KYBERNETIKA — VOLUME 28 (1992), NUMBER 4, PAGES 309–324

ON M-DIMENSIONAL UNIFIED (r,s)-JENSEN DIFFERENCE DIVERGENCE MEASURES AND THEIR APPLICATIONS

Maria L. Menndez, Leandro Pardo and Inder J. Taneja

During past years the Jensen difference divergence measure (Sibson [18], Rao [12]) has found its importance towards applications in various statistical areas. In this paper, we have presented three different ways to generalize this measure by using two scalar parameters. These generalizations have been put in unified expressions. Some connections with income inequality, generalized mutual information, Markov chains, deflation factor etc., have been made.

1. INTRODUCTION

Let

$$\Delta_n = \left\{ P = (p_1, .., p_n) \mid p_i > 0, \sum_{i=1}^{n} p_i = 1 \right\}$$

be the set of all complete finite discrete probability distributions. For all $P \in \Delta_n$, the Shannon's entropy is written as

$$H(P) = -\sum_{i=1}^{n} p_i \log_2 p_i. \tag{1}$$

Concavity of Shannon's entropy gives the following inequality :

$$\sum_{j=1}^{M} \lambda_j H(P_j) \le H\left(\sum_{j=1}^{M} \lambda_j P_j\right), \tag{2}$$

where $P_1, P_2, \ldots, P_M \in \Delta_n$, i.e., $P_j = (p_{1j}, p_{2j}, \ldots, p_{nj}) \in \Delta_n$, for each $j = 1, 2, \ldots, M$; and $\lambda_i \ge 0$, $\sum_{i=1}^{M} \lambda_i = 1$.

The Jensen difference divergence measure (cf. [12]) or Information radius (cf. [18]) for M-probability distribution is given by

$$R(P_1, P_2, \ldots, P_M) = H\left(\sum_{j=1}^{M} \lambda_j P_j\right) - \sum_{j=1}^{M} \lambda_j H(P_j). \tag{3}$$

We can write

$$R(P_1, P_2, \ldots, P_M) = \sum_{j=1}^{M} \lambda_j D\left(P_j \parallel \sum_{k=1}^{M} \lambda_k P_k\right), \tag{4}$$

where $D(P\|Q)$ is the Kullback–Leibler's directed divergence given by

$$D(P\|Q) = \sum_{i=1}^{n} p_i \log_2 \frac{p_i}{q_i}, \tag{5}$$

for all $P, Q \in \Delta_n$.

We shall call the measure (3) or (4), the *M-dimensional R-divergence*. We shall now present some different ways to generalize this measure. In order to do so, first we shall give a unified two parametric generalization of (5).

1.1. Unified (r,s)-directed divergence

Taneja [20] wrote some of the known generalizations of the measure (5) in a unified way. This unification is given by

$$\mathcal{F}_r^s(P\|Q) = \begin{cases} D_r^s(P\|Q) = \left(1-2^{1-s}\right)^{-1}\left\{\left(\sum\limits_{i=1}^n p_i^r q_i^{1-r}\right)^{\frac{s-1}{r-1}} - 1\right\}, & r \neq 1,\ s \neq 1 \\ D_1^s(P\|Q) = \left(1-2^{1-s}\right)^{-1}\left(2^{(s-1)\,D(P\|Q)} - 1\right), & r = 1,\ s \neq 1 \\ D_r^1(P\|Q) = \frac{1}{r-1}\log_2\left(\sum\limits_{i=1}^n p_i^r q_i^{1-r}\right), & r \neq 1,\ s = 1 \\ D(P\|Q) = -\sum\limits_{i=1}^n p_i \log_2 \frac{p_i}{q_i}, & r = 1,\ s = 1 \end{cases} \tag{6}$$

for all $r \in (0,\infty)$ and $s \in (-\infty,\infty)$. $\mathcal{F}_r^s(P\|Q)$ is called *unified* (r,s)-*directed divergence.* It includes in particular the measures studied by Sharma and Mittal [17], Rnyi [14] and Kullback and Leibler [7]. It has many interesting properties (cf. [21]). In particular, when $Q = U$, where $U = \left(\frac{1}{n},\ldots,\frac{1}{n}\right) \in \Delta_n$, then we can write

$$\mathcal{F}_r^s(P\|Q) = n^{s-1}\left(\mathcal{E}_r^s(U) - \mathcal{E}_r^s(P)\right), \tag{7}$$

where

$$\mathcal{E}_r^s(P) = \begin{cases} H_r^s(P) = \left(2^{1-s}-1\right)^{-1}\left\{\left(\sum\limits_{i=1}^n p_i^r\right)^{\frac{s-1}{r-1}} - 1\right\}, & r \neq 1,\ s \neq 1 \\ H_1^s(P) = \left(2^{1-s}-1\right)^{-1}\left(2^{(1-s)\,H(P)} - 1\right), & r = 1,\ s \neq 1 \\ H_r^1(P) = \frac{1}{r-1}\log_2\left(\sum\limits_{i=1}^n p_i^r\right), & r \neq 1,\ s = 1 \\ H(P) = -\sum\limits_{i=1}^n p_i \log_2 p_i, & r = 1,\ s = 1 \end{cases} \tag{8}$$

and

$$\mathcal{E}_r^s(U) = \begin{cases} \left(2^{1-s}-1\right)^{-1}\left(n^{1-s}-1\right), & s \neq 1 \\ \log n, & s = 1 \end{cases} \tag{9}$$

for all $r \in (0,\infty)$ and $s \in (-\infty,\infty)$. The measure $\mathcal{E}_r^s(P)$ is named as *unified* (r,s)-*entropy.*

1.2. M-dimensional unified (r,s)-Jensen difference divergence measures

This section deals with three different generalizations of M-dimensional R-divergence given by (4). The first generalization is based on the relations (4) and (6), while the second is obtained directly. The third is based on the inequality (2) and the unified (r,s)-entropy (8).

