Indian J. pure appl. Math., 13(6): 622–632, June 1982

TIME-DEPENDENT INFORMATIONS

SAMIR KUMAR MUKHERJEE
Department of Mathematics, Krishna Chandra College, P.O. Hetampur Rajbati, Dist. Birbhum, West Bengal

(*Received* 13 *May* 1981; *after revision* 22 *December* 1981)

Applying three methods, namely, substitution, summation and integration, in a few basic primitive functions constructed from two time-t functional equations, Sharma-Taneja entropy of type (a,b), Sharma-Mittal H^* entropies and Gupta-Sharma non-additive generalized inaccuracies are derived as particular cases of certain generalized information measures interpreted as information losses of the first and the second kinds over a closed time-interval $[a,b]$.

1. INTRODUCTION

Characterization processes of information measures become easy, if we can find certain basic primitive functions to be constructed from suitable functional equations containing a single power-parameter t and if certain convenient methods be applied in those functions. Further, some new interpretations of information measures may be obtained, if the parameter t be permitted to play the role of time

First, we are choosing a functional equation whose solution is the product of both additive and non-additive functions. Accordingly, let us consider

$$\phi(yv \parallel xu;t) = vu^{t-1}\phi(y \parallel x;t) + yx^{t-1}\phi(v \parallel u;t), \quad (x,y,u,v>0, -\infty<t<+\infty). \quad \ldots(1.1)$$

We call $\phi(y \parallel x;t)$ an elementary information of x relative to y at time t. A function $\phi(x;t) \equiv \phi(x \parallel x;t)$ is called self-relative information at time t. The solution of (1.1) is

$$\phi(y \parallel x;t) = Cyx^{t-1} \log_2 x = y\phi(x;t-1), \quad \ldots(1.2)$$

where C depends upon the initial values of x,y and t. Now,

$$\phi(x;t) \equiv \phi(x \parallel x;t) = x\phi(x;t-1) = Cx^t \log_2 x, \text{ (cf. Mittal 1975)} \quad \ldots(1.3)$$

Clearly, $\phi(y \parallel x;t) = \phi(x;0)w(y \parallel x;t)$ ...(1.4)

where, $\phi(x;0) = C \log_2 x$ (cf. Shannon 1948) ...(1.5)

and $w(y \parallel x;t) = yx^{t-1} = yw(x,t-1)$. ...(1.6)

$\therefore$ $w(x;t) \equiv w(x \parallel x;t) = x^t$ (cf. Sharma and Taneja 1975) ...(1.7)

But for variability of t, we modify $w(x;t)$ by

$$v(x;t) = Ax^t = Aw(x;t), \quad (A = \text{some constant}) \quad \ldots(1.8)$$

Second, we take a linear-exponential product function satisfying,

$$\psi(x+y;t) = \psi(x;t)2^{(1-t)y} + \psi(y;t)2^{(1-t)x}, \quad (x,y>0, \ -\infty<t<+\infty) \quad \ldots(1.9)$$

where $\psi(x;t) = Bx2^{(1-t)x}$ (B=some constant) ...(1.10)

which we call Cauchy-Renyi product function.

From above, we choose the following basic primitive functions:

$$\phi(y \parallel x;t);\ \phi(x;t);\ \phi(x;0);\ v(x;t);\ \psi(x;t) \quad \ldots(1.11)$$

in which we shall apply the following methods:

(i) Substitution for x,y which may be functions of several variables.

(ii) Summation over x,y.

(iii) Integration with respect to t over $[a,b]$.

Method (iii) defines a function called information loss which interpretes certain non-additive measures of information in a new way. Method (i) enables one to characterize an information measure quickly and conveniently. Method (ii) is common to (i) and (iii). Thus the process involves as a whole the three methods (i), (ii) and (iii).

By this process, as special studies, we have shown that Sharma-Taneja (1975) entropy of type (a,b), Sharma-Mittal (1975) H^* entropies and Gupta-Sharma (1976) non-additive inaccuracies could be derived as particular cases of certain generalized information losses over $[a,b]$ denoted by H_2 (. ; $[a,b]$), H_2^* (. ;$[a,b]$) and $H(x; [a,b])$. Furthermore, as natural consequences of these losses, some important properties and interpretations have been explored through certain results, theorems and inequalities.

2. Information Losses of the First Kind

In general, a definite integral over $[a,b]$ of any information function $\Phi(x,y;t)$ which is integrable on $a \leqslant t \leqslant b$ may be called an information loss of (x,y) over $[a,b]$.

We shall describe some particular types of information losses over $[a,b]$ referred to the basic primitives in (1.11) of which the first one we choose is $\phi(y \parallel x;t)$ and correspondingly the definite integral we have is called information loss of the first kind.

Definition 1—The loss of information of x relative to y over a closed time-interval $[a,b]$ referred to the basic primitive $\phi(y \parallel x;t)$ is defined by

$$F_2(y \parallel x;[a,b]) = \int_a^b \phi(y \parallel x;t)dt = Cy\log_2 x \int_a^b x^{t-1}dt$$
$$= (2^{1-a}-2^{1-b})^{-1}\, y\, (x^{a-1}-x^{b-1}),\ (0<x,y\leqslant 1,\ a\neq b) \qquad \ldots(2.1)$$

where, $F_2(\{\tfrac{1}{2}\} \parallel \{\tfrac{1}{2}\};[a,b]) = \tfrac{1}{2}$.

