APPROXIMATING CSISZÁR $f-$DIVERGENCE VIA A GENERALISED TAYLOR FORMULA

S.S. DRAGOMIR AND V. GLUŠCEVIĆ

Abstract. Some approximation of the Csiszár $f-$divergence by the use of a generalised Taylor formula and applications are given.

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 [1], Kullback and Leibler [2], Rényi [3], Havrda and Charvat [4], Kapur [5], Sharma and Mittal [6], Burbea and Rao [7], Rao [8], Lin [9], Csiszár [10], Ali and Silvey [12], Vajda [13], Shioya and Da-te [47] and others (see for example [5] and the references therein).

These measures have been applied in a variety of fields such as: anthropology [8], genetics [14], finance, economics, and political science [15], [16], [17], biology [18], the analysis of contingency tables [19], approximation of probability distributions [20], [21], signal processing [22], [23] and pattern recognition [24], [25]. A number of these measures of distance are specific cases of Csiszár f-divergence and so further exploration of this concept will have a flow on effect to other measures of distance and to areas in which they are applied.

Assume that a set Γ and the $\sigma-$finite measure μ are given. Consider the set of all probability densities on μ to be $\Omega := \left\{p | p : \Gamma \to \mathbb{R},\, p(x) \geq 0,\ \int_\Gamma p(x)\, d\mu(x) = 1\right\}$. The Kullback-Leibler divergence [2] is well known among the information divergences. It is defined as:

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

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 [40], χ^2-*divergence* D_{χ^2}, $\alpha-$*divergence* D_α, *Bhattacharyya distance* D_B [41], *Harmonic distance* D_{Ha}, *Jeffrey's distance* D_J [1], *triangular discrimination* D_Δ [35], etc... They are defined as follows:

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

Date: October 25, 2000.
1991 *Mathematics Subject Classification.* Primary 94Xxx; Secondary 26D15.
Key words and phrases. Csiszár $f-$divergence, Taylor formula.

$$D_{H}(p,q):=\int_{\Gamma}\left|\sqrt{p(x)}-\sqrt{q(x)}\right|d\mu(x),\quad p,q\in\Omega; \tag{1.3}$$

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

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

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

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

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

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

For other divergence measures, see the paper [5] by Kapur or the book on line [42] by Taneja. For a comprehensive collection of preprints available on line, see the RGMIA web site `http://rgmia.vu.edu.au/papersinfth.html`

Csiszár $f-$divergence is defined as follows [10]

$$I_{f}(p,q):=\int_{\Gamma}p(x)f\left[\frac{q(x)}{p(x)}\right]d\mu(x),\quad p,q\in\Omega, \tag{1.10}$$

where f is convex on $(0,\infty)$. It is assumed that $f(u)$ is zero and strictly convex at $u=1$. By appropriately defining this convex function, various divergences are derived. All the above distances $(1.1)-(1.9)$, are particular instances of Csiszár $f-$divergence. There are also many others which are not in this class (see for example [5] or [42]). For the basic properties of Csiszár $f-$divergence see [43]-[45].

In [46], Lin and Wong (see also [9]) introduced the following divergence

$$D_{LW}(p,q):=\int_{\Gamma}p(x)\log\left[\frac{p(x)}{\frac{1}{2}p(x)+\frac{1}{2}q(x)}\right]d\mu(x),\quad p,q\in\Omega. \tag{1.11}$$

This can be represented as follows, using the Kullback-Leibler divergence:

$$D_{LW}(p,q)=D_{KL}\left(p,\frac{1}{2}p+\frac{1}{2}q\right).$$

Lin and Wong have established the following inequalities

$$D_{LW}(p,q)\leq\frac{1}{2}D_{KL}(p,q); \tag{1.12}$$

$$D_{LW}(p,q)+D_{LW}(q,p)\leq D_{v}(p,q)\leq 2; \tag{1.13}$$

$$D_{LW}(p,q)\leq 1. \tag{1.14}$$

In [47], Shioya and Da-te improved $(1.12)-(1.14)$ by showing that

$$D_{LW}(p,q)\leq\frac{1}{2}D_{v}(p,q)\leq 1.$$

For classical and new results in comparing different kinds of divergence measures, see the papers [1]-[47] where further references are given.