1.2.1. First generalization

In (4) replace D by $\mathcal{F}_r^s$, we can write

$$^1\mathcal{V}_r^s\left(P_1, P_2, \ldots, P_M\right) = \sum_{j=1}^{M} \lambda_j \ \mathcal{F}_r^s\left(P_j \parallel \sum_{k=1}^{M} \lambda_k P_k\right), \tag{10}$$

for all $r \in (0, \infty)$ and $s \in (-\infty, \infty)$, where $\mathcal{F}_r^s$ is as given by (5). More clearly, the measure (10) stands as follows:

$$^1\mathcal{V}_r^s(P_1, \ldots, P_M) = \begin{cases} {}^1R_r^s(P_1, \ldots, P_M), & r \neq 1,\ s \neq 1 \\ {}^1R_1^s(P_1, \ldots, P_M), & r = 1,\ s \neq 1 \\ {}^1R_r^1(P_1, \ldots, P_M), & r \neq 1,\ s = 1 \\ R(P_1, \ldots, P_M), & r = 1,\ s = 1. \end{cases}$$

where

$$^1R_r^s(P_1, \ldots, P_M) = \left(1 - 2^{1-s}\right)^{-1} \left\{ \sum_{j=1}^{M} \lambda_j \left[\sum_{j=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}} - 1 \right\},$$

$$r \neq 1,\ s \neq 1$$

$$^1R_1^s(P_1, \ldots, P_M) = \left(1 - 2^{1-s}\right)^{-1} \left\{ 2^{(s-1)\, R(P_1, \ldots, P_M)} - 1 \right\}, \qquad s \neq 1,$$

$$^1R_r^1(P_1, \ldots, P_M) = (r-1)^{-1} \sum_{j=1}^{M} \lambda_j \log_2 \left[\sum_{i=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} \right], \qquad r \neq 1, \tag{11}$$

for all $r \in (0, \infty)$ and $s \in (-\infty, \infty)$.

1.2.2. <u>Second generalization</u>

In particular, when $r = s$, we have

$$^1R_s^s(P_1, P_2, \ldots, P_M) = \left(1 - 2^{1-s}\right)^{-1} \left(\sum_{i=1}^{n} \left(\sum_{j=1}^{M} \lambda_j p_{ij}^s \right) \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-s} - 1 \right),$$

$$s \neq 1,\ s > 0. \tag{12}$$

We shall use the expression appearing in (12) for defining the second generalization of M-dimensional R-divergence. It is given as follows

$$^2\mathcal{V}_r^s(P_1, P_2, \ldots, P_M) = \begin{cases} {}^2R_r^s(P_1, P_2, \ldots, P_M), & r \neq 1,\ s \neq 1 \\ {}^2R_1^s(P_1, P_2, \ldots, P_M), & r = 1,\ s \neq 1 \\ {}^2R_r^1(P_1, P_2, \ldots, P_M), & r \neq 1,\ s = 1 \\ R(P_1, P_2, \ldots, P_M), & r = 1,\ s = 1, \end{cases} \tag{13}$$

where

$$^2R_r^s(P_1, P_2, \ldots, P_M) = \left(1 - 2^{1-s}\right)^{-1} \left\{ \left[\sum_{i=1}^{n} \left(\sum_{j=1}^{M} \lambda_j p_{ij}^r \right) \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}} - 1 \right\},$$

$$r \neq 1,\ s \neq 1$$

$$^2R_1^s(P_1, P_2, \ldots, P_M) = \left(1 - 2^{1-s}\right)^{-1} \left\{ \exp_2\left[(s-1)\, R(P_1, \ldots, P_M)\right] - 1 \right\},\ s \neq 1,$$

$$^2R_r^1(P_1, P_2, \ldots, P_M) = (r-1)^{-1} \log_2 \left\{ \sum_{j=1}^{n} \left(\sum_{i=1}^{M} \lambda_j p_{ij}^r \right) \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} \right\},\ r \neq 1, \tag{14}$$

for all $r \in (0, \infty)$ and $s \in (-\infty, \infty)$.

In particular, when $r = s$, we have

$$ {}^{1}\mathcal{V}_s^s\left(P_1, P_2, \ldots, P_M\right) = {}^{2}\mathcal{V}_s^s\left(P_1, P_2, \ldots, P_M\right), \qquad s > 0. $$

1.2.3. Third generalization

In the inequality (2) if we replace H by $\mathcal{E}_r^s$ as of expression (8) we get

$$ \sum_{j=1}^{M} \lambda_j \mathcal{E}_r^s(P_j) \le \mathcal{E}_r^s\left(\sum_{j=1}^{M} \lambda_j P_j\right). $$

The validity of the above inequality depends upon the concavity of $\mathcal{E}_s^r$. This holds, when $(r, s) \in \Gamma$ (cf. [20]), where

$$ \Gamma = \left\{(r, s) \,|\, s \ge 2 - 1/r,\ r > 0\right\}. $$

Thus, the difference

$$ {}^{3}\mathcal{V}_r^s\left(P_1, P_2, \ldots, P_M\right) = \mathcal{E}_r^s\left(\sum_{j=1}^{M} \lambda_j P_j\right) - \sum_{j=1}^{M} \lambda_j \mathcal{E}_r^s(P_j), \tag{15} $$

for all $(r, s) \in \Gamma$ can be considered a third generalization of Jensen difference divergence measure (3). The particular case of (15), when $r = s$ has been extensively studied by Burbea and Rao [2, 3], Kapur [6], Sahoo and Wong [15]. And the case, when $s = 1$ has been studied by Rao [12]. We see that the nonnegativity of (15) is restrictive with respect to parameters, while this is not so for the measures (10) and (13). The measures (10) and (13) are presented for the first time in this paper.

In this paper, our aim is to study properties of the measure ${}^{\alpha}\mathcal{V}_r^s\left(P_1, P_2, \ldots, P_M\right)$ ($\alpha = 1$ and 2) such as convexity, Schur-convexity, monotonicity with respect to the parameters, generalized data processing inequalities etc. Some applications towards income inequality, deflation factor, generalized mutual information, Markov Chains etc. are specified.