In particular,

$$F_2(x;[a,b]) = (2^{1-a}-2^{1-b})^{-1}\,(x^a - x^b) \qquad \ldots(2.2)$$

(cf. Sharma and Taneja 1975 and Mittal 1975).

The following probability distributions are necessary:

$$\text{Let } \Delta = \left\{ G_u = (g_{u1}, g_{u2}, \ldots, g_{ui_u});\ 0<g_{uk}\leqslant 1;\ W(G_u) = \sum_{k=1}^{i_u} g_{uk} \leqslant 1;\ u=1,2,3,4,\ldots \right\}$$

where G_u may also be a variable element of Δ for any specified $u=n$. We shall use $G_u = P_u$, when G_u is complete.

Definition 2—The average information loss of $G_n \in \Delta$ relative to $Q = (q_1, q_2, \ldots, q_{i_n}) \in \Delta$ over $[a,b]$ referred to the basic primitive $\phi(y \parallel x;t)$ is defined by

$$H_2(Q \parallel G_n;[a,b]) = \sum_1^{i_n} F_2(q_k \parallel g_{nk};[a,b])/w(Q)$$

(*equation continued on p.* 624)

$$= (2^{1-a} - 2^{1-b})^{-1} \left(\sum_{1}^{i_n} q_k g_{nk}^{a-1} - \sum_{1}^{i_n} q_k g_{nk}^{b-1}\right)/w(Q), \; (a \neq b) \quad \ldots(2.3)$$

where F_2 (. ; $[a,b]$) is defined in (2.1). In particular, if $Q = G_n$, then

$$H_2(G_n; [a,b]) = (2^{1-a} - 2^{1-b})^{-1} \left(\sum_{1}^{i_n} g_{nk}^{a} - \sum_{1}^{i_n} g_{nk}^{b}\right) \Big/ w(G_n). \quad \ldots(2.4)$$

Clearly, Havrda-Charvat (1967) entropy or Daroczy (1970) entropy of type a or b is information loss over [a,1] or $[1,b]$ and Sharma-Taneja (1975) entropy of type (a,b) is also the same, but over $[a,b]$. All these measures are, therefore, the particular cases of (2.3) for $Q = G_n = P_n$ and $a > 0$, $b > 0$, $a \neq b \neq 1$; whereas, in our case, in general, $Q \neq G_n \neq P_n$ and $-\infty < a,b < +\infty$, $a \neq b$. Moreover, the time concept and the loss-concept reorient these previous measures completely in a new interpretation, while, we further interprete $t \gtrless 0$ as future time or past time correspondingly.

While exhibiting properties of information loss functions, we shall simply give the statements of theorems with hints for proofs where necessary. This becomes unavoidable to save space.

PROPERTIES: (i) WEIGHTED ADDITIVITY AND SYMMETRY—This is shown in the following :

Theorem 1—If $Q = (q_1, q_2, \ldots, q_{i_M}) \in \Delta$; $R = (r_1, r_2, \ldots, r_{i_N}) \in \Delta$; $G_M, G_N \in \Delta$;
$Q^*R = (q_1 r_1, q_1 r_2, \ldots, q_k r_j, \ldots, q_{i_M} r_{i_N}) \in \Delta$;
$G_M^* G_N = (g_{M1} g_{N1}, g_{M1} g_{N2}, \ldots, g_{Mk} g_{Nj}, \ldots, g_{M i_M} g_{N i_N}) \in \Delta$, then
$H_2(Q^*R \parallel G_M^* G_N; [o,b])$

$$= \left(\sum_{1}^{i_N} r_j g_{nj}^{a-1} \Big/ w(R)\right) H_2(Q \parallel G_M; [a,b]) + \left(\sum_{1}^{i_M} q_k g_{MK}^{b-1} \Big/ w(Q)\right) H_2(R \parallel G_N; [a,b]) \quad \ldots(2.5)$$

which is symmetrical not only in (a,b), but also, in (Q,R) and (G_M, G_N). This type of symmetry characterizes $H_2(Q \parallel G_n; [o,b])$ at once from (2.5). Its proof follows from simple computations.

(ii) LINEAR COMBINATION—This property comes out as a consequence of what follows: From (2.3), the information loss of G_n relative to Q over $[a,1]$ or $[1,b]$ is given by

$$H_1(Q \parallel G_n; \theta) = (2^{1-\theta} - 1)^{-1} \left(\sum_{1}^{i_n} q_k \; g_{nk}^{\theta-1} - w(Q)\right) / w(Q), \; (\theta = a,b; \; \theta \neq 1). \quad \ldots(2.6)$$

Let us now introduce the following:

$$C(\theta) = 2^{1-\theta} - 1, \; (\theta = a,b; \; \theta \neq 1); \; T(a,b) = [C(a) - c(b)]^{-1};$$

$$C(a,b) = \frac{C(a)}{C(a) - C(b)}; \; C(b,a) = \frac{C(b)}{C(b) - C(a)}, \; (a \neq b). \quad \ldots(2.7)$$

Then the linear combination property of $H_2(Q \parallel G_n; [a,b])$ is given by

$$H_2(Q \parallel G_n; [a,b]) = C(a,b) \; H_1(Q \parallel G_n; a) + C(b,a) \; H_1(Q \parallel G_n; b). \quad \ldots(2.8)$$

