QÜESTIIÓ, Vol 13, nº 1,2,3 pp. 47-73, 1989

ON GENERALIZED INFORMATION AND DIVERGENCE MEASURES AND THEIR APPLICATIONS: A BRIEF REVIEW

* TANEJA, I.J., PARDO, L.
MORALES, D. MENÉNDEZ, M.L.

The aim of this review is to give different two parametric generalizations of the following measures: directed divergence (Kullback and Leibler, 1951), Jensen difference divergence (Burbea and Rao 1982 a, b; Rao, 1982) and Jeffreys invariant divergence (Jeffreys, 1946). These generalizations are put in the unified expression and its properties are studied. The applications of generalized information and divergence measures towards comparison of experiments and the connections with Fisher information measure are also given.

Keywords: Shannon entropy, generalized information and divergence measures, inequalities, comparison of experiments, Fisher information.

1. INTRODUCTION

Let

$$\Delta_n = \left\{ P = (p_1, p_2, \dots, p_n),\ p_i \geq 0,\ \sum_{i=1}^{n} p_i = 1 \right\},\ n \geq 2$$

–Taneja, I.J. on leave from Departamento de Matemática, Universidad Federal de Santa Catarina, 88.049-Florianópolis, SC, Brazil.
–Taneja, I.J., Pardo, L. and Morales, D. - Departamento de Estadística e I.O. Facultad de Matemáticas. Universidad Complutense de Madrid. 28040-Madrid. SPAIN.
–Menéndez, M.L.- Departamento de Mat. Aplicada. E.T.S. de Arquitectura. U.P. de Madrid. 28040-Madrid. SPAIN.
–Article rebut l'octubre de 1989.

be the set of all complete finite discrete probability distributions. It is known that the Shannon's entropy satisfies the following inequalities

$$H(P) \leq H(P||Q) \tag{1}$$

and

$$\frac{H(P)+H(Q)}{2} \leq H\left(\frac{P+Q}{2}\right) \tag{2}$$

for all $P, Q \in \Delta_n$, with equality iff $P = Q$ i.e., $p_i = q_i$, $\forall i = 1, \dots, n$, where

$$H(P) = -\sum_{i=1}^{n} p_i \ln p_1 \tag{3}$$

and

$$H(P||Q) = -\sum_{i=1}^{n} p_i \ln q_i \tag{4}$$

It is understood that $0 \ln 0 = 0 \ln \frac{0}{0} = 0$ and $p_i = 0$ as and when q_i for some i, and vice-versa.

The measure $H(P)$ is the **Shannon's entropy** (Shannon, 1948) and the measure $H(P||Q)$ is the **inaccuracy** (Kerridge, 1961). The inequality (1) is known as **Shannon-Gibbs inequality**, and the inequality (2) arises due to **concavity** property of Shannon's entropy.

The difference

$$D(P||Q) = H(P||Q) \; - \; H(P) = \sum_{i=1}^{n} p_i \ln \frac{p_1}{q_i} \quad . \tag{5}$$

is known **directed divergence** (Kullback and Leibler, 1951). And the difference

$$R(P||Q) = H\left(\tfrac{P+Q}{2}\right) - \tfrac{H(P)+H(Q)}{2}$$

(6)

$$= \sum_{i=1}^{n} \left[\frac{p_i \ln p_i + q_i \ln q_i}{2} - \left(\frac{p_i + q_i}{2}\right) \ln \left(\frac{p_i + q_i}{2}\right) \right]$$

is known **information radius** (Sibson, 1969) or **Jensen differencen divergence measure** (Burbea and Rao, 1982a,b; Rao, 1982).

By simple calculations, we can write

$$R(P||Q) = \frac{1}{2}\left[D\left(P||\frac{P+Q}{2}\right) + D\left(Q||\frac{P+Q}{2}\right)\right] \tag{7}$$

The measure of directed divergence, $D(P||Q)$ is not symmetric in P and Q. Its simmetric version i.e., Jeffreys **invariant** (Jeffreys, 1946) is given by

$$J(P||Q) = D(P||Q) + D(Q||P) \tag{8}$$

The measure $J(P||Q)$ is known in the literature as **J-divergence**. For simplicity, we call the measure $R(P||Q)$ as **R-divergence**

Burbea and Rao (1982a) and Sgarro (1981) established an inequality between the measure (6) and (8) given by

$$J(P||Q) \geq 4\,R(P||Q) \tag{9}$$

The aim of this review is to give different two-parametric generalizations of the measures (5), (6) and (8) and to study their properties and applications. These generalizations are put in the form of unified expressions. Some inequalities generalizing the inequality (9) are also presented. In particular, entropy-type measures and measurement of generalized income inequelity are also specified. The applications of generalized information and divergence measures towards comparison of experiments and the connections with Fisher information measure are given. The work reported in this review is due to Taneja (1983, 1986c, 1987, 1989), Taneja et al. (1989b; 1990a, b), Pardo et al. (1989b) and Menéndez et al. (1990). Some applications of generalized measures to statistical pattern recognition can be seen in Taneja (1989, 1990).

We see the R-divergence (6) depends either on Shannon's entropy (3) or via Eq. (7) on directed divergence (5). While, the J-divergence (8) dependes only on directed divergence (3). In order to generalized the R & J-divergences first we need the generalizations of the directed divergence (3). While the generalizations of the Shannon's entropy (3) are obtained as a particular case of the generalizations of the directed divergence.

2. UNIFIED (r,s)-DIRECTED DIVERGENCE

Rényi (1961) fisrt presented a scalar parametric generalization of directed divergence (3) given by

$$D_r^1(P||Q) = (r-1)^{-1} \ln \left\{ \sum_{i=1}^{n} p_i^r \, q_i^{1-r} \right\}, \; r \neq 1, \; r > 0$$

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

Sharma and Mittal (1977) studied two parametric generalization of $D(P||Q)$ including $D_r^1(P||Q)$ as a limiting case given by

$$D_r^s(P||Q) = (s-1)^{-1} \left\{ \left[\sum_{i=1}^{n} p_i^r \, q_i^{1-r} \right]^{\frac{s-1}{r-1}} - 1 \right\}, \; r \neq 1, \; s \neq 1, \; r > 0, \; s > 0$$

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

$$D_s^s(P||Q) = (s-1)^{-1} \left\{ \sum_{i=1}^{n} p_i^s \, q_i^{1-s} - 1 \right\}, \; s \neq 1, \; s > 0$$

The measure $D_s^s(P||Q)$ has also been studied extensively by many authors. For a brief review refer Mathai and Rathie (1975) and Taneja (1979).

The following limiting cases are easy to check:

$$\lim_{s \to 1} D_r^s(P||Q) = D_r^1(P||Q); \; \lim_{r \to 1} D_r^s(P||Q) = D_1^s(P||Q);$$

$$\lim_{r \to 1} D_r^1(P||Q) = \lim_{s \to 1} D_s^s(P||Q) = \lim_{s \to 1} D_1^s(P||Q) = D(P||Q).$$

where

$$D_1^s(P||Q) = (s-1)^{-1} \left\{ \exp_e \left[(s-1) \sum_{i=1}^{n} p_i \ln \frac{p_i}{q_i} \right] - 1 \right\}, \; s \neq 1$$

Instead of studying the measures $D(P||Q)$, $D_r^1(P||Q)$, $D_s^s(P||Q)$, $D_1^s(P||Q)$ and $D_r^s(P||Q)$ separately, we can study them jointly. In order to do so Taneja (1989) put these measures in the unified expression and relaxed the condition of positivity of s. This unification is as follows:

$$\mathcal{F}_r^s(P||Q) = \begin{cases} D_r^s(P||Q), & r \neq 1,\ s \neq 1 \\ D_1^s(P||Q), & r = 1,\ s \neq 1 \\ D_r^1(P||Q), & r \neq 1,\ s = 1 \\ D(P||Q), & r = 1,\ s = 1 \end{cases}$$

for all $P, Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$. The measure $D_s^s(P||Q)$ don't appear in the unified expression (10), because it is a particular case of $D_r^s(P||Q)$, when $r = s$. Hence, it is already contained in it. The unified expression, $\mathcal{F}_r^s(P||Q)$ is called (Taneja, 1989), the **unified (r,s)-directed divegence**.

