APPROXIMATING CSISZÁR f-DIVERGENCE VIA AN OSTROWSKI TYPE IDENTITY FOR $n-$TIME DIFFERENTIABLE FUNCTIONS

N.S. BARNETT, P. CERONE, S. S. DRAGOMIR, AND A. SOFO

Abstract. Using an identity of Ostrowski type for $n-$differentiable functions, we point out an approximation of Csiszár $f-$divergence.

1. Introduction

One of the important issues in many applications of Probability Theory is finding an appropriate measure of *distance* (or *difference* or *discrimination*) between two probability distributions. A number of divergence measures for this purpose have been proposed and extensively studied by Jeffreys [25], Kullback and Leibler [34], Rényi [45], Havrda and Charvat [23], Kapur [28], Sharma and Mittal [47], Burbea and Rao [5], Rao [44], Lin [37], Csiszár [11], Ali and Silvey [1], Vajda [55], Shioya and Da-te [48] and others (see for example [28] and the references therein).

These measures have been applied in a variety of fields such as: anthropology [44], genetics [40], finance, economics, and political science [46], [50], [51], biology [42], the analysis of contingency tables [22], approximation of probability distributions [10], [29], signal processing [26], [27] and pattern recognition [3], [9].

Assume that a set χ and the $\sigma-$finite measure μ are given. Consider the set of all probability densities on μ to be $\Omega := \left\{p|p : \chi \to \mathbb{R},\, p(x) \geq 0,\ \int_\chi p(x)\, d\mu(x) = 1\right\}$. The Kullback-Leibler divergence [34] is well known among the information divergences. It is defined as:

$$D_{KL}(p,q) := \int_\chi p(x) \log\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \Omega, \tag{1.1}$$

where log is to base 2.

In Information Theory and Statistics, various divergences are applied in addition to the Kullback-Leibler divergence. These are the: *variation distance* D_v, *Hellinger distance* D_H [24], χ^2-*divergence* D_{χ^2}, $\alpha-$*divergence* D_α, *Bhattacharyya distance* D_B [4] , *Harmonic distance* D_{Ha}, *Jeffreys distance* D_J [25], *triangular discrimination* D_Δ [52], etc... They are defined as follows:

$$D_v(p,q) := \int_\chi |p(x) - q(x)|\, d\mu(x), \quad p,q \in \Omega; \tag{1.2}$$

$$D_H(p,q) := \int_\chi \left[\sqrt{p(x)} - \sqrt{q(x)}\right]^2 d\mu(x), \quad p,q \in \Omega; \tag{1.3}$$

Date: September 28, 2000.
1991 *Mathematics Subject Classification.* Primary 94A99; Secondary 26D15.
Key words and phrases. Csiszár $f-$divergence, Analytic Inequalities.

$$D_{\chi^2}(p,q) := \int_\chi p(x)\left[\left(\frac{q(x)}{p(x)}\right)^2 - 1\right] d\mu(x), \quad p,q \in \Omega; \tag{1.4}$$

$$D_\alpha(p,q) := \frac{4}{1-\alpha^2}\left[1 - \int_\chi [p(x)]^{\frac{1-\alpha}{2}} [q(x)]^{\frac{1+\alpha}{2}} d\mu(x)\right], \quad p,q \in \Omega; \tag{1.5}$$

$$D_B(p,q) := \int_\chi \sqrt{p(x)\,q(x)}\,d\mu(x), \quad p,q \in \Omega; \tag{1.6}$$

$$D_{Ha}(p,q) := \int_\chi \frac{2p(x)\,q(x)}{p(x)+q(x)} d\mu(x), \quad p,q \in \Omega; \tag{1.7}$$

$$D_J(p,q) := \int_\chi [p(x) - q(x)] \ln\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \Omega; \tag{1.8}$$

$$D_\Delta(p,q) := \int_\chi \frac{[p(x)-q(x)]^2}{p(x)+q(x)} d\mu(x), \quad p,q \in \Omega. \tag{1.9}$$

For other divergence measures, see the paper [28] by Kapur or the book on line [49] by Taneja. For a comprehensive collection of preprints available on line, see the RGMIA web site `http://rgmia.vu.edu.au/papersinfth.html`

2. Representation of Csiszár $f-$Divergence

In [8], the authors have pointed out the following integral identity generalising the trapezoid rule.

