Inder Jeet Taneja; Luis Pardo; D. Morales

# $(R, S)$-information radius of type $t$ and comparison of experiments

*Applications of Mathematics*, Vol. 36 (1991), No. 6, 440–455

Persistent URL: `http://dml.cz/dmlcz/104481`

# $(R, S)$-INFORMATION RADIUS OF TYPE $t$ AND COMPARISON OF EXPERIMENTS

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

(Received February 14, 1990)

*Summary:* Various information, divergence and distance measures have been used by researchers to compare experiments using classical approaches such as those of Blackwell, Bayesian etc. Blackwell's [1] idea of comparing two statistical experiments is based on the existence of stochastic transformations. Using this idea of Blackwell, as well as the classical bayesian approach, we have compared statistical experiments by considering unified scalar parametric generalizations of Jensen difference divergence measure.

*Keywords:* Divergence measures; Information radius; Statistical experiment; Sufficiency of experiments; Shannon's entropy.

*AMS subject classification:* 62B10, 94A15.

## 1. INTRODUCTION

Several measures have been introduced in literature on Information Theory and Statistics as measures of information. The most commonly used is Shannon's (1948) entropy. It gives the amount of uncertainty concerning the outcome of an experiment. Kullback and Leibler (1951) introduced a measure associated with two distributions of an experiment. It expresses the amount of information supplied by the data for discriminating among the distributions. As symmetric measure, Jeffreys-Kullback-Leibler's $J$-divergence is commonly used. Also, the measure arising due to concavity of Shannon's entropy known as information radius (Sibson [21]) is getting importance towards applications. Burbea and Rao [3, 4] called information radius of Jensen difference divergence measure and also presented its one real parametric generalization. Some properties and applications of the Jensen difference divergence measure and its generalization can be seen in [2]–[5], [16], [17], [19] and [27]. Recently, Taneja [26] introduced different ways how to generalize the Jensen difference divergence measure involving two real parameters, while two scalar parametric generalizations of $J$-divergence can be seen in Taneja [22, 25, 26].

*) On leave from Universidade Federal de Santa Catarina, Departamento de Matématica, 88.049 — Florianópolis, SC, BRAZIL.

Blackwell [1] gave an idea of comparing two statistical experiments that is based on the existence of a stochastic transformation. Using this idea many authors, [9], [10], [11], [12], [18], [23], [24] and [25], compared experiments (sometimes called sufficiency of experiments) using different information divergence and distance measures. This idea is very similar to majorization of probability distributions, originally introduced by Schur (ref. Marshall and Olkin [13]). The same is successfully carried out for Markov mappings using differential geometric approaches [6], [7] and [29].

In this paper, our aim is to compare experiments using generalized Jensen difference divergence measures adopting Blackwell's [1] approach. Some study using the classical bayesian approach for comparing experiments is also made. Before doing so, we present in the following section the generalized Jensen difference divergence measures.

## 2. JENSEN DIFFERENCE MEASURES AND THEIR GENERALIZATIONS

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

$$f(x) = \int_\Theta f(x/\theta)\,\mathrm{d}\xi\,.$$

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

$$f_i(x) = \int_\Theta f(x/\theta)\,\mathrm{d}\xi_i\,, \quad i = 1, 2\,.$$

The Shannon (1948) entropy for the marginal gpdf $f$ is given by

$$H(f) = -\int_{\mathscr{X}} f(x)\,\mathrm{Ln}\,f(x)\,\mathrm{d}\mu\,.$$

Using the concavity of Shannon's entropy, we can write

$$\frac{H(f_1) + H(f_2)}{2} \leqq H\left(\frac{f_1 + f_2}{2}\right).$$

The difference

$$(1) \qquad {}_XR(\xi_1 \,\|\, \xi_2) = H\left(\frac{f_1 + f_2}{2}\right) - \frac{H(f_1) + H(f_2)}{2} =$$

$$= \int_{\mathscr{X}} \left[\frac{f_1(x)\,\mathrm{Ln}\,f_1(x) + f_2(x)\,\mathrm{Ln}\,f_2(x)}{2} - \right.$$

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

is known as the information radius (Sibson [21]) or Jensen difference divergence measure (Burbea and Rao [3, 4]). For simplicity, we shall call ${}_X R(\xi_1 \parallel \xi_2)$ the $R$-DIVERGENCE.

Kullback and Leibler [11] were the first to introduce a measure of information between two probability distributions given by

$$(2) \qquad {}_X D(\xi_1 \parallel \xi_2) = \int_{\mathscr{X}} f_1(x) \operatorname{Ln} \frac{f_1(x)}{f_2(x)} \, d\mu .$$

Expression (2) is known in literature as a function of discrimination or relative information or directed divergence between the distributions.

By simple calculation, we can write

$$(3) \qquad {}_X R(\xi \parallel \xi_2) = \frac{1}{2}\left[ {}_X D\left(\xi_1 \,\middle\|\, \frac{\xi_1 + \xi_2}{2}\right) + {}_X D\left(\xi_2 \,\middle\|\, \frac{\xi_1 + \xi_2}{2}\right)\right].$$

Taneja [25] considered two ways how to generalize $R$-divergence, i.e., the Jensen difference divergence measure given by Eq. (1). In the first way, two parametric generalizations of (2) are considered and are substituted in (3). This generalization is as follows:

$$(4) \qquad {}^1_X\mathscr{V}^s_r(\xi_1 \parallel \xi_2) = \begin{cases} {}^1_X R^s_r(\xi_1 \parallel \xi_2), & r \neq 1, \quad s \neq 1, \quad r > 0 \\ {}^1_X R^s_1(\xi_1 \parallel \xi_2), & r = 1, \quad s \neq 1 \\ {}^1_X R^1_r(\xi_1 \parallel \xi_2), & r \neq 1, \quad s = 1, \quad r > 0 \\ {}_X R(\xi_1 \parallel \xi_2), & r = 1, \quad s = 1 \end{cases}$$

where

$$(5) \qquad {}^1_X R^s_r(\xi_1 \parallel \xi_2) = \frac{1}{2(s-1)}\left\{\left(\int_{\mathscr{X}} f_1(x)^r \left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} d\mu\right)^{(s-1)/(r-1)} + \right.$$
$$\left. + \left(\int_{\mathscr{X}} f_2(x)^r \left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} d\mu\right)^{(s-1)/(r-1)} - 2\right\}.$$
$$r \neq 1, \quad s \neq 1, \quad r > 0,$$

$$(6) \qquad {}^1_X R^s_1(\xi_1 \parallel \xi_2) = \frac{1}{2(s-1)}\left\{\exp\,(s-1)\, {}_X R\left(f_1 \,\middle\|\, \frac{f_1 + f_2}{2}\right) + \right.$$
$$\left. + \exp\,(s-1)\, {}_X R\left(f_2 \,\middle\|\, \frac{f_1 + f_2}{2}\right) - 2\right\}, \quad s \neq 1,$$

$$(7) \qquad {}^1_X R^1_r(\xi_1 \parallel \xi_2) = \frac{1}{2(r-1)}\left\{\operatorname{Ln}\left(\int_{\mathscr{X}} f_1(x)^r \left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} d\mu\right) + \right.$$
$$\left. + \operatorname{Ln}\left(\int_{\mathscr{X}} f_2(x)^r \left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} d\mu\right)\right\}, \quad r \neq 1, \quad r > 0 .$$