(iii) REPRESENTATION – By the 'representation property' we shall mean that an information loss may be expressed as a function of some other information measure. Here $H_2(Q \parallel G_n;[a,b])=T(a,b)[\{1+C(a)\}\, H^{a(Q\parallel G_n)}-\{1+C(b)\}H^{b(Q\parallel G_n)}]$ ...(2.9)

$$\text{where, } H^{\theta}(Q \parallel G_n) = (1-\theta)^{-1} \log_2\Big(\sum_{1}^{i_n} q_k g_{nk}^{\theta-1}/w(Q)\Big), \; (\theta=a,b;\, \theta>0,\; \theta\neq 1) \quad ...(2.10)$$

which is a known inaccuracy (ref. Nath 1968, 1970). In fact, the result (2.10) is a particular case of a time-t information of $(Q \parallel G_n)$ relative to time t_0, given by

$$I_n(Q \parallel G_n;(t,t_0)) = -\frac{1}{t-t_0}\log_2\left(\sum_{1}^{i_n} q_k\, g_{nk}^{t-t_0}\Big/ w(Q)\right), \; (-\infty<t>+\infty,\; t\neq t_0) \quad ...(2.11)$$

which is obtained by putting $x=\sum_{1}^{i_n} q_k\, g_{nk}^{t-t_0}/w(Q)$ in $\phi(x;0)$, where $0<x\lessgtr 1$, according as $t\gtrless t_0$ with $\phi[(\frac{1}{2})^{t-t_0};\, 0] = 1 \Rightarrow C = -\dfrac{1}{t-t_0}$, and clearly,

$$I_n(Q \parallel G_n) = \lim_{t\to t_0} I_n\,(Q \parallel G_n;\,(t,t_0)) = \lim_{a\to b=1} H_2(Q \parallel G_n;[a,b])$$

$$= \sum_{1}^{i_n} q_k \log_2 \frac{1}{g_{nk}}\Big/ w(Q) = H'\,(Q \parallel G_n) \quad ...(2.12)$$

which is a time-free information (ref. Shannon 1948 for $Q=G_n=P_n$; Kerridge 1961, for $Q=Q_n$ (complete). $G_n=P_n$; and Renyi 1961, for $Q=G_n$).

(iv) CONVEXITY—Let $L(Q \parallel G_n;\,[a,b])$ denote a general information loss function of G_n relative to Q over $[a,b]$. In proving restricted convexity-theorems on all possible forms of L, it is always assumed $a<b\Rightarrow T(a,b)>0$ and $p_k=q_k/w(Q)>0$, where $\sum_{1}^{i_n} p_k=1 \;\forall\; Q$ of size i_n. A suitable function $f(x)$, $x>0$, is assumed such that

$$L(Q \parallel G_n;[a,b]) = \sum_{1}^{i_n} p_k f(x_k),\; (x_k=g_{nk},\; g_{nk}^{\delta-1},\; \log_2\frac{1}{g_{nk}};\; \delta>0,\; \delta\neq 1). \quad ...(2.13)$$

Furthermore, $(\sum_{1}^{i_n} p_k x_k)^{\theta} \gtrless \sum_{1}^{i_n} p_k x_k^{\theta}$, according as $\theta \gtrless 1$. ...(2.14)

Thus for appropriate restricted conditions on a,b and δ, the loss L is said to be a convex $\cup$ or $\cap$ function of G_n relative to all Q of size in, according as

$$f\,(\sum_{1}^{i_n} p_k x_k) \lessgtr \sum_{1}^{i_n} p_k\, f\,(x_k). \quad ...(2.15)$$

In fact, G_n of size i_n is a variable element of Δ and hence L is convex $\cup$ or $\cap$ in the variables $g_{n1}, g_{n2},\ldots,g_{ni_n}$ with respect to any specified $P_n=(p_1, p_2,\ldots p_{i_n}) \in \Delta$, derived from any $Q\in\Delta$.

Theorem 2—H_2 ($Q \parallel G_n$; $[a,b]$) is a convex $\cap$ function of G_n for $a<2<b$.
Hint: Assume $f(x)= T(a,b)\,(x^{a-1}-x^{b-1})$, $(x>0;\ a\neq b)$. ...(2.16)

3. Information Losses of the Second Kind

Let us define

$$R_n(1)=H^{\delta}(Q \parallel G_n) \text{ or } H^{\delta}(G_n),\ (\delta>0,\ \delta\neq1),\ (\text{by } 2.10) \quad ...(3.1)$$

$$R_n(2)=H^{1}(Q \parallel G_n) \text{ or } H^{1}(G_n),\ (\delta=1),\ (\text{by } (2.12) \quad ...(3.2)$$

according as $Q\neq G_n$ or $Q=G_n$.

The definite integral over $[a,b]$ referred to the basic primitive $\nu(x;t)$ or $\phi(x;t)$ or $\psi(x;t)$ with suitable substitutions for x will be called information losses of the second kind.

Below we are showing how the same loss is characterized by three different ways.