2. Representation of Csiszár $f-$Divergence

We may state the following result which is a reformulation of Theorem 1 in [50]:

Theorem 1. *Let $\{S_n(\cdot,\cdot)\}_{n\in\mathbb{N}}$ be a sequence of polynomials of two variables satisfying the condition:*

$$\frac{\partial S_n(t,x)}{\partial t} = S_{n-1}(t,x), \quad S_0(t,x) = 1 \text{ for } x,t \in \mathbb{R} \text{ and } n \in \mathbb{N}. \tag{2.1}$$

Then we have the identity

$$\begin{aligned} f(x) &= f(a) + \sum_{k=1}^{n} (-1)^{k+1} \left[S_k(x,x) f^{(k)}(x) - S_k(a,x) f^{(k)}(a)\right] \\ &\quad + R_n(f;a,x), \end{aligned} \tag{2.2}$$

where

$$R_n(f;a,x) := (-1)^n \int_a^x S_n(t,x) f^{(n+1)}(t)\,dt$$

and $f : I \to \mathbb{R}$ is such that $f^{(n)}$ is absolutely continuous on I.

1. If in (2.2) we set $S_n(t,x) = \frac{1}{n!}(t-x)^n$ $(n \in \mathbb{N})$, then we get the Taylor's identity [48]:

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt, \ x \in I. \tag{2.3}$$

2. If in (2.2) we set $S_n(t,x) = \frac{1}{n!}\left(t - \frac{a+x}{2}\right)^n$, $(n \in \mathbb{N})$, then we have the identity [50]:

$$\begin{aligned} f(x) &= f(a) + \sum_{k=1}^{n} \frac{(x-a)^k}{2^k k!}\left[f^{(k)}(a) + (-1)^{k+1} f^{(k)}(x)\right] \\ &\quad + \frac{(-1)^n}{n!}\int_a^x \left(x - \frac{a+x}{2}\right)^n f^{(n+1)}(t)\,dt, \ x \in I. \end{aligned} \tag{2.4}$$

3. If in (2.2) we set $S_n(t,x) = \frac{(x-a)^n}{n!} B_n\left(\frac{t-a}{x-a}\right)$, $n \in \mathbb{N}$, $S_0(t,x) = 1$, where B_n denotes the Bernoulli polynomials and $B_n := B_n(0)$ are the Bernoulli numbers, then we have the following representation [50]

$$\begin{aligned} f(x) &= f(a) + \frac{x-a}{2}\left[f'(x) + f'(a)\right] \\ &\quad - \sum_{k=1}^{\left[\frac{n}{2}\right]} \frac{(x-a)^{2k}}{(2k)!} B_{2k}\left[f^{(2k)}(x) - f^{(2k)}(a)\right] \\ &\quad + (-1)^n \frac{(x-a)^n}{n!}\int_a^x B_n\left(\frac{t-a}{x-a}\right) f^{(n+1)}(t)\,dt. \end{aligned} \tag{2.5}$$

4. If in (2.2) we set $S_n(t,x) = \frac{(x-a)^n}{n!}E_n\left(\frac{t-a}{x-a}\right)$, $n \in \mathbb{N}$, $S_0(t,x) = 1$, where $E_n(t)$ denotes the Euler polynomials and B_n are the Bernoulli numbers, then we have the representation [50]

$$f(x) = f(a) + 2\sum_{k=1}^{\left[\frac{n+1}{2}\right]} \frac{(x-a)^{2k-1}\left(4^k-1\right)}{(2k)!} \times B_{2k}\left[f^{(2k-1)}(x) + f^{(2k-1)}(a)\right] + (-1)^n \frac{(x-a)^n}{n!}\int_a^x E_n\left(\frac{t-a}{x-a}\right) f^{(n+1)}(t)\,dt. \tag{2.6}$$

We are able now to point out the following representation for the Csiszár $f-$divergence.