Lemma 1. *Let $f : [a,b] \to \mathbb{R}$ be a mapping such that $f^{(n-1)}$ is absolutely continuous on $[a,b]$. Then for all $x \in [a,b]$, we have the identity:*

$$\begin{aligned} \int_a^b f(t)\,dt &= \sum_{k=0}^{n-1}\left[\frac{(b-x)^{k+1} + (-1)^k (x-a)^{k+1}}{(k+1)!}\right] f^{(k)}(x) \\ &\quad + (-1)^n \int_a^b K_n(x,t)\, f^{(n)}(t)\,dt \end{aligned} \tag{2.1}$$

where the kernel $K_n : [a,b]^2 \to \mathbb{R}$ is given by

$$K_n(x,t) := \begin{cases} \frac{(t-a)^n}{n!}, & a \le t \le x \le b \\ \\ \frac{(t-b)^n}{n!}, & a \le x < t \le b \end{cases} \tag{2.2}$$

and n is a natural number, $n \ge 1$.

In what follows, we need the identity (2.2) in the following equivalent form:

$$\begin{aligned} f(z) &= \frac{1}{b-a}\int_a^b f(t)\,dt \\ &\quad - \frac{1}{b-a}\sum_{k=1}^{n-1} \frac{1}{(k+1)!}\left[(b-z)^{k+1} + (-1)^k (z-a)^{k+1}\right] f^{(k)}(z) \\ &\quad + \frac{(-1)^{n+1}}{(b-a)\,n!}\left[\int_a^z (t-a)^n f^{(n)}(t)\,dt + \int_z^b (t-b)^n f^{(n)}(t)\,dt\right] \end{aligned} \tag{2.3}$$

for all $z \in [a,b]$.

Note that for $n=1$, the sum $\sum_{k=1}^{n-1}$ is empty and we obtain the known identity (see for example [6])

$$f(z) = \frac{1}{b-a}\int_a^b f(t)\,dt + \frac{1}{b-a}\left[\int_a^z (t-a) f^{(1)}(t)\,dt + \int_z^b (t-b) f^{(1)}(t)\,dt\right], \ x \in [a,b]. \tag{2.4}$$

In what follows, we assume that the probability distributions $p, q \in \Omega$ satisfy the standing condition:

$$0 \le r \le \frac{p(x)}{q(x)} \le R < \infty, \quad \text{for a.e. } x \in \Gamma. \tag{2.5}$$

Obviously $r \le 1 \le R$.

The following representation of Csiszár $f-$divergence holds.

Theorem 1. *Let $f : [r,R] \to \mathbb{R}$, where r, R are as above and $f^{(n-1)}$ is absolutely continuous on $[r,R]$. Then for all $p, q \in \Omega$ satisfying (2.5), we have the representation:*

$$\begin{aligned} & I_f(p,q) \\ = {} & \frac{1}{R-r}\int_r^R f(t)\,dt - \frac{1}{R-r}\sum_{k=1}^{n-1}\frac{1}{(k+1)!} I_{(R-\cdot)^{k+1} f^{(k)}(\cdot)}(p,q) \\ & + \frac{1}{R-r}\sum_{k=1}^{n-1}\frac{(-1)^{k+1}}{(k+1)!} I_{(\cdot-r)^{k+1} f^{(k)}(\cdot)}(p,q) \\ & + \frac{(-1)^{n+1}}{(R-r)\,n!}\int_\Gamma q(x)\left(\int_r^{\frac{p(x)}{q(x)}} (t-r)^n f^{(n)}(t)\,dt\right) d\mu(x) \\ & + \frac{(-1)^{n+1}}{(R-r)\,n!}\int_\Gamma q(x)\left(\int_{\frac{p(x)}{q(x)}}^R (t-R)^n f^{(n)}(t)\,dt\right) d\mu(x). \end{aligned} \tag{2.6}$$