In the following section, we give two different ways to generalize parametrically the R and J-divergences.

3. UNIFIED (r,s)-DIVERGENCE MEASURES

We see the R and J-divergences given by the Eqs. (7) and (8) respectively depend on the directed divergence, $D(P||Q)$. Based on the unified expression $\mathcal{F}_r^s(P||Q)$ and the Eqs. (7) and (8), we can generalize the R and J-divergences. This we have done in the first generalization. An alternative approach to generalize the R and J-divergence is also given, and it is based on an expression appearing in the particular case of the first generalization.

3.1 FIRST GENERALIZATIONS

By replacing $D(P||Q)$ by $\mathcal{F}_r^s(P||Q)$ in Eqs. (7) and (8), we get

(11) $$^1\mathcal{V}_r^s(P||Q) = \frac{1}{2}\left[\mathcal{F}_r^s\left(P||\frac{P+Q}{2}\right) + \mathcal{F}_r^s\left(Q||\frac{P+Q}{2}\right)\right]$$

and

(12) $$^1\mathcal{W}_r^s(P||Q) = \mathcal{F}_r^s(P||Q) + \mathcal{F}_r^s(Q||P)$$

respectively, for all $P, Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$.

The generalized Jensen difference divergence measures according to the Eq. (11) are given by the following unified expression:

$$
{}^1\mathcal{V}_r^s(P||Q) = \begin{cases} {}^1R_r^s(P||Q), & r \neq 1,\ s \neq 1 \\ {}^1R_1^s(P||Q), & r = 1,\ s \neq 1 \\ {}^1R_r^1(P||Q), & r \neq 1,\ s = 1 \\ R(P||Q), & r = 1,\ s = 1 \end{cases}
$$

where

$$
{}^1R_r^s(P||Q) = [2(s-1)]^{-1} \left\{ \left[\sum_{i=1}^n p_i^r \left(\tfrac{p_i+q_i}{2}\right)^{1-r} \right]^{\frac{s-1}{r-1}} + \right.
$$

$$
\left. + \left[\sum_{i=1}^n q_i^r \left(\tfrac{p_i+q_i}{2}\right)^{1-r} \right]^{\frac{s-1}{r-1}} - 2 \right\},\ r \neq 1, s \neq s
$$

$$
{}^1R_1^s(P||Q) = [2(s-1)]^{-1} \left\{ \exp_e \left[(s-1) \sum_{i=1}^n p_i \ln \left(\tfrac{2p_i}{p_i+q_i} \right) \right] + \right.
$$

$$
\left. + \exp_e \left[(s-1) \sum_{i=1}^n p_i \ln \left(\tfrac{2p_i}{p_i+q_i} \right) \right] - 2 \right\},\ s \neq 1
$$

and

$$
{}^1R_r^1(P||Q) = [2(r-1)]^{-1} \ln \left\{ \left[\sum_{i=1}^n p_i^r \left(\tfrac{p_i+q_i}{2}\right)^{1-r} \right] \left[\sum_{i=1}^n q_i^r \left(\tfrac{p_i+q_i}{2}\right)^{1-r} \right] \right\},
$$

$r \neq 1$

for all $P, Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$.

The generalized J-divergence measures according to expression (12) are given by the following unified expression:

$$
{}^1\mathcal{W}_r^s(P||Q) = \begin{cases} {}^1J_r^s(P||Q), & r \neq 1,\ s \neq 1 \\ {}^1J_1^s(P||Q), & r = 1,\ s \neq 1 \\ {}^1J_r^1(P||Q), & r \neq 1,\ s = 1 \\ J(P||Q), & r = 1,\ s = 1 \end{cases}
$$

where

$$^1J_r^s(P||Q) = (s-1)^{-1}\left\{\left[\sum_{i=1}^n p_i^r\, q_i^{1-r}\right]^{\frac{s-1}{r-1}} + \left[\sum_{i=1}^n p_i^{1-r}\, q_i^r\right]^{\frac{s-1}{r-1}} - 2\right\},$$

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

$$^1J_1^s(P||Q)= (s-1)^{-1}\left\{\exp_e\left[(s-1)\sum_{i=1}^n p_i \ln \frac{p_i}{q_i}\right] + \right.$$

$$\left. + \exp_e\left[(s-1)\sum_{i=1}^n q_i \ln \frac{q_i}{p_i}\right] - 2\right\},\ s \neq 1$$

and

$$^1J_r^1(P||Q) = (r-1)^{-1}\ln\left\{\left[\sum_{i=1}^n p_i^r\, q_i^{1-r}\right]\left[\sum_{i=1}^n q_i^r p_i^{1-r}\right]\right\},\ r \neq 1$$

for all $P, Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$.

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

(13) $$^1R_s^s(P||Q) = R_s^s(P||Q) =$$

$$= (s-1)^{-1}\left\{\sum_{i=1}^n \left(\frac{p_i^s+q_i^s}{2}\right)\left(\frac{p_i+q_i}{2}\right)^{1-s} - 1\right\},\ s \neq 1,\ s > 0$$

and

(14) $$^1J_s^s(P||Q) = J_s^s(P||Q) =$$

$$= (s-1)^{-1}\left\{\sum_{i=1}^n \left(p_i^s\, q_i^{1-s} + p_i^{1-s}\, q_i^s\right) - 2\right\},\ s \neq 1,\ s > 0$$

The expressions appearing in (13) and (14) are used to give an alternative way for generalizing the R and J-divergences respectively.

3.2 SECOND GENERALIZATIONS

The second generalizations of the Jensen difference divergence measure are based on an expression appearing in (13) and are given by

$$^2\mathcal{V}_r^s(P||Q) = \begin{cases} ^2R_r^s(P||Q), & r \neq 1,\ s \neq 1 \\ ^2R_1^s(P||Q), & r = 1,\ s \neq 1 \\ ^2R_r^1(P||Q), & r \neq 1,\ s = 1 \\ R(P||Q), & r = 1,\ s = 1 \end{cases}$$

where

$$^2R_r^s(P||Q) = (s-1)^{-1}\left\{\left[\sum_{i=1}^n \left(\frac{p_i^r+q_i^r}{2}\right)\left(\frac{p_i+q_i}{2}\right)^{1-r}\right]^{\frac{s-1}{r-1}} - 1\right\},$$

$r \neq 1, s \neq 1$

$$^2R_1^s(P||Q) = (s-1)^{-1}\{\exp_e[(s-1)R(P||Q)] - 1\}, \; s \neq 1$$

$$^2R_r^1(P||Q) = (r-1)^{-1}\ln\left\{\sum_{i=1}^n \left(\frac{p_i^r+q_i^r}{2}\right)\left(\frac{p_i+q_i}{2}\right)^{1-r}\right\}, \; r \neq 1$$

for all $P,Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$.

The second generalizations of J-divergence are based on an expression appearing in (14) and are given by

$$(16) \qquad ^2\mathcal{W}_r^s(P||Q) = \begin{cases} ^2J_r^s(P||Q), & r \neq 1, s \neq 1 \\ ^2J_1^s(P||Q), & r = 1, s \neq 1 \\ ^2J_r^1(P||Q), & r \neq 1, s = 1 \\ J(P||Q), & r = 1, s = 1 \end{cases}$$

where

$$^2J_r^s(P||Q) = 2(s-1)^{-1}\left\{\left[\sum_{i=1}^n \left(\frac{p_i^r q_i^{1-r} + p_i^{1-r} q_i^r}{2}\right)\right]^{\frac{s-1}{r-1}} - 1\right\}, \; r \neq 1, s \neq 1$$

$$^2J_1^s(P||Q) = 2(s-1)^{-1}\left\{\exp_e\left[\left(\frac{s-1}{2}\right) J(P||Q)\right] - 1\right\}, \; s \neq 1$$

and

$$^2J_r^1(P||Q) = 2(r-1)^{-1}\ln\left\{\sum_{i=1}^n \left(\frac{p_i^r q_i^{1-r} + p_i^{1-r} q_i^r}{2}\right)\right\}, \; r \neq 1$$

for all $P,Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$.

