TAMKANG JOURNAL OF MATHEMATICS
Volume 24, Number 4, Winter 1993

MEAN VALUE CHARACTERIZATION OF 'USEFUL' INFORMATION MEASURES

U. S. BHAKER AND D. S. HOODA

Abstract. In the present communication the generalized mean value characterization of 'useful' information and relative information measures has been studied. Some comparison theorems related to these measures have also been proved.

1. Introduction

Let (Ω, A, P) be a probability space of an experiment E with a finite measurable partition of events $\{E_1, E_2, \ldots, E_n\}$ $(n > 1)$ of Ω. Probabilities of these events are given by $p(E_i) = p_i > 0$ for every E_i such that $p \in \Delta_n$, where $\Delta_n = \{(p_1, p_2, \ldots, p_n);\ p_i > 0,\ \sum_{i=1}^{n} p_i = 1\}$. The different events $E_i s$ depend upon the experiment's goal or upon some qualitative characteristic of the physical system taken into consideration, that is, they have different weights or utilities. In order to distinguish the events $E_1, E_2, \ldots, E_n$ with respect to a given qualitative characteristic of physical system taken into account, ascribe to each event E_i a non-negative number $u(E_i) = u_i(> 0)$ directly proportional to its importance and call u_i, the utility of the event E_i. In general u_i is independent of p_i (seeLongo [5]).

Belis and Guiasu [2] characterized a quantitative-qualitative measure which was called the 'useful' information by [5] of the experiment E and is given as

$$H(p; u) = H(p_1, p_2, \ldots, p_n; u_1, u_2, \ldots, u_n)$$

Received December 6, 1991, revised February 10, 1993.

$$= -\sum_{i=1}^{n} u_i p_i \log p_i,\ u_i > 0,\ 0 < p_i \leq 1,\ \Sigma p_i = 1. \tag{1.1}$$

The measure (1.1) has been studied and generalized for complete probability distributions by many authors. We consider the following two measures of 'useful' information for generalized probability distribution

$$P = \{(p_1, p_2, \ldots, p_n),\ p_k > 0 \quad \text{and} \quad \sum_{k=1}^{n} p_k \leq 1\},$$

which is the probability distribution of a generalized random variable having utility distribution $U = \{(u_1, u_2, \ldots, u_n),\ u_i > 0\}$;

$$I(P;\mathcal{U}) = \frac{\sum_{k=1}^{n} u_k p_k \log \frac{1}{p_k}}{\sum_{k=1}^{n} u_k p_k}, \tag{1.2}$$

where $\sum_{k=1}^{n} u_k p_k$ is not necessarily ≤ 1 and

$$T_\alpha(P;\mathcal{U}) = \frac{1}{1-\alpha} \log \frac{\sum_{k=1}^{n} u_k p_k^\alpha}{\sum_{k=1}^{n} u_k p_k}. \tag{1.3}$$

It may be seen that (1.3) reduces to (1.2) when $\alpha \to 1$ and if utilities are ignored i.e. $u_i = 1$ for each i, the measures (1.2) and (1.3) reduce to Renyi's entropies of order 1 and α respectively. In section 2, we give mean value characterization of measures (1.2) and (1.3).

Further Taneja and Tuteja [9] considered two utility information schemes:-

$$S = \begin{bmatrix} E_1 & E_2 & \ldots & E_n \\ p_1 & p_2 & \ldots & p_n \\ u_1 & u_2 & \ldots & u_n \end{bmatrix}, \quad p_i > 0, \quad u_i > 0, \quad \sum_{i=1}^{n} p_i = 1,$$

of set of n events after experiment and

$$S^* = \begin{bmatrix} E_1 & E_2 & \ldots & E_n \\ q_1 & q_2 & \ldots & q_n \\ u_1 & u_2 & \ldots & u_n \end{bmatrix}, \quad q_i > 0, \quad u_i > 0, \quad \sum_{i=1}^{n} q_i = 1,$$

before experiment and characterized axiomatically a quantitative-qualitative measure of relative information as given below:

$$I(P|Q;U) = \sum_{i=1}^{n} u_i p_i \log p_i/q_i. \tag{1.4}$$

The measure (1.4) has been characterized and generalized for complete probability distribution by various authors. Here we consider the following two measures for P and Q posterior and priorior generalized probability distributions of an experiment having utility distribution U.