Theorem 2. *Let $\{S_n(t,z)\}_{n\in\mathbb{N}}$ and $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be as in Theorem 1. If $p, q \in \Omega$, then we have the representation*

$$I_f(p,q) = f(1) + \sum_{k=1}^{n} (-1)^{k+1}\left[I_{S_k(\cdot,\cdot)f^{(k)}(\cdot)}(p,q) - I_{S_k(1,\cdot)f^{(k)}(1)}(p,q)\right] + R_f(p,q), \tag{2.7}$$

where the remainder $R_f(p,q)$ can be given by

$$R_f(p,q) = (-1)^n \int_\Gamma q(x)\left(\int_1^{\frac{p(x)}{q(x)}} S_n\left(t, \frac{p(x)}{q(x)}\right) f^{(n+1)}(t)\,dt\right) d\mu(x). \tag{2.8}$$

Proof. From the representation (2.2), we may write that

$$f\left(\frac{p(x)}{q(x)}\right) = f(1) + \sum_{k=1}^{n}(-1)^{k+1}\left[S_k\left(\frac{p(x)}{q(x)}, \frac{p(x)}{q(x)}\right) f^{(k)}\left(\frac{p(x)}{q(x)}\right) - S_k\left(1, \frac{p(x)}{q(x)}\right) f^{(k)}(1)\right] + (-1)^n \int_1^{\frac{p(x)}{q(x)}} S_n\left(t, \frac{p(x)}{q(x)}\right) f^{(n+1)}(t)\,dt \tag{2.9}$$

for all $x \in \Gamma$.

If we multiply (2.9) by $q(x) \geq 0$ $(x \in \Gamma)$, integrate on Γ and take into account that $\int_\Gamma q(x)\,d\mu(x) = 1$, then we obtain the desired representation (2.7). ∎

The following particular cases are important in applications.

1. If we use the representation (2.3), we get

$$I_f(p,q) = f(1) + \sum_{k=1}^{n} \frac{f^{(k)}(1)}{k!} D_k(p,q) + \frac{1}{n!}\int_\Gamma q(x)\left[\int_1^{\frac{p(x)}{q(x)}} \left(\frac{p(x)}{q(x)} - t\right)^n f^{(n+1)}(t)\,dt\right] d\mu(x) \tag{2.10}$$

for all $p, q \in \Omega$, where

$$D_k(p,q) := \int_\Gamma q^{-k+1}(x)\,(p(x) - q(x))^k\,d\mu(x), \quad k = 1, \ldots, n.$$

2. From the identity (2.4), we may get that

$$
\begin{aligned}
(2.11)\qquad & I_f(p,q) \\
&= f(1)+\sum_{k=1}^{n}\frac{f^{(k)}(1)}{2^k k!}D_k(p,q)+\sum_{k=1}^{n}\frac{(-1)^{k+1}}{2^k k!}I_{(\cdot-1)^k f^{(k)}(\cdot)}(p,q) \\
&\quad +\frac{1}{n!}\int_\Gamma q(x)\left[\int_1^{\frac{p(x)}{q(x)}}\left(1-\frac{p(x)+q(x)}{q(x)}\right)^n f^{(n+1)}(t)\,dt\right]d\mu(x)
\end{aligned}
$$

for all $p,q\in\Omega$.

