Mathematical
Inequalities
& Applications
Volume 5, Number 3 (2002), 417–434

APPROXIMATING CSISZÁR f –DIVERGENCE BY THE USE OF TAYLOR'S FORMULA WITH INTEGRAL REMAINDER

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

Abstract. Some approximations of the Csiszár $f-$ divergence by the use of Taylor's formula and perturbed Taylor's formula and some applications for Kullback-Leibler distance 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 $\Omega := \left\{p | p : \Gamma \to \mathbb{R}, p(x) \geqslant 0, \int_\Gamma p(x)\, d\mu(x) = 1\right\}$. 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*

Mathematics subject classification (2000): 26D15.
Key words and phrases: Csiszár $f-$ divergence, Taylor's formula, Analytic inequalities.

Paper MIA-05-42

D_B [41], *Harmonic distance* D_{Ha}, *Jeffrey's distance* D_J [1], *triangular discrimination* D_Δ [35], etc... They are defined as follows:

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

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

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

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

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

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

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

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

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]

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

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

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

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

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

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

$$D_{LW}(p,q) \leqslant \frac{1}{2} D_v(p,q) \leqslant 1.$$

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

2. The results

We start with the following result.

THEOREM 1. *Let $f:(0,\infty)\to\mathbb{R}$ be $n-$ time differentiable and such that $f^{(n)}$ is absolutely continuous on $[r,R]$, where $0<r\leqslant 1\leqslant R<\infty$. Assume that the probability distributions p,q satisfy the condition*

$$r \leqslant \frac{p(x)}{q(x)} \leqslant R \quad a.e. \text{ on } \Gamma. \tag{2.1}$$

Then we have the inequalities:

$$\left| I_f(p,q) - f(1) - \sum_{k=1}^{n} \frac{f^{(k)}(1)}{k!} D_{\chi^k}(p,q) \right| \tag{2.2}$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_{\infty} D_{|\chi|^{n+1}}(p,q) & \text{if } f^{(n+1)} \in L_\infty[r,R]; \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}} \left\| f^{(n+1)} \right\|_{\alpha} D_{|\chi|^{n+\frac{1}{\beta}}}(p,q) & \text{if } f^{(n+1)} \in L_\alpha[r,R], \\ & \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\ \frac{1}{n!} \left\| f^{(n+1)} \right\|_1 D_{|\chi|^n}(p,q); \end{cases}$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_{\infty} (R-r)^{n+1} & \text{if } f^{(n+1)} \in L_\infty[r,R]; \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}} \left\| f^{(n+1)} \right\|_{\alpha} (R-r)^{n+\frac{1}{\beta}} & \text{if } f^{(n+1)} \in L_\alpha[r,R], \\ & \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\ \frac{1}{n!} \left\| f^{(n+1)} \right\|_1 (R-r)^n; \end{cases}$$

where

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

$$D_{|\chi|^s}(p,q) \ :\ = \int_\Gamma \frac{|p(x)-q(x)|^s}{q^{s-1}(x)} d\mu(x) \quad (k\in\mathbb{N},\ s\geqslant 0)$$

and $\|\cdot\|_\alpha$ *are the usual Lebesgue norms, i.e.,*

$$\left\|f^{(n+1)}\right\|_\alpha := \left(\int_r^R \left|f^{(n+1)}\right|^\alpha dt\right)^{\frac{1}{\alpha}}, \ \ \alpha \geqslant 1, \ \ \left\|f^{(n+1)}\right\|_\infty := ess \sup_{t\in[r,R]} \left|f^{(n+1)}(t)\right|.$$

Proof. We start with the Taylor representation with the integral remainder

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

for all $z, a \in (0,\infty)$.

Using the properties of the modulus, we have

$$\begin{aligned}\left|f(z) - f(a) - \sum_{k=1}^{n} \frac{(z-a)^k}{k!} f^{(k)}(a)\right| &\leqslant \frac{1}{n!}\left|\int_a^z |z-t|^n \left|f^{(n+1)}(t)\right| dt\right| \\ &: = M\left(f^{(n+1)}; a, z\right).\end{aligned} \tag{2.4}$$

Now, assume that $a, z \in [r,R]$. Then, obviously

$$\begin{aligned}M\left(f^{(n+1)}; a, z\right) &\leqslant \sup_{t\in[r,R]} \left|f^{(n+1)}(t)\right| \frac{1}{n!}\left|\int_a^z |z-t|^n dt\right| \\ &= \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_\infty |z-a|^{n+1},\end{aligned} \tag{2.5}$$

for all $a, z \in [r,R]$.