When $r = s$ in Eqs. (4) or (5), we have

$$(8) \qquad {}^1_X R^s_s(\xi_1 \parallel \xi_2) = \frac{1}{(s-1)} \left\{ \int_{\mathscr{X}} \left( \frac{f_1^s(x) + f_2^s(x)}{2} \right) \times \right.$$

$$\left. \times \left( \frac{f_1(x) + f_2(x)}{2} \right)^{1-s} d\mu - 1 \right\}, \quad s \neq 1, \ s > 0.$$

The second alternative way to generalize the measure (1) (Taneja [26]) is based on Eq. (8) and is given by

$$(9) \qquad {}^2_X \mathscr{V}^s_r(f_1 \parallel f_2) = \begin{cases} {}^2_X R^s_r(\xi_1 \parallel \xi_2), & r \neq 1, \quad s \neq 1, \quad r > 0 \\ {}^2_X R^s_1(\xi_1 \parallel \xi_2), & r = 1, \quad s \neq 1 \\ {}^2_X R^1_r(\xi_1 \parallel \xi_2), & r \neq 1, \quad s = 1, \quad r > 0 \\ {}_X R(\xi_1 \parallel \xi_2), & r = 1, \quad s = 1 \end{cases}$$

where

$$(10) \qquad {}^2_X R^s_r(\xi_1 \parallel \xi_2) = \frac{1}{s-1} \left\{ \left[ \int_{\mathscr{X}} \left( \frac{f_1^r(x) + f_2^r(x)}{2} \right) \times \right. \right.$$

$$\left. \left. \times \left( \frac{f_1(x) + f_2(x)}{2} \right)^{1-r} d\mu \right]^{(s-1)/(r-1)} - 1 \right\}, \quad r \neq 1, \quad s \neq 1, \quad r > 0,$$

$$(11) \qquad {}^2_X R^1_r(\xi_1 \parallel \xi_2) = \frac{1}{r-1} \operatorname{Ln} \left\{ \int_{\mathscr{X}} \left( \frac{f_1(x)^r + f_2(x)^r}{2} \right) \times \right.$$

$$\left. \times \left( \frac{f_1(x) + f_2(x)}{2} \right)^{1-r} d\mu \right\}, \quad r \neq 1, \quad r > 0,$$

$$(12) \qquad {}^2_X R^s_1(\xi_1 \parallel \xi_2) = (s-1)^{-1} \{ \exp((s-1)\, {}_X R(f_1 \parallel f_2) - 1 \}, \quad s \neq 1.$$

The following properties of ${}^1_X \mathscr{V}^s_r(\xi_1 \parallel \xi_2)$ and ${}^2_X \mathscr{V}^s_r(\xi_1 \parallel \xi_2)$ hold (Taneja [26]):

(i) ${}^t_X \mathscr{V}^s_r(\xi_1 \parallel \xi_2) \geqq 0 \quad (t = 1, 2)$ for all $r > 0$ and any $s$;

(ii) ${}^1_X \mathscr{V}^s_r(\xi_1 \parallel \xi_2) \begin{cases} \leqq {}^2_X \mathscr{V}^s_r(\xi_1 \parallel \xi_2), & s \leqq r \\ \geqq {}^2_X \mathscr{V}^s_r(\xi_1 \parallel \xi_2), & s \geqq r. \end{cases}$

We have equality sign in (ii) when $r = s$.

We call the measures ${}^t_X V^s_r(\xi_1 \parallel \xi_2)$ $(t = 1, 2)$, the $(r, s)$-*information radius of type t.*

## 3. THE UNIFIED $(r, s)$-JENSEN DIFFERENCE DIVERGENCE MEASURES AS A CRITERION FOR THE COMPARISON OF EXPERIMENTS

Consider two arbitrary experiments $\mathscr{E}_X = \{X, \beta_{\mathscr{X}}, P_\theta; \theta \in \Theta\}$ and $\mathscr{E}_Y = \{Y, \beta_{\mathscr{Y}}, Q_\theta; \theta \in \Theta\}$ with the same parameter space $\Theta$. Let $\Xi$ denote the class of all prior distributions on the space $\Theta$. We shall assume that there exist gpdf's $f(x/\theta)$ and $g(y/\theta)$ for the distributions $P_\theta$ and $Q_\theta$ with respect to some $\sigma$-finite measures $\mu$ and $\nu$, respectively. Given two prior distributions $\xi_1, \xi_2 \in \Xi$, let $f_1(x)$ denote the marginal gpdf

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

and let ${}^t_X V^s_r(\xi_1 \,\|\, \xi_2)$ $(t = 1, 2)$ denote the $(r, s)$-Jensen difference divergence measure of information contained in $\mathscr{E}_X$ for discriminating between $f_1(x)$ and $f_2(x)$. In this context, we say that experiment $\mathscr{E}_X$ is preferred to experiment $\mathscr{E}_Y$, denoted by $\mathscr{E}_X \overset{t}{\geqq} \mathscr{E}_Y$, if and only if

$${}^t_X V^s_r(\xi_1 \,\|\, \xi_2) \geqq {}^t_Y V^s_r(\xi_1 \,\|\, \xi_2) \text{ for all } \xi_1, \xi_2 \in \Xi.$$

We say that experiments $\mathscr{E}_X$ and $\mathscr{E}_Y$ are indifferent, denoted by $\mathscr{E}_X \overset{t}{\simeq} \mathscr{E}_y$, if and only if $\mathscr{E}_X \overset{t}{\geqq} \mathscr{E}_Y$ and $\mathscr{E}_Y \overset{t}{\geqq} \mathscr{E}_X$.

Based on the definition given above, we will prove interesting properties for $t = 1$ and 2.