2. PROPERTIES OF M-DIMENSIONAL UNIFIED (r, s)-JENSEN DIFFERENCEDIVERGENCE MEASURES

The definition of convexity for M-probability distributions is well known in the literature, while, the Schur-convexity for M-probability distributions is not very much known. It is defined as follows:

Definition 1. Let $P_j = (p_{1j}, \ldots, p_{nj}) \in \Delta_n$ and $Q_j = (q_{1j}, \ldots, q_{nj}) \in \Delta_n$, $j = 1, 2, \ldots, M$. A function $F : \Delta_n \times \Delta_n \times \cdots \times \Delta_n \longrightarrow \mathbb{R}$ (reals) is Schur-convex on $\Delta_n \times \Delta_n \times \cdots \Delta_n$ if $(P_1, \ldots, P_M) \prec (Q_1, \ldots, Q_M)$ implies $F\left(P_1, \ldots, P_M\right) \le F\left(Q_1, \ldots, Q_M\right)$, where $(P_1, \ldots, P_M) \prec (Q_1, \ldots, Q_M)$ means that there is a doubly stochastic matrix $\{a_{it}\}$, $i, t = 1, \ldots, n$, with

$$ \sum_{i=1}^{n} a_{it} = \sum_{t=1}^{n} a_{it} = 1 $$

such that

$$ p_{ij} = \sum_{t=1}^{n} a_{it} q_{tj}, \qquad \forall\ j = 1, 2, \ldots, M;\ i = 1, 2, \ldots, n. $$

Now we shall study some relations in the measures appearing in the expressions (10) and (13).

We can write

$$
\begin{aligned}
{}^{1}R_{r}^{s}\left(P_{1},\ldots,P_{M}\right) &= \sum_{j=1}^{M}\lambda_{j}G_{s}\left(D_{r}^{1}\left(P_{j}\,\|\,\sum_{k=1}^{M}\lambda_{k}P_{k}\right)\right) &(16)\\
{}^{2}R_{r}^{s}\left(P_{1},\ldots,P_{M}\right) &= G_{s}\left({}^{2}R_{r}^{1}\left(P_{1},\ldots,P_{M}\right)\right) &(17)\\
{}^{1}R_{1}^{s}\left(P_{1},\ldots,P_{M}\right) &= \sum_{j=1}^{M}\lambda_{j}\;G_{s}\left(D\left(P_{j}\,\|\,\sum_{k=1}^{M}\lambda_{k}P_{k}\right)\right) &(18)\\
{}^{2}R_{1}^{s}\left(P_{1},\ldots,P_{M}\right) &= G_{s}\left(R\left(P_{1},\ldots,P_{M}\right)\right) &(19)
\end{aligned}
$$

where

$$
G_{s}(x)=\begin{cases}\left(1-2^{1-s}\right)^{-1}\left(2^{(s-1)\,x}-1\right), & s\neq 1\\ x, & s=1.\end{cases} \qquad (20)
$$

The function G_s given by (20) satisfies many interesting properties given in the following result.

Result 1. For $x \geq 0,\ -\infty < s < \infty$, the followings are true:

(i) $G_s(x) \geq 0$ with equality iff $x = 0$;

(ii) $G_s(x)$ is an increasing function of x;

(iii) $G_s(x)$ is an increasing function of s;

(iv) $G_s(x)$ is a convex function of x for $s > 1$;

(v) $G_s(x)$ is a concave function of x for $s < 1$.

We shall now present some interesting properties of the M-dimensional unified (r,s)-Jensen difference divergence measures given by (10) and (13), i. e., for ${}^{\alpha}\mathcal{V}_{r}^{s}\left(P_{1},P_{2},\ldots,P_{M}\right)$
($\alpha = 1$ and 2). From now onwards, it is understood that $P_1, P_2, \ldots, P_M \in \Delta_n,\ r \in (0,\infty)$ and $s \in (-\infty,\infty)$.

Property 1. We have, ${}^{\alpha}\mathcal{V}_{r}^{s}\left(P_{1},P_{2},\ldots,P_{M}\right) \geq 0$ ($\alpha = 1$ and 2), with equality iff $p_{ij} = \sum_{j=1}^{M} p_{ij}\lambda_j$ for all $i = 1,\ldots,n,\ j = 1,\ldots,M$.

P r o o f. In view of the relations (16) – (19) and the result 1, it is sufficient to prove the nonnegativity of ${}^{2}R_{r}^{1}\left(P_{1},P_{2},\ldots,P_{M}\right)$, because the measures D_r^1, D and R are already nonnegative. The nonnegativity of ${}^{2}R_{r}^{1}\left(P_{1},P_{2},\ldots,P_{M}\right)$ can be proved by using Jensen's inequality. □

Property 2.

$$
{}^{1}\mathcal{V}_{r}^{s}\left(P_{1},P_{2},\ldots,P_{M}\right)\begin{cases}\leq\ {}^{2}\mathcal{V}_{r}^{s}\left(P_{1},\ldots,P_{M}\right), & s\leq r\\ \geq\ {}^{2}\mathcal{V}_{r}^{s}\left(P_{1},\ldots,P_{M}\right), & s\geq r.\end{cases}
$$

P r o o f. In view of the continuity of the measures ${}^{\alpha}\mathcal{V}_{r}^{s}$ ($\alpha = 1$ and 2) with respect to the parameters, it is sufficient to prove the result for ${}^{\alpha}R_{r}^{s}$ ($\alpha = 1$ and 2), $r \neq 1,\ s \neq 1$. The result for ${}^{\alpha}R_{r}^{s}$ ($\alpha = 1$ and 2) can be derived using Jensen's inequality. □

Property 3. ${}^{\alpha}\mathcal{V}_r^s\,(P_1, P_2 \ldots, P_M)$ $(\alpha = 1$ and $2)$ are increasing functions of r (s fixed) and of s (r fixed). In particular, when $r = s$, the result still holds.

Proof. In view of the relations (16) – (19) and the result 1 (iii), the measures ${}^{\alpha}\mathcal{V}_r^s\,(P_1, P_2, \ldots, P_M)$ $(\alpha = 1$ and $2)$ are increasing functions of s (r fixed). Now we shall prove the increasing character with respect to r. For all $P_1, P_2, \ldots, P_M \in \Delta_n$, let us consider

$$T_r\left(P_j \parallel \sum_{k=1}^{M} \lambda_k P_k\right) = \left[\sum_{i=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik}\right)^{1-r}\right]^{\frac{1}{r-1}}$$

$$= \left[\sum_{i=1}^{n} p_{ij}\left(\frac{p_{ij}}{\sum_{k=1}^{M} \lambda_k p_{ik}}\right)^{1-r}\right]^{\frac{1}{r-1}}, \qquad r \neq 1$$

for each $j = 1, 2, \ldots, M$.