$$I(P|Q;\mathcal{U}) = \frac{\sum_{i=1}^{n} u_i p_i \log p_i/q_i}{\sum_{i=1}^{n} u_i p_i}, \tag{1.5}$$

and

$$I_\alpha(P|Q;\mathcal{U}) = \frac{1}{\alpha - 1} \log \frac{\sum_{i=1}^{n} u_i p_i^\alpha / q_i^{\alpha-1}}{\sum_{i=1}^{n} u_i p_i}, \quad \alpha \neq 1. \tag{1.6}$$

It may be noted that (1.6) reduces to (1.5) when $\alpha \to 1$. In case utilities are ignored, the measures (1.5) and (1.6) reduce to the information measures characterized by Sharma [7]. Mean value characterization of measures (1.5) and (1.6) has been studied in section 3. In section 4 we derive some comparison theorems.

2. Characterization of 'Useful' Information Measures

By considering a set of postulates Renyi [6] characterized the following measure of information concerning event E_k having probability of occurrence as p_k.

$$I(\{p_k\}) = \log \frac{1}{p_k}. \tag{2.1}$$

Let us deline $W_k = \dfrac{u_k p_k}{\sum_{k=1}^{n} u_k p_k}$, $k = 1, 2, \cdots, n$, (2.2) then mean value of (2.1) taking (2.2) as weights is

$$I(P;\mathcal{U}) = \frac{\sum_{k=1}^{n} u_k p_k \log \frac{1}{p_k}}{\sum_{k=1}^{n} u_k p_k},$$

which is (1.2). It may be been that (1.2) is weighted entropy when weights are taken

$$W_k = \frac{u_k p_k}{\sum_{k=1}^{n} u_k p_k}, \quad k = 1, 2, \ldots, n.$$

Further we see that (1.2) satisfies the following postulates:

Postulate 1. $H(P;U)$ is a symmetric function of the elements of P and U.

Postulate 2. If $\{p\}$ and $\{u\}$ denote the generalized probability distributions consisting of the single probability p, single utility u of an event E, then $H(\{p\};\{u\})$ is continuous funct of p and u for $0 < p \leq 1$. It may be noted that the continuity of $H(\{p\},\{u\})$ is supposed only for $p > 0$ but not for $p = 0$.

Postulate 3. $H(\{1/2\};\{1\}) = 1$ and

$$H(\{1\};\{u\}) = 0,$$

that is, the measure of 'useful' information is unity when $p = 1/2$ and $u = 1$ and no useful information is conveyed when p.

Postulate 4.

$$H(P^*Q; U * V) = H(P;U) + H(Q;V),$$

where

$$P = \{(p_1, p_2, \ldots, p_n); \quad \sum_{i=1}^{n} p_i \leq 1\},$$

$$Q = \{(q_1, q_2, \ldots, q_m); \quad \sum_{j=1}^{m} q_j \leq 1\}.$$

$$
\begin{aligned}
U &= \{(u_1, u_2, \ldots, u_n),\ u_i > 0\}, \\
V &= \{(v_1, v_2, \ldots, v_m),\ v_j > 0\}, \\
P^*Q &= (p_1q_1, p_1q_2, \ldots, p_1q_m, \ldots, p_nq_1, p_nq_2, \ldots, p_nq_m), \\
U^*V &= (u_1v_1, u_1v_2, \ldots, u_1v_m, \ldots, u_nv_1, u_nv_2, \ldots, u_nv_m).
\end{aligned}
$$

Postulate 5. There exists a continuous and strictly increasing function $y = g(x)$ defined for real x, such that, its inverse function is given by $x = g^{-1}(y)$. If $P = (p_1 \cup p_2 \cup \ldots \cup p_n)$ and $\mathcal{U} = (u_1 \cup u_2 \cup \ldots \cup u_n)$, then

$$
H(P;U) = g^{-1}\left[\frac{\sum_{k=1}^{n} u_k p_k g(I\{p_k\})}{\sum_{k=1}^{n} u_k p_k}\right]. \tag{2.3}
$$

It is an open question which choices of function $g(x)$ are admissible such that postulate 5 is compatible with postulate 4. One form of $g(x)$ clearly is $g(x) = ax + b$ with $a \neq 0$, then the information measure satisfying postulate 1 to 5 with this form of $g(x)$ will be (1.2).