**Theorem 1.**

(a) *Let $\mathscr{E}_X$ be any experiment and $\mathscr{E}_N$ the null experiment (i.e., the distribution is independent of $\theta$ a.e. $\mu$), then $\mathscr{E}_X \overset{t}{\geqq} \mathscr{E}_N$.*

(b) *Given the compound experiment $(\mathscr{E}_X, \mathscr{E}_Y)$ where $\mathscr{E}_X$ and $\mathscr{E}_Y$ are the corresponding marginal experiments, then $(\mathscr{E}_X, \mathscr{E}_Y) \overset{t}{\geqq} \mathscr{E}_X$ (or $\mathscr{E}_Y$), with indifference iff $f(y/x, \theta)$ is independent of $\theta$ ($f(x/y, \theta)$ is independent of $\theta$) for almost every $(x, y)$.*

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

(d) *Let $\mathscr{E}_X = \{X, \mathscr{X}, f(x/\theta); \theta \in \Theta\}$ be an experiment and $\{E_i\}_{i \in \boldsymbol{N}}$ a measurable partition of $\mathscr{X}$. Let us consider another experimen $\mathscr{E}_Y = \{Y, \mathscr{Y}, Q_\theta; \theta \in \Theta\}$ with the σ-algebra generated by $\{E_i\}_{i \in \boldsymbol{N}}$ and with $Q_\theta(E_i) = \int_{E_i} f(x/\theta)\, \mathrm{d}\mu(x)$ $\forall i \in \boldsymbol{N}$. Then $\mathscr{E}_X \overset{t}{\geqq} \mathscr{E}_Y$, with indifference iff $f(x/\theta)$ is independent of $\theta$ for almost every $x$.*

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

Proof. First we consider $t = 1$.

(a) Let $\mathscr{E}_X$ be the null experiment, then $f(x/\theta) = f(x)$ and

$$f_k(x) = \int_\Theta f(x)\, \mathrm{d}\xi_k = f(x) \quad k = 1, 2\,.$$

Hence

$${}^1_N R^s_r(\xi_1 \,\|\, \xi_2) = \frac{1}{2(s-1)} \left\{ \left( \int_{\mathscr{X}} f_1(x)^r \left( \frac{f_1(x) + f_2(x)}{2} \right)^{1-r} \mathrm{d}\mu \right)^{(s-1)/(r-1)} + \right.$$
$$\left. + \left( \int_{\mathscr{X}} f_2(x)^r \left( \frac{f_1(x) + f_2(x)}{2} \right)^{1-r} \mathrm{d}\mu \right)^{(s-1)/(r-1)} - 2 \right\} = 0\,.$$

As ${}^{1}_{X}R^{s}_{r}(\xi_1 \| \xi_2) \geqq 0$, we have

$$ {}^{1}_{X}R^{s}_{r}(\xi_1 \| \varepsilon_2) \geqq {}^{1}_{N}R^{s}_{r}(\xi_1 \| \xi_2) \quad \text{for every} \quad \xi_1, \xi_2 \in \Xi \,. $$

(b) On the basis of

$$ f_k(x, y) = \int_\Theta f(x, y/\theta)\, \mathrm{d}\xi_k \quad k = 1, 2 $$

and

$$ f_k(x) = \int_{\mathscr{Y}} (f(x, y)\, \mathrm{d}\nu \quad k = 1, 2 $$

we can write

$$ (13) \qquad f_1^r(x)\left(\frac{f_1(x) + f_2(y)}{2}\right)^{1-r} = $$

$$ = \left[\int_{\mathscr{Y}} f_1(x, y)\, \mathrm{d}\nu\right]^r \left[\int_{\mathscr{Y}} \frac{f_1(x, y) + f_2(x, y)}{2}\, \mathrm{d}\nu\right]^{1-r} . $$

Applying Hölder's inequality on the right hand side of (13) we get

$$ f_1^r(x)\left(\frac{f_1(x) + f_2(y)}{2}\right)^{1-r} \geqq $$

$$ \geqq \int_{\mathscr{Y}} f_1^r(x, y)\left[\frac{f_1(x, y) + f_2(x, y)}{2}\right]^{1-r} \mathrm{d}\nu\,, \quad 0 < r < 1 $$

$$ \leqq \int_{\mathscr{Y}} f_1^r(x, y)\left[\frac{f_1(x, y) + f_2(x, y)}{2}\right]^{1-r} \mathrm{d}\nu\,, \quad r > 1\,. $$

Hence

$$ (14) \qquad \int_{\mathscr{X}} f_1^r(x)\left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu \geqq $$

$$ \geqq \int_X \int_{\mathscr{Y}} f_1^r(x, y)\left(\frac{f_1(x, y) + f_2(x, y)}{2}\right)^{1-r} \mathrm{d}\nu\, \mathrm{d}\mu\,, \quad 0 < r < 1 $$

$$ \leqq \int_{\mathscr{X}} \int_{\mathscr{Y}} f_1^r(x, y)\left(\frac{f_1(x, y) + f_2(x, y)}{2}\right)^{1-r} \mathrm{d}\nu, \mathrm{d}\mu\,, \quad r > 1. $$

As $\operatorname{sign}((s-1)/(r-1)) = \operatorname{sign}(s-1)$ for $r > 1$ and $\operatorname{sign}((s-1)/(r-1)) \neq \neq \operatorname{sign}(s-1)$ for $0 < r < 1$, where $\operatorname{sign}(x) = 1$ if $x > 0$ and $\operatorname{sign}(x) = -1$ if $x < 0$, from (14) we have

$$ (15) \qquad \frac{1}{s-1}\left\{\int_{\mathscr{X}} f_1^r(x)\left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\right\}^{(r-1)/(s-1)} \leqq $$

$$ \leqq \frac{1}{s-1}\left\{\int_{\mathscr{X}} \int_{\mathscr{Y}} f_1^r(x, y)\left(\frac{f_1(x, y) + f_2(x, y)}{2}\right)^{1-r} \mathrm{d}\nu\, \mathrm{d}\mu\right\}^{(r-1)/(s-1)} , $$

$$ r \neq 1, s \neq 1, r > 0\,. $$

Similarly we can obtain

(16) $$\frac{1}{s-1}\left\{\int_{\mathscr{X}} f_2^r(x)\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\right\}^{(r-1)/(s-1)} \leqq$$

$$\leqq \frac{1}{s-1}\left\{\int_{\mathscr{X}}\int_{\mathscr{Y}} f_2^r(x,y)\left(\frac{f_1(x,y)+f_2(x,y)}{2}\right)^{1-r} \mathrm{d}\nu\,\mathrm{d}\mu\right\}^{(r-1)/(s-1)},$$

$$r \neq 1\,, \quad s \neq 1\,, \quad r > 0\,.$$

Adding (15) and (16) subtracting $2(s-1)^{-1}(s \neq 1)$ and then dividing by 2, we get

$$ {}_X^1R_r^s(\xi_1 \parallel \xi_2) \geqq {}_{Y,X}^{\;\;1}R_r^s(\xi_1 \parallel \xi_2)\,, \quad r \neq 1\,, \quad s \neq 1\,, \quad r > 0\,.$$

Since the unified measure ${}_X^1\mathscr{V}_r^s(\xi_1 \parallel \xi_2)$ given in (4) is a continuous extension of ${}_X^1R_r^s(\xi_1 \parallel \xi_2)$ for the real parameters $r$ and $s$, we can immediately conclude that

(17) $${}_X^1\mathscr{V}_r^s(\xi_1 \parallel \xi_2) \geqq {}_{Y,X}^{\;\;1}\mathscr{V}_r^s(\xi_1 \parallel \xi_2) \quad \text{for all} \quad r > 0 \quad \text{and any } s.$$

As (17) holds for every $\xi_1, \xi_2 \in \Xi$, we get $(\mathscr{E}_X, \mathscr{E}_Y) \geqq^1 \mathscr{E}_X$.