3. If we use the identity (2.5), we may obtain:

$$
\begin{aligned}
(2.12)\quad I_f(p,q) &= f(1)+\int_\Gamma\left[\frac{p(x)-q(x)}{2}\right]f'\left(\frac{p(x)}{q(x)}\right)d\mu(x) \\
&\quad +\sum_{k=1}^{\left[\frac{n}{2}\right]}\frac{f^{(2k)}(1)}{(2k)!}D_{2k}(p,q)-\sum_{k=1}^{\left[\frac{n}{2}\right]}\frac{B_{2k}}{(2k)!}I_{(\cdot-1)^{2k}f^{(2k)}(\cdot)}(p,q) \\
&\quad +\frac{(-1)^n}{n!}\int_\Gamma q^{-n+1}(x)\,(p(x)-q(x))^n \\
&\quad \times\left[\int_1^{\frac{p(x)}{q(x)}}B_n\left(\frac{t-1}{\frac{p(x)}{q(x)}-1}\right)f^{(n+1)}(t)\,dt\right]d\mu(x)
\end{aligned}
$$

for all $p,q\in\Omega$.

4. Finally, by the use of identity (2.6), we may write that

$$
\begin{aligned}
(2.13)\qquad I_f(p,q) &= f(1)+2\sum_{k=1}^{\left[\frac{n+1}{2}\right]}\frac{\left(4^k-1\right)B_{2k}f^{(k)}(1)}{(2k)!}D_{2k-1}(p,q) \\
&\quad +2\sum_{k=1}^{\left[\frac{n+1}{2}\right]}\frac{B_{2k}\left(4^k-1\right)}{(2k)!}I_{(\cdot-1)^{2k-1}f^{(2k-1)}}(p,q) \\
&\quad +\frac{(-1)^n}{n!}\int_\Gamma q^{-n+1}(x)\,(p(x)-q(x))^n \\
&\quad \times\left[\int_1^{\frac{p(x)}{q(x)}}E_n\left(\frac{t-1}{\frac{p(x)}{q(x)}-1}\right)f^{(n+1)}(t)\,dt\right]d\mu(x)
\end{aligned}
$$

for all $p,q\in\Omega$.

3. Bounds for the Remainder

For $a,b\in\mathbb{R}$, we denote

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

and

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

It is obvious that the order $a < b$ or $a > b$ is irrelevant in the definitions of the above Lebesgue $p-$norms.

The following general theorem involving the estimation of the remainder $R_f(p,q)$ holds.

Theorem 3. *Assume that* $\{S_n(t,x)\}_{n\in\mathbb{N}}$ *and* f *are as in Theorem 1. If* $p,q\in\Omega$, *then we have the inequality*

(3.1)
$$|R_f(p,q)| \le \begin{cases} \displaystyle\int_\Gamma q(x)\left\|f^{(n+1)}\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\infty} \times \left\|S_n\left(\cdot,\frac{p(x)}{q(x)}\right)\right\|_{\left[1,\frac{p(x)}{q(x)}\right],1} d\mu(x), \\ \displaystyle\int_\Gamma q(x)\left\|f^{(n+1)}\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\alpha} \times \left\|S_n\left(\cdot,\frac{p(x)}{q(x)}\right)\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\beta} d\mu(x), \\ \qquad \text{if} \quad \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1, \\ \displaystyle\int_\Gamma q(x)\left\|f^{(n+1)}\right\|_{\left[1,\frac{p(x)}{q(x)}\right],1} \times \left\|S_n\left(\cdot,\frac{p(x)}{q(x)}\right)\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\infty} d\mu(x). \end{cases}$$

Proof. We have that

$$|R_f(p,q)| \le \int_\Gamma q(x)\left|\int_1^{\frac{p(x)}{q(x)}} S_n\left(t,\frac{p(x)}{q(x)}\right) f^{(n+1)}(t)\,dt\right| d\mu(x). \tag{3.2}$$