Proof. Using (2.3) for $z = \frac{p(x)}{q(x)}$, $x \in \Gamma$ and $a = r$, $b = R$, we may write

$$\begin{aligned} & f\left(\frac{p(x)}{q(x)}\right) \\ = {} & \frac{1}{R-r}\int_r^R f(t)\,dt - \frac{1}{R-r}\sum_{k=1}^{n-1}\frac{1}{(k+1)!}\left[\left(R - \frac{p(x)}{q(x)}\right)^{k+1} + (-1)^k\left(\frac{p(x)}{q(x)} - r\right)^{k+1}\right] f^{(k)}\left(\frac{p(x)}{q(x)}\right) \\ & + \frac{(-1)^{n+1}}{(R-r)\,n!}\left[\int_r^{\frac{p(x)}{q(x)}} (t-r)^n f^{(n)}(t)\,dt + \int_{\frac{p(x)}{q(x)}}^R (t-R)^n f^{(n)}(t)\,dt\right] \end{aligned} \tag{2.7}$$

for all $x \in \Gamma$.

If we multiply (2.7) by $q(x) \ge 0$, integrate on Γ and take into account that $\int_\Gamma q(x)\,d\mu(x) = 1$, we get the desired identity (2.6). ∎

Remark 1. *If $n=1$, then we have the representation:*

$$\begin{aligned} I_f(p,q) &= \frac{1}{R-r}\int_r^R f(t)\,dt \\ &+\frac{1}{R-r}\int_\Gamma q(x)\left(\int_r^{\frac{p(x)}{q(x)}}(t-r)\,f'(t)\,dt\right)d\mu(x) \\ &+\frac{1}{R-r}\int_\Gamma q(x)\left(\int_{\frac{p(x)}{q(x)}}^R (t-R)\,f'(t)\,dt\right)d\mu(x) \end{aligned} \tag{2.8}$$

if $n=2$, then we have the representation:

$$\begin{aligned} I_f(p,q) &= \frac{1}{R-r}\int_r^R f(t)\,dt+\frac{1}{2(R-r)}I_{(R-\cdot)^2 f^{(1)}(\cdot)}(p,q) \\ &+\frac{1}{R-r}I_{(\cdot-r)^2 f'(\cdot)}(p,q) \\ &-\frac{1}{2(R-r)}\int_\Gamma q(x)\left(\int_r^{\frac{p(x)}{q(x)}}(t-r)^2 f^{(2)}(t)\,dt\right)d\mu(x) \\ &-\frac{1}{2(R-r)}\int_\Gamma q(x)\left(\int_{\frac{p(x)}{q(x)}}^R (t-R)^2 f^{(2)}(t)\,dt\right)d\mu(x). \end{aligned} \tag{2.9}$$

3. Bounds for the Remainder

In formula (2.6), we consider the remainder $R_f(p,q)$ given by

$$\begin{aligned} R_f(p,q) \;:&= \frac{(-1)^{n+1}}{(R-r)\,n!}\left[\int_\Gamma q(x)\left(\int_r^{\frac{p(x)}{q(x)}}(t-r)^n f^{(n)}(t)\,dt\right)d\mu(x)\right. \\ &\left.+\int_\Gamma q(x)\left(\int_{\frac{p(x)}{q(x)}}^R (t-R)^n f^{(n)}(t)\,dt\right)d\mu(x)\right]. \end{aligned}$$

In this section, we are interested in obtaining some bounds for $R_f(p,q)$.

For this purpose, consider

$$I_1(z)=\int_r^z (t-r)^n f^{(n)}(t)\,dt$$

and

$$I_2(z):=\int_z^R (t-R)^n f^{(n)}(t)\,dt,$$

where $z\in[r,R]$.

For $a<b$, we also define the Lebesgue norm,

$$\|f\|_{[a,b],p}:=\left[\int_a^b |f(t)|^p\,dt\right]^{\frac{1}{p}},\quad p\geq 1$$

and

$$\|f\|_{[a,b],\infty}:=ess\sup_{t\in[a,b]}|f(t)|.$$

Now, we observe that

$$\begin{aligned}
|I_1(z)| &\le \int_r^z |t-r|^n \left|f^{(n)}(t)\right| dt \le \left\|f^{(n)}\right\|_{[r,z],\infty} \frac{(z-r)^{n+1}}{n+1}, \\
|I_1(z)| &\le \left\|f^{(n)}\right\|_{[r,z],\beta} \left(\int_r^z |t-r|^{\alpha n} dt\right)^{\frac{1}{\alpha}} = \left\|f^{(n)}\right\|_{[r,z],\beta} \left[\frac{(z-r)^{\alpha n+1}}{\alpha n+1}\right]^{\frac{1}{\alpha}} \\
&= \left\|f^{(n)}\right\|_{[r,z],\beta} \frac{(z-r)^{n+\frac{1}{\alpha}}}{(\alpha n+1)^{\frac{1}{\alpha}}} \text{ if } \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1,
\end{aligned}$$