(i) Putting $x=[R_n(s)]^{1/t(1/t-1)\ R_n(s)}>0$, $(t\neq0;\ s=1,2)$ in the basic primitive $\nu(x;t)$ and then using iii, we have

$$H_2^{(s)}(Q \parallel G_n;\ [a,b])=T(a,b)\,[\{1+c(a)\}^{R_n(s)}-\{1+c(b)\}^{R_n(s)}],(a\neq b;\ s=1,2) \quad ...(3.3)$$

$$=\frac{c(a,b)}{c(a)}[1+c(a)]^{R_n(s)}+\frac{c(b,a)}{c(b)}[1+c(b)]^{R_n(s)}, \quad ...(3.4)$$

where $H_2^{(s)}(\{\tfrac{1}{2}\} \parallel \{\tfrac{1}{2}\};\ [o,b])=1$. In particular, the loss over $[a,1]$ or $[1,b]$ is

$$H^{1(s)}(Q \parallel G_n;\ \theta)=C^{-1}(\theta)[\{1+C(\theta)\}^{R_n(s)}-1],\ (\theta=a,b;\ \theta\neq1;\ s=1,2) \quad ...(3.5)$$

which was previously obtained by Sharma and Mittal (1975) for $Q=G_n=P_n$ as H^* entropies in a different lengthy method and in a different form and later by Gupta and Sharma (1976) for $Q\neq G_n$ as generalized inaccuracies by their method.

(ii) Again, had we put $x=\sum_{1}^{in} q_k\, g_{nk}^{\delta-1}\big/ w(Q)\gtrless 1$, according as $\delta\lesseqgtr 2$, in $\phi(x;t)$ and used (iii), then, as before, we have

$$H_2^{*(s)}(Q \parallel G_n;[a,b])=[2^{(1-\delta)a}-2^{(1-\delta)b}]^{-1}\,[2^{(1-\delta)aR_n(s)}-2^{(1-\delta)bR_ns}]. \quad ...(3.6)$$

Putting $a^*=1-a(1-\delta)$, $b^*=1-b(1-\delta)$, we have

$$H_2^{*(s)}(Q \parallel G_n;[a,b])=H_2^{(s)}(Q \parallel G_n;[a^*,b^*]),\ (s=1,2) \quad ...(3.7)$$

which is self-explanatory.

(iii) In general, $x=F(Q,G_n)>0$ is any time-free information measure of at least two distributions Q,G_n. Then,

$$H(x;[a,b])=\int_a^b \psi(x;t)\,dt=Bx\int_a^b 2^{(1-t)x}\,dt$$

$$=T(a,b)\,[\{1+C(a)\}^x-\{1+C(b)\}^x]>0,\ (a\neq b) \quad ...(3.8)$$

where $H(1;[a,b])=1$. The general formula, then, becomes

$$H^*(Q,G_n;[o,b])=H[F(Q,G_n);\ [a,b]]$$

$$=T(a,b)[\{1+c(a)\}^{F(Q,G_n)}-\{1+c(b)\}^{F(Q,G_n)}] \quad ...(3.9)$$

whose importance depends on a useful choice of $F(Q,G_n)$. The choice $F(Q,G_n)=R_n(s)$, $(s=1,2)$ is advantageous in many respects. The basic reason is that $R_n(s)$ is a logarithmic function raised as the power of $1+C(\theta)$, $(\theta=a,b)$.

PROPERTIES: (i) WEIGHTED ADDITIVITY WITH SYMMETRY, LINEAR COMBINATION AND REPRESENTATION—These properties of $H_2^{(s)}(Q \parallel G_n;[a,b])$ are embedded in the following:

Theorem 3—Using distributions of Theorem 1, we have

$$H_2^{(s)}(Q^*R \parallel G_M^* G_N;[a,b]) = 2^{(1-a)R_N(s)} H_2^{(s)}(Q \parallel G_M;[a,b]) + 2^{(1-b)R_M(s)} H_2^{(s)}(R \parallel G_N;[a,b]) \quad (S=1,2); \qquad \ldots(3.10)$$

and if $Q=G_M$, $R=G_N$, then,

$$H_2^{(s)}(G_M * G_N;[a,b]) = 2^{(1-a)R_N(s)} H_2^{(s)}(G_M;[a,b]) + 2^{(1-b)R_M(s)} H_2^{(s)}(G_N;[a,b]) \qquad \ldots(3.11)$$

Also,

$$H_2^{(s)}(Q \parallel G_u;[a,b]) = C(a,b)H_1^{(s)}(Q \parallel G_u;a) + C(b,a)H_1^{(s)}(Q \parallel G_u;b), \quad (u\epsilon M,N) \qquad \ldots(3.12)$$

where,

$$H_1^{(s)}(Q \parallel G_u;\theta) = C^{-1}(\theta)\left[2^{(1-\theta)R_u(s)} - 1\right], (u=M,N;\ \theta=a,b;\ \theta\neq 1) \qquad \ldots(3.13)$$

and,

$$H_1^{(s)}(G_M^* G_N;\theta) = H_1^{(s)}(G_M;\theta) + H_1^{(s)}(G_N;\theta) + C(\theta)H_1^{(s)}(G_M;\theta)\,H_1^{(s)}(G_N;\theta). \qquad \ldots(3.14)$$