We can write,

$$T_r\,(P_j \| F_j) = \left[\sum_{i=1}^{n} p_{ij} f_{ik}^{r-1}\right]^{\frac{1}{r-1}}, \qquad j = 1, 2, \ldots, M,$$

where $F_j = (f_{1j}, \ldots, f_{nj})$ with $f_{ij} = \frac{p_{ij}}{\sum_{k=1}^{M} \lambda_k p_{ik}}$ for every $i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, M$. For each j, $T_r\,(P_j \| F_j)$ is an increasing function of r (cf. [1]). Since $\log_2(\cdot)$ is an increasing function this gives that

$$\frac{1}{r-1} \log_2\left(T_r(P_j \| F_j) = D_r^1\left(P_j \parallel \sum_{k=1}^{M} \lambda_k P_k\right)\right.$$

is an increasing function of r for each $j = 1, 2, \ldots, M$. In view of the relation (16), we conclude that the measure ${}^1R_r^s\,(P_1, \ldots, P_M)$ is increasing in r (s fixed). Again using the fact that $T_r(P_j \| F_j)$ is increasing in r, for each j, we conclude that $\sum_{j=1}^{M} \lambda_j T_r(P_j \| F_j)$ is increasing in r. Since $\log_2(\cdot)$ is increasing we get that

$$\frac{1}{r-1} \log_2\left(\sum_{j=1}^{M} \lambda_j T_r(P_j \| F_j)\right) = {}^2R_r^1(P_1, \ldots, P_M)$$

is increasing in r. In view of the relation (17) we conclude that ${}^2R_r^s(P_1, \ldots, P_M)$ is increasing in r (s fixed). Now we shall consider the particular case, i. e., when $r = s$. In this case, we have

$$\begin{aligned} {}^{\alpha}R_s^s\,(P_1, \ldots, P_M) &= \left(1 - 2^{1-s}\right)^{-1}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{M} \lambda_j p_{ij}^s\right)\left(\sum_{k=1}^{M} \lambda_k p_{ik}\right)^{1-s} - 1\right] = \\ &= \left(1 - 2^{1-s}\right)^{-1}\left[2^{(s-1)\,{}^2R_s^1(P_1, \ldots, P_M)} - 1\right], \qquad s \neq 1,\ \alpha = 1, 2. \end{aligned}$$

Using the result 1 (iii), we conclude that $R_s^s(P_1, \ldots, P_M)$ is increasing in s. □

Property 4. ${}^{\alpha}\mathcal{V}_r^s\,(P_1, P_2, \ldots, P_M)$ $(\alpha = 1$ and $2)$ are convex functions of $(P_1, P_2, \ldots, P_M)$ for all $s \geq r > 0$, with $P_1 \in \Delta_n,\ i = 1, \ldots, n$.

Proof. In view of continuity of ${}^{\alpha}\mathcal{V}_r^s$ $(\alpha = 1$ and $2)$ with respect to the parameters r and s, it is sufficient to show the convexity of ${}^{\alpha}R_r^s(P_1, P_2, \ldots, P_M)$ $(\alpha = 1$ and $2)$, for all $s \geq r > 0,\ r \neq 1,\ s \neq 1$.

For $\alpha = 1$. It can easily be checked that the function given by

$$K_r(p_{1j}, \ldots, p_{nj}) = \sum_{i=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r}$$

is convex for $r > 1$ and concave for $0 < r < 1$, for each $j = 1, 2, \ldots, M$. This is equivalent to say that the following inequalities hold

$$\left[\mu_1 \sum_{i=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} + \mu_2 \sum_{i=1}^{n} q_{ij}^r \left(\sum_{k=1}^{M} \lambda_k q_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}}$$

$$\begin{cases} \geq \left(\sum\limits_{i=1}^{n} (\mu_1 p_{ij} + \mu_2 q_{ij})^r \left[\mu_1 \left(\sum\limits_{k=1}^{M} \lambda_k p_{ik} \right) + \mu_2 \left(\sum\limits_{k=1}^{M} \lambda_k q_{ik} \right) \right]^{1-r} \right)^{\frac{s-1}{r-1}}, \\ \quad r > 1,\ \frac{s-1}{r-1} \text{ or } 0 < r < 1,\ \frac{s-1}{r-1} < 0 \\ \leq \left(\sum\limits_{i=1}^{n} (\mu_1 p_{ij} + \mu_2 q_{ij})^r \left[\mu_1 \left(\sum\limits_{k=1}^{M} \lambda_k p_{ik} \right) + \mu_2 \left(\sum\limits_{k=1}^{M} \lambda_k q_{ik} \right) \right]^{1-r} \right)^{\frac{s-1}{r-1}}, \\ \quad 0 < r < 1,\ \frac{s-1}{r-1} > 0 \ \text{ or } \ r > 1,\ \frac{s-1}{r-1} < 0 \end{cases} \tag{21}$$

for each $j = 1, 2, \ldots, M;\ \mu_1, \mu_2 \geq 0,\ \mu_1 + \mu_2 = 1$.
We know that the function $f(x) = x^t$ is convex for $t > 1$ or $t < 0$ and is concave for $0 < t < 1$. Using this, we have

$$\mu_1 \left[\sum_{i=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}} + \mu_2 \left[\sum_{i=1}^{n} q_{ij}^r \left(\sum_{k=1}^{M} \lambda_k q_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}}$$

$$\begin{cases} \geq \left[\mu_1 \sum\limits_{i=1}^{n} p_{ij} \left(\sum\limits_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} + \mu_2 \sum\limits_{i=1}^{n} q_{ij}^r \left(\sum\limits_{k=1}^{M} \lambda_k q_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}}, & \frac{s-1}{r-1} > 1 \text{ or } \frac{s-1}{r-1} < 0 \\ \leq \left[\mu_1 \sum\limits_{i=1}^{n} p_{ij} \left(\sum\limits_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} + \mu_2 \sum\limits_{i=1}^{n} q_{ij}^r \left(\sum\limits_{k=1}^{M} \lambda_k q_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}}, & 0 < \frac{s-1}{r-1} < 1. \end{cases} \tag{22}$$

for each $j = 1, 2, \ldots, M;\ \mu_1, \mu_2 \geq 0,\ \mu_1 + \mu_2 = 1$.
Joining the inequalities (21) and (22) and multiplying the resultant inequality by λ_j, adding for all $j = 1, 2, \ldots, M$, subtracting 1 on both sides and multiplying by $(1 - 2^{1-s})^{-1}$
$(s \neq 1)$, we get the convexity of ${}^1R_r^s(P_1, P_2, \ldots, P_M)$ for all $s > r > 0$. In particular when $r = s$, the inequalities (21) still hold. This completes the result for $\alpha = 1$.