Finally, equality holds iff

$$f_i(x,y) = K_i(x)\,\frac{f_1(x,y)+f_2(x,y)}{2} \quad \mu \times \nu \quad \text{a.e.}, \quad i = 1, 2\,,$$

i.e. iff

$$\frac{f_1(x,y)}{f_2(x,y)} = \frac{K_1(x)}{K_2(x)} = K(x) \quad \text{and} \quad K_1(x) + K_2(y) = 2 \quad \mu \times \nu \quad \text{a.e.}$$

So equality holds iff

(18) $$f_i(x,y) = h(x,y)\,K_i(x) = f(y/x)\,f_i(x) \quad \mu \times \nu \quad \text{a.e.}, \quad i = 1, 2\,.$$

Hence $(\mathscr{E}_X, \mathscr{E}_Y) \simeq \mathscr{E}_Y$ iff (18) holds for every $\xi_1, \xi_2 \in \Xi$; i.e. iff

$$f_\theta(x,y) = f(y/x)\,f_\theta(x) \quad \mu \times \nu \quad \text{a.e.}, \quad \forall \theta \in \Theta\,.$$

Dividing by $f_\theta(x)$ we obtain the condition we are looking for, i.e.

$$f(y/x, \theta) = f(y/x) \quad \mu \times \nu \quad \text{a.e.}, \quad \forall \theta \in \Theta\,.$$

(c) An immediate consequence of property (b).

(d) As

$$Q_k(E_i) = \int_\Theta Q_\theta(E_i)\,\mathrm{d}\xi_k = \int_\Theta \int_{E_i} f(x/\theta)\,\mathrm{d}\mu\,\mathrm{d}\xi_k = \int_{E_i} f_k(x)\,\mathrm{d}\mu \quad \text{for} \quad k = 1, 2\,,$$

we can write

$$Q_1(E_i)^r\left[\frac{Q_1(E_i)+Q_2(E_i)}{2}\right]^{1-r} =$$

$$= \left[\int_{E_i} f_1(x)\,\mathrm{d}\mu\right]^r\left[\int_{E_i} \frac{f_1(x)+f_2(x)}{2}\,\mathrm{d}\mu\right]^{1-r}.$$

Applying Hölder's inequality, we get

$$Q_1(E_i)^r\left[\frac{Q_1(E_i)+Q_2(E_i)}{2}\right]^{1-r} \geqq$$

$$\geqq \int_{E_i} f_1(x)^r\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad 0<r<1$$

$$\leqq \int_{E_i} f_1(x)^r\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad r>1\,.$$

Hence

$$\sum_{i\in N} Q_1(E_i)^r\left[\frac{Q_1(E_i)+Q_2(E_i)}{2}\right]^{1-r} \geqq$$

$$\geqq \int_{\mathscr{X}} f_1(x)^r\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad 0<r<1$$

$$\leqq \int_{\mathscr{X}} f_1(x)^r\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad r>1\,;$$

i.e.

$$(19) \qquad \frac{1}{s-1}\left\{\sum_{i\in N} Q_1(E_i)^r\left[\frac{Q_1(E_i)+Q_2(E_i)}{2}\right]^{1-r}\right\} \leqq$$

$$\leqq \frac{1}{s-1}\left\{\int_{\mathscr{X}} f_1(x)^r\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\right\}^{(s-1)/(r-1)}.$$

Similarly we can obtain

$$(20) \qquad \frac{1}{s-1}\left\{\sum_{i\in N} Q_2(E_i)^r\left[\frac{Q_1(E_i)+Q_2(E_i)}{2}\right]^{1-r}\right\}^{(s-1)/(r-1)} \leqq$$

$$\leqq \frac{1}{s-1}\left\{\int_{\mathscr{X}} f_2(x)^r\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\right\}^{(s-1)/(r-1)}.$$

Adding (19) and (20), subtracting 2 and then dividing by 2, we get

$${}^1_X R^s_r(\xi_1 \parallel \xi_2) \geqq {}^1_Y R^s_r(\xi_1 \parallel \xi_2) \quad \text{for every} \quad \xi_1, \xi_2 \in \Xi\,.$$

As the unified measure given in (4) is a continuous extension of ${}^1_X R^s_r(\xi_1 \parallel \xi_2)$ for the real parameters $r$ and $s$, we can conclude that $\mathscr{E}_X \geqq^1 \mathscr{E}_Y$.

Finally, equality holds iff

$$f_1(x) = K_{1,i}\,\frac{f_1(x)+f_2(x)}{2} \quad \forall x\in E_i \quad \text{a.e.} \quad \mu \quad \forall i \in N$$

and

$$f_1(x) = K_{2,i}\,\frac{f_1(x)+f_2(x)}{2} \quad \forall x\in E_i \quad \text{a.e.} \quad \mu \quad \forall i \in N$$

i.e. iff

$$f_1(x) = f_2(x) \quad \forall x \in E_i \quad \text{a.e.} \quad \mu \quad \forall \boldsymbol{i} \in \boldsymbol{N} .$$

So indifference holds iff

$$f_1(x) = f_2(x) \quad \forall x \in \mathscr{X} \quad \text{a.e.} \quad \mu \quad \forall \xi_1, \xi_2 \in \Xi ,$$

i.e.

$$\forall \xi \in \Xi \quad \int_\Theta f(x/\theta)\, \mathrm{d}\xi = f(x) \quad \text{a.e.} \quad \mu ,$$

i.e.

$$f(x/\theta) = f(x) \quad \text{a.e.} \quad \mu \quad \forall \theta \in \Theta .$$