Now, observe that

$$\begin{aligned} &\left|\int_1^{\frac{p(x)}{q(x)}} S_n\left(t,\frac{p(x)}{q(x)}\right) f^{(n+1)}(t)\,dt\right| \\ \le\ & \left\|f^{(n+1)}\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\infty} \times \left\|S_n\left(\cdot,\frac{p(x)}{q(x)}\right)\right\|_{\left[1,\frac{p(x)}{q(x)}\right],1} \end{aligned} \tag{3.3}$$

and, by Hölder's inequality for $\alpha>1$, $\frac{1}{\alpha}+\frac{1}{\beta}=1$,

$$\begin{aligned} &\left|\int_1^{\frac{p(x)}{q(x)}} S_n\left(t,\frac{p(x)}{q(x)}\right) f^{(n+1)}(t)\,dt\right| \\ \le\ & \left|\int_1^{\frac{p(x)}{q(x)}} \left|S_n\left(t,\frac{p(x)}{q(x)}\right)\right|^\beta dt\right|^{\frac{1}{\beta}} \times \left|\int_1^{\frac{p(x)}{q(x)}} \left|f^{(n+1)}(t)\right|^\alpha dt\right|^{\frac{1}{\alpha}} \\ =\ & \left\|S_n\left(\cdot,\frac{p(x)}{q(x)}\right)\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\beta} \times \left\|f^{(n+1)}\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\alpha}. \end{aligned} \tag{3.4}$$

Finally,

$$\begin{aligned} &\left|\int_1^{\frac{p(x)}{q(x)}} S_n\left(t,\frac{p(x)}{q(x)}\right) f^{(n+1)}(t)\,dt\right| \\ \le\ & \left\|f^{(n+1)}\right\|_{\left[1,\frac{p(x)}{q(x)}\right],1} \times \left\|S_n\left(\cdot,\frac{p(x)}{q(x)}\right)\right\|_{\left[1,\frac{p(x)}{q(x)}\right],\infty} \end{aligned} \tag{3.5}$$

for all $x\in\Gamma$.

Using (3.2) and (3.3) – (3.5), we deduce (3.1). ∎

Remark 1. *If we assume that* $0 \le r \le \frac{p(x)}{q(x)} \le R < \infty$ *for all* $x \in \Gamma$, *then obviously* $r \le 1 \le R$ *and the right side of the inequality (3.1) may be upper bounded by*

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

If we choose some particular instances of polynomials $S_n(\cdot,\cdot)$ we may compute the Lebesgue norm $\left\|S_n\left(\cdot, \frac{p(x)}{q(x)}\right)\right\|_s$, $s \in [1,\infty]$, obtaining more explicit bounds for the remainder $R_f(p,q)$.

1. If we choose $S_n(t,z) = \frac{1}{n!}(t-z)^n$, then

$$\|S_n(\cdot,z)\|_{[1,z],1} = \frac{1}{n!}\left|\int_1^z |t-z|^n \, dt\right| = \frac{1}{(n+1)!}|z-1|^{n+1},$$

$$\begin{aligned} \|S_n(\cdot,z)\|_{[1,z],\alpha} &= \frac{1}{n!}\left|\int_1^z |t-z|^{\alpha n} \, dt\right|^{\frac{1}{\alpha}} \\ &= \frac{1}{n!}\left[\frac{|z-1|^{\alpha n+1}}{\alpha n+1}\right]^{\frac{1}{\alpha}} = \frac{|z-1|^{n+\frac{1}{\alpha}}}{n!(\alpha n+1)^{\frac{1}{\alpha}}}, \end{aligned}$$

and

$$\|S_n(\cdot,z)\|_{[1,z],\infty} = \frac{1}{n!}|z-1|^n.$$

Consequently, we may state the following corollary which is useful in practice.

Corollary 1. *Let* f *be as in Theorem 1. Then, for* $p,q \in$,

$$I_f(p,q) = f(1) + \sum_{k=1}^n \frac{f^{(k)}(1)}{k!} D_k(p,q) + R_f(p,q), \tag{3.6}$$

where

$$D_k(p,q) := \int_\Gamma q^{-k+1}(x)(p(x)-q(x))^k \, d\mu(x)$$

and the remainder $R_f(p,q)$ *satisfies the bound*

$$|R_f(p,q)| \tag{3.7}$$

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

Moreover, if $0 \leq r \leq \frac{p(x)}{q(x)} \leq R < \infty$ for all $x \in \Gamma$, then the right hand side of (3.7) can be upper bounded by