and, finally,

$$|I_1(z)| \le \left\|f^{(n)}\right\|_{[r,z],1} \sup_{t\in[r,z]} |t-r|^n = \left\|f^{(n)}\right\|_{[r,z],1} (z-r)^n .$$

Consequently, we have

$$|I_1(z)| \le \begin{cases} \frac{(z-r)^{n+1}}{n+1} \left\|f^{(n)}\right\|_{[r,z],\infty} \\ \frac{(z-r)^{n+\frac{1}{\alpha}}}{(\alpha n+1)^{\frac{1}{\alpha}}} \left\|f^{(n)}\right\|_{[r,z],\beta}, \quad \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\ (z-r)^n \left\|f^{(n)}\right\|_{[r,z],1} . \end{cases} \tag{3.1}$$

In a similar fashion, we may point out that

$$|I_2(z)| \le \begin{cases} \frac{(R-z)^{n+1}}{n+1} \left\|f^{(n)}\right\|_{[z,R],\infty} \\ \frac{(R-z)^{n+\frac{1}{\alpha}}}{(\alpha n+1)^{\frac{1}{\alpha}}} \left\|f^{(n)}\right\|_{[z,R],\beta}, \quad \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\ (R-z)^n \left\|f^{(n)}\right\|_{[z,R],1} . \end{cases} \tag{3.2}$$

The following bound for the remainder $R_f(p,q)$ holds.

Theorem 2. *Let f, r, R, p and q be as in Theorem 1. Then we have the inequality*

$$|R_f(p,q)| \le \frac{1}{n!\,(R-r)} \tag{3.3}$$

$$
\times \begin{cases}
\frac{1}{n+1}\int_{\Gamma} q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^{n+1}\left\|f^{(n)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],\infty} \right. \\
\qquad \left. +\left(\frac{p(x)}{q(x)}-r\right)^{n+1}\left\|f^{(n)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],\infty}\right] d\mu(x) \\
\\
\frac{1}{(\alpha n+1)^{\frac{1}{\alpha}}}\int_{\Gamma} q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^{n+\frac{1}{\alpha}}\left\|f^{(n)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],\beta} \right. \\
\qquad \left. +\left(\frac{p(x)}{q(x)}-r\right)^{n+\frac{1}{\alpha}}\left\|f^{(n)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],\beta}\right] d\mu(x) \\
\text{if } \ \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\
\\
\int_{\Gamma} q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^{n}\left\|f^{(n)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],1} \right. \\
\qquad \left. +\left(\frac{p(x)}{q(x)}-r\right)^{n}\left\|f^{(n)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],1}\right] d\mu(x)
\end{cases}
$$

$$
\leq \frac{1}{n!(R-r)} \times \begin{cases}
\frac{\left\|f^{(n)}\right\|_{[r,R],\infty}}{n+1}\int_{\Gamma} q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^{n+1}+\left(\frac{p(x)}{q(x)}-r\right)^{n+1}\right] d\mu(x) \\
\\
\frac{\left\|f^{(n)}\right\|_{[r,R],\beta}}{(\alpha n+1)^{\frac{1}{\alpha}}}\int_{\Gamma} q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^{n+\frac{1}{\alpha}}+\left(\frac{p(x)}{q(x)}-r\right)^{n+\frac{1}{\alpha}}\right] d\mu(x) \\
\\
\left\|f^{(n)}\right\|_{[r,R],1}\int_{\Gamma} q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^{n}+\left(\frac{p(x)}{q(x)}-r\right)^{n}\right] d\mu(x)
\end{cases}
$$

$$
\leq \begin{cases}
\frac{\left\|f^{(n)}\right\|_{[r,R],\infty}(R-r)^{n}}{(n+1)!} \\
\\
\frac{\left\|f^{(n)}\right\|_{[r,R],\beta}(R-r)^{n+\frac{1}{\alpha}-1}}{n!(\alpha n+1)^{\frac{1}{\alpha}}} \quad \text{if } \ \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\
\\
\frac{\left\|f^{(n)}\right\|_{[r,R],1}(R-r)^{n-1}}{n!}
\end{cases}
$$