(e) Let $\xi_1$ and $\xi_2$ be two arbitrary prior distributions and let us consider experiments

$$\mathscr{E}_X^{(n)} = \{X^n, \beta_{X^n}, f(x_1, x_2, \ldots, x_n); \theta \in \Theta\}$$

and

$$\mathscr{E}_T = \{\mathscr{T} \subseteq \boldsymbol{R}, \beta_{\mathscr{T}}, g(t_1, t_2, \ldots, t_n/\theta); \theta \in \Theta\}$$

with

(21) $$g(t/\theta) = g(t_1, \ldots, t_m/\theta) = \\ = \int_{\mathscr{X}^{n-m}} f(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n/\theta)\, |J|\, \mathrm{d}x_{m+1}, \ldots, \mathrm{d}x_n$$

where

$$x_1 = h_1(t_1, \ldots, t_2)\,, \quad x_2 = h_2(t_1, \ldots, t_2), \ldots, \quad x_m = h_m(t_1, \ldots, t_2)\,,$$

$$x_{m+1} = x_{m+1}, \ldots, x_n = x_n$$

and

$$J = \frac{\partial(x_1, \ldots, x_m)}{\partial(t_1, \ldots, t_n)} .$$

Since

$$g_k(t) = \int_\Theta g(t/\theta)\, \mathrm{d}\xi_k = \\ = \int_{\mathscr{X}^{n-m}} f_k(h_1(t), \ldots, h_m(t), x_{m+1} \ldots, x_n)\, |J|\, \mathrm{d}x_{m+1}, \ldots, \mathrm{d}x_n ,$$

we have

$$g_1^r(t) \left(\frac{g_1(t) + g_2(t)}{2}\right)^{1-r} = \\ = \left[\int_{\mathscr{X}^{n-m}} f_k(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n)\, |J|\, \mathrm{d}x_{m+1}, \ldots, \mathrm{d}x_n\right]^r . \\ . \left[\int_{\mathscr{X}^{n-m}} \frac{f_1(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n) + f_2(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n)}{2} . \right. \\ \left. . |J|\, \mathrm{d}x_{m+1}, \ldots, \mathrm{d}x_n\right]^{1-r} .$$

Applying Hölder inequality, we have

$$g_1^r(t)\left(\frac{g_1(t)+g_2(t)}{2}\right)^{1-r} \geqq$$

$$\geqq \int_{\mathscr{X}^{n-m}} \left(\frac{f_1(h_1(t),\ldots,h_m(t),x_{m+1},\ldots,x_n)+f_2(h_1(t),\ldots,h_m(t),x_{m+1},\ldots,x_n)}{2}\right)^{1-r} .$$

$$.\,|J|\,(f_1\,(h_1(t),\ldots,h_m(t),x_{m+1},\ldots,x_n))\,\mathrm{d}x_{m+1},\ldots,\mathrm{d}x_n \quad \text{if} \quad 0<r<1\,,$$

$$\leqq \int_{\mathscr{X}^{n-m}} \left(\frac{f_1(h_1(t),\ldots,h_m(t),x_{m+1},\ldots,x_n)+f_2(h_1(t),\ldots,h_m(t),x_{m+1},\ldots,x_n)}{2}\right)^{1-r} .$$

$$.\,|J|\,(f_1(h_1(t),\ldots,h_m(t),x_{m+1},\ldots,x_n))\,\mathrm{d}x_{m+1},\ldots,\mathrm{d}x_n \quad \text{if} \quad r>1\,.$$

Integrating over $\mathscr{T}^m$ and applying the transformation (21), we have

$$\int_{\mathscr{T}^m} g_1^r(t)\left(\frac{g_1(t)+g_2(t)}{2}\right)^{1-r} \mathrm{d}t$$

$$\geqq \int_{\mathscr{X}^n} f_1^r\,(x_1,\ldots,x_n)\left(\frac{f_1(x_1,\ldots,x_n)+f_2(x_1,\ldots,x_n}{2}\right)^{1-r} \mathrm{d}x_1,\ldots,\mathrm{d}x_n$$

if $0<r<1$

$$\leqq \int_{\mathscr{X}^n} f_1^r(x_1,\ldots,x_n)\left(\frac{f_1(x_1,\ldots,x_n)+f_2(x_1,\ldots,x_n)}{2}\right)^{1-r} \mathrm{d}x_1,\ldots,\mathrm{d}x_n$$

if $r>1$.

Hence

$$\text{(22)} \qquad \frac{1}{s-1}\left[\int_{\mathscr{T}^m} g_1^r(t)\left(\frac{g_1(t)+g_2(t)}{2}\right)^{1-r} \mathrm{d}t\right]^{(s-1)/(r-1)} \leqq$$

$$\leqq \frac{1}{s-1}\left[\int_{\mathscr{X}^n} f_1^r(x_1,\ldots,x_n)\,\times\right.$$

$$\left.\times\left(\frac{f_1(x_1,\ldots,x_n)+f_2(x_1,\ldots,x_n)}{2}\right)^{1-r} \mathrm{d}x_1,\ldots,\mathrm{d}x_n\right]^{(s-1)/(r-1)}.$$

Similarly we can obtain

$$\text{(23)} \qquad \frac{1}{s-1}\left[\int_{\mathscr{T}^m} g_2^r(t)\left(\frac{g_1(t)+g_2(t)}{2}\right)^{1-r} \mathrm{d}t\right]^{(s-1)/(r-1)} \leqq$$

$$\leqq \frac{1}{s-1}\left[\int_{\mathscr{X}^n} f_2^r(x_1,\ldots,x_n)\,\times\right.$$

$$\left.\times\left(\frac{f_1(x_1,\ldots,x_n)+f_2(x_1,\ldots,x_n)}{2}\right)^{1-r} \mathrm{d}x_1,\ldots,\mathrm{d}x_n\right]^{(s-1)/(r-1)}.$$

Adding (22) and (23), subtracting $2(s-1)^{-1}(s+1)$ and then dividing by 2, we get

$$_{X^{(n)}}^{\ 1}R_r^s(\xi_1 \,\|\, \xi_2) \geqq {}_T^1R_r^s(\xi_1 \,\|\, \xi_2) \quad \text{for every} \quad \xi_1, \xi_2 \in \Xi \,.$$

Finally, equality holds iff

$$f_i(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n) = G_i(t) \,.$$
$$\cdot \left( \frac{f_1(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n) + f_2(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n)}{2} \right)$$
$$\text{a.e.} \quad \mu \quad \text{for} \quad i = 1, 2 \,,$$

i.e. iff

$$G_1(t) + G_2(t) = 2 \quad \text{and}$$
$$\frac{f_1(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n)}{f_2(h_1(t), \ldots, h_m(t), x_{m+1}, \ldots, x_n)} = \frac{G_1(t)}{G_2(t)} = G(t) \quad \text{a.e.} \quad \mu \,;$$

i.e. iff

(24) $$f_i(x_1, \ldots, x_n) = G_i(t)\, K(x_1, \ldots, x_n) \quad \text{a.e.} \quad \mu \quad \text{for} \quad i = 1, 2 \,.$$

So indifference holds iff

$$f_\xi(x_1, \ldots, x_n) = \int_\Theta f(x_1, \ldots, x_n/\theta)\, \mathrm{d}\xi = G_\xi(t)\, K(x_1, \ldots, x_n) \quad \text{a.e.} \quad \mu$$
$$\forall \xi \in \Xi \,.$$

This last condition is equivalent to

$$f_\theta(x_1, \ldots, x_n/\theta) = G_\theta(t)\, K(x_1, \ldots, x_n) \quad \text{a.e.} \quad \mu \quad \forall \theta \in \Theta \,,$$

which is the factorization formula given by Halmos-Savage's theorem to establish the sufficiency of a statistic.