Another choice of $g(x)$ which is admissible is an expenential function. If $g(x) = g_\alpha(x)$, where $\alpha > 0$, $\alpha \neq 1$ and $g_\alpha(x) = 2^{(1-\alpha)x}$. Then postulates 1 to 5 characterize the weighted entropy of order α. Thus we give the above result as theorem given below:

Theorem 1. *If $H(P;U)$ is defined for all $P = \{(p_1, p_2, \ldots, p_n);\ p_i > 0$ and $\Sigma p_i \leq 1\}$ and $U = \{(u_1, u_2, \ldots, u_n),\ u_i > 0\}$ and satisfies the postulates 1 to 5 with $g(x) = g_\alpha(x)$ where $g_\alpha(x) = 2^{(1-\alpha)x}$, $\alpha > 0$ and $\alpha \neq 1$, then*

$$
H_\alpha(P;U) = \frac{1}{1-\alpha}\log\left(\frac{\sum_{k=1}^{n} u_k p_k^\alpha}{\sum_{k=1}^{n} u_k p_k}\right), \quad \alpha \neq 1. \tag{2.4}
$$

3. Characterization of 'Useful' Relative Information Measures

Let $P = \{(p_1, p_2, \ldots, p_n),\ p_k > 0$ for $k = 1, 2, \ldots, n$ and $\sum_{k=1}^{n} p_k \leq 1\}$ and $Q = \{(q_1, q_2, \ldots, q_n),\ q_k > 0$ for $k = 1, 2, \ldots, n$ and $\sum_{k=1}^{n} q_k \leq 1\}$ be posterior and priorior generalized probability distributions of a random variable in an experiment. Let $U = \{(u_1, u_2, \ldots, u_n), u_k > 0\}$ be utility distribution such that u_i is only value or importance of event E_i in reference to some specific goal.

By considering a set of postulates Rényi [6] characterized the following measure of amount of information concerning an event E_k having posterior and priorior probability as p_k and q_k.

$$I(\{p_k\}/\{q_k\}) = \log_2 \frac{p_k}{q_k}. \tag{3.1}$$

Let us define $W_k = \dfrac{u_k p_k}{\sum_{k=1}^{n} u_k p_k}$, $k = 1, 2, \cdots, n$ as weights then (1.5) can be written as weighted amount of information as

$$I(P/Q; U) = \frac{\sum_{k=1}^{n} u_k p_k \log_2 p_k/q_k}{\sum_{k=1}^{n} u_k p_k}.$$

It implies that (1.5) is nothing but weighted 'useful' relative information.

Now we assume that (1.5) satisfies the following postulates:

Postulate 6. $I(P/Q; U)$ remains unchanged if the elements of P, Q and U are rearranged in the same way so that one-one correspondance between them is not disturbed.

Postulate 7. $I(P/Q; U)$ is a continuous function of p_k, q_k and u_k for $k = 1, 2, \cdots, n$.