Proof. Using (3.1) and (3.2), we may write:

$$\begin{aligned}
\left|R_f(p,q)\right| &\le \frac{1}{n!\,(R-r)}\left[\int_\Gamma q(x)\left|\int_r^{\frac{p(x)}{q(x)}}(t-r)^n f^{(n)}(t)\,dt\right|d\mu(x)\right.\\
&\quad\left.+\int_\Gamma q(x)\left|\int_{\frac{p(x)}{q(x)}}^R (t-R)^n f^{(n)}(t)\,dt\right|d\mu(x)\right]\\
&\le \frac{1}{(R-r)\,n!}\times\begin{cases}
\frac{1}{n+1}\int_\Gamma q(x)\left(\frac{p(x)}{q(x)}-r\right)^{n+1}\left\|f^{(n)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],\infty}d\mu(x)\\
\frac{1}{(\alpha n+1)^{\frac{1}{\alpha}}}\int_\Gamma q(x)\left(\frac{p(x)}{q(x)}-r\right)^{n+\frac{1}{\alpha}}\left\|f^{(n)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],\beta}d\mu(x)\\
\text{if }\ \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1;\\
\int_\Gamma q(x)\left(\frac{p(x)}{q(x)}-r\right)^{n}\left\|f^{(n)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],1}d\mu(x)
\end{cases}\\
&\quad+\frac{1}{(R-r)\,n!}\times\begin{cases}
\frac{1}{n+1}\int_\Gamma q(x)\left(R-\frac{p(x)}{q(x)}\right)^{n+1}\left\|f^{(n)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],\infty}d\mu(x)\\
\frac{1}{(\alpha n+1)^{\frac{1}{\alpha}}}\int_\Gamma q(x)\left(R-\frac{p(x)}{q(x)}\right)^{n+\frac{1}{\alpha}}\left\|f^{(n)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],\beta}d\mu(x)\\
\text{if }\ \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1;\\
\int_\Gamma q(x)\left(R-\frac{p(x)}{q(x)}\right)^{n}\left\|f^{(n)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],1}d\mu(x)
\end{cases}\\
&\le \frac{1}{(R-r)\,n!}\times\begin{cases}
\frac{\left\|f^{(n)}\right\|_{[r,R],\infty}}{n+1}\int_\Gamma q(x)\left[\left(\frac{p(x)}{q(x)}-r\right)^{n+1}+\left(R-\frac{p(x)}{q(x)}\right)^{n+1}\right]d\mu(x)\\
\frac{\left\|f^{(n)}\right\|_{[r,R],\beta}}{(\alpha n+1)^{\frac{1}{\alpha}}}\int_\Gamma q(x)\left[\left(\frac{p(x)}{q(x)}-r\right)^{n+\frac{1}{\alpha}}+\left(R-\frac{p(x)}{q(x)}\right)^{n+\frac{1}{\alpha}}\right]d\mu(x)\\
\left\|f^{(n)}\right\|_{[r,R],1}\int_\Gamma q(x)\left[\left(\frac{p(x)}{q(x)}-r\right)^{n}+\left(R-\frac{p(x)}{q(x)}\right)^{n}\right]d\mu(x)
\end{cases}\\
&\le \frac{1}{(R-r)\,n!}\times\begin{cases}
\frac{\left\|f^{(n)}\right\|_{[r,R],\infty}}{n+1}(R-r)^{n+1}\\
\frac{\left\|f^{(n)}\right\|_{[r,R],\beta}(R-r)^{n+\frac{1}{\alpha}}}{(\alpha n+1)^{\frac{1}{\alpha}}}\\
\left\|f^{(n)}\right\|_{[r,R],1}(R-r)^{n}
\end{cases}
=\begin{cases}
\frac{\left\|f^{(n)}\right\|_{[r,R],\infty}(R-r)^{n}}{(n+1)!}\\
\frac{\left\|f^{(n)}\right\|_{[r,R],\beta}(R-r)^{n+\frac{1}{\alpha}-1}}{n!(\alpha n+1)^{\frac{1}{\alpha}}}\\
\frac{\left\|f^{(n)}\right\|_{[r,R],1}(R-r)^{n-1}}{n!}
\end{cases}
\end{aligned}$$

