{n^2}{\sum_{i=1}^{n} p_i} - \lambda \left( \sum_{i=1}^{n} p_i - 1 \right)$$

$$\frac{\partial L}{\partial p_i} = \frac{1}{2^n p_i} + \frac{n^2}{\left( \sum_{i=1}^{n} p_i \right)^2} \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)^2 - \lambda = \frac{p_i + 2^n n^2}{2^n p_i \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)^2} - \lambda$$

Similarly,

$$\frac{\partial L}{\partial p_2} = \frac{p_2 + 2^n n^2}{2^n p_2 \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)^2} - \lambda$$

$$\ldots$$

$$\frac{\partial L}{\partial p_i} = \frac{p_i + 2^n n^2}{2^n p_i \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)^2} - \lambda$$
Thus \( \frac{\partial L}{\partial p_1} = \frac{\partial L}{\partial p_2} = \frac{\partial L}{\partial p_3} = \ldots \frac{\partial L}{\partial p_n} = 0 \), gives

\[
\frac{p_i + 2^n n^2}{2^n p_i^2 \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)} - \lambda = \frac{p_2 + 2^n n^2}{2^n p_2^2 \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)} - \lambda = \ldots = \frac{p_n + 2^n n^2}{2^n p_n^2 \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)} - \lambda
\]

which further gives

\[
\frac{1}{p_i} + \frac{2^n n^2}{p_i^2} = \frac{1}{p_2} + \frac{2^n n^2}{p_2^2} = \ldots = \frac{1}{p_n} + \frac{2^n n^2}{p_n^2} \quad \text{or} \quad p_1 = p_2 = \ldots = p_n.
\]

Also using \( \sum_{i=1}^{n} p_i = 1 \), we get \( p_i = \frac{1}{n} \). Thus, the maximum value arises when the distribution is uniform. Further, the maximum value is given by

\[
[\phi_n(P)]_{\text{max}} = \frac{1}{2^n} \times n \log \frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n} = - \frac{n \log n}{2^n} + n^2 = f(n), \text{say.}
\]

Also \( f''(n) = n + (n-1)(1+ \log n) > 0 \forall n \geq 2 \)

Hence, the maximum value is an increasing function of \( n \).

(v) We have \( \phi_n^\prime (p) = \frac{1}{2^n p_i} + \frac{n^2}{p_i^2} \left( \sum_{i=1}^{n} \frac{1}{p_i} \right)^2 \) and

\[
\phi_n^\prime (p) = - \left[ \frac{1}{2^n p_i^2} + \frac{2n^2}{p_i^4} \left( \sum_{i=1}^{n} \frac{1}{p_i} \right) \right] \left[ p_i \sum_{i=1}^{n} \frac{1}{p_i} - 1 \right]
\]

Now, since \( p_i \sum_{i=1}^{n} \frac{1}{p_i} - 1 > 0 \) always, we see that \( \phi_n^\prime (p) < 0 \)

Thus \( \phi_n(P) \) is a concave function of \( p_1, p_2, \ldots, p_n \). This is very useful property since a local maximum will also be the global maximum for a concave function. In fact, in this case, a stationary point, if it exists, will give maximum rather than minimum. Thus, we see that \( \phi_n(P) \) introduced in equation (2.2.9) satisfies all the
essential properties of an information measure. Hence, we conclude from equation (2.2.8) that with the known values of geometric mean $G$, arithmetic mean $M$ and harmonic mean $H$, we can find the information content. Consequently, we have proved that $\phi_n(P)$ is a new measure of information.

Next, on fixing the value $n = 2$ and with different probabilities, we have computed different values of the measure $\phi_n(P)$ as shown in table-2.2.1:

**Table-2.2.1: Computations of $\phi_n(P)$**

<table>
<thead>
<tr>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$\phi_n(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.90</td>
<td>0.0986</td>
</tr>
<tr>
<td>0.20</td>
<td>0.80</td>
<td>0.4410</td>
</tr>
<tr>
<td>0.30</td>
<td>0.70</td>
<td>0.6726</td>
</tr>
<tr>
<td>0.40</td>
<td>0.60</td>
<td>0.8050</td>
</tr>
<tr>
<td>0.50</td>
<td>0.50</td>
<td>0.8494</td>
</tr>
<tr>
<td>0.60</td>
<td>0.40</td>
<td>0.8050</td>
</tr>
<tr>
<td>0.70</td>
<td>0.30</td>
<td>0.6726</td>
</tr>
<tr>
<td>0.80</td>
<td>0.20</td>
<td>0.4410</td>
</tr>
<tr>
<td>0.90</td>
<td>0.10</td>
<td>0.0986</td>
</tr>
</tbody>
</table>
Next, we have presented the values of $\phi_n(P)$ and obtained the Fig.-2.2.1 which shows that the measure introduced in equation (2.2.9) is a concave function.

![Fig.-2.2.1: Concavity of $\phi_n(P)$](image)

### 2.2.2. Information Measure in terms of Measures of Dispersion:

The variance of a discrete distribution of $n$ observations $(x_1, x_2, ..., x_n)$ is defined as

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \left( \frac{1}{n} \sum_{i=1}^{n} x_i \right)^2$$

(2.2.10)

The above equation (2.2.10) can be rewritten as

$$\sigma^2 = \frac{1}{n} \left[ \left( x_1^2 + x_2^2 + ... + x_n^2 \right) - \frac{1}{n} \left( \sum_{i=1}^{n} x_i \right)^2 \right] \left( \sum_{i=1}^{n} x_i \right)^2$$

$$= \frac{1}{n} \left[ \left( p_1^2 + p_2^2 + ... + p_n^2 \right) - \frac{1}{n} \right] \left( nM \right)^2$$

where $p_i = \frac{x_i}{\sum_{i=1}^{n} x_i}$ and $M$ is the arithmetic mean.
Thus, we have
\[ \frac{\sigma^2 + M^2}{2^n n M^2} = \frac{1}{2^n} \sum_{i=1}^{n} p_i^2 \]  \hspace{1cm} (2.2.11)

Adding equations (2.2.6) and (2.2.11), we get
\[ \frac{\sigma^2 + M^2}{2^n n M^2} + \frac{H}{M} = \frac{1}{2^n} \sum_{i=1}^{n} p_i^2 + \frac{n^2}{\sum_{i=1}^{n} 1 p_i} \]  \hspace{1cm} (2.2.12)

Now, we introduce an information theoretic measure depending upon arithmetic mean M, harmonic mean H and variance \( \sigma^2 \). This measure is given by
\[ \psi_n(P) = \frac{1}{2^n} \sum_{i=1}^{n} p_i^2 + \frac{n^2}{\sum_{i=1}^{n} 1 p_i}, \quad n \geq 2, \quad 0 < p_i < 1 \]  \hspace{1cm} (2.2.13)

We shall prove that the R.H.S.of equation (2.2.13) is an information measure. To prove this, we study its following properties:

(i) Obviously \( \psi_n(P) \geq 0 \)

(ii) \( \psi_n(P) \) is continuous function of \( p_i \forall 0 < p_i < 1 \)

(iii) \( \psi_n(P) \) is permutationally symmetric function of \( p_i \forall 0 < p_i < 1 \)

(iv) **Maximum Value**: To obtain, the maximum value of the entropy measure (2.2.13), we consider the following Lagrange’s function:
\[
L = \frac{1}{2^n} \sum_{i=1}^{n} p_i^2 + \frac{n^2}{\sum_{i=1}^{n} 1 p_i} - \lambda \left( \sum_{i=1}^{n} p_i - 1 \right). \text{ Thus, we have}
\]
\[
\frac{\partial L}{\partial p_i} = \frac{2 p_i}{2^n} + \frac{n^2}{(p_i)^2} \left( \sum_{i=1}^{n} 1 p_i \right)^{-2} - \lambda
\]