On representation, beside (3.12) expanded by means of (3.13), we give some additional results as follow:

$$H_2^{(1)}(Q \parallel G_n;[a,b]) = T(a,b)\left[\left(\sum_1^{n} q_k g_{nk}^{\delta-1}\Big/ w(Q)\right)^{(1-a)/(1-\delta)} - \left(\sum_1^{n} q_k g_{nk}^{\delta-1}\Big/ w(Q)\right)^{(1-b)/1-\delta)}\right], (\delta>0,\ \delta\neq 1) \qquad \ldots(3.15)$$

$$= T(a,b)[\{1+C(\delta)H_1(Q \parallel G_n;\delta)\}^{(1-a/1-\delta)} - \{1+C(\delta)H_1(Q \parallel G_n;\delta)\}^{(1-b)/(1-\delta)} \qquad \ldots(3.16)$$

and,

$$\{1+C(\theta)H_1^{(s)}(Q \parallel G_n;\theta\}^{1-\delta} = \{1+C(\delta)H_1(Q \parallel G_n;\delta)\}^{1-\theta}, (\delta>0,\ \delta\neq 1;\ \theta=a,b). \qquad \ldots(3.17)$$

A little inspection will make (3.16) and (3.17) self-explanatory.

(ii) CONVEXITY—*Theorem* 4 — $H_2^{(1)}(Q \parallel G_n;[a,b])$ is a convex $\cap$ or $\cup$ function of G_n according as $1<a<\delta<b$ or $a<\delta<b<1$.

Hint: $f(x) = T(a,b)\,(x^{(1-a)/(1-\delta)} - x^{(1-b)/(1-\delta)}),\ (x>0,\ (a\neq b)).$ ...(3.18)

The convexity of the function shows also the corresponding inequalities

$$H_2^{(1)}(Q \parallel G_n;[a,b]) \gtrless H_2(Q \parallel G_n;[a.b]) \qquad \ldots(3.19)$$

according as the first function is convex $\cap$ or $\cup$.

Theorem 5 — $H_2^{(2)}(Q \parallel G_n;[a,b])$ is a convex $\cap$ function of G_n for $a<1<b$.

Hint: $f(x) = T(a,b)\,[2^{(1-a)x} - 2^{(1-b)x}]$, $(x>0,\ a\neq b)$, (see 3.8) ...(3.20)

Put $x_k = p_k$ and $y_k = p_k g_{nk}$ in Jensen's inequality

$\Sigma x_k \log_2 x_k/y_k \geqslant (\Sigma x_k) \log_2 (\Sigma x_k/\Sigma y_k)$ and obtain

$$\exp_2 \left\{(1-a) \sum_1^{i_n} p_k \log_2 \frac{1}{g_{nk}}\right\} \geqslant \left(\sum_1^{i_n} p_k g_{nk}\right)^{a-1} > \sum_1^{i_n} p_k g_{nk}^{a-1}, (a<1) \quad \ldots(3.21)$$

$$\text{evp}_2 \left\{(1-b) \sum_1^{i_n} p_k \log_2 \frac{1}{g_{nk}}\right\} \leqslant \left(\sum_1^{i_n} p_k g_{nk}\right)^{b-1} < \sum_1^{i_n} p_k g_{nk}^{b-1}, (b>1) \quad \ldots(3.22)$$

4. Information Losses of Bivariate Distributions

Let X and Y be two discrete generalized random variables describing respectively the source and the sink of a communication process. We define;

(i) *The joint probability distribution*

$$P_n Q_m = \{p_{kj} \mid p_{kj} = P(X=x_k,\ Y=y_j),\ 0 < \sum_1^n \sum_1^m p_{kj} = w_{nm} \leqslant 1;\ k=1,2,\ldots,n; j=1,2,\ldots,m\}$$

(ii) *The product distribution*

$$P_n^* Q_m = \{p_k q_j \mid p_k = \sum_1^m p_{kj};\ q_j = \sum_1^n p_{kj};\ \sum_1^n p_k = \sum_1^m q_j = w_{nm} \leqslant 1\}$$

in which the marginal and conditional probability functions of (X, Y) are (p_k, q_j) and $(p_{k|j}, p_{j|k})$ are respectively, where $k=1,2,3,\ldots n$ and $j=1,2,3,\ldots m$.

Let $L(Z; [a,b])$ denote any non-additive joint or marginal or conditional information loss over $[a,b]$ of $Z=(X,Y)$ or X/Yor $X|Y/Y|X$ Following Shannon (1948), one can easily construct analogous meanings of these losses. In our present studies L is of two kinds.

Case 1—Bivariate Losses of the First Kind.

Following (2.4), we have

$$H_2((X,Y); [a,b]) = T(a,b) w_{nm}^{-1}\left(\sum_1^n \sum_1^m p_{kj}^a - \sum_1^n \sum_1^m p_{kj}^b\right) \quad \ldots(4.1)$$

$$H_2(X\ ;\ [a,b]) = T(a,b)\ \omega_{nm}^{-1}\left(\sum_1^n p_k^a - \sum_1^n p_k^b\right) \quad \ldots(4.2)$$

$$H_2(Y\ ;\ [a,b]) = T(a,b)\ w_{nm}^{-1}\left(\sum_1^m q_j^a - \sum_1^m q_j^b\right) \quad \ldots(4.3)$$

$$H_2(X|Y\ ;\ [a,b]) = T(a,b)\ w_{nm}^{-1}\left(\sum_1^m \sum_1^n q_j\, p_{k|j}^a - \sum_1^m \sum_1^n q_j\, p_{k|j}^b\right) \quad \ldots(4.4)$$

$$H_2(Y|X;\ [a,b]) = T(a,b) w_{nm}^{-1}\left(\sum_1^n \sum_1^m p_k p_{j|k}^a - \sum_1^n \sum_1^m p_k p_{j|k}^b\right),\ (a \neq b) \quad \ldots(4.5)$$