Proof for $t = 2$ is similar.

## 4. DIVERGENCE MEASURES AND SUFFICIENCY OF EXPERIMENTS

Blackwell's [1] definition of comparing two experiments states that experiment $\mathscr{E}_X$ is sufficient for experiment $\mathscr{E}_Y$. denoted $\mathscr{E}_X \geqq \mathscr{E}_Y$, if there exists a stochastic transformation of $X$ to a random variable $Z(X)$ such that for each $\theta \in \Theta$ the random variables $Z(X)$ and $Y$ have identical distributions. By $\mathscr{E}_Y = \{Y, \mathscr{Y}, Q_\theta; \theta \in \Theta\}$ we denote a second statistical experiment for which there exists a gpdf $g(y/\theta)$ for the distribution $Q$ with respect to a $\sigma$-finite measure $\nu$. According to this definition, if $\mathscr{E}_X \geqq \mathscr{E}_Y$, then there exists a nonnegative function $h$ satisfying (cf. DeGroot [8], p. 434).

(25) $$g(y/\theta) = \int_{\mathscr{X}} h(y/x)\, f(x/\theta)\, \mathrm{d}\mu$$

and

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

If we have two prior distributions $\xi_1, \xi_2 \in \Xi$, after integrating over $\Xi$ and changing the order of integration in (25) we get

(26) $$g_i(y/\theta) = \int_{\mathscr{X}} h(y/x) f_i(x)\, \mathrm{d}\mu\,, \quad 1 = 1, 2\,.$$

Let $I$ be any measure of information contained in an experiment. If $\mathscr{E}_X \geqq \mathscr{E}_Y$ implies $I_X \geqq I_Y$, then we say that $\mathscr{E}_X$ is at least as informative as $\mathscr{E}_Y$ in terms of measure I. This approach was successfully carried out by Lindley [13] and Sakaguchi [18] for Shannon's entropy. Goel and DeGroot [10] applied it to Kullback-Leibler's [11] discrimination function. Ferentinos and Papaioannou [9] aplied it to the $\alpha$-order generalization of Kullback and Leibler's discrimination function and some generalizations of Fisher's measure of information. Taneja [25] extended it to different generalizations of $J$-divergence measure having two scalar parameters. Recently, the authors [14, 15, 28] extended it to $\lambda$-measures of hypoentropy and $\phi$-measures of Jensen difference, where Bayesian and Lehmann's approaches are also adopted. Now, we will compare the experiments for the generalized divergence measures given in (4) and (9).

**Theorem 2.** If $\mathscr{E}_X \geqq \mathscr{E}_Y$, then ${}^1_X\mathscr{V}^s_r(\xi_1 \parallel \xi_2) \geqq {}^1_Y\mathscr{V}^s_r(\xi_1 \parallel \xi_2)$ for every $\xi, \xi_2 \in \Xi$, for all $r > 0$ and any $s$.

Proof. Since $\mathscr{E}_X \geqq \mathscr{E}_Y$, there exists a function $h$ satisfying (25) and (26), and we can write

(27) $$g_1^r(y)\left(\frac{g_1(y) + g_2(y)}{2}\right)^{1-r} = \\ = \left[\int_{\mathscr{X}} h(y/x) f_1(x)\, \mathrm{d}\mu\right]^r \left[\int_{\mathscr{X}} h(y/x)\, \frac{f_1(x) + f_2(x)}{2}\, \mathrm{d}\mu\right]^{1-r}.$$

Applying Hölder's inequality on the right hand side of (27) we get

$$g_1^r(y)\left(\frac{g_1(y) + g_2(y)}{2}\right)^{1-r} \geqq \\ \geqq \int_{\mathscr{X}} [h(y/x) f_1(x)]^r \left[h(y/x)\, \frac{f_1(x) + f_2(x)}{2}\right]^{1-r} \mathrm{d}\mu\,, \quad 0 < r < 1 \\ \leqq \int_{\mathscr{X}} [h(y/x) f_1(x)]^r \left[h(y/x)\, \frac{f_1(x) + f_2(x)}{2}\right]^{1-r} \mathrm{d}\mu\,, \quad r > 1\,.$$

Hence

(28) $$\int_{\mathscr{Y}} g_1^r(y)\left(\frac{g_1(y) + g_2(y)}{2}\right)^{1-r} \mathrm{d}\nu \geqq \\ \geqq \int_{\mathscr{Y}} f_1^r(x)\left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad 0 < r < 1 \\ \leqq \int_{\mathscr{X}} f_1^r(x)\left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad r > 1\,.$$

As $\operatorname{sign}((s-1)/(r-1)) = \operatorname{sign}(s-1)$ for $r > 1$ and $\operatorname{sign}((s-1)/(r-1)) \neq \operatorname{sign}(s-1)$ for $0 < r < 1$, where $\operatorname{sign}(x) = 1$ if $x > 0$ and $\operatorname{sign}(x) = -1$, from (28) we have

$$(29)\qquad \frac{1}{s-1}\left[\int_{\mathscr{Y}} g_1^r(y)\left(\frac{g_1(y)+g_2(y)}{2}\right)^{1-r} \mathrm{d}\nu\right]^{(s-1)/(r-1)} \leqq$$
$$\leqq \frac{1}{s-1}\left[\int_{\mathscr{X}} f_1^r(x)\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\right]^{(s-1)/(r-1)},$$
$$r \neq 1\,, \quad s \neq 1\,, \quad r > 0\,.$$

Similarly we can obtain

$$(30)\qquad \frac{1}{s-1}\left[\int_{\mathscr{Y}} g_2^r(y)\left(\frac{g_1(y)+g_2(y)}{2}\right)^{1-r} \mathrm{d}\nu\right]^{(s-1)/(r-1)} \leqq$$
$$\leqq \frac{1}{s-1}\left[\int_{\mathscr{X}} f_2^r(x)\left(\frac{f_1(x)+f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\right]^{(s-1)/(r-1)},$$
$$r \neq 1\,, \quad s \neq 1\,, \quad r > 0\,.$$

Adding (29) and (30), subtracting $2(s-1)^{-1}$ $(s \neq 1)$ and then dividing by 2, we get

$$ {}_X^1R_r^s(\xi_1 \,\|\, \xi_2) \geqq {}_Y^1R_r^s(\xi_1 \,\|\, \xi_2)\,, \quad r \neq 1\,, \quad s \neq 1\,, \quad r > 0\,.$$