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

$$ {}^1\mathcal{V}_s^s(P||Q) = {}^2\mathcal{V}_s^s(P||Q) \text{ and } {}^1\mathcal{W}_s^s(P||Q) = {}^2\mathcal{W}_s^s(P||Q) $$

The measures ${}^t\mathcal{V}_r^s(P||Q)$ ($t = 1$ and 2) are called (Taneja, 1989) the **unified (r,s)-Jensen difference divergence measures** and the measures ${}^t\mathcal{W}_r^s(P||Q)$ ($t = 1$ and 2) are called (Taneja, 1989) the **unified (r,s)-J-divergence measures.**

Remarks

Measures appearing in the unified expressions (11) and (15) i.e., ${}^1\mathcal{V}_r^s(P||Q)$ and ${}^2\mathcal{V}_r^s(P||Q)$ are due to Taneja (1989). Most of the measures appearing in the unified expressions (12) and (16) i.e., ${}^1\mathcal{W}_r^s(P||Q)$ and ${}^2\mathcal{W}_r^s(P||Q)$, are due to Taneja (1983, 1987, 1989), except the measures ${}^1J_r^1(P||Q)$ and $J_s^s(P||Q)$. The measure ${}^1J_r^1(P||Q)$ is due to Burbea (1983) and the measure $J_s^s(P||Q)$ is due to Rathie and Sheng (1981), Rao (1982), and Burbea and Rao (1982a,b).

4. PROPERTIES OF THE GENERALIZED DIVERGENCES

We have five unified expressions given by the Eqs. (10), (11), (12), (15) and (16). For simplicity, let us write them as follow:

$$ {}^1\Phi_r^s(P||Q) = \mathcal{F}_r^s(P||Q) $$
$$ {}^2\Phi_r^s(P||Q) = {}^1\mathcal{V}_r^s(P||Q) $$
$$ {}^3\Phi_r^s(P||Q) = {}^2\mathcal{V}_r^s(P||Q) $$
$$ {}^4\Phi_r^s(P||Q) = {}^1\mathcal{W}_r^s(P||Q) $$

and

$$ {}^5\Phi_r^s(P||Q) = {}^2\mathcal{W}_r^s(P||Q) $$

for all $P, Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$.

Our aim is to present properties like: convesity, Schur-convesity, monotonicity with respect to the parameter, generalized data processing inequality etc... of the unified measures ${}^t\Phi_r^s(P||Q)$ ($t =$ 1,2,3,4 and 5). The definition of convexity for the pair of probability is well known in the literature, while, the Schur-convexity for the pair of distributions is not very well known. Now, we shall define the Schur-convexity for a single probability distribution and a pair of probability distributions and a pair of probability distributions by the concept of majorization.

DEFINITION 1 (Majorization)

For all $P, Q \in \Delta_n$ we say that P is majorized by Q, i.e., $P < Q$ if

(i) $p_1 \geq p_2 \geq \ldots \geq p_n$ and $q_1 \geq q_2 \geq \ldots \geq q_n$ with

$$\sum_{i=1}^{\sigma} p_i \leq \sum_{i=1}^{\sigma} q_i \quad 1 \leq \sigma \leq n$$

or equivalently,

(ii) there is a doubly stochastic matrix $A = \{a_{ik}\}$, $a_{ik} \geq 0$, $i, k = 1, 2, \ldots, n$ with

$$\sum_{i=1}^{n} a_{ik} = \sum_{k=1}^{n} a_{ik} = 1,$$

such that

$$p_i = \sum_{k=1}^{n} a_{ik}\, q_k, \quad i = 1, 2, \ldots, n.$$

DEFINITION 2

A function $F : \Delta_n \rightarrow \Re$ (reals) is Schur-concave on Δ_n if $P < Q$ implies $F(P) \geq F(Q)$. For Schur-convexity, the last inequality is reversed, i.e., $P < Q$ implies $F(P) \leq F(Q)$.

These definitions can be seen in Marshall and Olkin (1979). Their extension for two variables (Taneja, 1986a) is as follows:

DEFINITION 3

If there is a doubly stochastic matrix $A = \{a_{ik}\}$, $a_{ik} \geq 0$, $i, k = 1, 2, \ldots, n$ with

$$\sum_{i=1}^{n} a_{ik} = \sum_{k=1}^{n} a_{ik} = 1,$$

such that

$$P(A) = \left(\sum_{k=1}^{n} p_k a_{1k}, \ldots, \sum_{k=1}^{n} p_k\, a_{nk} \right) \in \Delta_n$$

and

$$Q(A)=\left(\sum_{k=1}^{n} q_k a_{1k}, \ldots, \sum_{k=1}^{n} q_k a_{nk}\right) \in \Delta_n,$$

Then we say the pair $(P(A), Q(A))$ is majorized by the pair (P,Q) in $\Delta_n \times \Delta_n$ i.e., $(P(A), Q(A)) < (P,Q)$ in $\Delta_n \times \Delta_n$.

DEFINITION 4

A function $G : \Delta_n \times \Delta_n \rightarrow \Re$ (reals) is Schur-convex in $\Delta_n \times \Delta_n$ if $(P(A), Q(A)) < (P,Q)$ on $\Delta_n \times \Delta_n$ implies $G(P(A)||Q(A)) \leq G(P||Q)$. For Schur-concavity the last inequality is reversed, i.e,

$$(P(A), Q(A)) < (P,Q) \text{ implies } G(P(A)||Q(A)) \geq G(P||Q)$$

The definitions 3 and 4 were considered by Taneja (1986a) and some of their interpretations can be seen in Vadja [23, pp. 265-267].

The following theorem gives the properties of the unified measures ${}^t\Phi_r^s(P||Q)$ (t = 1,2,3,4 and 5).

THEOREM 1

For all $P, Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$, we have

(i) **(Nonnegativity)** ${}^t\Phi_r^s(P||Q) \geq 0$ (t = 1,2,3,4 and 5) with equality iff $P = Q$.

(ii) **(Continuity)** ${}^t\Phi_r^s(P||Q)$ (t = 1,2,3,4 and 5) are continous functions of the pair (P,Q) and are also continous with respect to the parameters r and s.

(iii) **(Symmetry)** ${}^t\Phi_r^s(P||Q)$ (t = 1,2,3,4 and 5) are symmetric function of their arguments in pair, i.e.,

$${}^t\Phi_r^s(p_1, \ldots, p_n||q_1, \ldots, q_n) = {}^t\Phi_r^s(p_{\tau(1)}, \ldots, p_{\tau(n)}||q_{\tau(1)}, \ldots, q_{\tau(n)})$$

(t = 1,2,3,4 and 5), where τ is any permutation from 1 to n.