So, corresponding to (4.1)—(4.5) one can easily write the expressions for $H_1((X,Y);\theta)$, $H_1(X;\theta)$, $H_1(Y;\theta)$, $H_1(X\mid Y;\theta)$ and $H_1(Y\mid X;\theta)$ $(\theta=a,b;\ \theta\neq 1)$ as losses over $[a,1]$ or $[1,b]$.

Taking the meaning of Z given as above, we have

$$H_2(Z;[a,b])=C(a,b)H_1(Z;a)+C(b,a)H_1(Z;b) \qquad ...(4.6)$$

in which for $Z=(X,Y)$,

$$H_1((X,Y);\theta)=w_{nm}^{-1}\sum_1^n p_k^\theta H_1(Y\mid x_k;\theta)+H_1(X;\theta) \qquad ...(4.7)$$

$$=w_{nm}^{-1}\sum_1^m q_j^\theta H_1(X\mid y_j;\theta)+H_1(Y;\theta),\ (\theta=a,b) \qquad ...(4.8)$$

giving two interpretations.

Again the information loss over $[a,b]$ of the product variate XY is

$$H_2(XY;[a,b]) = C(a,b)H_1(XY;a)+C(b,a)H_1(XY;b) \qquad ...(4.9)$$

where, $H_1(XY;\theta)=H_1(X;\theta)+H_1(Y;\theta)+C(\theta)H_1(X;\theta)H_1(Y_j;\theta)$, $(\theta=a,b)$ ...(4.10)

But if the variates X, Y are statistically independent, then

$$H_2((X,Y);[a,b])=H_2(XY;[a,b]). \qquad ...(4.11)$$

Now following Shannon (1948), if one intends to define transinformation loss over $[a,b]$ by

$$I((X,Y);[a,b])=L(X;[a,b])-L(X\mid Y;[a,b])\geqslant 0, \qquad ...(4.12)$$

then in the present case one needs the following

Theorem 6—If X, Y are not statistically independent, then

$H_2(X;[a,b]) > H_2(X\mid Y;[a,b])$, for $a<1<b$; and, if X, Y are statistically independent, then,

$H_2(X;[a,b]) \geqslant H_2(X\mid Y;[a,b])$. $\forall\ a<b$.

PROOF : $H_2(X;[a,b])-H_2(X\mid Y;[a,b])$

$$=T(a,b)w_{nm}^{-1}\Big[\Big(\sum_1^n p_k^a-\sum_1^m\sum_1^n q_j p_{k\mid j}^a\Big)-\Big(\sum_1^n p_k^b-\sum_1^m\sum_1^n q_j p_{k\mid j}^b\Big)\Big] \qquad ...(4.13)$$

$$\text{But, } \sum_1^n p_k^a=\sum_1^n\Big(\sum_1^m q_j p_{k\mid j}\Big)^a>\sum_1^n\sum_1^m q_j p_{k\mid j}^a\ ,\ (\text{if } a<1) \qquad ...(4.14)$$

$$\text{and, } \sum_1^n p_k^b=\sum_1^n\Big(\sum_1^m q_j p_{k\mid j}\Big)^b<\sum_1^n\sum_1^m q_j p_{k\mid j}^b\ ,\ (\text{if } b>1) \qquad ...(4.15)$$

(ref. Sharma and Mittal 1975).

Using (4.14) and (4.15) ln (4.13), the first part is proved. Again, if X,Y are statistically independent, then

$$H_2(X\mid Y;[a,b])=w_{nm}H_2(X;[a,b]) \leqslant H_2(X;[a,b]),\ \forall a<b \qquad ...(4.16)$$

and thus the second part is proved.

Case 2—Bivariate Losses of the second kind.
Sharma-Mittal (1975) bivariate H^* entropies and Gupta-Sharma (1976) bivariate non-additive inaccuracies are interpreted by us as bivariate information losses over $[a,1]$ or $[1,b]$ only, whereas these should be over a more general interval $[a,b]$.

Here, for simplicity, we are showing the generalizations of bivariate H^* entropies as generalized information losses over $[a, b]$.

For two variates X, Y, it is assumed that one is acquainted with the expressions for the following Renyi's entropies :

$H^\delta(x_k)$, $H^\delta(X)$, $H^\delta(y_j)$, $H^\delta(Y)$, $H^\delta(Y|x_k)$, $H^\delta(Y|X)$, $H^\delta(X|y_j)$,

$H^\delta(X|Y)$, $H^\delta(x_k, Y)$, $H^\delta(X, y_j)$ and $H^\delta(X, Y)$, $(\delta > 0,\ \delta \neq 1)$; and also, correspondingly with those for $\delta = 1$. The expressions could be of interest. Secondly, we are using $R(1)$ for $H^\delta(\cdot)$, $(\delta > 0,\ \delta \neq 1)$ and $R(2)$ for $H^1(\cdot)$, $(\delta = 1)$.