Similarly,
\[
\frac{\partial L}{\partial p_2} = \frac{2 p_2}{2^n} + \frac{n^2}{(p_2)^2} \left( \sum_{i=1}^{n} 1 p_i \right)^{-2} - \lambda, \ldots, \frac{\partial L}{\partial p_i} = \frac{2 p_i}{2^n} + \frac{n^2}{(p_i)^2} \left( \sum_{i=1}^{n} 1 p_i \right)^{-2} - \lambda
\]

For maximum value, we put
\[\frac{\partial L}{\partial p_1} = \frac{\partial L}{\partial p_2} = \frac{\partial L}{\partial p_3} = \ldots = \frac{\partial L}{\partial p_n} = 0, \text{ which gives} \]

\[\frac{2p_1}{2^n} + \frac{n^2}{(p_1)^2} \left( \sum_{i=1}^n \frac{1}{p_i} \right)^{-2} = \frac{2p_2}{2^n} + \frac{n^2}{(p_2)^2} \left( \sum_{i=1}^n \frac{1}{p_i} \right)^{-2} = \ldots = \frac{2p_n}{2^n} + \frac{n^2}{(p_n)^2} \left( \sum_{i=1}^n \frac{1}{p_i} \right)^{-2} \]

This further gives

\[2^{1-n} p_1 + \frac{n^2}{(p_1)^2} \left( \sum_{i=1}^n \frac{1}{p_i} \right)^{-2} = 2^{1-n} p_2 + \frac{n^2}{(p_2)^2} \left( \sum_{i=1}^n \frac{1}{p_i} \right)^{-2} = \ldots = 2^{1-n} p_n + \frac{n^2}{(p_n)^2} \left( \sum_{i=1}^n \frac{1}{p_i} \right)^{-2} \text{ which is possible only if } p_1 = p_2 = \ldots = p_n.

Also using \( \sum_{i=1}^n p_i = 1 \), we get \( p_i = \frac{1}{n} \). Thus, the maximum value arises when the distribution is uniform. Further, the maximum value is

\[f(n) = \frac{1}{n^{2^n}} + n^2 \]

Also

\[f'(n) = 2n - \frac{1+\log n}{n^{2^n}} = 2n - \frac{1}{2^n} \left( \frac{\log 2}{n} + \frac{1}{n^2} \right) > 0 \forall n \geq 2\]

Hence, the maximum value is an increasing function of \( n \).

**(v) Concavity:** To study its concavity, we have

\[\psi_n'(P) = 2^{1-n} p_1 + \frac{n^2}{(p_1)^2} \left( \sum_{i=1}^n \frac{1}{p_i} \right)^{-2} \]

Also,

\[\psi_n''(P) = -2 \left[ \frac{n^2}{p_1^4 \left( \sum_{i=1}^n \frac{1}{p_i} \right)} \left\{ p_1 \left( \sum_{i=1}^n \frac{1}{p_i} \right) - 1 \right\} - \frac{1}{2^n} \right] < 0\]
which shows that $\psi_n(P)$ is concave. Thus, we see that the measure introduced in equation (2.2.13) satisfies all the essential properties of being an information measure. Hence, we conclude that $\psi_n(P)$ is another new measure of information.

**Table-2.2.2: Computations of $\psi_n(P)$**

<table>
<thead>
<tr>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$\psi_n(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.9</td>
<td>0.5650</td>
</tr>
<tr>
<td>0.2</td>
<td>0.8</td>
<td>0.8100</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7</td>
<td>0.9850</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6</td>
<td>1.0900</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>1.2500</td>
</tr>
<tr>
<td>0.6</td>
<td>1.4</td>
<td>1.0900</td>
</tr>
<tr>
<td>0.7</td>
<td>0.3</td>
<td>0.9850</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2</td>
<td>0.8100</td>
</tr>
<tr>
<td>0.9</td>
<td>0.1</td>
<td>0.5650</td>
</tr>
</tbody>
</table>

Next, on fixing the value $n = 2$ and with the set of different probabilities, we have computed different values of $\psi_n(P)$ which are shown in the above table-2.2.2.
Then, we have presented the values of $\psi_n(P)$ and obtained the Fig.-2.2.2 which shows that the measure introduced in equation (2.2.13) is a concave function.

![Concavity of $\psi_n(P)$](image)

**Fig.-2.2.2: Concavity of $\psi_n(P)$**

### 2.2.3. Information Measure in terms of Power Mean and Quadratic Mean:

We know that the power measure $M_p$ for $n$ observations $x_1, x_2, ..., x_n$ is given by the following expression:

$$M_p = \left( \frac{1}{n} \sum_{i=1}^{n} x_i^r \right)^{\frac{1}{r}} , r > 0$$  \hspace{1cm} (2.2.14)

$$= \left[ \frac{1}{n} \left\{ x_1^r + x_2^r + \ldots + x_n^r \right\} \right]^{\frac{1}{r}}$$

$$= \left[ \frac{1}{n} \left( \sum_{i=1}^{n} x_i \right)^r \left( \sum_{i=1}^{n} x_i \right)^r + \ldots + \left( \sum_{i=1}^{n} x_i \right)^r \right]^{\frac{1}{r}}$$

54
\[
\left[ \frac{1}{n} \left( \sum_{i=1}^{n} x_i \right) \left( \sum_{i=1}^{n} p_i^\prime \right) \right]^{\frac{1}{2}}
\]

where \( p_i = \frac{x_i}{\sum_{i=1}^{n} x_i} \)

Thus \( M_p = \left[ \frac{1}{n} \left( nM \right)^\prime \left( \sum_{i=1}^{n} p_i^\prime \right) \right]^{\frac{1}{2}} \)

or \( M_p^\prime = \frac{1}{n} \left( nM \right)^\prime \left( \sum_{i=1}^{n} p_i^\prime \right) \)

or \( 2^n \frac{M_p^\prime}{M^\prime} = 2^n n^{\prime-1} \sum_{i=1}^{n} p_i^\prime \) \hspace{1cm} (2.2.15)

We also know that the quadratic mean of \( n \) observations \( x_1, x_2, \ldots, x_n \) is defined as

\[
M_q = \left[ \frac{1}{n} \sum_{i=1}^{n} x_i^2 \right]^{\frac{1}{2}}
\]

\[
= \left[ \frac{1}{n} \left( x_1^2 + x_2^2 + \ldots + x_n^2 \right) \right]^{\frac{1}{2}}
\]

\[
= \left[ \frac{1}{n} \left\{ \left( \frac{x_1}{\sum_{i=1}^{n} x_i} \right)^2 \left( \sum_{i=1}^{n} x_i \right)^2 + \left( \frac{x_2}{\sum_{i=1}^{n} x_i} \right)^2 \left( \sum_{i=1}^{n} x_i \right)^2 + \ldots + \left( \frac{x_n}{\sum_{i=1}^{n} x_i} \right)^2 \left( \sum_{i=1}^{n} x_i \right)^2 \right\} \right]^{\frac{1}{2}}
\]

\[
= \left[ \frac{1}{n} \left\{ p_1^2 (nM)^2 + p_2^2 (nM)^2 + \ldots + p_n^2 (nM)^2 \right\} \right]^{\frac{1}{2}}
\]

\[
= \left( nM^2 \sum_{i=1}^{n} p_i^2 \right)^{\frac{1}{2}}
\]

\[
= M \sqrt{n \sum_{i=1}^{n} p_i^2}
\]

Thus \( \frac{M_q^2}{M^2} = n \sum_{i=1}^{n} p_i^2 \) \hspace{1cm} (2.2.16)

Subtracting equations (2.2.15) and (2.2.16), we get the following result:
\[ 2^n \frac{M'_r}{M^r} - \frac{M^2}{M^2} = 2^n n^{r-1} \sum_{i=1}^{n} p_i^r - n \sum_{i=1}^{n} p_i^2 \]  \hfill (2.2.17)