and the theorem is proved. ∎

Remark 2. *For $n=1$, we obtain the estimate*

$$|R_f(p,q)| \le \frac{1}{R-r}\times\begin{cases}\frac{1}{2}\int_\Gamma q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^2\left\|f^{(1)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],\infty}\right.\\ \qquad\left.+\left(\frac{p(x)}{q(x)}-r\right)^2\left\|f^{(1)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],\infty}\right]d\mu(x)\\ \\ \frac{1}{(\alpha+1)^{\frac{1}{\alpha}}}\int_\Gamma q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^{1+\frac{1}{\alpha}}\left\|f^{(1)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],\beta}\right.\\ \qquad\left.+\left(\frac{p(x)}{q(x)}-r\right)^{1+\frac{1}{\alpha}}\left\|f^{(1)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],\beta}\right]d\mu(x)\\ \text{if } \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1;\\ \\ \int_\Gamma q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)\left\|f^{(1)}\right\|_{\left[\frac{p(x)}{q(x)},R\right],1}\right.\\ \qquad\left.+\left(\frac{p(x)}{q(x)}-r\right)\left\|f^{(1)}\right\|_{\left[r,\frac{p(x)}{q(x)}\right],1}\right]d\mu(x)\end{cases}$$

$$\le \frac{1}{R-r}\times\begin{cases}\frac{\left\|f^{(1)}\right\|_{[r,R],\infty}}{2}\int_\Gamma q(x)\left[\left(R-\frac{p(x)}{q(x)}\right)^2+\left(\frac{p(x)}{q(x)}-r\right)^2\right]d\mu(x)\\ \\ \frac{\left\|f^{(1)}\right\|_{[r,R],\beta}}{(\alpha+1)^{\frac{1}{\alpha}}}\int_\Gamma q(x)\left[\left(\frac{p(x)}{q(x)}-r\right)^{1+\frac{1}{\alpha}}+\left(R-\frac{p(x)}{q(x)}\right)^{1+\frac{1}{\alpha}}\right]d\mu(x)\\ \text{if } \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1;\\ \\ (R-r)\left\|f^{(1)}\right\|_{[r,R],1}\end{cases}$$

$$\le \begin{cases}\frac{\left\|f^{(1)}\right\|_{[r,R],\infty}(R-r)}{2}\\ \\ \frac{\left\|f^{(1)}\right\|_{[r,R],\beta}(R-r)^{\frac{1}{\alpha}}}{(\alpha+1)^{\frac{1}{\alpha}}} \quad \text{if } \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1;\\ \\ \left\|f^{(1)}\right\|_{[r,R],1}\end{cases}$$

which improves some results from [18] *and* [19].

References

[1] S.M. ALI and S.D. SILVEY, A general class of coefficients of divergence of one distribution from another, *J. Roy. Statist. Soc. Sec B,* **28** (1966), 131-142.

[2] N.S. BARNETT, S.S. DRAGOMIR and A. SOFO, Better bounds for an inequality of the Ostrowski type with applications, *RGMIA Research Report Collection,* **3** (1) (2000), Article 1.

[3] M. BETH BASSAT, f-entropies, probability of error and feature selection, *Inform. Control*, **39** (1978), 227-242.

[4] A. BHATTACHARYYA, On a measure of divergence between two statistical populations defined by their probability distributions, *Bull. Calcutta Math. Soc.,* **35** (1943), 99-109.

[5] I. BURBEA and C.R. RAO, On the convexity of some divergence measures based on entropy function, *IEEE Trans. Inf. Th.,* **28** (3) (1982), 489-495.

[6] P. CERONE and S.S. DRAGOMIR, Midpoint type rules from an inequalities point of view, *Handbook of Analytic-Computational Methods in Applied Mathematics*, Editor: G. Anastassiou, CRC Press, N.Y., (2000), 135-200.