(iv) **(Nonadditivity)** For all

$$P_1 = (p_{11}, p_{12}, \ldots, p_{1n}) \in \Delta_n,\ P_2 = (p_{21}, p_{22}, \ldots, p_{2n}) \in \Delta_n,$$

$$Q_1 = (q_{11}, q_{12}, \ldots, q_{1m}) \in \Delta_m,\ Q_2 = (q_{21}, q_{22}, \ldots, q_{2m}) \in \Delta_m,$$

$$P_1 * Q_1 = (p_{11}q_{11}, \ldots, p_{11}q_{1m}, p_{12}q_{11}, \ldots, p_{12}q_{1m}, \ldots, p_{1n}q_{11}, \ldots, p_{1n}q_{1m}) \in \Delta_{nm}, \text{ and}$$

$$P_2 * Q_2 = (p_{21}q_{21}, \ldots, p_{21}q_{2m}, p_{22}q_{21}, \ldots, p_{22}q_{2m}, \ldots, p_{2n}q_{21}, \ldots, p_{2n}q_{2m}) \in \Delta_{nm}$$

we have

$$\begin{aligned} {}^t\Phi_r^s(P_1 * Q_1 || P_2 * Q_2) = \\ &= {}^t\Phi_r^s(P_1||P_2) + {}^t\Phi_r^s(Q_1||Q_2) + \\ &+ (s-1)\, {}^t\Phi_r^s(P_1||P_2)\, {}^t\Phi_r^s(Q_1||Q_2), \end{aligned}$$

for $t = 1,2,3,4$ and 5.

(v) **Monotonicity)** ${}^t\Phi_r^s(P||Q)$ ($t=$ 1,2,3,4 and 5) are increasing functions of r (s fixed) and of s (r fixed). In particular, when $r = s$ the result still holds.

(vi) **(Convexity)** ${}^t\Phi_r^s(P||Q)$ ($t=$ 1,2,3,4 and 5) are convex functions of the pair of probability distributions $(P, Q) \in \Delta_n \times \Delta_n$ for all $s \geq r > 0$.

(vii) **(Schur-Convexity)** ${}^t\Phi_r^s(P||Q)$ ($t=$ 1,2,3,4 and 5) are Schur-convex functions of the pair of probability distributions $(P, Q) \in \Delta_n \times \Delta_n$, i.e., $(P(A), Q(A)) < (P, Q)$ implies

$${}^t\Phi_r^s(P(A)||Q(A)) \leq {}^t\Phi_r^s(P||Q) \quad (t = 1, 2, 3, 4 \text{ and } 5)$$

(viii) **(Generalized data processing inequalities)** Let $P \in \Delta_m$, $Q \in \Delta_m$,

$$P(B) = \left(\sum_{j=1}^{m} p_j b_{1j}, \ldots, \sum_{j=1}^{m} p_j b_{nj}\right) \in \Delta_n$$

and

$$Q(B) = \left(\sum_{j=1}^{m} q_j b_{1j}, \ldots, \sum_{j=1}^{m} q_j b_{nj}\right) \in \Delta_n,$$

where $B = \{b_{ij}\}$, $b_{ij} \geq 0$, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$ is a stochastic matrix with $\sum_{i=1}^{n} b_{ij} = 1$ for each $j = 1, 2, \ldots, m$, then

$${}^t\Phi_r^s(P(B)||Q(B)) \leq {}^t\Phi_r^s(P||Q) \quad (t = 1, 2, 3, 4 \text{ and } 5)$$

(ix) **(Strong generalized data processing inequalities)** If the stochastic matrix B given in part (viii) is such that there exists an i_0 for which $b_{i_0 j} \geq c > 0$, $\forall j = 1, 2, \ldots, m$, then for $P, Q \in \Delta_m$, we have

$${}^t\Phi_r^s(P(B)||Q(B)) \leq (1-c)\,{}^t\Phi_r^s(P||Q) \quad (t = 1, 2, 3, 4 \text{ and } 5)$$

for $s \geq r > 0$.

5. ENTROPY-TYPE MEASURES AND GENERALIZED MEASUREMENT OF INCOME INEQUALITY

Let $U = (1/n, 1/n, \ldots, 1/n) \in \Delta_n$ be a uniform probability distribution. Then for all $P \in \Delta_n$, we can write

$$\mathcal{F}_r^s(P||U) = n^{s-1}\left[\mathcal{E}_r^s(U) - \mathcal{E}_r^s(P)\right] \tag{17}$$

where

$$\mathcal{E}_r^s(P) = \begin{cases} (1-s)^{-1}\left[\left(\sum_{i=1}^n p_i^r\right)^{\frac{s-1}{r-1}} - 1\right], & r \neq 1,\ s \neq 1 \\ (1-s)^{-1}\left[e^{(s-1)\sum_{i=1}^n p_i \ln p_i} - 1\right], & r = 1,\ s \neq 1 \\ (1-r)^{-1}\ln\left(\sum_{i=1}^n p_i^r\right), & r \neq 1,\ s \neq 1 \\ -\sum_{i=1}^n p_i \ln p_i, & r = 1,\ s = 1 \end{cases} \tag{18}$$

and

$$\mathcal{E}_r^s(U) = \begin{cases} \dfrac{n^{1-s} - 1}{1-s}, & s \neq 1 \\ \ln n, & s = 1 \end{cases} \tag{19}$$

for all $P \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$.

The measure given in (18) are the well known generalizations of Shannon's entropy. The unified expression, $\mathcal{E}_r^s(P)$, is called (Taneja, 1989), the **unified (r,s)-entropy**.

Similar to theorem 1, we have the following theorem:

THEOREM 2

For all $P \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$, the **unified (r,s)-entropy**, $\mathcal{E}_r^s(P)$ satisfies the following properties:

(i) **(Nonnegativity)** $\mathcal{E}_r^s(P) \geq 0$ with equality iff $P = P^0$, where $P^0 \in \Delta_n$ is a probability distribution such that one of the probability is 1 and all others are zero.

(ii) **(Continuity)** $\mathcal{E}_r^s(P)$ is a continous function of P, and is also continous function with respect to the parameters r and s.

(iii) **(Symmetry)** $\mathcal{E}_r^s(P)$ is a symmetric function of its arguments i.e.,

$$\mathcal{E}_r^s(p_1, \ldots, p_n) = \mathcal{E}_r^s(p_{\tau(1)}, \ldots, p_{\tau(n)}),$$

where τ is any permutation from 1 to n.

(iv) **(Nonadditivity)** For

$$P = (p_1, \ldots, p_n) \in \Delta_n,\ Q = (q_1, \ldots, q_m) \in \Delta_m,$$

and

$$P * Q = (p_1q_1, \ldots, p_1q_m, \ldots, p_nq_1, \ldots, p_nq_m) \in \Delta_{nm}$$

we have

$$\mathcal{E}_r^s(P * Q) = \mathcal{E}_r^s(P) + \mathcal{E}_r^s(Q) + (1-s)\mathcal{E}_r^s(P)\mathcal{E}_r^s(Q).$$

(v) **(Monotonicity)** $\mathcal{E}_r^s(P)$ is a decreasing function of r (s fixed) and of s (r fixed). In particular, when $r = s$ the result still holds.

(vi) **(Concavity)** $\mathcal{E}_r^s(P)$ is a concave function of the probability distribution P for $(r,s) \in \Gamma$, where

$$\Gamma = \{(r,s)/r > 0 \text{ with } s \geq r \text{ or } s \geq 2 - 1/r\}$$

(vii) **(Schur-concavity)** $\mathcal{E}_r^s(P)$ is a Schur-concave function of P, i.e.,

$$\mathcal{E}_r^s(P(A)) \geq \mathcal{E}_r^s(P)$$

where $P(A) \in \Delta_n$ is as given in Definition 3.

Remarks

The proof of parts (ii), (iii), (iv) and (vii) of Th. 2 follows directly from Th.1 and Eq. (17). While, the parts (v) and (vi) of Th. 1 follow partially from Th. 1. Thus the details of parts (i), (v) and (vi) are as follows:
(i) From Th. 1 and (17), we have

$$\mathcal{E}_r^s(P) \leq \mathcal{E}_r^s(U)$$

for all $P \in \Delta_n$. This proves that $\mathcal{E}_r^s(P)$ is maximum when the distribution is uniform. But it don't guarantee the nonnegativity of $\mathcal{E}_r^s(P)$. From Th. 2 (vii) and Marshall and Olkin (1979, pp. 7), we have

$$\mathcal{E}_r^s(P^0) \leq \mathcal{E}_r^s(P) \leq \mathcal{E}_r^s(U)$$

with equality on the L.H.S. iff $P = P^0$ and on the R.H.S. iff $P = U$. But $\mathcal{E}_r^s(P^0) = 0$. This proves the nonnegativity of $\mathcal{E}_r^s(P)$ for all $P \in \Delta_n$.