For $\alpha = 2$. To prove the convexity of ${}^1R_r^s(P_1, \ldots, P_M)$ we used the functions $K_r(p_{1j}, \ldots, p_{nj})$ $(j = 1, 2, \ldots, M)$. Instead, using it again, if we use the fact that the function

$$\sum_{j=1}^{M} \lambda_j K_r(p_{1j}, \ldots, p_{nj}) = \sum_{j=1}^{M} \lambda_j \sum_{i=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} \tag{23}$$

is convex in Δ_n^M for $r > 1$ and is concave in Δ_n^M for $0 < r < 1$, and proceeding on the similar lines as before we get the required result. □

Property 5. ${}^{\alpha}\mathcal{V}_r^s\,(P_1, P_2, \ldots, P_M)$ ($\alpha = 1$ and 2) are Schur-convex functions of $(P_1, P_2, \ldots, P_M) \in \Delta_n^M$, i. e., $(P_1, P_2, \ldots, P_M) \prec (Q_1, Q_2, \ldots, Q_M)$ implies

$$ {}^{\alpha}\mathcal{V}_r^s\,(P_1, P_2, \ldots, P_M) \leq {}^{\alpha}\, \mathcal{V}_r^s\,(Q_1, Q_2, \ldots, Q_M) \qquad (\alpha = 1 \text{ and } 2) $$

P r o o f. By the definition of $(P_1, P_2, \ldots, P_M) \prec (Q_1, Q_2, \ldots, Q_M)$ implies that

$$ p_{ij} = \sum_{t=1}^{n} a_{it}\, q_{tj} \qquad \forall\, j = 1, 2, \ldots, M;\ i = 1, 2, \ldots, n, $$

where a_{it}, are as given in Definition 1. This gives,

$$ p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} = \left(\sum_{t=1}^{n} a_{it}\, q_{tj} \right)^r \left(\sum_{k=1}^{M} \sum_{t=1}^{n} a_{it}\, \lambda_k\, q_{tk} \right)^{1-r}, \tag{24} $$

for all $j = 1, 2, \ldots, M;\ i = 1, 2, \ldots, n$.

For $\alpha = 1$. From Hlder inequality, we have

$$ p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k p_{ik} \right)^{1-r} \begin{cases} \geq \sum\limits_{t=1}^{n} a_{it}\, q_{tj}^r \left(\sum\limits_{k=1}^{M} \lambda_k\, q_{tk} \right)^{1-r}, & 0 < r < 1, \\ \leq \sum\limits_{t=1}^{n} a_{it}\, q_{tj}^r \left(\sum\limits_{k=1}^{M} \lambda_k\, q_{tk} \right)^{1-r}, & r > 1, \end{cases} $$

for all $j = 1, 2, \ldots, M$ and $i = 1, 2, \ldots, n$.

Summing over all $i = 1, 2, \ldots, n$, using the fact that $\sum_{i=1}^{n} a_{it} = 1$ for all $t = 1, 2, \ldots, n$ and raising both sides of the resultant inequality by $\frac{s-1}{r-1}$, we have

$$ \left[\sum_{i=1}^{n} p_{ij}^r \left(\sum_{k=1}^{M} \lambda_k\, p_{ik} \right)^{1-r} \right]^{\frac{s-1}{r-1}} \begin{cases} \geq \left[\sum\limits_{t=1}^{n} q_{tj}^r \left(\sum\limits_{k=1}^{M} \lambda_k\, q_{tk} \right)^{1-r} \right]^{\frac{s-1}{r-1}}, \\ \qquad \frac{s-1}{r-1} > 0,\ 0 < r < 1, \text{ or } \frac{s-1}{r-1} < 0,\ r > 1 \\ \leq \left[\sum\limits_{t=1}^{n} q_{tj}^r \left(\sum\limits_{k=1}^{M} \lambda_k\, q_{tk} \right)^{1-r} \right]^{\frac{s-1}{r-1}}, \\ \qquad \frac{s-1}{r-1} < 0,\ 0 < r < 1 \text{ or } \frac{s-1}{r-1} > 0,\ r > 1 \end{cases} $$

for each $j = 1, 2, \ldots, M$.

Multiplying by λ_j, summing over all $j = 1, 2, \ldots, M$, subtracting 1 on both sides, multiplying by $(1 - 2^{1-s})^{-1}$ ($s \neq 1$) and simplifying, we get

$$ {}^{1}R_r^s\,(P_1, P_2, \ldots, P_M)) \leq\ {}^{1}R_r^s\,(Q_1, Q_2, \ldots, Q_M)\,, \qquad r \neq 1,\ s \neq 1. $$

For $\alpha = 2$. From relation (24) proceeding on the similar lines as before we get the required result. □

Property 6. If $P_j(B) = \left(\sum_{k=1}^{M} p_{kj}\, b_{1k}, \ldots, \sum_{k=1}^{M} p_{kj}\, b_{nk} \right) \in \Delta_n$ for each $j = 1, 2, \ldots, M$, where $B = \{b_{ik}\},\ b_{ik} \geq 0,\ i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, M$ is a stochastic matrix with $\sum_{i=1}^{n} b_{ik} = 1$ for each $k = 1, 2, \ldots, M$, then

$$ {}^{\alpha}\mathcal{V}_r^s\,(P_1(B), \ldots, P_M(B)) \leq\ {}^{\alpha}\mathcal{V}_r^s\,(P_1, \ldots, P_M) \qquad (\alpha = 1 \text{ and } 2). $$

P r o o f. Follows on the lines similar to Property 5. □

Property 7. If the stochastic matrix B given in Property 6 is such that exists an i_0 for which $b_{i_0k} \geq c > 0,\ \forall\, k = 1, 2, \ldots, M$, then

$$^{\alpha}\mathcal{V}_r^s\,(P_1(B), \ldots, P_M(B)) \leq (1-c)\ ^{\alpha}\mathcal{V}_r^s\,(P_1, \ldots, P_M) \qquad (\alpha = 1 \text{ and } 2),$$

for all $s \geq r > 0$.