[7] P. CERONE and S.S. DRAGOMIR, Trapezoidal type rules from an inequalities point of view, *Handbook of Analytic-Computational Methods in Applied Mathematics*, Editor: G. Anastassiou, CRC Press, N.Y., (2000), 65-134.

[8] P. CERONE, S.S. DRAGOMIR, and J. ROUMELIOTIS, Some Ostrowski type inequalities for $n-$time differentiable mappings and applications, *Demonstratio Math.,* **32**(2) (1999), 697-712.

[9] C.H. CHEN, *Statistical Pattern Recognition,* Rocelle Park, New York, Hoyderc Book Co., 1973.

[10] C.K. CHOW and C.N. LIN, Approximating discrete probability distributions with dependence trees, *IEEE Trans. Inf. Th.,* **14** (3) (1968), 462-467.

[11] I. CSISZÁR, Information-type measures of difference of probability distributions and indirect observations, *Studia Math. Hungarica,* **2** (1967), 299-318.

[12] I. CSISZÁR, On topological properties of $f-$divergences, *Studia Math. Hungarica,* **2** (1967), 329-339.

[13] I. CSISZÁR, A note on Jensen's inequality, *Studia Sci. Math. Hung.,* **1** (1966), 185-188.

[14] I. CSISZÁR, On topological properties of $f-$divergences, *Studia Math. Hungarica,* **2** (1967), 329-339.

[15] I. CSISZÁR, A note on Jensen's inequality, *Studia Sci. Math. Hung.,* **1** (1966), 185-188.

[16] I. CSISZÁR and J. KÖRNER, *Information Theory: Coding Theorem for Discrete Memoryless Systems,* Academic Press, New York, 1981.

[17] D. DACUNHA-CASTELLE, *Ecole d'ete de Probability de Saint-Flour*, III-1977, Berlin, Heidelberg: Springer 1978.

[18] S.S. DRAGOMIR, V. GLUŠCEVIĆ and C.E.M. PEARCE, Csiszar f-divergence, Ostrowski inequality and mutual information, submitted.

[19] S.S. DRAGOMIR, V. GLUŠCEVIĆ and C.E.M. PEARCE, The approximation of Csiszár f-Divergence for Absolutely Continuous Mappings , submitted

[20] S.S. DRAGOMIR and S. MABIZELA, Some error estimates in the trapezoidal quadrature rule, *RGMIA Research Report Collect.,* **2** (1999), No. 5, Article 6 (`http://rgmia.vu.edu.au/v2n5.html`)

[21] S.S. DRAGOMIR and S. WANG, A new inequality of Ostrowski-Grüss type and its applications to the estimation of error bounds for special means and for some numerical quadrature rules, *Comp. Math. Appl.,* **33** (1997), No. 11, 15-20.

[22] D.V. GOKHALE and S. KULLBACK, *Information in Contingency Tables*, New York, Marcel Dekker, 1978.

[23] J.H. HAVRDA and F. CHARVAT, Quantification method classification process: concept of structural α-entropy, *Kybernetika*, **3** (1967), 30-35.

[24] E. HELLINGER, Neue Bergrüirdung du Theorie quadratisher Formerus von uneudlichvieleu Veränderlicher, *J. für reine and Augeur. Math.,* **36** (1909), 210-271.

[25] H. JEFFREYS, An invariant form for the prior probability in estimating problems, *Proc. Roy. Soc. London*, **186** A (1946), 453-461.

[26] T.T. KADOTA and L.A. SHEPP, On the best finite set of linear observables for discriminating two Gaussian signals, *IEEE Trans. Inf. Th.,* **13** (1967), 288-294.

[27] T. KAILATH, The divergence and Bhattacharyya distance measures in signal selection, *IEEE Trans. Comm. Technology.,* Vol COM-15 (1967), 52-60.

[28] J.N. KAPUR, A comparative assessment of various measures of directed divergence, *Advances in Management Studies*, **3** (1984), 1-16.

[29] D. KAZAKOS and T. COTSIDAS, A decision theory approach to the approximation of discrete probability densities, *IEEE Trans. Perform. Anal. Machine Intell.,* **1** (1980), 61-67.

[30] J.H.B. KEMPERMAN, On the optimum note of transmitting information, *Ann. Math. Statist.,* **40** (1969), 2158-2177.