We shall show that the R.H.S. of equation (2.2.17) is an information measure. Thus with the known values of arithmetic mean, power mean and quadratic mean, the information model can be constructed. We propose such a model by the following mathematical expression:

\[ H_n(P) = 2^n n^{r-1} \sum_{i=1}^{n} p_i^r - n \sum_{i=1}^{n} p_i^2, \quad r < 1, n \geq 2, 0 \leq p_i \leq 1 \]  \hfill (2.2.18)

Next, we have the following properties:

I  Obviously \( H_n(P) \geq 0 \)

II  \( H_n(P) \) is permutationally symmetric function of \( p_i \)

III  \( H_n(P) \) is continuous function of \( p_i \)

IV  Maximum Value:

Consider the Lagrange’s function given by

\[ L = 2^n n^{r-1} \sum_{i=1}^{n} p_i^r - n \sum_{i=1}^{n} p_i^2 - \lambda \left( \sum_{i=1}^{n} p_i - 1 \right), r < 1, n \geq 2 \]

\[ \frac{\partial L}{\partial p_1} = 2^n n^{r-1} p_1^{r-1} - 2np_1 - \lambda \]

\[ \frac{\partial L}{\partial p_2} = 2^n n^{r-1} p_2^{r-1} - 2np_2 - \lambda \]

\[ \frac{\partial L}{\partial p_n} = 2^n n^{r-1} p_n^{r-1} - 2np_n - \lambda \]

For maximum value, putting

\[ \frac{\partial L}{\partial p_1} = \frac{\partial L}{\partial p_2} = \frac{\partial L}{\partial p_3} = \ldots \frac{\partial L}{\partial p_n} = 0, \text{ we get} \]
\[2^n n^{-1} r p_i^{-1} - 2np_i = 2^n n^{-1} r p_2^{-1} - 2np_2 = \cdots = 2^n n^{-1} r p_n^{-1} - 2np_n\]

which is possible only if \( p_1 = p_2 = \ldots = p_n \).

Also using \( \sum_{i=1}^{n} p_i = 1 \), we get \( p_i = \frac{1}{n} \).

Thus, the maximum value arises when the distribution is uniform.

Further, the maximum value is

\[f(n) = 2^n n^{-1} \sum_{i=1}^{n} \frac{1}{n} - n \sum_{i=1}^{n} \frac{1}{n^2} = 2^n - 1\]

Also,

\[f(n) = 2^n \log 2 > 0 \forall n\]

Hence, the maximum value is an increasing function of \( n \).

V. Concavity: We have

\[\frac{\partial H_n(P)}{\partial p_i} = 2^n r n^{-1} p_i^{-1} - 2np_i\]

Also

\[\frac{\partial^2 H_n(P)}{\partial p_i^2} = 2^n r (r-1) n^{-1} p_i^{-2} - 2n < 0 \quad \text{as} \quad r < 1\]

Thus \( H_n(P) \) is a concave function of \( p_i \).

Hence, the measure proposed in (2.2.18) satisfies all the conditions of being a measure of information. Thus, we conclude that \( H_n(P) \) is a correct information measure.

Next, to check the validity of the proposed measure, we have fixed the value of \( n = 2 \), \( r = 1/2 \) and with the set of different probabilities, we have computed different values of \( H_n(P) \) which are shown in table-2.2.3.
Further, we have presented $H_n(P)$ graphically and obtained Fig.-2.2.3 which shows that the measure introduced in equation (2.2.18) is a concave function.

**Table-2.2.3: Computations of $H_n(P)$**

<table>
<thead>
<tr>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$H_n(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1.0</td>
<td>0.8284</td>
</tr>
<tr>
<td>0.1</td>
<td>0.9</td>
<td>1.9376</td>
</tr>
<tr>
<td>0.2</td>
<td>0.8</td>
<td>2.4345</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7</td>
<td>2.7553</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6</td>
<td>2.9395</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>2.9999</td>
</tr>
<tr>
<td>0.6</td>
<td>1.4</td>
<td>2.9395</td>
</tr>
<tr>
<td>0.7</td>
<td>0.3</td>
<td>2.7553</td>
</tr>
<tr>
<td>0.8</td>
<td>0.2</td>
<td>2.4345</td>
</tr>
<tr>
<td>0.9</td>
<td>0.1</td>
<td>1.9376</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0</td>
<td>0.8284</td>
</tr>
</tbody>
</table>
2.2.4. Information Measure In Terms of Exponential and Quadratic Mean:

The exponential mean of $n$ observations $x_1, x_2, \ldots, x_n$ is defined as

$$M_e = \log \left( \frac{e^{x_1} + e^{x_2} + \ldots + e^{x_n}}{n} \right)$$

$$= \log \left( \frac{e^{p_1 nM} + e^{p_2 nM} + \ldots + e^{p_n nM}}{n} \right)$$

$$= \log \left( \frac{\left( e^{nM} \right)^{p_1} + \left( e^{nM} \right)^{p_2} + \ldots + \left( e^{nM} \right)^{p_n}}{n} \right)$$

$$= \log \left( \frac{\lambda_1^{p_1} + \lambda_2^{p_2} + \ldots \lambda_n^{p_n}}{n} \right)$$
To develop a new information model, we make a simple adjustment in the value of exponential mean. This adjustment gives the following expression:

$$\log\left(\frac{\lambda + n - 1}{n}\right) - M_e = \log\left(\frac{\lambda + n - 1}{n}\right) - \log\left(\frac{\sum_{i=1}^{n} \lambda p_i}{n}\right)$$  \hspace{1cm} (2.2.19)

Multiplying equation (2.2.19) both sides by $4^2$ and then subtracting equation (2.2.16) from the resulting equation, we get

$$4^2 \left[ \log\left(\frac{\lambda + n - 1}{n}\right) - M_e \right] - \frac{M^2}{2} = 4^2 \left[ \log\left(\frac{\lambda + n - 1}{n}\right) - \log\left(\frac{\sum_{i=1}^{n} \lambda p_i}{n}\right) \right]$$

$$- n \sum_{i=1}^{n} p_i^2$$  \hspace{1cm} (2.2.20)

We shall show that the R.H.S. of equation (2.2.20) is an information measure. Thus with the known values of arithmetic mean, exponential mean and quadratic mean, information model can be constructed. We propose such a model by the following mathematical expression:

$$\xi_n(P) = 4^2 \left[ \log\left(\frac{\lambda + n - 1}{n}\right) - \log\left(\frac{\sum_{i=1}^{n} \lambda p_i}{n}\right) \right] - n \sum_{i=1}^{n} p_i^2 ,$$

$$\lambda > 0, \lambda \neq 1 \text{ and } 0 < p_i < 1$$  \hspace{1cm} (2.2.21)

Next, we study the essential properties of the function $\xi_n(P)$.

(i) Obviously $\xi_n(P) \geq 0$

(ii) $\xi_n(P)$ is permutationally symmetric function of $p_i$

(iii) $\xi_n(P)$ is continuous function of $p_i$

(iv) **Maximum Value:**

Consider the Lagrange’s function given by
\[ L = 4^n \left[ \log \left( \frac{\lambda + n - 1}{n} \right) - \log \left( \frac{\sum_{i=1}^{n} \lambda^{p_i}}{n} \right) \right] - n \sum_{i=1}^{n} p_i^2 - \lambda_i \left( \sum_{i=1}^{n} p_i - 1 \right) \]

Thus, we have the following results:

\[ \frac{\partial L}{\partial p_1} = -4^n \lambda^{p_1} \frac{\lambda}{\sum_{i=1}^{n} \lambda^{p_i}} - 2np_1 - \lambda_1, \quad \frac{\partial L}{\partial p_2} = -4^n \lambda^{p_2} \frac{\lambda}{\sum_{i=1}^{n} \lambda^{p_i}} - 2np_2 - \lambda_2, \ldots, \]

\[ \frac{\partial L}{\partial p_n} = -4^n \lambda^{p_n} \frac{\lambda}{\sum_{i=1}^{n} \lambda^{p_i}} - 2np_n - \lambda_1 \]

For maximum value, put

\[ \frac{\partial L}{\partial p_1} = \frac{\partial L}{\partial p_2} = \frac{\partial L}{\partial p_3} = \ldots = \frac{\partial L}{\partial p_n} = 0, \]

we get

\[ \frac{4^n \lambda^{p_1}}{\sum_{i=1}^{n} \lambda^{p_i}} + 2np_1 = \frac{4^n \lambda^{p_2}}{\sum_{i=1}^{n} \lambda^{p_i}} + 2np_2 = \ldots = \frac{4^n \lambda^{p_n}}{\sum_{i=1}^{n} \lambda^{p_i}} + 2np_n \]

which is possible if \( p_1 = p_2 = \ldots = p_n \).