$$
(3.8)\qquad \begin{cases} \dfrac{\left\| f^{(n+1)} \right\|_{[r,R],\infty}}{(n+1)!} \displaystyle\int_\Gamma |p(x) - q(x)|^{n+1} [q(x)]^{-n} d\mu(x), \\[2ex] \dfrac{\left\| f^{(n+1)} \right\|_{[r,R],\beta}}{n!(\alpha n+1)^{\frac{1}{\alpha}}} \displaystyle\int_\Gamma |p(x) - q(x)|^{n+\frac{1}{\alpha}} [q(x)]^{-n-\frac{1}{\alpha}+1} d\mu(x), \\ \text{if} \quad \alpha > 1, \ \frac{1}{\alpha} + \frac{1}{\beta} = 1, \\[2ex] \dfrac{\left\| f^{(n+1)} \right\|_{[r,R],1}}{n!} \displaystyle\int_\Gamma |p(x) - q(x)|^{n} [q(x)]^{-n+1} d\mu(x), \end{cases}
$$

$$
\leq \begin{cases} \dfrac{\left\| f^{(n+1)} \right\|_{[r,R],\infty} (R-r)^{n+1}}{(n+1)!}, \\[2ex] \dfrac{\left\| f^{(n+1)} \right\|_{[r,R],\beta} (R-r)^{n+\frac{1}{\alpha}}}{n!(\alpha n+1)^{\frac{1}{\alpha}}}, \\[2ex] \dfrac{\left\| f^{(n+1)} \right\|_{[r,R],1} (R-r)^{n}}{n!}. \end{cases}
$$

2. If we choose $S_n(t,z) = \frac{1}{n!}\left(t - \frac{1+z}{2}\right)^n$, then

$$
\|S_n(\cdot,z)\|_{[1,z],1} = \frac{1}{n!}\left| \int_1^z \left| t - \frac{1+z}{2} \right|^n dt \right|.
$$

If we assume that $z \geq 1$, then

$$
\begin{aligned}
\int_1^z \left| t - \frac{1+z}{2} \right|^n dt &= \int_1^{\frac{1+z}{2}} \left(\frac{1+z}{2} - t \right)^n dt + \int_{\frac{1+z}{2}}^z \left(t - \frac{1+z}{2} \right)^n dt \\
&= \frac{1}{n+1}\left(\frac{z-1}{2} \right)^{n+1} + \frac{1}{n+1}\left(\frac{z-1}{2} \right)^{n+1} \\
&= \frac{(z-1)^{n+1}}{(n+1)2^n}.
\end{aligned}
$$

If we assume that $z \leq 1$, then

$$
\int_z^1 \left| t - \frac{1+z}{2} \right|^n dt = \frac{(1-z)^{n+1}}{(n+1)2^n}
$$

and thus, we may state that

$$
\|S_n(\cdot,z)\|_{[1,z],1} = \frac{1}{(n+1)!} \cdot \frac{|z-1|^{n+1}}{2^n}.
$$

Similarly, we have

$$\begin{aligned}\|S_n(\cdot,z)\|_{[1,z],\alpha} &= \frac{1}{n!}\left|\int_1^z\left|t-\frac{1+z}{2}\right|^{n\alpha}dt\right|^{\frac{1}{\alpha}} = \frac{1}{n!}\left[\frac{|z-1|^{n\alpha+1}}{(n\alpha+1)\,2^{n\alpha}}\right]^{\frac{1}{\alpha}} \\ &= \frac{1}{n!}\cdot\frac{|z-1|^{n+\frac{1}{\alpha}}}{(n\alpha+1)^{\frac{1}{\alpha}}\,2^n},\quad \alpha\geq 1\end{aligned}$$

and

$$\|S_n(\cdot,z)\|_{[1,z],\infty} = \frac{1}{n!}\cdot\frac{|z-1|^n}{2^n}.$$

Consequently, we may state the following corollary which is useful in practice.