[31] C. KRAFT, Some conditions for consistency and uniform consistency of statistical procedures, *Univ. of California Pub. in Statistics*, **1** (1955), 125-142.

[32] S. KULLBACK, A lower bound for discrimination information in terms of variation, *IEEE Trans. Inf. Th.,* **13** (1967), 126-127.

[33] S. KULLBACK, Correction to a lower bound for discrimination information in terms of variation, *IEEE Trans. Inf. Th.,* **16** (1970), 771-773.
[34] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Ann. Math. Stat.*, **22** (1951), 79-86.
[35] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Annals Math. Statist.,* **22** (1951), 79-86.
[36] L. LECAM, *Asymptotic Methods in Statistical Decision Theory,* New York: Springer, 1986.
[37] J. LIN, Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Th.,* **37** (1) (1991), 145-151.
[38] J. LIN and S.K.M. WONG, A new directed divergence measure and its characterization, *Int. J. General Systems,* **17** (1990), 73-81.
[39] H.P. McKEAN, JR., Speed of approach to equilibrium for Koc's caricature of a Maximilian gas, *Arch. Ration. Mech. Anal.,* **21** (1966), 343-367.
[40] M. MEI, The theory of genetic distance and evaluation of human races, *Japan J. Human Genetics*, **23** (1978), 341-369.
[41] C.E.M. PEARCE, J. PEČARIĆ, N. UJEVIĆ and S. VAROŠANEC, Generaliasations of some inequalities of Ostrowski-Grüss type, *Math. Ineq. & Appl.,* **3** (2000), No. 1, 25-34.
[42] E.C. PIELOU, *Ecological Diversity,* Wiley, New York, 1975.
[43] M.S. PINSKER, Information and Information Stability of Random variables and processes, (in Russian), Moscow: Izv. Akad. Nouk, 1960.
[44] C.R. RAO, Diversity and dissimilarity coefficients: a unified approach, *Theoretic Population Biology*, **21** (1982), 24-43.
[45] A. RÉNYI, On measures of entropy and information, *Proc. Fourth Berkeley Symp. Math. Stat. and Prob., University of California Press*, **1** (1961), 547-561.
[46] A. SEN, *On Economic Inequality,* Oxford University Press, London 1973.
[47] B.D. SHARMA and D.P. MITTAL, New non-additive measures of relative information, *Journ. Comb. Inf. Sys. Sci.,* **2** (4)(1977), 122-132.
[48] H. SHIOYA and T. DA-TE, A generalisation of Lin divergence and the derivative of a new information divergence, *Elec. and Comm. in Japan*, **78** (7) (1995), 37-40.
[49] I. J. TANEJA, *Generalised Information Measures and their Applications* (http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html).
[50] H. THEIL, *Economics and Information Theory*, North-Holland, Amsterdam, 1967.
[51] H. THEIL, *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.
[52] F. TOPSOE, Some inequalities for information divergence and related measures of discrimination, *Res. Rep. Coll., RGMIA,* **2** (1) (1999), 85-98.
[53] G.T. TOUSSAINT, Sharper lower bounds for discrimination in terms of variation, *IEEE Trans. Inf. Th.,* **21** (1975), 99-100.
[54] I. VAJDA, Note on discrimination information and variation, *IEEE Trans. Inf. Th.,* **16** (1970), 771-773.
[55] I. VAJDA, *Theory of Statistical Inference and Information,* Dordrecht-Boston, Kluwer Academic Publishers, 1989.
[56] V.A. VOLKONSKI and J. A. ROZANOV, Some limit theorems for random function -I, (English Trans.), *Theory Prob. Appl., (*USSR), **4** (1959), 178-197.

School of Communications and Informatics, Victoria University of Technology, PO Box 14428, Melbourne City MC 8001, Australia
E-mail address: neil@matilda.vu.edu.au
URL: http://sci.vu.edu.au/staff/neilb.html

E-mail address: pc@matilda.vu.edu.au
URL: http://sci.vu.edu.au/staff/peterc.html

E-mail address: sever@matilda.vu.edu.au
URL: http:// rgmia.vu.edu.au/SSDragomirWeb.html

E-mail address: sofo@matilda.vu.edu.au
URL: http://cams.vu.edu.au/staff/anthonys.html