Then following (3.3) with $Q = Gn$, replacing $Rn(s)$ by $R(s)$, $(s=1,2)$ and using the forms of Z, we have

$$H_2^{(1)}((X, Y); [a, b]) = T(a, b)\left[2^{(1-a)H^\delta(X,Y)} - 2^{(1-b)H^\delta(X,Y)}\right] \quad \ldots(4.17)$$

$$= T(a, b)\left[\left(\sum_1^n p_k\, 2^{(1-\delta)H^\delta(X_k,Y)} \Big/ w_{nm}\right)^{(1-a)/(1-\delta)}\right.$$

$$\left. - \left(\sum_1^n p_k\, 2^{(1-\delta)H^\delta(X_k,Y)} \Big/ w_{nm}\right)^{(1-b)/(1-\delta)}\right] \quad \ldots(4.18)$$

$$= T(a, b)\left[\left(\frac{\sum_1^n \sum_1^m p_{kj}^\delta}{w_{nm}}\right)^{(1-a)/(1-\delta)} - \left(\frac{\sum_1^n \sum_1^m p_{kj}^\delta}{w_{nm}}\right)^{(1-b)/(1-\delta)}\right] \quad \ldots(4.19)$$

$$H_2^{(1)}(X; [a, b]) = T(a, b)\left[\left(\sum_1^n p_k^\delta \Big/ w_{nm}\right)^{(1-a)/(1-\delta)} - \left(\sum_1^n p_k^\delta \Big/ w_{nm}\right)^{(1-b)/(1-\delta)}\right] \quad \ldots(4.20)$$

$$H_2^{(1)}(Y: [a, b]) = T(a, b)\left[\left(\sum_1^m q_j^\delta \Big/ w_{nm}\right)^{(1-a)/(1-\delta)} - \left(\sum_1^m q_j^\delta \Big/ w_{mn}\right)^{(1-b)/(1-\delta)}\right] \quad \ldots(4.21)$$

$$H_2^{(1)}(X \mid Y; [a,b]) = T(a, b)\left[\left(\sum_1^m \sum_1^n q_j\, p_{k/j}^\delta \Big/ w_{nm}\right)^{(1-a)/(1-\delta)}\right.$$

$$\left. - \left(\sum_1^m \sum_1^n q_j\, p_{k/j}^\delta \Big/ w_{nm}\right)^{(1-b)/(1-\delta)}\right] \quad \ldots(4.22$$

$$H_2^{(1)}(Y \mid X; [a, b]) = T(a, b)\left[\left(\sum_1^n \sum_1^m p_k\, p_{jk}^\delta \Big/ w_{nm}\right)^{(1-a)/(1-\delta)}\right.$$

$$- \left(\sum_{1}^{n} \sum_{1}^{m} p_k \, p_{j/k}^{\delta} \Big/ w_{nm} \right)^{(1-b)/(1-\delta)} \Big] \qquad \ldots(4.23)$$

from which the bivariate H^* entropies can be obtained as particular cases interpreted as information losses over $[a, 1]$ or $[1, b]$ by putting $b = 1$ or $a = 1$ in (4.18)—(4.24)

For above formulae become simpler when $\delta = 1$:

$$H_2^{(2)}\Big(z;\ [a, b]\Big) = T\,(a, b) \Big[\, 2^{(1-a)H^1(Z)} - 2^{(1-b)H^1(Z)} \,\Big], \ (a \neq b), \qquad \ldots(4.24)$$

where z is already described.

Properties of linear combination, decomposition, relationships between joint, marginal and conditional losses, representations and convexities are exhibited through the following theorems :

Theorem 7 — $H_2^{(1)}\Big((x_k,\ Y);\ [a, b]\Big) = C\,(a, b)\ H_1^{(1)}\Big((x_k, Y);\ a\Big)$

$$+ \, C\,(b, a)\, H_1^{(1)}\Big((x_k, Y);\ b\Big) \qquad \ldots(4.25)$$

and

$$H_2^{(2)}\Big((X,\ Y);\ [a, b]\Big) = C\,(a, b)\ H_1^{(2)}\Big((X,\ Y);\ a\Big)$$

$$+ \, C\,(b, a)\ H_1^{(2)}\Big(\ (X,\ Y);\ b\ \Big) \qquad \ldots(4.26)$$

where the losses over $[a, 1]$ or $[1, b]$ are given by

$$H_1^{(1)}\Big((x_k,\ Y);\ \theta\Big) = H_1^{(1)}\,(x_k;\ \theta) + H_1^{(1)}\,(Y \mid x_k;\ \theta)$$

$$+ \, C\,(\theta)\ H_1^{(1)}\,(x_k;\ \theta)\ H^{(1)}\,(Y \mid x_k;\ \theta) \qquad \ldots(4.27)$$

$$H_1^{(2)}\Big((X,\ Y);\ \theta\Big) = H_1^{(2)}\,(X;\ \theta) + H_1^{(2)}\,(Y \mid X;\ \theta)$$

$$+ \, C\,(\theta)\ H_1^{(2)}\,(X;\theta)\, H_1^{(2)}\,(Y \mid X;\ \theta) \qquad \ldots(4.28)$$

$\theta = a, b;\ \theta \neq 1.$

The linearities in (4.25) and (4.26) also possess decompositions through (4.27) and (4.28). One may take ordinary averages with respect to p_k as weights of each quantity in (4.25) and (4.27), $(K = 1, 2, \ldots, n)$ and thereby a question will arise as to the nature of the new formulae connecting the averages. This remains open.