Also using \( \sum_{i=1}^{n} p_i = 1 \), we get \( p_i = \frac{1}{n} \). Thus, the maximum value arises when the distribution is uniform.

Further the maximum value is

\[ f(n) = 4^n \left( \log (\lambda + n - 1) - \log n - \frac{\log \lambda}{n} \right) - 1 \]

Also

\[ f'(n) = 4^n \left( \frac{4n \log \left( \frac{\lambda + n - 1}{n} \right) + \frac{\log \lambda}{n^2} \left( 4n^2 - 1 \right) - \left( 1 - \frac{1}{\lambda + n - 1} \right) }{4n^2 - 1} \right) > 0 \]

Hence, the maximum value is an increasing function of \( n \).

(v) **Concavity:** We have \( \frac{\partial^2 L}{\partial p_i^2} = -4^n \frac{1}{\sum_{i=1}^{n} \lambda^{p_i}} \lambda^{p_i} \log \lambda - 2np_i \)
Also \[ \frac{\partial^2 \bar{\xi}_n(P)}{\partial p_i^2} = -4p^{2} \lambda^{pi} \left\{ \log \lambda \right\}^{2} \left[ \sum_{i=1}^{n} \lambda^{pi} - \lambda^{pi} \right] - 2n < 0 \]

Thus \( \bar{\xi}_n(P) \) is a concave function of \( p_i \).

Hence, the measure proposed in (2.2.21) satisfies all the conditions of being a measure of information. Thus, we conclude that \( \bar{\xi}_n(P) \) is a correct information measure.

\[ \begin{array}{|c|c|c|}
\hline
p_1 & p_2 & \bar{\xi}_n(P) \\
\hline
0.10 & 0.90 & 0.68960 \\
\hline
0.20 & 0.80 & 2.81024 \\
\hline
0.30 & 0.70 & 4.34144 \\
\hline
0.40 & 0.60 & 5.25504 \\
\hline
0.50 & 0.50 & 5.54720 \\
\hline
0.60 & 0.40 & 5.25504 \\
\hline
0.70 & 0.30 & 4.34144 \\
\hline
0.80 & 0.20 & 2.81024 \\
\hline
0.90 & 0.10 & 0.68960 \\
\hline
\end{array} \]
Next, to check the validity of the proposed measure, we have fixed the value of \( n = 2, \lambda = 2 \) and with the set of different probabilities, we have computed different values of \( \xi_n(P) \) which are shown in the above table-2.2.4. Further, we have presented the values of \( \xi_n(P) \) graphically and obtained the Fig.-2.2.4 which shows that the measure introduced in equation (2.2.21) is a concave function.

![Fig.-2.2.4: Concavity of \( \xi_n(P) \)](image)

### 2.3. Some New Information Measures Based Upon Sampling Distributions

In this section, we introduce some new probabilistic measures of entropy based upon \( \chi^2, t \) and \( F \) distributions, and study their important properties:

#### 2.3.1. Measure Of Entropy Based Upon \( \chi^2 \)-Distribution:

We know that \( \chi^2 \) variate is defined as

\[
\chi^2 = \sum_{i=1}^{a} \frac{(x_i - M)^2}{M}
\]  

(2.3.1)
where

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{x_i}{\sum_{i=1}^{n} x_i}$$

$$= n^2 M \sum_{i=1}^{n} \left( p_i - \frac{1}{n} \right)^2$$

$$= n^3 M \left( \sum_{i=1}^{n} \frac{p_i^2}{n} - \frac{1}{n} \right)$$

$$\frac{\chi^2 + nM}{n^2 M} = \sum_{i=1}^{n} p_i^2 \quad (2.3.2)$$

The R.H.S. of equation (2.3.2) is a standard measure of information known as measure of energy and has been introduced by Onicescu (1966). Thus, we conclude that measure of information can be calculated for known values of $\chi^2$ variate and arithmetic mean and consequently, we have

$$H^1 (P) = \frac{\chi^2 + nM}{n^2 M} \quad (2.3.3)$$

which is a new measure of information.

2.3.2. Measure Of Entropy Based Upon t-Distribution:

Let $x_1, x_2, ..., x_n$ be a random sample of size $n$ drawn from a normal population with mean $\mu$. Then, we know that $t$-statistic is defined as
$$t = \frac{|\bar{x} - \mu|}{s/\sqrt{n}}$$  \hspace{1cm} (2.3.4)$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ and $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Now $s^2 = \frac{1}{n-1} \left[ \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right]$

$$= \frac{1}{n-1} \left[ \left\{ \sum_{i=1}^{n} p_i^2 \right\}(n\bar{x})^2 - n\bar{x}^2 \right] \text{ where } p_i = \frac{x_i}{\sum x_i}$$

$$= \frac{n\bar{x}^2}{n-1} \left[ n \sum_{i=1}^{n} p_i^2 - 1 \right], n \neq 1 \hspace{1cm} (2.3.5)$$

Using (2.3.5) in (2.3.4), we get

$$t^2 = \frac{(\bar{x} - \mu)^2}{\bar{x}^2 \left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}}$$

Taking logarithm both sides, we get

$$\log \frac{t^2}{n-1} = \log \left( \frac{\bar{x} - \mu}{\bar{x}} \right)^2 - \log \left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}$$

or

$$-\frac{1}{2} \log \frac{t^2}{n-1} = -\log \left| \frac{\bar{x} - \mu}{\bar{x}} \right| + \log \left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}^{1/2}$$

Hence, we observe that for different values of $n > 2$, the information model for t-distribution becomes

$$I_n(P) = -\frac{1}{2} \log \frac{t^2}{n-1} = -C + \log \left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}^{1/2}, n > 2 \hspace{1cm} (2.3.6)$$

where $C = \log \left| \frac{\bar{x} - \mu}{\bar{x}} \right|$.

We choose $C$ such that $\left| \frac{\bar{x} - \mu}{\bar{x}} \right| < 1$. 

65
that is, \(-\bar{x} < \bar{x} - \mu < \bar{x}\) or \(\bar{x} > \frac{\mu}{2}\) and \(\mu > 0\)

Without any loss of generality, we can assume that 

\[ C < \left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}^{1/2} \]

Now, we show that the R.H.S. of equation (2.3.6) is a theoretical measure of information. For this purpose, we have studied the following properties:

(i) Obviously \(I_n(P) \geq 0\) as \(C < 0\) and \(\left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}^{1/2} < 1 \forall p_i\)

(ii) \(I_n(P)\) is a continuous function of \(p_i\).

(iii) \(I_n(P)\) is permutationally symmetric function of \(p_i\).

(iv) **Concavity:** To study its concavity, we proceed as follows:

We have

\[ \frac{\partial I_n(P)}{\partial p_i} = \frac{np_i}{n \sum_{i=1}^{n} p_i^2 - 1} \]

Also

\[ \frac{\partial^2 I_n(P)}{\partial p_i^2} = \frac{n}{\left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}^2} \left[ n \sum_{i=1}^{n} p_i^2 - 2np_i^2 - 1 \right] \]

It has been verified numerically that

\[ \left[ n \sum_{i=1}^{n} p_i^2 - 2np_i^2 - 1 \right] < 0 \forall n, p_i \]

Thus, we observe that \(\frac{\partial^2 I_n(P)}{\partial p_i^2} < 0\)

which shows that \(I_n(P)\) is a concave function of \(p_i\).