(v) From the expression (19) we see that the measure $\mathcal{E}_r^s(U)$ depends only on s, whatever r may be. Thus the monotonicity of $\mathcal{E}_r^s(P)$ with respect to r follows from theorem 1 (v). And the monotonicity of $\mathcal{E}_r^s(P)$ with respect to s can't be concluded, but it can be proved directly by taking derivatives with respect to s.

(vi) The concavity of $\mathcal{E}_r^s(P)$ for $s \geq r > 0$ follows from the Th. 1 (vi), Van der Pyl (1977) proved that $H_r^s(P)$ is concave for $s \geq 2 - 1/r$, $r > 0$, $r \neq 1$, $s \neq 1$. Thus the concavity of $\mathcal{E}_r^s(P)$ for $s \geq 2 - 1/r$, $r > 0$ follows in view of property (ii).

Apart from the seven properties, the **unified (r,s)-entropy**, $\mathcal{E}_r^s(P)$ enjoys many others, specially with the maximum probability. These are summarized in the following theorem:

THEOREM 3

Let $P = (p_1, \ldots, p_n) \in \Delta_n$, $p_{\max} = \max\{p_1, \ldots, p_n\}$, $0 < r < \infty$. Then the following results hold:

(a)

(i) $\mathcal{E}_r^s(p_{\max}, 1 - p_{\max}) \leq \mathcal{E}_r^s(P)$

(ii) $\mathcal{E}_r^s(P) \leq \mathcal{E}\left(\frac{1 - p_{\max}}{n-1}, \ldots, \frac{1 - p_{\max}}{n-1}, p_{\max}\right)$

(iii)

$$\lim_{r \to \infty} \mathcal{E}_r^s(P) = \begin{cases} (1-s)^{-1}\left[p_{\max}^{s-1} - 1\right], & s \neq 1 \\ -\ln p_{\max}, & s = 1 \end{cases}$$

(b) $1 - p_{\max} \leq \frac{k(s)}{2}\,\mathcal{E}_r^s(P)$,

where (b) holds under the conditions that either $(r > 0, s \leq 2 - 1/n \leq p_{\max} \leq 1/2)$ or $(r > 0, s \geq 2, p_{\max} \geq 1/2)$ or $((r,s) \in \Gamma, p_{\max} \geq 1/2)$, where

$$k(s) = \begin{cases} \frac{1-s}{2^{1-s} - 1}, & s \neq 1 \\ \frac{1}{\ln 2}, & s = 1 \end{cases}$$

(c) For all $P \in \Delta_n, Q \in \Delta_m, n > m, 0 < r < \infty, -\infty < s < \infty$, we have $\mathcal{E}_r^s(P) > \mathcal{E}_r^s(R)$ provided that $H(Q) < -\ln p_{\max}$ and that $\ln m < H(P)$ hold together.

For $P \in \Delta_n$, and $U = (1/n, \ldots, 1/n) \in \Delta_n$, let us define the measure

$$\mathcal{G}_r^s(P\|U) = \frac{\mathcal{E}_r^s(U) - \mathcal{E}_r^s(P)}{\mathcal{E}_r^s(U)} \tag{20}$$

for $0 < r < \infty$ and $-\infty < s < \infty$. Them we can easily check that

$$0 \leq \mathcal{G}_r^s(P\|U) \leq 1 \tag{21}$$

The measures (17) and (20) are the generalizations of the measurement of income inequality (ref. Theil, 1972, 1980). We shall call them, the **unified**

(r,s)-income inequality measures. Some particular cases of (20) can be seen in Kapur (1986).

7. INEQUALITIES AMONG GENERALIZED DIVERGENCE MEASURES

The following results, which give inequalities among the generalized divergence measures, hold:

RESULT 1

For all $P, Q \in \Delta_n$, t = 1,2,3,4 and 5, we have

(i)

$$ {}^t\Phi_r^s(P||Q)\begin{cases} \leq {}^t\Phi_r^1(P||Q), & -\infty < s \leq 1 \\ \geq {}^t\Phi_r^1(P||Q), & 1 \leq s < \infty \end{cases} $$

(ii)

$$ {}^t\Phi_r^s(P||Q)\begin{cases} \leq {}^t\Phi_1^s(P||Q), & 0 < r \leq 1 \\ \geq {}^t\Phi_1^s(P||Q), & 1 \leq r < \infty \end{cases} $$

(iii)

$$ \mathcal{E}_r^s(P)\begin{cases} \geq \mathcal{E}_r^1(P), & -\infty < s \leq 1 \\ \leq \mathcal{E}_r^1(P), & 1 \leq s < \infty \end{cases} $$

(iv)

$$ \mathcal{E}_r^s(P)\begin{cases} \geq \mathcal{E}_1^s(P), & 0 < r \leq 1 \\ \leq \mathcal{E}_1^s(P), & 1 \leq r < \infty \end{cases} $$

RESULT 2

For all $P, Q \in \Delta_n$, $0 < r < \infty$ and $-\infty < s < \infty$, we have

(i)

$$ {}^2\Phi_r^s(P||Q)\begin{cases} \leq {}^3\Phi_r^s(P||Q), & s \leq r \\ \geq {}^3\Phi_r^s(P||Q), & s \geq r \end{cases} $$

(ii)

$$ {}^4\Phi_r^s(P||Q)\begin{cases} \leq {}^5\Phi_r^s(P||Q), & s \leq r \\ \geq {}^5\Phi_r^s(P||Q), & s \geq r \end{cases} $$

(iii) ${}^4\Phi_r^s(P||Q) \geq 2\,{}^2\Phi_r^s(P||Q)$

(iv) ${}^5\Phi_r^s(P||Q) \geq 2\,{}^3\Phi_r^s(P||Q)$

(v) ${}^4\Phi_s^s(P||Q) = {}^5\Phi_s^s(P||Q) \geq 4\,{}^2\Phi_s^s(P||Q) = 4\,{}^3\Phi_s^s(P||Q)$

(vi) ${}^4\Phi_r^1(P||Q) \geq 4\,{}^2\Phi_r^1(P||Q),\ 0 < r \leq 1$

(vii) ${}^5\Phi_r^1(P||Q) \geq 4\,{}^3\Phi_r^1(P||Q),\ 0 < r \leq 1$

8. STATISTICAL APPLICATIONS

8.1 COMPARISON OF EXPERIMENTS

Let $\mathcal{E}_X = \{X, \beta_X, P_\theta;\, \theta \in \Theta\}$ denote a statistical experiment in which a random variable or random vector X defined on some sample space $\mathcal{X}$ is to be observed and the distribution P_θ of X depends on the parameter θ whose values are unknown and lie in some parameter space Θ. We shall assume that there exists a generalized probability density function $f(x/\theta)$ for the distribution P_θ with respect to a σ-finite measure μ. Let also Ξ denote the class of all prior distribution $\xi \in \Xi$, and let $f(x)$ denote the corresponding marginal generalized probability density function ($gpdf$) given by

$$f(x) = \int_\Theta f(x/\theta)\, d\xi$$

Similarly, if we have two prior distributions $\xi_1, \xi_2 \in \Xi$, the corresponding $gpdf$'s are

$$f_i(x) = \int_\Theta f(x/\theta)\, d\xi_i,\ i = 1, 2$$

In this context, the directed divergence, the J-divergence and the Jensen difference divergence measure are given by:

Directed divergence

$${}_X D(\xi_1||\xi_2) = \int_{\mathcal{X}} f_1(x) \ln \frac{f_1(x)}{f_2(x)}\, d\mu$$

J-divergence

$$_X J(\xi_1||\xi_2) = \int_{\mathcal{X}} [f_1(x) - f_2(x)] \ln \frac{f_1(x)}{f_2(x)} \, d\mu$$

Jensen difference divergence measure

$$_X R(\xi_1||\xi_2) = \int_{\mathcal{X}} \left\{ \left[\frac{f_1(x) \ln f_1(x) + f_2(x) \ln f_2(x)}{2} \right] - \right.$$

$$\left. - \left[\left(\frac{f_1(x)+f_2(x)}{2} \right) \ln \left(\frac{f_1(x)+f_2(x)}{2} \right) \right] \right\} d\mu$$

In a similar way, we can write the corresponding unified measures in the integral forms, such as

$^1_X\Phi^s_r(\xi_1||\xi_2) = {}_X\mathcal{F}^s_r(\xi_1||\xi_2)$: unified (r,s)-directed divergence

$$\left. \begin{array}{l} ^2_X\Phi^s_r(\xi_1||\xi_2) = {}^1_X\mathcal{V}^s_r(\xi_1||\xi_2) \\ ^3_X\Phi^s_r(\xi_1||\xi_2) = {}^2_X\mathcal{V}^s_r(\xi_1||\xi_2) \end{array} \right\} : \text{ unified (r,s)-Jensen difference divergence measure}$$

$$\left. \begin{array}{l} ^4_X\Phi^s_r(\xi_1||\xi_2) = {}^1_X\mathcal{W}^s_r(\xi_1||\xi_2) \\ ^5_X\Phi^s_r(\xi_1||\xi_2) = {}^2_X\mathcal{W}^s_r(\xi_1||\xi_2) \end{array} \right\} : \text{ unified (r,s)-J-divergence measure}$$

Consider two arbitrary experiments $\mathcal{E}_X = \{X, \beta_X, P_\theta; \theta \in \Theta\}$ and $\mathcal{E}_Y = \{Y, \beta_Y, Q_\theta; \theta \in \Theta\}$ with the same parameter space Θ. Let Ξ denote the class of all prior distributions on the space Θ. We shall assume that there exist *gpdf*'s $f(x/\theta)$ and $g(y/\theta)$ for the distributions P_θ and Q_θ, with respect to some σ-finite measures μ and ν respectively. Given two prior distributions $\xi_1, \xi_2 \in \Xi$, let $f_i(x)$ denote the marginal *gpdf*

$$\int_\Theta f(x/\theta) \, d\xi_i, \; i = 1 \text{ and } 2$$

and let $^t_X\Phi^s_r(\xi_1||\xi_2)$ ($t = 1,2,3,4$ and 5) denote the generalized divergence measures of information contained in $\mathcal{E}_X$ for discriminating between $f_1(x)$ and $f_2(x)$. In this context, we give the following definition.

DEFINITION

We say that experiment $\mathcal{E}_X$ is prefered to experiment $\mathcal{E}_Y$, denoted by $\mathcal{E}_X \overset{t}{\geq} \mathcal{E}_Y$, if and only if

$$ {}^{t}_{X}\Phi^{s}_{r}(\xi_1||\xi_2) \geq {}^{t}_{Y}\Phi^{s}_{r}(\xi_1||\xi_2) \text{ for all } \xi_1, \xi_2 \in \Xi $$

We say that experiment $\mathcal{E}_X$ and $\mathcal{E}_Y$ are indifferent, denoted by $\mathcal{E}_X \overset{t}{\simeq} \mathcal{E}_Y$, if and only if $\mathcal{E}_X \overset{t}{\geq} \mathcal{E}_Y$ and $\mathcal{E}_Y \overset{t}{\geq} \mathcal{E}_X$.

Some studies towards this direction, including bayesian and Lehmann approaches, has been undertaken by Pardo et al. (1989a), Taneja (1986b), Taneja et al. (1989a) and Morales et al. (1989). Based on the above definition the following theorem gives interesting properties for the unified measures ${}^{t}_{X}\Phi^{s}_{r}(\xi_1||\xi_2)$ (t = 1,2,3,4 and 5).

THEOREM 4

(a) Let $\mathcal{E}_X$ be any experiment and $\mathcal{E}_N$ be the null experiment (i.e., the distribution is independent of θ a.e. μ), then $\mathcal{E}_X \overset{t}{\geq} \mathcal{E}_N$.

(b) Given the compound experiment $(\mathcal{E}_X, \mathcal{E}_Y)$, where $\mathcal{E}_X$ and $\mathcal{E}_Y$ are the corresponding marginal experiments. Then $(\mathcal{E}_X, \mathcal{E}_Y) \overset{t}{\geq} \mathcal{E}_X$ (or $\mathcal{E}_Y$), with indifference iff $f(y/x,\theta)$ is independent of θ (respectively $f(x/y,\theta)$ is independent of θ) for almost every (x,y), where $f(y/x,\theta)$ is the conditional *gpdf* of Y given $X = x$ and $\theta \in \Theta$.

(c) Let $\mathcal{E}_X^{(n)}$ be the resulting experiment after observing $\mathcal{E}_X$ n-times, then $\mathcal{E}_X^{(n)} \overset{t}{\geq} \mathcal{E}_X^{(n-1)}$.

(d) Let $\mathcal{E}_X = \{X, \mathcal{X}, f(x/\theta); \theta \in \Theta\}$ be and experiment and $\{E_i\}_{i \in N}$ be a measurable partition of $\mathcal{X}$. Let us consider another experiment $\mathcal{E}_Y = \{Y, \mathcal{Y}, Q_\theta; \theta \in \Theta\}$ with the σ-algebra generated by $\{E_i\}_{i \in N}$ and with $Q_\theta(E_i) = \int_{E_i} f(x/\theta)\, d\mu(x) \, \forall\, i \in N$. Then $\mathcal{E}_X \overset{t}{\geq} \mathcal{E}_Y$ with indifference iff $T = (\mathcal{E}_X^{(n)})$ based on the experiment $\mathcal{E}_X^{(n)}$, it is verfied that $\mathcal{E}_X^{(n)} \overset{t}{\geq} \mathcal{E}_T$ with indifference iff $f(x/\theta)$ is independent of θ for almost every x.

(e) For all statistic $T = T(\mathcal{E}_X^{(n)})$ based on the experiment $\mathcal{E}_X^{(n)}$, it is verified that $\mathcal{E}_X^{(n)} \overset{t}{\geq} \mathcal{E}_T$ with indifference iff T is a sufficient statistic.

Now we give the relation between the above criterion and Blackwell's criterion. Blackwell's (1951) definition of comparing two experiments states that experiment $\mathcal{E}_X$ is sufficient for experiment $\mathcal{E}_Y$, denoted $\mathcal{E}_X \geq \mathcal{E}_Y$, if there exists a stochastic transformation of X to a random variable $Z(X)$ such that for each $\theta \in \Theta$ the random variable $Z(X)$ and Y have identical distributions. By $\mathcal{E}_Y = \{Y, \beta_{\mathcal{Y}}, Q_\theta;\ \theta \in \Theta\}$ we shall denote a second statistical experiment for which there exists a *gpdf* $g(y/\theta)$ for the distribution Q with respect to a σ-finite measure ν. According to this definition, if $\mathcal{E}_X \geq \mathcal{E}_Y$, then there exists a nonnegative function h satisfying (cf. DeGroot (1970), p. 434).

$$g(y/\theta) = \int_{\mathcal{X}} h(y/x)\, f(x/\theta)\, d\mu \tag{22}$$

and

$$\int_{\mathcal{Y}} h(y/x)\, d\nu = 1$$

If we have two prior distributions $\xi_1, \xi_2 \in \Theta$, after integrating over Θ and changing the order of integration in (22), we get

$$g_i(y) = \int_{\mathcal{X}} h(y/x)\, f_i(x)\, d\mu,\ i = 1, 2 \tag{23}$$

Let I be any measure of information contained in an experiment. If $\mathcal{E}_X \geq \mathcal{E}_Y$ implies $I_X \geq I_Y$, then we say that $\mathcal{E}_X$ is at least as informative as $\mathcal{E}_Y$ in terms of measure I. Goel and DeGroot (1979) applied it for directed divergence. Ferentinos and Papaionnou (1982) applied it for α-order generalization of directed divergence. Taneja (1987) extended it to different generalizations of J-divergence measure having one and two scalar parameters. According to this approach, the results for the unified measures ${}^t_X\Phi^s_r(\xi_1||\xi_2)$ (t = 1,2,3,4 and 5) are summarized in the following theorem.