Corollary 2. *Let f be as in Theorem 1. Then, for $p,q\in\Omega$,*

$$\begin{aligned}I_f(p,q) &= f(1)+\sum_{k=1}^{n}\frac{f^{(k)}(1)}{2^k k!}D_k(p,q) \\ &\quad +\sum_{k=1}^{n}\frac{(-1)^{k+1}}{2^k k!}I_{(\cdot-1)^k f^{(k)}(\cdot)}(p,q)+\tilde{R}_f(p,q)\end{aligned}\tag{3.9}$$

and the remainder $\tilde{R}_f(p,q)$ satisfies the bound

$$\left|\tilde{R}_f(p,q)\right|\tag{3.10}$$

$$\leq\begin{cases}\dfrac{1}{(n+1)!2^n}\displaystyle\int_\Gamma[q(x)]^{-n}|p(x)-q(x)|^{n+1}\left\|f^{(n+1)}\right\|_{[1,\frac{p(x)}{q(x)}],\infty}d\mu(x), \\[2ex] \dfrac{1}{n!2^n(\alpha n+1)^{\frac{1}{\alpha}}}\displaystyle\int_\Gamma[q(x)]^{-n+1-\frac{1}{\alpha}}|p(x)-q(x)|^{n+\frac{1}{\alpha}}\left\|f^{(n+1)}\right\|_{[1,\frac{p(x)}{q(x)}],\beta}d\mu(x), \\ \text{if}\quad \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1, \\[2ex] \dfrac{1}{n!2^n}\displaystyle\int_\Gamma[q(x)]^{-n+1}|p(x)-q(x)|^{n}\left\|f^{(n+1)}\right\|_{[1,\frac{p(x)}{q(x)}],1}d\mu(x).\end{cases}$$

Moreover, if $0\leq r\leq\frac{p(x)}{q(x)}\leq R<\infty$ for all $x\in\Gamma$, then the right hand side of (3.10) can be upper bounded by:

$$\begin{cases}\dfrac{\left\|f^{(n+1)}\right\|_{[r,R],\infty}}{(n+1)!2^n}\displaystyle\int_\Gamma|p(x)-q(x)|^{n+1}[q(x)]^{-n}d\mu(x), \\[2ex] \dfrac{\left\|f^{(n+1)}\right\|_{[r,R],\beta}}{2^n n!(\alpha n+1)^{\frac{1}{\alpha}}}\displaystyle\int_\Gamma|p(x)-q(x)|^{n+\frac{1}{\alpha}}[q(x)]^{-n-\frac{1}{\alpha}+1}d\mu(x), \\ \text{if}\quad \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1, \\[2ex] \dfrac{\left\|f^{(n+1)}\right\|_{[r,R],1}}{2^n n!}\displaystyle\int_\Gamma|p(x)-q(x)|^{n}[q(x)]^{-n+1}d\mu(x),\end{cases}\tag{3.11}$$

$$\leq \begin{cases} \dfrac{\left\|f^{(n+1)}\right\|_{[r,R],\infty}(R-r)^{n+1}}{(n+1)!2^n}, \\ \\ \dfrac{\left\|f^{(n+1)}\right\|_{[r,R],\beta}(R-r)^{n+\frac{1}{\alpha}}}{n!\,(\alpha n+1)^{\frac{1}{\alpha}}\,2^n}, \quad \alpha>1, \ \ \frac{1}{\alpha}+\frac{1}{\beta}=1, \\ \\ \dfrac{\left\|f^{(n+1)}\right\|_{[r,R],1}(R-r)^{n}}{n!}. \end{cases}$$

References

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

[2] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Ann. Math. Stat.*, **22** (1951), 79-86.

[3] A. RÉNYI, On measures of entropy and information, *Proc. Fourth Berkeley Symp. Math. Stat. and Prob., University of California Press*, **1** (1961), 547-561.