Also, by the use of Hölder's integral inequality, we have:

$$\begin{aligned}M\left(f^{(n+1)}; a, z\right) &\leqslant \frac{1}{n!}\left|\int_a^z |x-t|^{n\beta} dt\right|^{\frac{1}{\beta}} \left[\int_a^z \left|f^{(n+1)}(t)\right|^\alpha dt\right]^{\frac{1}{\alpha}} \\ &\leqslant \frac{1}{n!}\left\|f^{(n+1)}\right\|_\alpha \left[\frac{|z-a|^{n\beta+1}}{n\beta+1}\right]^{\frac{1}{\beta}} \\ &= \frac{1}{n!\,(n\beta+1)^{\frac{1}{\beta}}}\left\|f^{(n+1)}\right\|_\alpha |z-a|^{n+\frac{1}{\beta}}, \ \ \alpha > 1, \frac{1}{\alpha}+\frac{1}{\beta} = 1\end{aligned} \tag{2.6}$$

and, obviously,

$$\begin{aligned}M\left(f^{(n+1)}; a, z\right) &\leqslant \frac{1}{n!}|z-a|^n \left|\int_a^z \left|f^{(n+1)}(t)\right| dt\right| \\ &\leqslant \frac{1}{n!}|z-a|^n \left\|f^{(n+1)}\right\|_1\end{aligned} \tag{2.7}$$

for all $z, a \in [r,R]$.

Consequently, by (2.4)–(2.7), we may write (see also [48] for a similar inequality)

$$\left| f(z) - f(a) - \sum_{k=1}^{n} \frac{(z-a)^k}{k!} f^{(k)}(a) \right| \tag{2.8}$$

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

for all $z, a \in [r, R]$.

If in (2.8) choose $z = \frac{p(x)}{q(x)}$, $a = 1$, then we obtain

$$\left| f\left(\frac{p(x)}{q(x)}\right) - f(1) - \sum_{k=1}^{n} \frac{\left(\frac{p(x)}{q(x)} - 1\right)^k}{k!} f^{(k)}(1) \right| \tag{2.9}$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_{\infty} \left| \frac{p(x)}{q(x)} - 1 \right|^{n+1} \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}} \left\| f^{(n+1)} \right\|_{\alpha} \left| \frac{p(x)}{q(x)} - 1 \right|^{n+\frac{1}{\beta}} \\ \frac{1}{n!} \left\| f^{(n+1)} \right\|_{1} \left| \frac{p(x)}{q(x)} - 1 \right|^{n}, \end{cases}$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_{\infty} (R-r)^{n+1} \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}} \left\| f^{(n+1)} \right\|_{\alpha} (R-r)^{n+\frac{1}{\beta}} \\ \frac{1}{n!} \left\| f^{(n+1)} \right\|_{1} (R-r)^{n}, \end{cases}$$

for a.e. $x \in \Gamma$.

If we multiply (2.9) by $q(x) \geqslant 0$ and integrate on Γ, and then use the generalised triangle inequality, we may state that

$$\left|I_f(p,q) - f(1) - \sum_{k=1}^{n} \frac{f^{(k)}(1)}{k!} \cdot \int_\Gamma \frac{(p(x)-q(x))^k}{q^{k-1}(x)} d\mu(x)\right|$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!} \left\|f^{(n+1)}\right\|_\infty \int_\Gamma \frac{|p(x)-q(x)|^{n+1}}{q^n(x)} d\mu(x) \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}} \left\|f^{(n+1)}\right\|_\alpha \int_\Gamma \frac{|p(x)-q(x)|^{n+\frac{1}{\beta}}}{q^{n+\frac{1}{\beta}-1}(x)} d\mu(x) \\ \frac{1}{n!} \left\|f^{(n+1)}\right\|_1 \int_\Gamma \frac{|p(x)-q(x)|^n}{q^{n-1}(x)} d\mu(x), \end{cases}$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!} \left\|f^{(n+1)}\right\|_\infty (R-r)^{n+1} \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}} \left\|f^{(n+1)}\right\|_\alpha (R-r)^{n+\frac{1}{\beta}} \\ \frac{1}{n!} \left\|f^{(n+1)}\right\|_1 (R-r)^n, \end{cases}$$

and the inequality (2.1) is proved. □

The following theorem also holds.