THEOREM 5

If $\mathcal{E}_X \geq \mathcal{E}_Y$, then ${}^t_X\Phi^s_r(\xi_1||\xi_2) \geq {}^t_Y\Phi^s_r(\xi_1||\xi_2)$ (t = 1,2,3,4,5) for every $\xi_1, \xi_2 \in \Xi$, $0 < r < \infty$ and $-\infty < s < \infty$.

8.2 CONNECTIONS WITH FISHER MEASURE OF INFORMATION

Consider a family $M = \{P_\theta;\ \theta \in \Theta\}$ of probability measures on a measurable space $(\mathcal{X}, \beta_{\mathcal{X}})$ dominated by a finite or σ-finite measure μ. The parameter space Θ can either be an open subset of the real line or an open subset of a n-dimensional Euclidean space $\Re^n$. Let $f(x/\theta) = \dfrac{dP_\theta}{d\mu}$. Then the Fisher (1925) measure of information is given by

$$I_X^F(\theta) = \begin{cases} E_\theta\left[\frac{\partial}{\partial\theta}\log f(x/\theta)\right], & \text{if } \theta \text{ is univariate} \\ \\ E_\theta\left\|\left[\frac{\partial}{\partial\theta_i}\log f(x/\theta)\cdot\frac{\partial}{\partial\theta_j}\log f(x/\theta)\right]\right\|_{n\times n}, & \text{if } \theta \text{ is n-variate} \end{cases}$$

where $\| \ \|_{n\times n}$ denotes an $n \times n$ matrix and E_θ denotes expectation with respect to $f(x/\theta)$.

Some studies towards Fisher information measure applying differential geometric approach has been succesfully carried out by Rao (1945, 1973, 1987), Atkinson and Mitchell (1981), Burbea and Rao (1982), Amari (1984, 1985), Cuadras et al. (1985), Campbell (1985, 1986, 1987), Burbea (1986), Burbea and Oller (1988), Oller (1989), Cuadras (1988), etc. A direct approach has been undertaken by Kagan (1963), Vadja (1973, 1989), Aggarwal (1974), Boekee (1978), Ferentinos and Papaioannou (1981), Taneja (1987), Pardo et al. (1989b), Salicrú (1990), etc.

Let us suppose that the following regularity conditions are satisfied

(a) $f(x/\theta) > 0$ for all $x \in \mathcal{X}$, $\theta \in \Theta$.

(b) $\frac{\partial}{\partial\theta_i} f(x/\theta)$ exists for all $x \in \mathcal{X}$, all $\theta \in \Theta$ and all $i = 1, 2, \ldots n$.

(c) for any $A \in \beta_{\mathcal{X}}$, $\dfrac{\partial}{\partial\theta_i}\displaystyle\int_A f(x/\theta)\,d\mu = \int_A \dfrac{\partial}{\partial\theta_i} f(x/\theta)\,d\mu$ for all i.

Define

$${}^t\Phi_r^s(\theta) = \liminf_{\Delta\theta\to 0}\frac{1}{(\Delta\theta)^2}\,{}^t\Phi_r^s(f(x/\theta)|f(x/\theta+\Delta\theta)),\ t = 1,2,3,4 \text{ and } 5$$

Then the following theorem holds.

THEOREM 6

Let Θ be univariate and let the regularity conditions of the Fisher information measure be satisfied. Also, suppose that $\int_{\mathcal{X}} |\frac{\partial^2}{\partial\theta^2} f(x/\theta)|\,d\mu < \infty$ for all $\theta \in \Theta$ and that the third order partial derivative of $f(x/\theta)$ with respect to θ exists for all $\theta \in \Theta$ and for all $x \in \mathcal{X}$. Then

$$^1\Phi_r^s(\theta) = \frac{r}{2} I_X^F(\theta)$$

$$^2\Phi_r^s(\theta) = {}^3\Phi_r^s(\theta) = \frac{r}{8} I_X^F(\theta)$$

$$^4\Phi_r^s(\theta) = {}^5\Phi_r^s(\theta) = r I_X^F(\theta),$$

for $\theta \in \Theta$, $0 < r < \infty$ and $-\infty < s < \infty$.

9. ACKNOWLEDGEMENT

One of the author (I.J. Taneja) is thankful to the University Complutense de Madrid, Departamento de Estadística e I.O., Madrid, for providing facilities and financial support.

b10. REFERENCES

[1] **Aggarwal, N.L.** (1974). "Sur l'information de Fisher". In: J. Kampe de Feriet, Ed., Theories de l'information, Springer-Verlag, Berlin, 111-117.

[2] **Amari, S.I.** (1984). "Differential Geometry of Statistics: Towards new Developments". In: NATO Workshop on Differential Geometry in Statistical Inference, London, 9-11 April.

[3] **Amari, S.I.** (1985). "Differential-Geometric Methods in Statistic". Lecture Notes in Statistics, Springer-Verlag, Berlin.

[4] **Atkinson, C.** and **Mitchell, A.F.S.** (1981). "Rao's Distance Measure". Sankhÿa, 43, A, 345-65.

[5] **Blackwell, D.** (1951). "Comparison of Experiments". In: Proc. 2nd Berkeley Symp. Math. Statist. Probl. Univ. of Californ. Press, Berkeley, California, 93-103.

[6] **Boekee, D.E.** (1978). "The D_f-Information of Order s". Trans. 8th Prague Conf. on Inform. Th. Vol C, N., 55-66.

[7] **Burbea, J.** (1984). "The Bose-Einstein Entropy of Degree α and Its Jensen Difference". Utilitas Mathematica 25, 225-240.

[8] **Burbea, J.** (1986). "Information Geometry of Probability Spaces". Exposit. Math., 4, 347-378.

[9] **Burbea, J.** and **Oller, J.M.** (1988). "The Information Metric for Univariate Linear Elliptic Models". Statistics & Decision, 6, 209-221.

[10] **Burbea, J.** and **Rao, C.R.** (1982a). "Entropy Differential Metric, Distance and Divergence Measures in Probability Spaces: A unified Approach". J. Multi. Analy., 12, 575-596.

[11] **Burbea, J.** and **Rao, C.R.** (1982b). "On the Convexity of Some Divergence Measures Based on Entropy Functions". IEEE Trans. on Information Theory, IT-28, 489-495.

[12] **Campbell, L.L.** (1985). "The Relation Between Information Theory and the Differential Geometry Approach to Statistics". Information Sciences, 35, 199-210.

[13] **Campbell, L.L.** (1986). "An Extended Cencov Characterization of the Information Metric". Proc. Am. Math. Soc., 98, 135-141.

[14] **Campbell, L.L.** (1987). "Information Theory and Differential Geometry". Department of Math. & Statist., Queen's University Preprint # 1987-12.

[15] **Cuadras, C.M.**, **Oller, J.M.**, **Arcas, A.** and **Rios, M.** (1985). "Métodos Geométricos de la Estadística". Qüestiió, 9(4), 219-250.