P r o o f. For given B, fix B_1 such that

$$B = (1-c)B_1 + cB_2,$$

where

$$^2b_{ik} = \begin{cases} 1, & \text{if } i = i_0, \\ 0, & \text{otherwise.} \end{cases}$$

Using convexity property of $^{\alpha}\mathcal{V}_r^s(P_1, \ldots, P_M)$ ($\alpha = 1$ and 2) and the property 6, we have

$$\begin{aligned} ^{\alpha}\mathcal{V}_r^s\,(P_1(B), \ldots, P_M(B)) &\leq (1-c)\,^{\alpha}\mathcal{V}_r^s\,(P_1(B_1), \ldots, P_M(B_1)) + {}^{\alpha}\mathcal{V}_r^s\,(P_1(B_2), \ldots, P_M(B_2)) \\ &\leq (1-c)\ ^{\alpha}\mathcal{V}_r^s\,(P_1, \ldots, P_M) \quad (\alpha = 1 \text{ and } 2) \end{aligned}$$

for all $s \geq r > 0,$ since $^{\alpha}\mathcal{V}_r^s\,(P_1(B_2), \ldots, P_M(B_2)) = 0 \quad (\alpha = 1$ and 2).

3. APPLICATIONS

In this section, we shall specify some applications of the unified (r,s)-divergence measures given in Section 1. The applications are given towards income inequality, deflation factor, generalized mutual information and Markov chains.

3.1. Generalized measures of income inequality

Following the approach of Nayak and Gastwirth [10], the generalized measures of income inequality are defined as:

$$^{\alpha}\mathcal{I}_r^s\,(P_1, P_2, \ldots, P_M) = \frac{^{\alpha}\mathcal{V}_r^s(P_1, \ldots, P_M)}{\mathcal{E}_r^s\left(\sum_{j=1}^{M} \lambda_j P_j\right)} \tag{25}$$

for all $r \in (0, \infty)$ and $s \in (-\infty, \infty)$ when $\alpha = 1$ and 2, and $(r,s) \in \Gamma$, when $\alpha = 3$.

Following the approach of Theil [23, 24], the generalized measure of income inequality is written as

$$\mathcal{I}_r^s(P\|U) = \frac{\mathcal{E}_r^s(U) - \mathcal{E}_r^s(P)}{\mathcal{E}_r^s(U)}, \tag{26}$$

where U is uniform distribution and $P \in \Delta_n$. Some particular cases of measure (26) are studied by Kapur [5].

3.2. General mutual information

Let us consider a bidimensional random variable (X, Y) taking the values $(x_i,\, y_j)$, $i = 1, \ldots, n;\ j = 1, 2, \ldots, M$ with joint and marginal probability distributions given by

$$P_{XY} = \{p(x_i,\, y_j)\}\,, \quad P_X = \{p(x_i)\} \quad \text{and} \quad P_Y = \{p(y_j)\}$$

for all $i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, M$.
The conditional probability distributions are given by

$$P_{X|Y=y_j} = \{p(x_i \,|\, y_j)\} \quad \text{and} \quad P_{Y|X=x_i} = \{p(y_j \,|\, x_i)\}$$

for all $i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, M$.
Let us also denote

$$P_X \times P_Y = \{p(x_i)p(y_j)\}, \qquad i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, M.$$

Let us take $\lambda_j = p(y_j)$ and $p_{ij} = p(x_i \,|\, y_j)$, then from (11), we have

$${}^1R_r^s(P_1, \ldots, P_M) = \sum_{j=1}^{M} p(y_j)\ D_r^s\left(P_{X/Y=y_j} \,\|\, P_X\right).$$

Hence

$${}^1\mathcal{V}_r^s(X;Y) = \sum_{j=1}^{M} p(y_j)\, \mathcal{F}_r^s\left(P_{X\,|\,Y=y_j} \,\|\, P_X\right),$$

for all $r \in (0, \infty)$ and $s \in (-\infty, \infty)$, where in this particular case ${}^1\mathcal{V}_r^s(X;Y) = {}^1\mathcal{V}_r^s(P_1, \ldots, P_M)$, and $\mathcal{F}_r^s$ is as given in (6).
Similarly, we can write

$${}^2\mathcal{V}_r^s(X;Y) = \mathcal{F}_r^s\left(P_{XY} \,\|\, P_X \times P_Y\right)$$

for all $r \in (0, \infty)$ and $s \in (-\infty, \infty)$.
Again making the same substitutions as above, we have

$$\begin{aligned} &{}^3R_r^s(P_1, \ldots, P_M) = \\ = \ &\left(2^{1-s} - 1\right)^{-1} \left\{ \left[\sum_{i=1}^{n} \left(\sum_{j=1}^{M} p(y_j)\, p(x_i \,|\, y_j) \right)^r \right]^{\frac{s-1}{r-1}} - 1 \right\} = H_r^s(X) - H_r^s(X|Y). \end{aligned}$$

Hence

$${}^3\mathcal{V}_r^s(X;Y) = \mathcal{E}_r^s(X) - \mathcal{E}_r^s(X|Y), \tag{27}$$

for all $(r, s) \in \Gamma$.

In particular, when $r = s = 1$, we have

$$\begin{aligned} {}^1\mathcal{V}_1^1(X;Y) = {}^2\mathcal{V}_1^1(X;Y) &= {}^3\mathcal{V}_1^1(X;Y) = R(X;Y) = \sum_{j=1}^{M} p(y_j) D\big(P_{X|Y=y_j} \,\big\|\, P_X\big) \\ &= D\left(P_{XY} \,\|\, P_X \times P_Y\right) = H(X) - H(X|Y), \end{aligned}$$

where $H(X)$ and $H(X|Y)$ are the Shannon's entropy and Shannon's conditional entropy respectively.