Theorem 8 — If it is defined

$$L\,(\,(X,\ Y);\ \theta) = H_1^{(1)}(X;\ \theta) - H_1^{(1)}\,(Y \mid X;\ \theta)$$

$$+ \, C\,(\theta)\ H_1^{(1)}\,(X;\ \theta)\, H_1^{(1)}\,(Y \mid X;\ \theta),\ (\theta = a, b) \qquad \ldots(4.29)$$

Then, $T\,(a, b)\ \ [2^{(1-a)\{H^{\delta}(x) + H^{\delta}(Y \mid X)\}} - 2^{(1-b)\{H^{\delta}(x) + H^{\delta}(Y \mid X)\}}]$

$$= C\,(a, b)\ L\,(\,(X,\ Y);\ a) + C\,(b, a)\ L\,((X,\ Y);\ b) = L\,(\,(X,\ Y);\ [a, b]). \qquad \ldots(4.30)$$

This theorem is purely a conjecture in regard to the existence of an appropriate mathematical form of the information loss function L as in (4.29) and (4.30).

Theorem 9 — For certain power-parametric restrictions :

(i) $H_2^{(1)}(X; [a, b]) \geqslant H_2^{(1)}(X \mid Y; [a, b])$,

whenever $a < \delta < 1 < b$, $\delta < a < 1 < b$, $a < 1 < b < \delta$, $a < 1 < \delta < b$; but, (i) is not so sure for $\delta < 1 < a < b$ or $a < b < 1 < \delta$, unless some additional different kinds of condition are imposed.

(ii) $H_2^{(2)}(X; [a, b]) \geqslant H_2^{(2)}(X \mid Y; [a, b])$, whenever $a < 1 < b$.

Hint : $H_2^{(1)}(X; [a, b]) - H_2^{(1)}(X|Y;[a, b])$

$$= T(a, b)\Big[\Big\{\Big(\sum_1^n p_k^\delta \Big/ w_{nm}\Big)^{(1-a)/(1-\delta)} - \Big(\sum_4^n \sum_1^m q_j p_{k/j}^\delta \Big/ w_{nm}\Big)^{(1-a)/(1-\delta)}\Big\}$$

$$-\Big\{\Big(\sum_1^n p_k^\delta \Big/ w_{nm}\Big)^{(1-b)/(1-\delta)} - \Big(\sum_1^n \sum_1^m q_i p_{k/j}^\delta \Big/ w_{nm}\Big)^{(1-b)/(1-x)}\Big\}\Big] \quad \ldots(4.31)$$

$$\text{Also,} \quad 1 \lesseqgtr \sum_1^n p_k^\delta \Big/ w_{nm} = \sum_1^n \Big(\sum_1^m q_j p_{k/j}\Big)_k^\delta \Big/ w_{nm} \lesseqgtr \sum_1^n \sum_1^m q_j \, p_{k/j}^\delta \Big/ w_{nm} \quad \ldots(4.32)$$

according to as $\delta \lesseqgtr 1$.

Again, $H_2^{(2)}\Big(X; [a, b]\Big(- H_2^{(2)}\Big) X \mid Y; [a, b]\Big)$

$$= T(a, b)\Big[\Big(2^{(1-a)H^1(X)} - 2^{(1-a)H^1(X|Y)}\Big) - \Big(2^{(1-b)H^1(X)} - 2^{(1-b)H^1(X|Y)}\Big)\Big]. \quad \ldots(4.33)$$

For generalized X, Y, use $H^1(X) \geqslant H^1(X|Y)$ and $a < 1 < b$.

References

Daroczy, Z. (1970). Generalized information and functions. *Information and Control*, **16**, 35–51.

Gupta, H. C., and B. D. Sharma (1976). On non-additive measures of inaccuracy. *Czech. Math. J.*, **26**, (101), 584–95.

Havrda, J., and Charvat, F. (1967). Quantification method of classification processes. The concept of structural α-entropy. *Kybernetika*, **3**, 30–35.

Kerridge, D. F. (1961). Inaccuracy and inference. *J. R. Stat. Soc. Ser.* B 23, 184–94.

Mittal D. P. (1975). On some functional equations concerning entropy, directed divergence and inaccuracy. *Metrika*, 22, 35–45.

Nath, P. (1968). On the measures of errors information. *J. Math. Sci*, **3**, No. 1, 1–16.

——— (1970). An axiomatic characterization of inaccuracy for finite generalized distributions. *Opsearch*, **7**, 115–33.

Renyi, A. (1961). On measures of entropy and information. Proc. 4th. Berkely Symp. Math. Stat. and Prob., **1**, 547–61.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, **27**, 378–423, 623–56.

Sharma, B. D., and Mittal, D. P. (1975). New non-additive measures of entropy for discrete probability distributions. *J. Math. Sci.*, **10**, 28–40.

Sharma, B. D. and Taneja, I. J. (1975). Entropy of type (α, β) and other generalized measures in information theory. *Metrika*, **22**, 205–15.