Since the unified measure ${}_X^1\mathscr{V}_r^s(\xi_1 \,\|\, \xi_2)$ given in (4) is a continuous extension of ${}_X^1R_r^s(\xi_1 \,\|\, \xi_2)$ for the real parameters $r$ and $s$, we can immediately conclude that

$$ {}_X^1\mathscr{V}_r^s(\xi_1 \,\|\, \xi_2) \geqq {}_Y^1\mathscr{V}_r^s(\xi_1 \,\|\, \varepsilon_2) \quad \text{for all} \quad r > 0$$

and any $s$ whenever $\mathscr{E}_X \geqq \mathscr{E}_Y$. ■

**Theorem 3.** *If $\mathscr{E}_X \geqq \mathscr{E}_Y$, then ${}_X^2\mathscr{V}_r^s(\xi_1 \,\|\, \xi_2) \geqq {}_Y^2\mathscr{V}_r^s(\xi_1 \,\|\, \xi_2)$ for every $\xi_1, \xi_2 \in \Xi$, for all $r > 0$ and any $s$.*

Proof. Since $\mathscr{E}_X \geqq \mathscr{E}_Y$, there exists a function $h$ satisfying (25) and (26), and we can write

$$(31)\qquad \frac{g_1^r(y)+g_2^r(y)}{2}\left(\frac{g_1(y)+g_2(y)}{2}\right)^{1-r} =$$
$$= \frac{1}{2}\left[\int_{\mathscr{X}} h(y/x) f_1(x)\,\mathrm{d}\mu\right]^r \left[\int_{\mathscr{X}} h(y/x)\frac{f_1(x)+f_2(x)}{2}\,\mathrm{d}\mu\right]^{1-r} +$$
$$+ \frac{1}{2}\left[\int_{\mathscr{X}} h(y/x) f_2(x)\,\mathrm{d}\mu\right]^r \left[\int_{\mathscr{X}} h(y/x)\frac{f_1(x)+f_2(x)}{2}\,\mathrm{d}\mu\right]^{1-r}.$$

Applying Hölder's inequality on the right hand side of (31), integrating over $\mathscr{Y}$, and using the fact that $\int_{\mathscr{Y}} h(y/x)\,\mathrm{d}v = 1$, we get

$$\int_{\mathscr{Y}} \frac{g_1^r(y) + g_2^r(y)}{2} \left(\frac{g_1(y) + g_2(y)}{2}\right)^{1-r} \mathrm{d}v \geqq$$

$$\geqq \int_{\mathscr{X}} \frac{f_1^r(x) + f_2^r(x)}{2} \left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad 0 < r < 1$$

$$\leqq \int_{\mathscr{X}} \frac{f_1^r(x) + f_2^r(x)}{2} \left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\,, \quad r > 1\,.$$

As $\operatorname{sign}((s-1)/(r-1)) = \operatorname{sign}(s-1)$ for $r > 1$ and $\operatorname{sign}((s-1)/(r-1)) \neq \operatorname{sign}(s-1)$ for $0 < r < 1$, we have

$$\frac{1}{s-1}\left[\int_{\mathscr{Y}} \frac{g_1^r(y) + g_2^r(y)}{2} \left(\frac{g_1(y) + g_2(y)}{2}\right)^{1-r} \mathrm{d}v\right]^{(s-1)/(r-1)} \leqq$$

$$\leqq \frac{1}{s-1}\left[\int_{\mathscr{X}} \frac{f_1^r(x) + f_2^r(x)}{2} \left(\frac{f_1(x) + f_2(x)}{2}\right)^{1-r} \mathrm{d}\mu\right]^{(s-1)/(r-1)},$$

$$r \neq 1\,, \quad s \neq 1\,, \quad r > 0\,.$$

Thus, ${}^2_X\mathscr{V}^s_r(\xi_1 \| \xi_2) \geqq {}^2_Y\mathscr{V}^s_r(\xi_1 \| \xi_2)$, $r \neq 1$, $s \neq 1$, $r > 0$. Again using the fact that the unified measure ${}^2_X\mathscr{V}^s_r(\xi_1 \| \xi_2)$ given in (9) is a continuous extension of ${}^2_XR^s_r(\xi_1 \| \xi_2)$ for the real parameters $r$ and $s$, we can immediately conclude that

$${}^2_X\mathscr{V}^s_r(\xi_1 \| \xi_2) \geqq {}^2_Y\mathscr{V}^s_r(\xi_1 \| \xi_2) \quad \text{for all} \quad r > 0$$

and any $s$ whenever $\mathscr{E}_X \geqq \mathscr{E}_Y$.

**Referee's remarks**

There are two composite forms to write the measures ${}^1_X\mathscr{V}^s_r(\xi_1 \| \xi_2)$ and ${}^2_X\mathscr{V}^s_r(\xi \| \xi_2)$.

*First form.* We can write

$${}^1_XR^s_r(\xi_1 \| \xi_2) = \tfrac{1}{2}[\Phi_s(f(\xi_1 \| \xi_2)^{1/(r-1)}) + \Phi_s(f^*(\xi_1 \| \xi_2)^{1/(r-1)})]$$

and

$${}^2_XR^s_r(\xi_1 \| \xi_2) = \Phi_s(\bar{f}(\xi_1 \| \xi_2)^{1/(r-1)})\,,$$

where

$$\Phi_s(y) = (s-1)^{-1}\,[y_s^{-1} - 1]\,, \quad s \neq 1\,,$$

and $f(\xi_1 \| \xi_2)$, $f^*(\xi_1 \| \xi_2)$ and $\bar{f}(\xi_1 \| \xi_2)$ are the $f$-divergences of $\xi_1$, $\xi_2$ in the notation of Vajda [30] with

$$f(x) = \frac{1+x}{2} f_r\left(\frac{2x}{1+x}\right), \quad f_r(x) = x^r\,, \quad x > 0\,,$$

$$f^*(x) = x f\left(\frac{1}{x}\right) \quad \text{and} \quad \bar{f}(x) = \tfrac{1}{2}[f(x) + f^*(x)]\,.$$

The function $f_r(x) = x^r$ is convex in $x$ for $r > 1$ and is concave in $x$ for $0 < r < 1$.

*Second form.* We can write

$$ {}^1_X R^s_r(\xi_1 \parallel \xi_2) = \tfrac{1}{2}[\Psi_s(R_r(\xi_1 \parallel \bar{\xi})) + \Phi_s(R_r(\xi_2 \parallel \bar{\xi}))] $$

and

$$ {}^2_X R^s_r(\xi_1 \parallel \xi_2) = \Psi_s(\tfrac{1}{2}[R_r(\xi_1 \parallel \bar{\xi}) + R_r(\xi_2 \parallel \bar{\xi}))],), $$

where

$$ \Psi_s(y) = (s-1)^{-1}\left[e^{(s-1)y} - 1\right], \quad s \neq 1, $$

$$ \bar{\xi} = \frac{\xi_1 + \xi_2}{2}, $$

and

$$ R_r(\xi \parallel \eta) = (r-1)^{-1} \log_e \left[\int_{\mathbf{R}} p(x)^r q(x)^{1-r} \, d\mu\right], \quad r \neq 1, \quad r > 0 $$

with

$$ f_r(\xi \parallel \eta) = \exp\{R_r(\xi \parallel \eta)\}, \quad r \neq 1, \quad r > 0. $$

Some of the results given in theorems 1, 2 and 3 can be simplified by using the approach of $f$-divergences given in Vajda [30], chapter 9.