The measure $R(X;Y)$ is famous in the literature on Information Theory as *mutual information* between the random variables X and Y. We call the measures ${}^\alpha\mathcal{V}_r^s(X;Y)$ ($\alpha = 1, 2$ and 3), the *unified (r, s)-mutual information.*

For the three discrete random variables X, Y and W, let us define the expressions ${}^\alpha\mathcal{V}_r^s$ ($\alpha = 1, 2$ and 3) as follows:

$${}^\alpha\mathcal{V}_r^s(X;Y|W) = \sum_{l=1}^{t} p(w_l)\ {}^\alpha\mathcal{V}_r^s(X;Y \,|\, W = w_l),$$

where for each value w_l of W, we have

$$\begin{aligned} {}^1\mathcal{V}_r^s(X;Y \,|\, W = w_l) &= \sum_{j=1}^{M} p(y_j \,|\, w_l)\, \mathcal{F}_r^s\left(P_{X\,|\,Y=y_j, W=w_l}, \big\|\, P_{X|W=w_l}\right), \\ {}^2\mathcal{V}_r^s(X;Y \,|\, W = w_l) &= \mathcal{F}_r^s\left(P_{XY\,|\,W=w_l} \,\big\|\, P_{X\,|\,W=w_l} \times P_{X\,|\,W=w_l}\right), \end{aligned}$$

and

$$^{3}\mathcal{V}_r^s\left(X;Y\,|\,W=w_l\right)=\mathcal{E}_r^s\left(X\,|\,W=w_l\right)-\mathcal{E}_r^s\left(X\,|\,Y,W=w_l\right),$$

with

$$\mathcal{E}_r^s\left(X\,|\,Y,W=w_l\right)=\sum_{j=1}^{M}p\left(y_j\,|\,w_l\right)\;\mathcal{E}_r^s\left(P_{X|Y=y_j,W=w_l}\right)$$

for all $r\in(0,\infty)$ and $s\in(-\infty,\infty)$ when $\alpha=1$ and 2, and $(r,s)\in\Gamma$ when $\alpha=3$. The expressions $^{2}\mathcal{V}_r^s$ and $^{3}\mathcal{V}_r^s$ can be also understood as follows :

$$^{2}\mathcal{V}_r^s\left(X;Y\,|\,W\right)=\mathcal{F}_r^s\left(P_{XY|W}\,\middle\|\,P_{X|W}\times P_{Y|W}\right)$$

and

$$^{3}\mathcal{V}_r^s(X;Y\,|\,W)=\mathcal{E}_r^s(X\,|\,W)-\mathcal{E}_r^s(X\,|\,Y,W).$$

The following proposition holds.

Proposition 1.

(i) For all $r\in(0,\infty)$ and $s\in(-\infty,\infty)$, we have
- (a) $^{\alpha}\mathcal{V}_r^s(X;Y)\geq 0$ ($\alpha=1$ and 2) with equality iff X and Y are independent;
- (b) $^{\alpha}\mathcal{V}_r^s(X;Y\,|\,W)\geq 0$ ($\alpha=1$ and 2) with equality iff X and Y are independent given W.

(ii) For all $(r,s)\in\Gamma$, we have
- (a) $^{\alpha}\mathcal{V}_r^s(X;Y)\geq 0$ with equality iff X and Y are independent;
- (b) $^{\alpha}\mathcal{V}_r^s(X;Y\,|\,W)\geq 0$ with equality iff X and Y are independent given W.

Proof. Part (i) (a) and (b) follows from the Property 1. In order to prove part (ii) (a) and (b) it is sufficient to prove (b) part, i. e., equivalent to prove the following:

$$\mathcal{E}_r^s(X\,|\,Y,W)\leq\mathcal{E}_r^s(X\,|\,Y)$$

with equality iff X and Y are independent given W. It can be proved by using concavity of $\mathcal{E}_r^s$ for $(r,s)\in\Gamma$ (cf. [20]).

3.3. Markov chain

We shall now apply the concept of unified (r,s)-mutual information discussed above to Markov Chains.

Definition (Markov chain). A sequence of random variables X_1, $X_2,\ldots$ forms a Markov chain denoted by $X_1\ominus X_2\ominus\ldots$ if for every i, the random variable X_{i+1} is conditionally independent of $(X_1,X_2,\ldots,X_{i-1})$ given X_i.

Proposition 2. The random variables X, Y and W form a Markov chain, i. e., $X\ominus Y\ominus W$ iff $^{\alpha}\mathcal{V}_r^s(X;W\,|\,Y)=0$ ($\alpha=1$, 2 and 3).

The proof is obvious from the definitions and Proposition 1.

Proposition 3. If $X \ominus Y \ominus W$, then

(a) $^{\alpha}\mathcal{V}_r^s(X;W) \leq \begin{cases} ^{\alpha}\mathcal{V}_r^s(X;Y) \\ ^{\alpha}\mathcal{V}_r^s(X;W) \end{cases}$

for all $r \in (0,\infty)$, and $s \in (-\infty,\infty)$ when $\alpha = 1$ and 2, and $(r,s) \in \Gamma$, when $\alpha = 3$.

(b) $\mathcal{E}_r^s(X\,|\,Y) \leq \mathcal{E}_r^s(X\,|\,W)$, for all $(r,s) \in \Gamma$.

Proof. (a) For $\alpha = 1$ and 2 the result follows from Property 6. For $\alpha = 3$, we have the following identity :

$$^{3}\mathcal{V}_r^s(X;W) + \,^{3}\mathcal{V}_r^s(X;Y\,|\,W) = \,^{3}\mathcal{V}_r^s(X;Y) + \,^{3}\mathcal{V}_r^s(X;W\,|\,Y).$$

Since X, Y and W form a Markov chain, then by Proposition 2, $^{3}\mathcal{V}_r^s(X;W\,|\,Y) = 0$. Also, $^{3}\mathcal{V}_r^s(X;Y\,|\,W) \geq 0$. Thus, the required result follows immediately from the above identity.

(b) From Proposition 2, we have

$$^{3}\mathcal{V}_r^s(X;W\,|\,Y) = 0$$

for $(r,s) \in \Gamma$. This implies that

$$\mathcal{E}_r^s(X\,|\,Y) = \mathcal{E}_r^s(X\,|\,Y,W) \leq \mathcal{E}_r^s(X\,|\,W), \qquad \text{(from Prop. 1 (b))}$$

for all $(r,s) \in \Gamma$, whenever X, Y and W forms a Markov chain. □

Proposition 4. If $X \ominus Y \ominus W \ominus T$, then

$$^{\alpha}\mathcal{V}_r^s(X;T) \leq \,^{\alpha}\mathcal{V}_r^s(Y;W)$$

for all $r \in (0,\infty)$, $s \in (-\infty,\infty)$ when $\alpha = 1$ and $(r,s) \in \Gamma$, when $\alpha = 3$.

