On Certainty and Generalized Information Measures

P. K. Bhatia

Department of Applied Sciences
DCR University of Science & Technology
Murthal – 131 039(Sonepat) India
bhatiapk@rediffmail.com

Abstract

By considering the relation between certainty and information, three generalized classes of information measures have been derived, which include the information measures studied by Van der Lubbe et al. [12] as their special cases.

1. INTRODUCTION

In information theory the quantification of information plays an important role. Since Shannon [11], defined his probabilistic information measure, many other authors, among them Renyi [10], Daroczy [5], Arimoto [2], Rathie et al. [9], and Kapur [7], have introduced new measures of information.

It has been customary to define measures of information by relating information to uncertainty. But Van der Lubbe et al. [12] have followed a different approach and chosen the relation between information and certainty as a starting point. They have introduced information measures with the help of certainty measures within this framework. This approach leads to three generalized classes of information measures and unifies the known measures of information into one generalized probabilistic theory of discrete information measures.

They have defined a generalized measure of average certainty as follows:

Consider a probabilistic experiment X, which has possible outcomes
$x_i \in \chi$, $i = 1, 2, \ldots, n$, $n > 1$. Furthermore, let $P = (p_1, p_2, \ldots, p_n)$ denote the set of probabilities associated with the outcomes of $X$. That means that $p_i$ is the probability that the outcome of the experiment $X$ will be $x_i \in \chi$. $P$ is an element of the set $A_n$ of all complete discrete probability distributions:

$$A_n = \{P = (p_1, p_2, \ldots, p_n) \mid p_i \geq 0, i=1, \ldots, n, \sum_{i=1}^{n} p_i = 1\}. \quad (1.1)$$

Any function $\Phi(p_i)$ can be used as a measure of certainty with respect to $x_i$ if $\Phi(p_i)$ satisfies the following properties:

1. $\Phi(p_i)$ is continuous on $[0, 1]$,
2. $\Phi(p_i)$ is monotonically increasing on $[0, 1]$,
3. for any $i$ and $j$,

$$\Phi(p_i, q_j) = \Phi(p_i) \cdot \Phi(q_j) \quad (1.2)$$

where $q_j$ is the probability that experiment $Y$ has outcome $y_j \in \gamma$, $j = 1, \ldots, m$.

The general solution of (1.2) is given by, Hardy et al. [6] is

$$\Phi(p_i) = p_i^a, \quad (1.3)$$

where $a$ is a real constant and is greater than zero. A measure of average certainty with respect to the probabilistic experiment $X, P \in A_n$, can be introduced by setting

$$f(P) = \Psi^{-1} \left[ \sum_{i=1}^{n} p_i \Psi(p_i^a) \right], \quad a > 0, \quad (1.4)$$

where $\Psi(.)$ is a strictly monotonic continuous function. However $\Psi(.)$ should also be such that $f(P)$ exists for all $P \in A_n$ and that it satisfies the property of multiplicativity.

In this communication, in Section 2 we define a generalized measure of average certainty and characterize it axiomatically. The properties and special
cases of this measure have been mentioned in Section 3. Section 4 deals with the relation between information and generalized measure of average certainty.

2. A GENERALIZED MEASURE OF AVERAGE CERTAINTY

We denote by \( P = (p_1, p_2, \ldots, p_n) \) the set of probabilities associated with the outcomes of a probabilistic experiment \( X \). \( P \) is an element of the set \( A_n \) of all generalized discrete probability distributions:

\[
A_n = \{ P = (p_1, p_2, \ldots, p_n) \mid p_i \geq 0, \ i=1, \ldots, n, \sum_{i=1}^{n} p_i = 1 \}.
\]

If \( \Phi(p_i) = p_i^\alpha \) be the measure of certainty with respect to an event \( x_i \). Then we shall consider measure of average certainty as mean of these certainty measures with weights as functions of \( p_i \), i.e.

\[
f(P) = \Psi^{-1} \left[ \frac{\sum_{i=1}^{n} \alpha(p_i) \Psi(p_i^\alpha)}{\sum_{i=1}^{n} \alpha(p_i)} \right], \quad (2.1)
\]

where \( \Psi \) is a strictly monotonic and continuous function and \( \alpha \) is a weight function. In particular, here we shall consider the case \( \alpha(p_i) = p_i^\gamma, \ \gamma > 0 \). The function \( \Psi \) should also be such that \( f(P) \) exists for all \( P \in A_n \) and that it satisfies the so-called property of multiplicativity i.e. for all \( X \) and \( Y \), where \( X \) and \( Y \) are independent stochastic experiment, it should hold that

\[
f(PQ) = f(P) \cdot f(Q), \quad (2.2)
\]

We now give the following characterization theorem:

**THEOREM 2.1.** Let \( f(P) = f(p_1, \ldots, p_n) \) satisfy the following conditions for all \( P \in A_n \)

\[
f(P) = \Psi^{-1} \left[ \frac{\sum_{i=1}^{n} p_i^\gamma \Psi(p_i^\alpha)}{\sum_{i=1}^{n} p_i^\gamma} \right], \quad \alpha > 0, \gamma > 0, \quad (2.3)
\]

where \( \Psi(.) \) is a continuous and strictly monotonic function;

\[
f(\prod) = f(P) \cdot f(Q), \quad (2.4)
\]
in the case that X and Y are stochastically independent experiments. Then \( f(P) \) is uniquely determined by

\[
\begin{align*}
\Psi^{-1} & \left[ \sum_{i=1}^{n} p_i^{a+b} \Psi \left( p_i^{a} n^{-a} \right) \right] = n^{-a} \Psi^{-1} \left[ \sum_{i=1}^{n} p_i^{\gamma} \Psi \left( p_i^{a} \right) \right] \\
\end{align*}
\] (2.7)

If we set, in (2.7), for fixed \( n \) and \( a \)

\[
\Psi \left( p_i^{a} n^{-a} \right) = \chi \left( p_i^{a} \right). \quad (2.8)
\]

We find

\[
\chi^{-1} \left[ \sum_{i=1}^{n} p_i^{\gamma} \chi \left( p_i^{-a} \right) \right] = \Psi^{-1} \left[ \sum_{i=1}^{n} p_i^{\gamma} \Psi \left( p_i^{a} \right) \right]. \quad (2.9)
\]
Following Hardy et al. [6], it is necessary and sufficient in order for (2.9) to be satisfied
\[ \chi \left( p_i^a \right) = \Psi \left( p_i^a n^{-a} \right) = A \left( n^{-a} \right) \Psi \left( p_i^a \right) + B \left( n^{-a} \right), A \neq 0. \quad (2.10) \]
Substitution of \( x = p_i^a \) and \( y = n^{-a} \) leads to
\[ \Psi \left( xy \right) = A \left( y \right) \Psi \left( x \right) + B \left( y \right). \quad (2.11) \]
The strictly monotonic and continuous solutions of (2.11) are given by Hardy et al. [6],
\[ \Psi \left( x \right) = \frac{x^b - 1}{c_i} + c_2, b \neq 0, c_i \neq 0 \text{ and } x \neq 0 \quad \text{for } b < 0, \quad (2.12) \]
\[ \Psi \left( x \right) = c_i \log x + c_2, x > 0, c_i \neq 0. \quad (2.13) \]

Now substituting (2.12) into (2.3), we get
\[ f(P) = \left[ \sum_{i=1}^{n} p_i^{a + b + \gamma} \right]^{1/b}, \text{ where } a > 0, \gamma > 0, \text{ and } b \neq 0, \]
similarly by substituting (2.13) into (2.3), we get
\[ f^*(P) = \prod_{i=1}^{n} p_i^{a\sigma + \gamma}, \text{ where } a > 0, \gamma > 0. \]

To make notations simple, we substitute \( ab + \gamma = \rho \) and \( 1/b = \sigma = \frac{a}{\rho - \gamma} \) and define generalized \((\rho, \sigma, \gamma)\) - measure as :

**Definition 2.1.** For \( P \in A_n \) and \((\rho, \sigma, \gamma) \in D\), where
\[ D = \{ (\rho, \sigma, \gamma) | 0 < \rho < \sigma, \sigma < 0, \rho > \gamma, \sigma > 0 \text{ and } \gamma > 0 \} \]
The generalized measures of average certainty with respect to \( X \) are defined by
\[ G_n(P; \rho, \sigma, \gamma) = \left[ \frac{\sum_{i=1}^{n} p_i^\rho}{\sum_{i=1}^{n} p_i^\gamma} \right]^\sigma \] (2.14)

and

\[ G_n^*(P; \rho, \sigma, \gamma) = \prod_{i=1}^{n} p_i^\gamma (\rho^\gamma - \gamma) \sigma \] (2.15)

3. PROPERTIES OF THE \((\rho, \sigma, \gamma)\) – CERTAINTY MEASURE

For some particular values of \(\rho, \sigma\), and \(\gamma\), the generalized measure of average certainty \(G_n(P; \rho, \sigma, \gamma)\) reduces to some special measures:

1. For \(\gamma = 1\), we obtain

\[ G_n(P; \rho, \sigma, 1) = \left[ \sum_{i=1}^{n} p_i^\rho \right]^\sigma \] (3.1)

which is \((\rho, \sigma)\) – certainty measure introduced by Van der Lubbe et al. [12].

2. For \(\rho > 1, \sigma = 1, \gamma = 1\), we find

\[ G_n(P; \rho, 1, 1) = \sum_{i=1}^{n} p_i^\rho \] (3.2)

This function is called as the certainty measure of type \(\rho\) [12]. For \(\rho = 2\), this reduces to measure of information energy [4,8].