**Author's remarks**

The approach suggested by the referee is economical but works with the convexity of $f$-divergence measures. With the approach presented the proofs are direct. One of the drawbacks in using the composition relation is in proving the convexity of the measures ${}^1_X\mathscr{V}^s_r$ and ${}^2_X\mathscr{V}^s_r$, because by this approach, we get the convexity of these measures in the pair $(\xi_1, \xi_2)$ for either $0 < r \leqq 2 \leqq s$ (first form) of $0 < r \leqq 1 \leqq s$ (second form), while by the direct approach we arrive at $0 < r \leqq s$ (ref. Taneja [26]). This last condition is much better than those obtained by composition.

**Acknowledgements.** One of the author (I. J. TANEJA) is thankful to the "Universidad Complutense de Madrid, Departamento de Estadística e I.O." for providing facilities and financial support. This work was partially supported by the Dirección General de Investigación Cientifica y Técnica (DGCYT) under the contract P589—0019.

*References*

[1] *Blackwell, D.* (1951): Comparison of experiments. Proc. 2nd Berkeley Symp. Berkeley: University of California Press, 93—102.

[2] *Burbea J.* (1984): The Bose-Einstein Entropy of degree $\alpha$ and its Jensen Difference. Utilitas Math. 25, 225—240.

[3] *Burbea, J.* and *Rao, C. R.* (1982): Entropy Differential Metric, Distance and Divergence Measures in Probability Spaces: A Unified Approach. J. Multi. Analy. 12, 575—596.

[4] *Burbea, J.* and *Rao, C. R.* (1982): On the Convexity of some Divergence Measures based on Entropy Functions. IEEE Trans. on Inform. Theory IT-28, 489—495.

[5] *Capocelli, R. M.* and *Taneja, I. J.* (1984): Generalized Divergence Measures and Error Bounds. Proc. IEEE Internat. Conf. on Systems, man and Cybernetics, Oct. 9—12, Halifax, Canada, pp. 43—47.

[6] *Campbell, L. L.* (1986): An extended Čencov characterization of the Information Metric. Proc. Ann. Math. Soc., 98, 135—141.

[7] *Čencov, N. N.* (1982): Statistical Decisions Rules and Optimal Inference. Trans. of Math. Monographs, 53, Am. Math. Soc., Providence, R. I..
[8] *De Groot, M. H,* (1970): Optimal Statistical Decisions. McGraw-Hill. New York.
[9] *Ferentinos, K.* and *Papaioannou, T.* (1982): Information in experiments and sufficiency. J. Statist. Plann. Inference 6, 309—317.
[10] *Goel, P. K.* and *De Groot* (1979): Comparison of experiments and information measures. Ann. Statist. 7, 1066—1077.
[11] *Kullback, S.* and *Leibler, A.* (1951): On information and sufficiency. Ann. Math Stat. 27, 986—1005.
[12] *Lindley, D. V.* (1956): On a measure of information provided by an experiment. Ann. Math. Statis. 27, 986—1005.
[13] *Marshall, A. W.* and *Olkin, I.* (1979): Inequalities: Theory of Majorization and its Applications. Academic Press. New York.
[14] *Morales, D., Taneja, I. J.* and *Pardo, L.*: Comparison of Experiments based on $\phi$-Measures of Jensen Difference. Communicated.
[15] *Pardo, L., Morales, D.* and *Taneja, I. J.*: $\lambda$-measures of hypoentropy and comparison of experiments: Bayesian approach. To appear in Statistica.
[16] *Rao, C. R.* (1982): Diversity and Dissimilarity Coefficients: A Unified Approach. J. Theoret. Pop. Biology, 21, 24—43.
[17] *Rao, C. R.* and *Nayak, T. K.* (1985): Cross Entropy, Dissimilarity Measures and characterization of Quadratic Entropy. IEEE Trans. on Inform. Theory, IT-31(5), 589—593.
[18] *Sakaguchi, M.* (1964): Information Theory and Decision Making. Unpublished Lecture Notes, Statist. Dept., George Washington Univ., Washington DC.
[19] *Sant'anna, A. P.* and *Taneja, I. J.*: Trigonometric Entropies, Jensen Difference Divergence Measures and Error Bounds. Infoımation Sciences 25, 145—156.
[20] *Shannon, C. E.* (1948): A Mathematical Theory of Communications. Bell. Syst. Tech. J. 27, 379—423.
[21] *Sibson, R.* (1969): Information Radius. Z. Wahrs. und verw. Geb. 14, 149—160.
[22] *Taneja, I. J.*: 1(983): On characterization of $J$-divergence and its generalizations. J. Combin. Inform. System Sci. 8, 206—212.
[23] *Taneja, I. J.* (1986): $\lambda$-measures of hypoentropy and their applications. Statistica, anno XLVI, n. 4, 465—478.
[24] *Taneja, I. J.* (1986): Unified Measure of Information applied to Markov Chains and Sufficiency, J. Comb. Inform. & Syst. Sci., 11, 99—109.
[25] *Taneja, I. J.* (1987): Statistical aspects of Divergence Measures. J. Statist. Plann. & Inferen., 16, 137—145.
[26] *Taneja, I. J.* (1989): On Generalized Information Measures and their Applications. Adv. Elect. Phys. 76, 327—413. Academic Press.
[27] *Taneja, I. J.* (1990): Bounds on the Probability of Error in Terms of Generalized Information Radius. Information Sciences. 46.
[28] *Taneja, I. J., Morales, D.* and *Pardo, L.* (1991): $\lambda$-measures of hypoentropy and comparison of experiments: Blackwell and Lehemann approach. Kybernetika, 27, 413—420.
[29] *Vajda, I.* (1968): Bounds on the Minimal Error Probability and checking a finite or countable number of Hypothesis. Inform. Trans. Problems 4, 9—17.
[30] *Vajda, I.* (1989): Theory of Statistical Inference and Information. Kluwer Academic Publishers, Dordrecht/Boston/London/.

*Authors' address:* Prof. *I. J. Taneja*, Prof. *L. Pardo*, Prof. *D. Morales*, Departamento de Estadística e I. O. Facultad de Matemáticas Universidad Complutense de Madrid, 28040 Madrid, Spain.