(v) For obtaining maximum value, we consider the Lagrange function given by

\[ g(P) = -\frac{1}{2} \log \frac{\lambda^2}{n-1} = -C + \log \left\{ n \sum_{i=1}^{n} p_i^2 - 1 \right\}^{1/2} - \lambda \left( \sum_{i=1}^{n} p_i - 1 \right) \]

Thus
\[ \frac{\partial g(P)}{\partial p_1} = np_1 - \lambda \]
\[ \frac{\partial g(P)}{\partial p_2} = np_2 - \lambda \]
\[ \frac{\partial g(P)}{\partial p_n} = np_n - \lambda \]
\[ n \sum_{i=1}^{n} p_i^2 - 1 \]
For maximum value, putting
\[ \frac{\partial g(P)}{\partial p_1} = \frac{\partial g(P)}{\partial p_2} = \frac{\partial g(P)}{\partial p_3} = \ldots = \frac{\partial g(P)}{\partial p_n} = 0 , \]
which gives
\[ \frac{np_1}{n \sum_{i=1}^{n} p_i^2 - 1} = \frac{np_2}{n \sum_{i=1}^{n} p_i^2 - 1} = \ldots = \frac{np_n}{n \sum_{i=1}^{n} p_i^2 - 1} \]
which further gives \( p_1 = p_2 = \ldots = p_n \).

Also using \( \sum_{i=1}^{n} p_i = 1 \), we get \( p_i = \frac{1}{n} \). Thus the maximum value arises when the distribution is uniform.

Under the above properties, we see that \( I_n(P) \) will be an information measure and consequently, we conclude that if \( t \) is any \( t \)-statistic, then \( -\frac{1}{2} \log \frac{t^2}{n-1} \) will act as an information measure.

2.3.3. Measure Of Entropy Based Upon F-Distribution:

The F-statistic is given by
\[ F = \frac{s_x^2}{s_y^2} , \quad s_x^2 > s_y^2 , \quad \text{where} \]
\[ s_x^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2 \] and \[ s_y^2 = \frac{1}{n_2} \sum_{j=1}^{n_2} (y_j - \bar{y})^2, \] \( n_1 > 2, n_2 > 2 \)

Now \[ s_x^2 = \frac{1}{n_1 - 1} \left[ \sum_{i=1}^{n_1} x_i^2 - n_1 \bar{x}^2 \right] \]
\[ = \frac{1}{n_1 - 1} \left[ \left\{ \sum_{i=1}^{n_1} p_i^2 \right\} (n_1 \bar{x})^2 - n_1 \bar{x}^2 \right] \] where \( p_i = \frac{x_i}{\sum_{i=1}^{n_1} x_i} \)
\[ = \frac{n_1 \bar{x}^2}{n_1 - 1} \left[ n_1 \sum_{i=1}^{n_1} p_i^2 - 1 \right] \]

Similarly, we have \[ s_y^2 = \frac{n_2 \bar{y}^2}{n_2 - 1} \left[ n_2 \sum_{j=1}^{n_2} p_j^2 - 1 \right] \]

Thus, \[ F = \frac{s_x^2}{s_y^2} = \frac{n_1 (n_2 - 1) \bar{x}^2}{n_2 (n_1 - 1) \bar{y}^2} \frac{n_1 \sum_{i=1}^{n_1} p_i^2 - 1}{n_2 \sum_{j=1}^{n_2} p_j^2 - 1}, n_2 > n_1 \]

Taking logarithm both sides, we get \[ \log F = \log \left( \frac{n_1 (n_2 - 1) \bar{x}^2}{n_2 (n_1 - 1) \bar{y}^2} \right) + \log \left( \frac{n_1 \sum_{i=1}^{n_1} p_i^2 - 1}{n_2 \sum_{j=1}^{n_2} p_j^2 - 1} \right) \]
or \[ \log F = C + \log A \]

where \( C = \log \left( \frac{n_1 (n_2 - 1) \bar{x}^2}{n_2 (n_1 - 1) \bar{y}^2} \right), \)
\[ n_1 \sum_{i=1}^{n_1} p_i^2 - 1 \]
and \( A = \frac{\sum_{j=1}^{n_2} p_j^2 - 1}{n_2 \sum_{j=1}^{n_2} p_j^2 - 1} \)

We, now propose a new measure of information based on F-distribution, given by
\[ n\psi(P) = C + \log \left( \frac{n_1 \sum_{i=1}^{n_1} p_i^2 - 1}{n_2 \sum_{j=1}^{n_2} p_j^2 - 1} \right), \quad n_2 > n_1 \] (2.3.7)

Now, we show that the R.H.S. of equation (2.3.7) is a theoretical measure of information. For this purpose, we study the following properties:

(i) Obviously, \( n\psi(P) \geq 0 \) as \( C > 0 \), \( \left\{ n_1 \sum_{i=1}^{n_1} p_i^2 - 1 \right\}^{1/2} < 1 \) \( \forall p_i \)

and \( \left\{ n_2 \sum_{j=1}^{n_2} p_j^2 - 1 \right\}^{1/2} < 1 \) \( \forall p_j \)

(ii) \( n\psi(P) \) is a continuous function of \( p_i \) and \( p_j \).

(iii) \( n\psi(P) \) is a symmetric function of \( p_i \) and \( p_j \).

(vi) **Concavity:** To study its concavity, we have

\[ n\psi(P) = C + \log \left( n_1 \sum_{i=1}^{n_1} p_i^2 - 1 \right) - \log \left( n_2 \sum_{j=1}^{n_2} p_j^2 - 1 \right) \]

Take \( n\psi'(P) = \log \left( n_1 \sum_{i=1}^{n_1} p_i^2 - 1 \right) \) and \( n\psi'(P) = \log \left( n_2 \sum_{j=1}^{n_2} p_j^2 - 1 \right) \)

Thus

\[ \frac{\partial_n \psi'(P)}{\partial p_i} = \frac{2n_1 p_i}{n_1 \sum_{i=1}^{n_1} p_i^2 - 1} \]

Also,

\[ \frac{\partial^2_n \psi'(P)}{\partial p_i^2} = \frac{2n_1}{\left\{ n_1 \sum_{i=1}^{n_1} p_i^2 - 1 \right\}^2} \left[ n_1 \sum_{i=1}^{n_1} p_i^2 - 2n_1 p_i^2 - 1 \right] < 0 \forall n_i > 2 \]

Thus, \( n\psi'(P) \) is a concave function of \( p_i \)

69
Similarly, \( \frac{\partial^2 \psi^2(P)}{\partial p_j^2} = \frac{2n_2}{n_2} \left[ n_2 \sum_{j=1}^{n_2} p_j^2 - 2n_2 p_j^2 - 1 \right] < 0 \forall n_2 > 2 \)

Hence, \( \psi^2(P) \) is a concave function of \( p_j \). As difference of two concave functions is also a concave function, we conclude that \( \psi(P) \) is a concave function.