[16] **Cuadras, C.M.** (1988). "Distancias Estadísticas". Estadística Española, 30(119), 295-378.

[17] **Degroot, M.H.** (1970). "Optimal Statistical Decisions". Mc-Graw-Hill, New York.

[18] **Ferentinos, K.** and **Papaioannou, T.** (1981). "New Parametric Measures Information". Inform. and Cont., 51, 193-208.

[19] **Ferentinos, K.** and **Papaioannou, T.** (1982). "Information in Experiments and Sufficiency". J. Statist. Plann. & Inferen., 6, 309-317.

[20] **Ferentinos, K.** and **Papaioannou, T.** (1983). "Convexity of Measures of Information and Loss of Information due to Grouping of Observations". J. Comb. Inform. & Syst. Sci., 4, 286-294.

[21] **Fisher, R.A.** (1925). "Theory of Statistical Estimation". Proc. Cambridge Phil. Soc., 22, 700-725.

[22] **Goel, P.K.** and **Degroot, M.H.** (1979). "Comparison of Experiments and Information Measures". Ann. Statist., 7, 1066-1077.

[23] **Jeffreys, H.** (1946). "An Invariant form of the Prior Probability in Estimation Problems". Proc. Royal Soc., Ser A, 186, 453-561.

[24] **Kagan, A.M.** (1963). "On the Theory of Fisher's Amount of Information (in Russian)". Dokl. Acad. Nauk SSSR, 151, 277-278.

[25] **Kapur, J.N.** (1986). "Entropic Measures of Economic Inequality". Indian J. Pure & Appl. Maths., 17 (3), 273-285.

[26] **Kerridge, D.F.** (1961). "Inaccuracy and Inference". J. Royal Stat. Soc. Ser. B, 23 (1), 184-194.

[27] **Kullback, S.** and **Leibler, A.** (1951). "On the Information and Sufficiency". Ann. Math. Statist., 22, 79-86.

[28] **Lindley, D.V.** (1956). "On a Measure of Information provided by an Experiment". Ann. Math. Statist., 27, 986-1005.

[29] **Marshall, A.W.** and **Olkin, I.** (1979). "Inequalities: Theory of Majorization ant its Applications". Academic Press, New York.

[30] **Mathai, A.M.** and **Rathie, P.N.** (1975). "Basic Concepts of Information Theory and Statistics". Wiley, New York.

[31] **Menéndez, M.L.**, **Taneja, I.J.** and **Pardo, L.** (1989). "On Generalized Information Radii and Their Properties". Communicated.

[32] **Morales, D.**, **Taneja, I.J.** and **Pardo, L.** (1989). "Comparison of Experiments Based on ϕ-Measures of Jensen Difference". Communicated.

[33] **Oller, J.M.** (1989). "Some Geometrical Aspects of Data Analysis and Statistics". Statistical Data Analysis and Inference, Ed. Y. Dodge, Elsevier Science Publishers B.V. (North-Holland), 41-58.

[34] **Pardo, L.**, **Morales, D.** and **Taneja, I.J.** (1989a). "λ-Measures of Hypoentropy and Comparison of Experiments: Bayesian Approach". Communicated.

[35] **Pardo, L.**, **Morales, D.** and **Taneja, I.J.** (1989b). "Generalized Jensen Difference Divergence Measures and Fisher Measure of Information ". Communicated.

[36] **Rao, C.R.** (1945). "Information and Accuracy Attainable in the Estimation of Statistical Parameters". Bull. Calcutta Math. Soc., 37, 81-91.

[37] **Rao, C.R.** (1973). "Linear Statistical Inference and its Applications". 2nd Ed., Wiley. New York.

[38] **Rao, C.R.** (1982). "Diversity and Dissimilarity Coefficients: A Unified Approach". J. Theoret. Popul. Biology, 21, 24-43.

[39] **Rao, C.R.** (1987). "Differential Metrics in Probability Spaces". S.S. Gupta (Ed.). Differential Geometry in Statistical Inference, IMS Lecture Notes-Monograph Series 10, Hayward, California, 217-240.

[40] **Rathie, P.N.** and **Sheng, L.T.** (1981). "The J-Divergence of Order α". J. Comb. Inform. & Syst. Sci., 6, 197-205.

[41] **Renyi, A.** (1961). "On Measures of Entropy and Information". Proc. 4th Berkeley Symp. Math. Statist. and Prob., 1, 547-561.

[42] **Salicrú, M.** (1989). "Measures of Information Associated with Csiszar's Divergences". Communicated.

[43] **Sgarro, A.** (1981). "Information Divergence and the Dissimilarity of Probability Distributions". Estratto di Calcolo, XVIII, 293-302.

[44] **Shannon, C.E.** (1948). "A Mathematical Theory of Communication". Bell. Syst. Tech. J,. 27, 379-423.

[45] **Sharma, B.D.** and **Mittal, D.P.** (1977). "New Non-additive Measures of Relative Information". J. Comb. Inform. & Syst. Sci., 2, 122-133.

[46] **Sibson, R.** (1969). "Information Radius". Z. Wahrs und verw Geb., 14, 149-160.

[47] **Taneja, I.J.** (1979). "Some Contributions to Information Theory I (A Survey): On Measures of Information". J. Comb., Inform. & Syst. Sci., 4 (4), 253-274.

[48] **Taneja, I.J.** (1983). "On a Characterization of J-Divergence and its Generalizations". J. Comb., Inform. & Syst. Sci., 8 (3), 206-212.

[49] **Taneja, I.J.** (1986a). "On the Convexity and Schur-Convexity of Burbea and Rao's Divergence Measures and Their Generalizations". 7th National Symposium on Porb. and Statis., Capminas, SP, Brazil.

[50] **Taneja, I.J.** (1986b). "λ-Measures of Hypoentropy and Their Applications". Statistica, XLVI, 465-478.

[51] **Taneja, I.J.** (1986c). "Unified Measures of Information Applied to Markov Chains and Sufficiency". J. Comb., Inform. & Syst. Sci., 11, 99-109.

[52] **Taneja, I.J.** (1987). "Statistical Aspects of Divergence Measures". J. Statist. Planning & Inference, 16, 136-145.

[53] **Taneja, I.J.** (1989). "On Generalized Information Measures and their Applications". Ad. Electronics and Electron Physics, 76, 327-413.

[54] **Taneja, I.J.** (1990). "Bounds on the Probability of Error in Terms of Generalized Information Radii". Information and Sciencie, 46.

[55] **Taneja, I.J.**, **Pardo, L.** and **Morales, D.** (1989a). "λ-Measures of Hypoentropy and Comparison of Experiments: Blackwell and Lhemann Approach". Communicated.

[56] **Taneja, I.J.**, **Pardo, L.** and **Morales, D.** (1989b). "Generalized Jensen Difference Divergence Measures and Comparison of Experiments". Communicated.

[57] **Taneja, I.J.**, **Pardo, L.** and **Morales, D.** (1990a). "Some Inequalities Among Generalized Divergence Measures". Communicated.

[58] **Taneja, I.J., Pardo, L. and Morales, D.** (1990b). "On Unified (r,s)-J-Divergence Measures". Communicated.

[59] **Theil, H.** (1972). "Statistical Descomposition Analysis". North-Holland Pub. Co., Amsterdam.

[60] **Theil, H.** (1980). "The increased use of Statistical Concepts in Economics Analysis". Developments in Statistics, Ed. P.R. Krishnaiah, Academic Press, Vol 3, Chapter 3, 159-215.

[61] **Vadja, I.** (1973). "χ^a-Divergence and Generalized Fisher's Information". Trans. 6th Prague Conf. on Inform. Th., 873-886.

[62] **Vadja, I.** (1989). "Theory of Statistical Inference and Information". Kluvwer Academic Press, Dordrecht, The Netherlands.