3. Putting \(\sigma = 1/\rho, \rho > 1\) and \(\gamma = 1\), we obtain

\[ G_n(P; \rho, 1/\rho, 1) = \left[ \sum_{i=1}^{n} p_i^{\rho^\rho} \right]^{1/\rho} \] (3.3)

this measure is called the \(\rho\)-norm certainty measure [12].

4. Let \(\sigma = 1/\rho - 1, \rho > 0, \rho \neq 1\), and \(\gamma = 1\), then we find

\[ G_n(P; \rho, 1/\rho - 1, 1) = \left[ \sum_{i=1}^{n} p_i^{\rho^{\rho^\rho}} \right]^{1/\rho - \rho} \] (3.4)
On certainty and generalized information measures

the so called $\rho$ – mean certainty measure.

Now, we mention some properties of the $(\rho, \sigma, \gamma)$ – certainty measure, which are direct consequences of the definition of the certainty measure:

(i) $G_n(P; \rho, \sigma, \gamma)$ is a symmetric function of $P \in A_n$.

(ii) $G_n(P; \rho, \sigma, \gamma)$ is expansive for all $n > 1$,

$G_n(p_1, \ldots, p_n; \rho, \sigma, \gamma) = G_{n+1}(p_1, \ldots, p_n, 0; \rho, \sigma, \gamma)$ (3.5)

(iii) $G_n(P; \rho, \sigma, \gamma)$ is continuous on $A_n$.

4. RELATION BETWEEN INFORMATION AND THE GENERALIZED CERTAINTY MEASURES

We know, refer Van der Lubbe et al. [12], an information measure should be a monotonically decreasing function of the certainty measure and should be equal to zero if and only if the corresponding certainty measure is equal to one.

Let $G_n(P; \rho, \sigma, \gamma)$ be the $(\rho, \sigma, \gamma)$– certainty measure as given in Definition 2.1, then the three possible relations between (fulfilling the above mentioned condition) certainty measure and information measure can be given as

$1H_n(P; \rho, \sigma, \gamma) = - \log_2 [G_n(P; \rho, \sigma, \gamma)]$, (4.1)

$2H_n(P; \rho, \sigma, \gamma) = 1 - G_n(P; \rho, \sigma, \gamma)$, (4.2)

$3H_n(P; \rho, \sigma, \gamma) = \left[ \frac{1}{G_n(P; \rho, \sigma, \gamma)} \right] - 1$ (4.3)

By suitable substitutions of $\rho$, $\sigma$, and $\gamma$ in the definition of $1H_n(P; \rho, \sigma, \gamma)$ a large number of well-known measures of information can be obtained, a few of which are listed below:

(1) Since the information measure introduced by Van der Lubbe et al. [12] are particular cases of (4.1), (4.2) and (4.3), therefore all the particular classes of information measures obtained by them can be obtained here.

(2) Substitution of $\rho = \alpha$, $\rho = 1/\alpha$, and $\gamma = \beta$ in $1H_n(P; \rho, \sigma, \gamma)$ leads to

$1H_n(P; \alpha, 1/\alpha, \beta) = H_{\alpha,\beta}(P)$
\[(1/\alpha - \beta) \log \left( \frac{\sum_{i=1}^{n} P_i^\alpha}{\sum_{i=1}^{n} P_i^\beta} \right), \]

\(\alpha \neq \beta \) and \(\alpha, \beta > 0\). This measure is called the information measure of order \((\alpha, \beta)\). It was discussed by Aczél and Daroczy [3].

(3) Substituting \(\sigma = 1/1-\alpha, \gamma = \beta\) and \(\rho = \alpha + \beta - 1\) in \(1H_n(P; \rho, \sigma, \gamma)\) leads to Kapur’s [7] entropy of order \(\alpha\) and type \(\beta\)

\[1H_n(P; \alpha+\beta-1, 1/\alpha-1, \beta) = (1/1 - \alpha) \log \left( \frac{\sum_{i=1}^{n} P_i^{\alpha+\beta-1}}{\sum_{i=1}^{n} P_i^\beta} \right), \alpha \neq 1.\]

(4) For \(\sigma = 1/\alpha-1, \gamma = 1, \rho = \alpha\), we obtain Rényi’s [10] entropy for generalized probability distributions i.e.

\[1H_n(P; \alpha, 1/\alpha-1, 1) = \left( \frac{\sum_{i=1}^{n} P_i^\alpha}{\sum_{i=1}^{n} P_i} \right), \alpha \neq 1.\]

(5) For \(\sigma = 1/\alpha-1, \gamma = \beta_1 - 1, \rho = \alpha + \beta_1 - 1\), \(1H_n\) reduces to Rathie’s entropy [9],

\[1H_n(P; \alpha + \beta_1 - 1, 1/\alpha - 1, \beta_1 - 1) = (1/1 - \alpha) \log \left( \frac{\sum_{i=1}^{n} P_i^{\alpha+\beta_1-1}}{\sum_{i=1}^{n} P_i^{\beta_1-1}} \right), \beta_1 \geq 1, \alpha \neq 1.\]

**REFERENCES**


Received: December, 2009