(v) For obtaining maximum value, we consider the Lagrange function given by
\[
f(p) = C + \log \left( \sum_{i=1}^{n_1} p_i^2 - 1 \right) - \log \left( \sum_{j=1}^{n_2} p_j^2 - 1 \right) - \lambda \left( \sum_{i=1}^{n_1} p_i - 1 \right) - \mu \left( \sum_{j=1}^{n_2} p_j - 1 \right)
\]

Thus, we have
\[
\frac{\partial f(p)}{\partial p_1} = \frac{2n_1 p_1}{n_1} - \frac{2n_2 p_1}{n_2} - \lambda - \mu
\]
\[
\frac{\partial f(p)}{\partial p_2} = \frac{2n_1 p_2}{n_1} - \frac{2n_2 p_2}{n_2} - \lambda - \mu
\]
\[
\frac{\partial f(p)}{\partial p_n} = \frac{2n_1 p_n}{n_1} - \frac{2n_2 p_n}{n_2} - \lambda - \mu
\]

For maximum value, putting
\[
\frac{\partial f(p)}{\partial p_1} = \frac{\partial f(p)}{\partial p_2} = \frac{\partial f(p)}{\partial p_3} = \ldots = \frac{\partial f(p)}{\partial p_n} = 0,
\] we get
\[ 2p_1 \left( \frac{n_1}{n_1 \sum_{i=1}^n p_i^2 - 1} - \frac{n_2}{n_2 \sum_{j=1}^n p_j^2 - 1} \right) = 2p_2 \left( \frac{n_1}{n_1 \sum_{i=1}^n p_i^2 - 1} - \frac{n_2}{n_2 \sum_{j=1}^n p_j^2 - 1} \right) \]
\[ \ldots = 2p_n \left( \frac{n_1}{n_1 \sum_{i=1}^n p_i^2 - 1} - \frac{n_2}{n_2 \sum_{j=1}^n p_j^2 - 1} \right) \]
which gives \( p_1 = p_2 = \ldots = p_n \).

Also using \( \sum_{i=1}^n p_i = 1, \sum_{j=1}^n p_j = 1 \), we get
\[ p_i = \frac{1}{n_1} \quad \text{and} \quad p_j = \frac{1}{n_2} \]
Thus, we see that the maximum value arises when the distribution is uniform.

Under the above properties, we see that \( \psi(P) \) introduced above is an information theoretic measure and consequently, we conclude that if \( F \) is any \( F \)-statistic, then \( \log F \) will act as an information measure.

**2.4. Some Generalized Information Theoretic Measures Based Upon Probability Distributions:**

In this section, we propose some new generalized information measures for a probability distribution \( P = \left\{ (p_1, p_2, ..., p_n), p_i \geq 0, \sum_{i=1}^n p_i = 1 \right\} \) and study their essential and desirable properties.

**2.4.1 Generalized Probabilistic Entropy Involving Three Parameters \( \alpha, \beta \) and \( \gamma \):**

We propose the generalized entropy depending upon three real parameters \( \alpha, \beta \) and \( \gamma \) as given by the following mathematical expression:
\[ H_{\alpha,\beta,\gamma}(P) = \frac{1}{1-\alpha} \sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1} \left( 1 - \sum_{i=1}^{n} p_i^{\beta+\gamma} \right) \]  

(2.4.1)

\[ \alpha \neq 1, \beta \geq 0, \gamma \geq 0, \alpha + \beta + \gamma - 1 > 0, \beta + \gamma - 1 \geq 0 \]

As \( \alpha \to 1 \), the measure (2.4.1) reduces to

\[ H_{\beta,\gamma}(P) = -\sum_{i=1}^{n} p_i^{\beta+\gamma} \log p_i \sum_{i=1}^{n} p_i^{\beta+\gamma} \]  

(2.4.2)

If \( \beta = 0 \) and \( \gamma = 1 \) then the above measure (2.4.2) reduces to

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

which is Shannon’s (1948) measure of entropy.

Similarly, if \( \gamma = 0 \) and \( \beta = 1 \), again we observe that the measure (2.4.2) reduces to Shannon’s (1948) measure of entropy. Further for \( \beta + \gamma = 1 \), (2.4.1) becomes

\[ H_{\alpha}(P) = \frac{1}{1-\alpha} \log \sum_{i=1}^{n} p_i^{\alpha}, \alpha \neq 1, \alpha > 0, \]  

which is Renyi’s (1961) entropy.

Thus, we observe that the expression introduced in equation (2.4.1) is a generalized probabilistic measure of entropy. Next, we study some important properties of this generalized measure.

The measure (2.4.1) satisfies the following properties:

(i) It is continuous function of \( p_1, p_2, ..., p_n \), so that it changes by a small amount when \( p_1, p_2, ..., p_n \) change by small amounts.

(ii) It is permutationally symmetric function of \( p_1, p_2, ..., p_n \), that is, it does not change when \( p_1, p_2, ..., p_n \) are permuted among themselves.

(iii) **Concavity of** \( H_{\alpha,\beta,\gamma}(P) \): We have
\[ \frac{\partial}{\partial p_i} H_{\alpha,\beta,\gamma}(P) = \frac{\left(\alpha + \beta + \gamma - 1\right) p_i^{\alpha+\beta+\gamma-2} - \left(\beta + \gamma\right) p_i^{\beta+\gamma-1}}{\left(1-\alpha\right) \sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1} - \left(1-\alpha\right) \sum_{i=1}^{n} p_i^{\beta+\gamma}} \]

\[ \frac{\partial^2}{\partial p_i^2} H_{\alpha,\beta,\gamma}(P) = \frac{\left(\alpha + \beta + \gamma - 1\right)}{\left(1-\alpha\right)} \left[ \frac{\left(\alpha + \beta + \gamma - 2\right) \left(\sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1}\right) p_i^{\alpha+\beta+\gamma-3} - \left(\alpha + \beta + \gamma - 1\right) p_i^{2\alpha+\beta+\gamma-2}}{\left(\sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1}\right)^2} \right] - \frac{\left(\beta + \gamma\right)}{1-\alpha} \left[ \frac{\left(\beta + \gamma - 1\right) \left(\sum_{i=1}^{n} p_i^{\beta+\gamma}\right) p_i^{\beta+\gamma-2} - \left(\beta + \gamma\right) p_i^{2\beta+\gamma-1}}{\left(\sum_{i=1}^{n} p_i^{\beta+\gamma}\right)^2} \right] \]

Clearly \( \frac{\partial^2}{\partial p_i^2} H_{\alpha,\beta,\gamma}(P) < 0 \forall i \)

Hence, \( H_{\alpha,\beta,\gamma}(P) \) is a concave function of \( p_i \)

(iv) To find the maximum value of (2.4.1), we proceed as follows:

Let \( f(P) = \frac{1}{1-\alpha} \log \sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1} - \frac{1}{1-\alpha} \log \sum_{i=1}^{n} p_i^{\beta+\gamma} - \lambda \left(\sum_{i=1}^{n} p_i - 1\right) \)

Thus

\[ \frac{\partial f}{\partial p_1} = \frac{\left(\alpha + \beta + \gamma - 1\right) p_1^{\alpha+\beta+\gamma-2} - \left(\beta + \gamma\right) p_1^{\beta+\gamma-1} - \lambda}{\left(1-\alpha\right) \sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1} - \left(1-\alpha\right) \sum_{i=1}^{n} p_i^{\beta+\gamma}} \]

\[ \frac{\partial f}{\partial p_2} = \frac{\left(\alpha + \beta + \gamma - 1\right) p_2^{\alpha+\beta+\gamma-2} - \left(\beta + \gamma\right) p_2^{\beta+\gamma-1} - \lambda}{\left(1-\alpha\right) \sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1} - \left(1-\alpha\right) \sum_{i=1}^{n} p_i^{\beta+\gamma}} \]

...
\[
\frac{\partial f}{\partial p_i} = \frac{\alpha + \beta + \gamma - 1}{(1 - \alpha)} p_i^{\alpha + \beta + \gamma - 2} - \frac{\beta + \gamma}{(1 - \alpha)} \sum_{i=1}^{n} p_i^{\beta + \gamma} - \lambda
\]