Postulate 8. $I(\{1\}/\{1/2\}; \{1\}) = 1$

Postulate 9.

$$I(P * P'/Q^* Q'; U * U') = I(P/Q; U) + I(P'/Q'; U'). \tag{3.2}$$

where

$$P * P' = (p_1p'_1, p_1p'_2, \ldots, p_1p'_m, \ldots, p_np'_1, p_np'_2, \ldots, p_np'_m),$$
$$Q^*Q' = (q_1q'_1, q_1q'_2, \ldots, q_1q'_m, \ldots, q_nq'_1, q_nq'_2, \ldots, q_nq'_m), \text{ etc.}$$

Postulate 10. There exists a continuous and strictly increasing function $y = g(x)$ defined for all real x such that its inverse function is given by $x = g^{-1}(y)$. If

$$P = (P_1 \cup P_2 \cup \ldots \cup P_n), \quad Q = (Q_1 \cup Q_2 \cup \ldots \cup Q_n)$$

and $U = (U_1 \cup U_2 \cup \ldots \cup U_n)$, then

$$I(P/Q;U) = g^{-1}\left(\frac{\sum_{k=1}^{n} u_kp_kg(I[\{p_k\}/\{q_k\}])}{\sum_{k=1}^{n} u_kp_k}\right), \tag{3.3}$$

where $I(\{p_k\}/\{q_k\}) = \log_2 \frac{p_k}{q_k}$ for all $k = 1, 2, \cdots, n$.

Next, we consider what possible choices of the function $g(x)$ are compatible with postulate 9. It follows from postulate 9 that for any $\lambda \geq 0$ and $\mu \geq 0$ we have

$$I(P^*\{Q^{-\lambda}\}/Q^*\{Q^{-\mu}\};U) = I(P/Q;U) + \mu - \lambda. \tag{3.4}$$

Thus putting $\mu - \lambda = y$, we see that for an arbitrary real y, we have

$$g^{-1}\left[\frac{\sum_{k=1}^{n} u_kp_kg(\log_2 \frac{p_k}{q_k} + y)}{\sum_{k=1}^{n} u_kp_k}\right] = g^{-1}\left[\frac{\sum_{k=1}^{n} u_kp_kg(\log_2 \frac{p_k}{q_k})}{\sum_{k=1}^{n} u_kp_k}\right] + y. \tag{3.5}$$

If $\frac{u_kp_k}{\sum_{k=1}^{n} u_kp_k} = W_k$ and $\log_2 \frac{p_k}{q_k} = x_k$, for $k = 1, 2, \cdots, n$, (3.6)

then $W_1, W_2, \ldots, W_n$ is a sequence of positive numbers such that $\sum_{k=1}^{n} W_k = 1$ and $x_1, x_2, \ldots, x_n$ is any sequence of real numbers.

On substituting (3.6) in (3.5) we have

$$g^{-1}\left[\sum_{k=1}^{n} W_k g(x_k+y)\right] = g^{-1}\left[\sum_{k=1}^{n} W_k g(x_k)\right] + y. \tag{3.7}$$

If

$$g_y(x) = g(x+y), \tag{3.8}$$

then (3.7) can be expressed in the following form

$$g_y^{-1}\left[\sum_{k=1}^{n} W_k g_y(x_k)\right] = g^{-1}\left[\sum_{k=1}^{n} W_k g(x_k)\right]. \tag{3.9}$$

It implies that $g(x)$ and $g_y(x)$ generate the same mean value and this is possible only if $g_y(x)$ is a linear function of $g(x)$ refer (Theorem 8[3]) i.e. there exists constants $a(y) \neq 0$ and $b(y)$ such that

$$g_y(x) = g(x+y) = a(y)g(x) + b(y). \tag{3.10}$$

Without restricting the generality we may suppose $g(0) = 0$. Thus we obtain $b(y) = g(y)$ and

$$g(x+y) = a(y)g(x) + g(y). \tag{3.11}$$

Since (3.11) is true for any x and y therefore we may interchange the roles of x and y. Thus we get

$$g(x+y) = a(x)g(y) + g(x). \tag{3.12}$$

If $x \neq 0$ and $y \neq 0$ then (3.11) and (3.12) together give

$$\frac{a(y)-1}{g(y)} = \frac{a(x)-1}{g(x)} = K(\text{say})$$

It implies

$$a(x) - 1 = K\, g(x), \tag{3.13}$$

for all real x. Two cases arise:

Case (i). When $K = 0$, (3.13) gives $a(x) = 1$ and from (3.11) we obtain

$$g(x+y) = g(x) + g(y), \tag{3.14}$$

which is Cauchy's functional equation and has the solution $g(x) = cx$, where $c \neq 0$ is a constant. In this case from (3.3) we have

$$I(P/Q;U) = \frac{\sum_{k=1}^{n} u_k p_k \log_2 \frac{p_k}{q_k}}{\sum_{j=1}^{n} u_j p_j}.$$

Cases (ii). When $K \neq 0$, the substitution of (3.13) into (3.11) yields

$$a(x+y) = a(x)a(y), \tag{3.15}$$

for any real x and y.