Proof. Since X, Y, W and T forms a Markov chain, then X, Y and T and Y, W and T also form Markov chains. Applying Proposition 3 (a) over these two sub-Markov chains, we get the required result. □

3.4. Deflation factor

Nayak [9], considered the following decomposition for the *entropy of degree s*

$$\mathcal{E}_s^s(X,Y) = \mathcal{E}_s^s(X) + \sum_{i=1}^{n} p(x_i)\, w_s^s(p(x_i))\, \mathcal{E}_s^s(Y\,|\,X = x_i), \qquad s > 0 \tag{28}$$

where $w_s^s\,(p(x_i))$ is the "deflation factor" (cf. [11]) given by

$$w_s^s(p(x_i)) = p(x_i)^{s-1}.$$

The expression (28) given in [9] is for one parameter. This can be generalized for two parameter family of measures in the following way:

$$\mathcal{E}_r^s(X,Y) \begin{cases} \leq \mathcal{E}_r^s(X) + \sum\limits_{i=1}^{n} p(x_i)\, w_r^s(p(x_i))\, \mathcal{E}_r^s(Y\,|\,X = x_i) & r \geq s \geq 2 - 1/r \geq 1 \\ \geq \mathcal{E}_r^s(X) + \sum\limits_{i=1}^{n} p(x_i)\, w_r^s(p(x_i))\, \mathcal{E}_r^s(Y\,|\,X = x_i) & 1 \geq r \geq s \geq 2 - 1/r \end{cases} \tag{29}$$

where

$$w_s^s(p(x_i)) = p(x_i)^{r\frac{s-1}{r-1}-1}, \qquad r \neq 1.$$

As specified in [9], here also the above expression (29) does not applies in the case of Rnyi's entropy of order r. In particular, when $r = s$, the expression (29) reduces to (28). For the proof of inequalities (29) refer to [13].

ACKNOWLEDGEMENT

This work has been done during one of the authors (I. J. Taneja) stay with the "Departamento de Estadstica e I.Ȯ., Universidad Complutense de Madrid, Madrid, Spain", for that he is thankful to the above mentioned University for providing facilities and financial support. The authors also thanks to the referee for valuable comments on an earlier version of the paper. This work was partially supported by the Direccion General de Investigacin Cientfica y Tcnica (DGCYT) under the contract PS89-0019.

(Received December 29, 1990.)

REFERENCES

[1] E. F. Beckenbach and R. Bellman: Inequalities. Springer-Verlag, Berlin – Heidelberg – New York 1971.
[2] J. Burbea and C. R. Rao: Entropy differential metric, distance and divergence measures in probability spaces: a unified approach. J. Multivariate Analysis *12* (1982), 575–596.
[3] J. Burbea and C. R. Rao: On the convexity of some divergence measures based on entropy functions. IEEE Trans. Inform. Theory *IT-28* (1982), 489–495.
[4] T. S. Ferguson: Mathematical Statistics. Academic Press, New York 1967.
[5] J. N. Kapur: Entropic measures of economic inequality. Indian J. Pure Appl. Math. *17* (1986), 3, 273–285.
[6] J. N. Kapur: On measures of divergence based on Jensen difference. Nat. Acad. Sci. Lett. *11* (1988), 1, 23–27.
[7] S. Kullback and A. Leibler: On the information and sufficiency. Ann. Math. Statist. *22* (1951), 79–86.
[8] A. W. Marshall and I. Olkin: Inequalities: Theory of Majorization and its Applications. Academic Press, New York 1979.
[9] T. K. Nayak: Applications of Entropy Functions in Measurement and Analysis of Diversity. Ph. D. Thesis, Univ. of Pittsburg 1983.
[10] T. K. Nayak and J. L. Gastwirth: The use of diversity analysis to assess the relative influence of factors affecting the income distributions. J. Business & Economic Statist. *7* (1989), 4, 453–460.
[11] G. P. Patil and C. Taillie: Diversity as a concept and its measurement. J. Amer. Statist. Assoc. *7* (1982), 1, 548–567.
[12] C. R. Rao: Diversity and dissimilarity coefficients: a unified approach. J. Theoret. Popul. Biology *21* (1982), 24–43.
[13] P. N. Rathie and I. J. Taneja: Unified (r,s)-entropy and its bivariate measures. Inform. Sci. *46* (1990), 114–121.
[14] A. Renyi: On measures of entropy and information. In: Proc. 4th Berkeley Symp. Math. Statist. Prob. *1* (1961), 547–561.
[15] P. K. Sahoo and A. K. C. Wong: Generalized Jensen difference based on entropy functions. Kybernetika *24* (1988), 4, 241–250.
[16] C. E. Shannon: A mathematical theory of communication. Bell System Tech. J. *27* (1948), 379–423.
[17] B. D. Sharma and D. P. Mittal: New non-additive measures of relative information. J. Combin. Inform. System Sci. *2* (1977), 122–133.
[18] R. Sibson: Information radius. Z. Wahrsch. verw. Geb. *14* (1969), 149–160.
[19] I. J. Taneja: Unified measures of information applied to Markov chains and sufficiency. J. Combin. Inform. System Sci. *11* (1986), 99–109.
[20] I. J. Taneja: On generalized information measures and their applications. Ad. Electronics and Electron Physics *76* (1989), 327–413.
[21] I. J. Taneja, L. Pardo, D. Morales and M. L. Menendez: On generalized information and divergence measures and their applications: a brief review. Qestii *13* (1989), 47–74.
[22] I. J. Taneja, L. Pardo and M. L. Menendez: Generalized divergence measures and the probability of error. Communicated.
[23] H. Theil: Economics and Information Theory. North-Holland, Amsterdam 1967.
[24] H. Theil: Statistical Decomposition Analysis. North-Holland, Amsterdam 1972.

Prof. Dr. Maria Luisa Menndez, Departamento de Matemtica Aplicada, Escuela Tcnica Superior de Arquitectura, Universidad Politcnica de Madrid, 28040-Madrid. Spain.

Prof. Dr. Leandro Pardo, Departamento de Estadstica e I. O., Facultad de Matemticas, Universidad Complutense de Madrid, 28040-Madrid. Spain.

Prof. Dr. Inder Jeet Taneja, Departamento de Matematica, Universidade Federal de Santa Catarina, 88049-Florianpolis, SC. Brasil.