Putting \( \frac{\partial f}{\partial p_i} = 0 \) \( \forall i = 1, 2, \ldots, n \), we get

\[
\frac{1}{(1 - \alpha)} p_1^{\beta + \gamma - 1} \left[ \frac{\alpha + \beta + \gamma - 1}{\sum_{i=1}^{n} p_i^{\alpha + \beta + \gamma - 1}} - \frac{\beta + \gamma}{\sum_{i=1}^{n} p_i^{\beta + \gamma}} \right]
\]

\[
= \frac{1}{(1 - \alpha)} p_2^{\beta + \gamma - 1} \left[ \frac{\alpha + \beta + \gamma - 1}{\sum_{i=1}^{n} p_i^{\alpha + \beta + \gamma - 1}} - \frac{\beta + \gamma}{\sum_{i=1}^{n} p_i^{\beta + \gamma}} \right]
\]

\[
= \frac{1}{(1 - \alpha)} p_n^{\beta + \gamma - 1} \left[ \frac{\alpha + \beta + \gamma - 1}{\sum_{i=1}^{n} p_i^{\alpha + \beta + \gamma - 1}} - \frac{\beta + \gamma}{\sum_{i=1}^{n} p_i^{\beta + \gamma}} \right]
\]

which is possible only if \( p_1 = p_2 = \ldots = p_n \)

Thus, \( \sum_{i=1}^{n} p_i = 1 \) gives \( p_1 = p_2 = \ldots = p_n = \frac{1}{n} \)

Hence, we see that the generalized entropy measure (2.4.1) possesses its maximum value and this maximum value subject to natural constraint arises when \( p_1 = p_2 = \ldots = p_n = \frac{1}{n} \). This result is most desirable.

(v) The maximum value of \( H_{\alpha, \beta, \gamma}(P) \) is given by

\[
f(n) = \frac{1}{1 - \alpha} \log \left[ \sum_{i=1}^{n} \left( \frac{1}{n} \right)^{\alpha + \beta + \gamma - 1} \right] - \frac{1}{1 - \alpha} \log \left[ \sum_{i=1}^{n} \left( \frac{1}{n} \right)^{\beta + \gamma} \right]
\]
\[
= \frac{1}{1-\alpha} \left[ -((\alpha + \beta + \gamma - 1) \log n) \right] - \frac{1}{1-\alpha} \left[ -((\beta + \gamma) \log n) \right] = \log n
\]

Thus, we have

\[
f'(n) = \frac{1}{n} > 0. \text{ Also, } \]

\[
f''(n) = -\frac{1}{n^2} < 0 \quad \forall n
\]

Thus, we conclude that the maximum value is a concave function of \(n\).

(vi) The measure (2.4.1) is additive in nature. To prove this additivity property, we consider the following joint entropy:

\[
H_{nm}(P \ast Q) = \frac{1}{1-\alpha} \log \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} p_i^{\alpha+\beta+\gamma-1} q_j^{\alpha+\beta+\gamma-1}}{\sum_{i=1}^{m} \sum_{j=1}^{n} p_i^{\beta+\gamma} q_j^{\beta+\gamma}}
\]

\[
= \frac{1}{1-\alpha} \log \left[ \frac{\sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1} \sum_{j=1}^{m} q_j^{\alpha+\beta+\gamma-1}}{\sum_{i=1}^{n} p_i^{\beta+\gamma} \sum_{j=1}^{m} q_j^{\beta+\gamma}} \right]
\]

\[
= \frac{1}{1-\alpha} \log \frac{\sum_{i=1}^{n} p_i^{\alpha+\beta+\gamma-1}}{\sum_{i=1}^{n} p_i^{\beta+\gamma}} + \frac{1}{1-\alpha} \log \frac{\sum_{j=1}^{m} q_j^{\alpha+\beta+\gamma-1}}{\sum_{j=1}^{m} q_j^{\beta+\gamma}}
\]

\[
= H_n(P) + H_m(Q)
\]

Thus, the joint entropy is given by

\[
H_{nm}(P \ast Q) = H_n(P) + H_m(Q)
\]

which proves the additivity property of the proposed measure (2.4.1).

Next, with the help of the data, we have presented the generalized measure (2.4.1) graphically. For this purpose, we have fixed \(\beta = 2\) and \(\gamma = 3\). Then, for different values of \(\alpha\), we have computed different values of \(H_{\alpha,\beta,\gamma}(P)\) as shown in the following table- 2.4.1:
Table-2.4.1: Computations $H_{a,\beta,\gamma}(P)$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$H_{a,\beta,\gamma}(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.9</td>
<td>0.1054</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>0.2239</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.7</td>
<td>0.3655</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.6</td>
<td>0.5514</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>0.6931</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.4</td>
<td>0.5514</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.3</td>
<td>0.3655</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.2</td>
<td>0.2239</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.1</td>
<td>0.1054</td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
<td>0.9</td>
<td>0.1054</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>0.2236</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.7</td>
<td>0.3611</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.6</td>
<td>0.5402</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>0.6932</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.4</td>
<td>0.5402</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.3</td>
<td>0.3611</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.2</td>
<td>0.2236</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.1</td>
<td>0.1054</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>0.9</td>
<td>0.1053</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>0.2233</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.7</td>
<td>0.3591</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.6</td>
<td>0.5276</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>0.6955</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.4</td>
<td>0.5276</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.3</td>
<td>0.3591</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.2</td>
<td>0.2233</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.1</td>
<td>0.1053</td>
</tr>
</tbody>
</table>
Next, we have presented the values of $H_{\alpha,\beta,\gamma}(P)$ graphically for $\alpha=2$ and obtained the above Fig.-2.4.1 which shows that the measure introduced in equation (2.4.1) is a concave function. It is further added that different values of $\alpha$, we get similar curves of concave functions.

![Concavity of $H_{\alpha,\beta,\gamma}(P)$](image)

**Fig.-2.4.1: Concavity of $H_{\alpha,\beta,\gamma}(P)$**

### 2.4.2 Generalized Probabilistic Entropy Involving Two Parameters $\alpha$ and $\beta$

Now, we introduce another parametric measure of entropy depending upon two parameters $\alpha$ and $\beta$. This measure is given by

$$H_{\alpha,\beta}(P) = \sum_{i=1}^{n} p_i^{\alpha+\beta} - 1, \quad \alpha + \beta \neq 1, \alpha + \beta > 0$$  \hspace{1cm} (2.4.3)

If $\beta = 0$ and $\alpha \neq 1$, the above measure (2.4.3) reduces to

$$H_{\alpha}(P) = \sum_{i=1}^{n} p_i^\alpha - 1, \quad \alpha \neq 1, \alpha > 0$$ \hspace{1cm} (2.4.4)
which is Havrada and Charvat’s (1967) measure of entropy. Further as $\alpha \to 1$, (2.4.4) reduces to Shannon’s [1948] measure of entropy.

Further as $\alpha = 0$, $\beta \neq 1$, again we see that (2.4.3) reduces to Havrada and Charvat’s (1967) measure of entropy. Thus, we observe that the measure (2.4.3) is a generalized parametric measure of entropy. Next, we study the properties of the measure (2.4.3).

The measure (2.4.3) satisfies the following properties:

(I) It is continuous function of $p_1, p_2, \ldots, p_n$ so that it changes by a small amount when $p_1, p_2, \ldots, p_n$ change by small amounts.

(II) It is permutationally symmetric function of $p_1, p_2, \ldots, p_n$, that is, it does not change when $p_1, p_2, \ldots, p_n$ are permuted among themselves.

(III) Concavity of $H_{n}^{\alpha,\beta}(P)$:

We have

\[
\frac{\partial}{\partial p_i} H_{n}^{\alpha,\beta}(P) = \frac{(\alpha + \beta) p_i^{\alpha + \beta - 1}}{2^{2^{1-i^2} - \beta - 1}}
\]

Also

\[
\frac{\partial^2}{\partial p_i^2} H_{n}^{\alpha,\beta}(P) = \frac{(\alpha + \beta)(\alpha + \beta - 1) p_i^{\alpha + \beta - 2}}{2^{2^{1-i^2} - \beta - 1}}
\]

\[
= -\frac{2^{\alpha + \beta} \left[(\alpha + \beta)^2 - (\alpha + \beta)\right]}{2^{\alpha + \beta - 2}} < 0 \forall i
\]

Hence, $H_{n}^{\alpha,\beta}(P)$ is a concave function of $p_i$.