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

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

[6] B.D. SHARMA and D.P. MITTAL, New non-additive measures of relative information, *Journ. Comb. Inf. Sys. Sci.,* **2** (4)(1977), 122-132.

[7] 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.

[8] C.R. RAO, Diversity and dissimilarity coefficients: a unified approach, *Theoretic Population Biology*, **21** (1982), 24-43.

[9] J. LIN, Divergence measures based on the Shannon entropy, *IEEE Trans. Inf. Th.,* **37** (1) (1991), 145-151.

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

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

[12] 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.

[13] I. VAJDA, *Theory of Statistical Inference and Information,* Dordrecht-Boston, Kluwer Academic Publishers, 1989.

[14] M. MEI, The theory of genetic distance and evaluation of human races, *Japan J. Human Genetics*, **23** (1978), 341-369.

[15] A. SEN, *On Economic Inequality,* Oxford University Press, London 1973.

[16] H. THEIL, *Economics and Information Theory*, North-Holland, Amsterdam, 1967.

[17] H. THEIL, *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.

[18] E.C. PIELOU, *Ecological Diversity,* Wiley, New York, 1975.

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

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

[21] 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.

[22] 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.

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

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

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

[26] V.A. VOLKONSKI and J. A. ROZANOV, Some limit theorems for random function -I, (English Trans.), *Theory Prob. Appl., (*USSR), **4** (1959), 178-197.

[27] M.S. PINSKER, Information and Information Stability of Random variables and processes, (in Russian), Moscow: Izv. Akad. Nouk, 1960.

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

[29] 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.

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

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

[32] S. KULLBACK, Correction to a lower bound for discrimination information in terms of variation, *IEEE Trans. Inf. Th.,* **16** (1970), 771-773.

[33] I. VAJDA, Note on discrimination information and variation, *IEEE Trans. Inf. Th.,* **16** (1970), 771-773.

[34] G.T. TOUSSAINT, Sharper lower bounds for discrimination in terms of variation, *IEEE Trans. Inf. Th.,* **21** (1975), 99-100.

[35] F. TOPSOE, Some inequalities for information divergence and related measures of discrimination, *Res. Rep. Coll., RGMIA,* **2** (1) (1999), 85-98.

[36] L. LECAM, *Asymptotic Methods in Statistical Decision Theory,* New York: Springer, 1986.

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

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

[39] S. KULLBACK and R.A. LEIBLER, On information and sufficiency, *Annals Math. Statist.,* **22** (1951), 79-86.

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

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

[42] I. J. TANEJA, *Generalised Information Measures and their Applications* (`http://www.mtm.ufsc.br/~taneja/bhtml/bhtml.html`).

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

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

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

[46] J. LIN and S.K.M. WONG, A new directed divergence measure and its characterization, *Int. J. General Systems,* **17** (1990), 73-81.

[47] 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.

[48] S.S. DRAGOMIR, New estimation of the remainder in Taylor's formula using Grüss' type inequalities and applications, *Math. Ineq. and Appl.*, **2** (2) (1999), 183-193.

[49] S.S. DRAGOMIR, An improvement of the remainder estimate in the generalised Taylor's formula, *RGMIA Res. Rep. Coll.,* **3** (2000), *No. 1, Article 1.*

[50] M. MATIĆ, J. PEČARIĆ and N. UJEVIĆ, On new estimation of the remainder in generalised Taylor's formula, *Math. Ineq. and Appl.*, **2** (3) (1999), 343-361.

School of Communications and Informatics, Victoria University of Technology, PO Box 14428, Melbourne City MC, Victoria 8001, Australia.

E-mail address: `sever@matilda.vu.edu.au`

URL: `http://rgmia.vu.edu.au/SSDragomirWeb.html`

Royal Australian Airforce, ARDU, PO Box 1500, Salisbury SA 5108, Australia.

E-mail address: `vgluscev@spam.adelaide.edu.au`