Now (3.13) shows that $a(x)$ is monotonic and so from (3.15) it follows that $a(x)$ is an exponential function and can be written in the following form:

$$a(x) = c\,2^{(\alpha-1)x}, \tag{3.16}$$

where $\alpha > 0$ ($\neq 1$) and $c \neq 0$ are constants. It follows from (3.13) that

$$g(x) = \frac{c\,2^{(\alpha-1)x} - 1}{K}. \tag{3.17}$$

On substituting (3.17) in (3.3) we obtain (1.6), thus we have proved a theorem:

Theorem 2. *The useful relative information measures satisfying postulates* 6 *to* 10 *are only of the form given by* (1.5) *and* (1.6).

4. Comparison Theorems

Let $P^\alpha = (p_1^\alpha, p_2^\alpha, \ldots, p_k^\alpha)$, $p_k^\alpha > 0$ for $\alpha > 0$, be the power distribution of p. We shall derive a comparison result involving the useful information measure of order α given by (1.3).

Theorem 3. *With the usual notations the following inequality holds:*

$$I(P^\alpha;U) \le \alpha I_\alpha(P;U) \quad for \quad \alpha \ge, \tag{4.1}$$

Proof. We have the inequality (refer Beckenbach and Bellman [2]; p.17)

$$\log \frac{\sum_{j=1}^{n} a_j b_j}{\sum_{j=1}^{n} b_j} \le \frac{\sum_{j=1}^{n} a_j b_j \log a_j}{\sum_{j=1}^{n} a_j b_j}, \tag{4.2}$$

for $a_j > 0$ and $b_j > 0$, with equality iff all a_j's are equal.

Setting $a_j = p_j^{\alpha-1}$ and $b_j = u_j p_j$, u_j, $p_j > 0$ for each j, in (4.2) we get

$$\log \frac{\sum_{j=1}^{n} u_j p_j^\alpha}{\sum_{j=1}^{n} u_j p_j} \le \frac{\sum_{j=1}^{n} u_j p_j^\alpha \log p_j^{\alpha-1}}{\sum_{j=1}^{n} u_j p_j^\alpha}$$

or

$$\log \frac{\sum_{j=1}^{n} u_j p_j^\alpha}{\sum_{j=1}^{n} u_j p_j} \le (\alpha-1) \sum_{j=1}^{n} u_j p_j^\alpha \log p_j$$

or

$$\frac{1}{\alpha-1} \log \frac{\sum_{j=1}^{n} u_j p_j^\alpha}{\sum_{j=1}^{n} u_j p_j} \le \frac{1}{\alpha} I(P^\alpha;U).$$

or

$$I(P^\alpha;U) \le \alpha I_\alpha(P;U).$$

Hence the theorem is proved.

Let $U^\alpha = (u_1^\alpha, u_2^\alpha, \ldots, u_n^\alpha)$, $u_j^\alpha > 0$ for $\alpha > 0$, be the power distribution of U. Then (1.3) becomes

$$I_\alpha(P;U^\alpha) = \frac{1}{1-\alpha} \log \frac{\sum_{j=1}^{n} u_j^\alpha p_j^\alpha}{\sum_{j=1}^{n} u_j^\alpha p_j}. \tag{4.3}$$

Now we shall obtain a comparison result involving 'useful' information measure of order α of power distribution U^α and 'useful' information measure $H(P;U) = -2^{\alpha-1}\sum_{j=1}^{n}(u_j p_j)^\alpha \log p_j$ of type $\alpha(>0)$ studied by Hooda and Tuteja [4].