(IV) To find the maximum value of (2.4.3), we apply Lagrange’s method of maximum multipliers as follows:

Let

\[
f(P) = \sum_{i=1}^{n} p_i^{\alpha + \beta - 1} - \lambda \left[\sum_{i=1}^{n} p_i - 1\right]
\]

Thus

\[
\frac{\partial f}{\partial p_i} = \frac{(\alpha + \beta)}{2^{2^{1-i^2} - \beta - 1}} p_i^{\alpha + \beta - 1} - \lambda
\]
\[
\frac{\partial f}{\partial p_2} = \frac{(\alpha + \beta)}{2^{1-a-\beta} - 1} p_2^{\alpha+\beta-1} - \lambda
\]

.

.

.

\[
\frac{\partial f}{\partial p_n} = \frac{(\alpha + \beta)}{2^{1-a-\beta} - 1} p_n^{\alpha+\beta-1} - \lambda
\]

Putting \( \frac{\partial f}{\partial p_i} = 0 \quad \forall \quad i = 1, 2, 3, \ldots, n \), we get

\[
\frac{(\alpha + \beta)}{2^{1-a-\beta} - 1} p_i^{\alpha+\beta-1} = \frac{(\alpha + \beta)}{2^{1-a-\beta} - 1} p_2^{\alpha+\beta-1} = \ldots = \frac{(\alpha + \beta)}{2^{1-a-\beta} - 1} p_n^{\alpha+\beta-1}
\]

Since \( \alpha + \beta \neq 1 \) and \( \alpha + \beta > 0 \), the above result is possible only if

\[p_1 = p_2 = \ldots = p_n.\]

Thus, \( \sum_{i=1}^{n} p_i = 1 \) gives \( p_1 = p_2 = \ldots = p_n = \frac{1}{n} \)

Hence, we observe that the entropy measure (2.4.3) possesses its maximum value
and its maximum value subject to the constraint \( \sum_{i=1}^{n} p_i = 1 \) arises when \( p_i = \frac{1}{n} \quad \forall \quad i \),
that is, for the most uniform distribution. This result is again most desirable.

(iv) The maximum value of \( H_{n}^{\alpha,\beta}(P) \) is given by

\[
f(n) = \frac{n^{1-a-\beta} - 1}{2^{1-a-\beta} - 1}
\]

Thus, we have

\[
f'(n) = \frac{(1 - \alpha - \beta)n^{-(\alpha+\beta)}}{2^{1-a-\beta} - 1} > 0 \quad \forall \quad \alpha + \beta \neq 1, \quad \alpha + \beta > 0
\]

Also

\[
f''(n) = \frac{[1 - (\alpha + \beta)](\alpha + \beta)}{n^{1+a+\beta} \left[ 2^{1-a-\beta} - 1 \right]}
\]
Thus, we conclude that the maximum value is a concave function of $n$.

(v) Non-additive property of $H_n^{\alpha,\beta}(P)$:

The joint entropy of the two probability distributions $P$ and $Q$ is given by

$$H_{nm}^{\alpha,\beta}(P \ast Q) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_i^{\alpha+\beta} q_j^{\alpha+\beta} - 1$$

$$= \left\{ \sum_{i=1}^{n} p_i^{\alpha+\beta} - 1 \right\} \left\{ \sum_{j=1}^{m} q_j^{\alpha+\beta} - 1 \right\} + \left\{ \sum_{i=1}^{n} p_i^{\alpha+\beta} - 1 \right\} + \left\{ \sum_{j=1}^{m} q_j^{\alpha+\beta} - 1 \right\}$$

Thus, we have

$$H_{nm}^{\alpha,\beta}(P \ast Q) = H_n^{\alpha,\beta}(P) \left\{ 2^{1-\alpha-\beta} - 1 \right\} H_m^{\alpha,\beta}(Q) + H_n^{\alpha,\beta}(P) + H_m^{\alpha,\beta}(Q)$$  \hspace{1cm} (2.4.5)$$

The result (2.4.5) clearly shows that the entropy measure (2.4.3) is non-additive.

(vi) Recursivity Property:

We rewrite the measure (2.4.3) as follows:

$$H_{n-1}^{\alpha,\beta}(p_1 + p_2, p_3, p_4, \ldots, p_n) = \frac{(p_1 + p_2)^{\alpha+\beta} + \sum_{i=1}^{n} p_i^{\alpha+\beta} - 1}{2^{1-\alpha-\beta} - 1}$$

$$= \frac{(p_1 + p_2)^{\alpha+\beta} + \sum_{i=1}^{n} p_i^{\alpha+\beta} - p_1^{\alpha+\beta} - p_2^{\alpha+\beta} - 1}{2^{1-\alpha-\beta} - 1}$$

$$= \frac{(p_1 + p_2)^{\alpha+\beta} - p_1^{\alpha+\beta} - p_2^{\alpha+\beta} + \sum_{i=1}^{n} p_i^{\alpha+\beta} - 1}{2^{1-\alpha-\beta} - 1}$$
The equation (2.4.6) clearly shows that the generalized entropy measure (2.4.3) possesses recursivity property.  

**Note:** From equation (2.4.6), we have  

$$H_{n-1}^{\alpha,\beta}(p_1, p_2, \ldots, p_n) \leq H_n^{\alpha,\beta}(p_1 + p_2, p_3, \ldots, p_n)$$  

This shows that if two outcomes are combined, the entropy is reduced. This property is much desirable since when outcomes are combined, the uncertainty should not increase.  

Next, with the help of data, we have presented the measure (2.4.3) graphically. For this purpose, we have fixed the value of the parameter $\beta = 2$. Then, for different values of $\alpha$, we have computed different values of the probabilistic measure $H_n^{\alpha,\beta}(P)$ as shown in table-2.4.2:
Table-2.4.2: Computations of $H_n^{\alpha,\beta}(P)$

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$H_n^{\alpha,\beta}(P)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.9</td>
<td>0.3929</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>0.6729</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.7</td>
<td>0.8592</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.6</td>
<td>0.9654</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.4</td>
<td>0.9654</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.3</td>
<td>0.8592</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.2</td>
<td>0.6729</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.1</td>
<td>0.3929</td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
<td>0.9</td>
<td>0.4836</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>0.7615</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.7</td>
<td>0.9100</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.6</td>
<td>0.9798</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.4</td>
<td>0.9798</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.3</td>
<td>0.9100</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.2</td>
<td>0.7615</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.1</td>
<td>0.4836</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>0.9</td>
<td>0.5260</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>0.8</td>
<td>0.8944</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>0.7</td>
<td>0.9736</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>0.6</td>
<td>0.9958</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.5</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.4</td>
<td>0.9958</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.3</td>
<td>0.9736</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.2</td>
<td>0.8944</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>0.1</td>
<td>0.5260</td>
</tr>
</tbody>
</table>
Next, we have presented the values of $H_n^{\alpha,\beta}(P)$ graphically for $\alpha = 4$ and $\alpha = 2$ and obtained the above Fig.-2.4.2 which shows that the measure introduced in equation (2.4.3) is a concave function. It is further added that for different values of $\alpha$, we get similar curves of concave functions.
Concluding Remarks: In the existing literature of information theory, we find many probabilistic, non-probabilistic, parametric and non-parametric measures of entropy, each with its own merits, demerits and limitations. But, we have to develop those measures which can be successfully applied to a variety of disciplines. In Biological Sciences, we have observed that researchers frequently use Shannon’s (1948) measure for measuring diversity in different species. But, if we have a variety of information measures, then we shall be more flexible in applying a standard measure depending upon the situation. Keeping this idea in mind, we have developed some measures and concluded that for the known values of arithmetic mean, geometric mean, harmonic mean, power mean, and other measures of dispersion, the information content of a discrete frequency distribution can be calculated and consequently, new probabilistic information theoretic measures can be investigated and developed.