Theorem 4. *With usual notations the following inequality holds:*

$$I^\alpha(P;U) \le \overline{U}_\alpha I_\alpha(P;U), \quad \alpha > 0, \tag{4.4}$$

$$\textit{where} \quad \overline{U}_\alpha = 2^{\alpha-1}\sum_{j=1}^{n} u_j^\alpha p_j^\alpha.$$

Proof. Setting $a_j = p_j^{\alpha-1}$ and $b_j = u_j^\alpha p_j$, $u_j, p_j > 0$, in the inequality (4.2) we get

$$\log \frac{\sum_{j=1}^{n} u_j^\alpha p_j^\alpha}{\sum_{j=1}^{n} u_j^\alpha p_j} \le \frac{\sum_{j=1}^{n} u_j^\alpha p_j^\alpha \log p_j^{\alpha-1}}{\sum_{j=1}^{n} u_j^\alpha p_j^\alpha}.$$

Or

$$\log \frac{\sum_{j=1}^{n} u_j^\alpha p_j^\alpha}{\sum_{j=1}^{n} u_j^\alpha p_j} \le \frac{(\alpha-1)\sum_{j=1}^{n} u_j^\alpha p_j^\alpha \log p_j}{\sum_{j=1}^{n} u_j^\alpha p_j^\alpha}$$

or

$$\frac{1}{1-\alpha}\log \frac{\sum_{j=1}^{n} u_j^\alpha p_j^\alpha}{\sum_{j=1}^{n} u_j^\alpha p_j} \ge \frac{-2^{\alpha-1}\sum_{j=1}^{n} u_j^\alpha p_j^\alpha \log p_j}{2^{\alpha-1}\sum_{j=1}^{n} u_j^\alpha p_j^\alpha}$$

or

$$I_\alpha(P;U^\alpha) \ge \frac{I^\alpha(P;U)}{\overline{U}_\alpha}$$

or

$$I^\alpha(P;U) \le \overline{U}_\alpha I_\alpha(P;U^\alpha),$$

which is (4.4). This completes the proof.

Corollary. *If $u_j = 1$ for each j, then*

$$I_\alpha(P) \le \overline{U}_\alpha I_\alpha(P), \tag{4.5}$$

where $\overline{U}_\alpha = 2^{\alpha-1} \sum_{i=1}^{n} p_j^\alpha$. *This is a comparison result between Renyi's* [6] *entropy and the generalized entropy studied by Sharma and Taneja* [8].

References

[1] E. F. Beckenbach and R. Bellman, "Inequalities", *Springer-Verlag*, New York, 1971.

[2] M. Belis and S. Guiasu, "A quantitative-qualitative Measure of Information in Cybernetic Systems", *IEEE Trans. Information Theory*, 14 (1968), 593-594.

[3] G. H. Hardy, I. E. Littlewood and G. Polya, "Inequalities", *Combridge University Press*, Cambridge, 1934.

[4] D. S. Hooda and R. K. Tuteja, "Two generalized measures of 'useful' Information", *Information Sciences*, Vol. 23, (1981), 1, 11-19.

[5] G. Longo, "Quantitative-Qualitative Measure of Information", *Springer-Verlag*, New York, (1972).

[6] A. Renyi, "On measure of Entropy and Information", *Proc. 4th Berkeley Symp. on Math. Stat. and Probability*, University of California Press (1966), 547-561.

[7] B. D. Sharma, "The Mean Value of quantities in Information Theory", *Ph. D. Thesis*, University of Delhi, (1970).

[8] B. D. Sharma and I. J. Taneja, "Three generalized additive measures of Entropy", *Elecktronische Informations-verabeitung und Kybernetik*, 18, 13, (1977), 419-433.

[9] H. C. Taneja and R. K. Tuteja, "On characterization of a quantitative-qualitative measure of relative Information", *Information Sciences*, 33 (1984), 217-222.

Department of Mathematics, Goverment College, Bhiwani-125021, India.

Department of Mathematics and Statistics, Haryana Agricultural University, Hisar-125004, India.