THEOREM 2. *Let f be as in Theorem 1. If $p^{(j)}$, $q^{(j)}$ $(j=1,2)$ are probability distributions such that*

$$r \leqslant \frac{p^{(j)}(x)}{q^{(j)}(x)} \leqslant R \text{ a.e. on } \Gamma \text{ and } j=1,2. \tag{2.10}$$

Then

$$r \leqslant \frac{\lambda p^{(1)}(x) + (1-\lambda) p^{(2)}(x)}{\lambda q^{(1)}(x) + (1-\lambda) q^{(2)}(x)} \leqslant R \text{ a.e. on } \Gamma \text{ and } \lambda \in [0,1] \tag{2.11}$$

and we have the inequality, for $f^{(n+1)} \in L_\alpha[r,R]$,

$$\begin{aligned} &\Big| I_f\left(\lambda p^{(1)} + (1-\lambda) p^{(2)}, \lambda q^{(1)} + (1-\lambda) q^{(2)}\right) \\ &- \lambda I_f\left(p^{(1)}, q^{(1)}\right) - (1-\lambda) I_f\left(p^{(2)}, q^{(2)}\right) \\ &- \lambda \sum_{k=1}^{n} \frac{1}{k!} \int_\Gamma \frac{(1-\lambda)^k}{\left[q^{(1)}(x)\right]^{k-1}} (-1)^k \eta^k(x) f^{(k)}\left(\frac{p^{(1)}(x)}{q^{(1)}(x)}\right) d\mu(x) \\ &- (1-\lambda) \sum_{k=1}^{n} \frac{1}{k!} \int_\Gamma \frac{\lambda^k}{\left[q^{(2)}(x)\right]^{k-1}} \eta^k(x) f^{(k)}\left(\frac{p^{(2)}(x)}{q^{(2)}(x)}\right) d\mu(x) \Big| \end{aligned} \tag{2.12}$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{\infty}\lambda\,(1-\lambda)\int_{\Gamma}|\eta\,(x)|^{n+1}\left[\frac{(1-\lambda)^{n}}{\left[q^{(1)}(x)\right]^{n}}+\frac{\lambda^{n}}{\left[q^{(2)}(x)\right]^{n}}\right]d\mu\,(x)\,; \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}}\left\|f^{(n+1)}\right\|_{\alpha}\lambda\,(1-\lambda) \\ \qquad\times\int_{\Gamma}|\eta\,(x)|^{n+\frac{1}{\beta}}\left[\frac{(1-\lambda)^{n+\frac{1}{\beta}-1}}{\left[q^{(1)}(x)\right]^{n+\frac{1}{\beta}-1}}+\frac{\lambda^{n+\frac{1}{\beta}-1}}{\left[q^{(2)}(x)\right]^{n+\frac{1}{\beta}-1}}\right]d\mu\,(x)\,; \\ \frac{1}{n!}\left\|f^{(n+1)}\right\|_{1}\lambda\,(1-\lambda)\int_{\Gamma}|\eta\,(x)|^{n}\left[\frac{(1-\lambda)^{n-1}}{\left[q^{(1)}(x)\right]^{n-1}}+\frac{\lambda^{n-1}}{\left[q^{(2)}(x)\right]^{n-1}}\right]d\mu\,(x)\,, \end{cases}$$

where

$$\eta\,(x)=\eta\left(\lambda,p^{(1)},p^{(2)},q^{(1)},q^{(2)}\right)(x)=\frac{\det\begin{bmatrix} p^{(1)}\,(x) & p^{(2)}\,(x) \\ q^{(1)}\,(x) & q^{(2)}\,(x)\end{bmatrix}}{\lambda q^{(1)}\,(x)+(1-\lambda)\,q^{(2)}\,(x)},$$

for all $\lambda\in[0,1]$ *and* $x\in\Gamma$.

Proof. We use the inequality (2.8) to write

$$\left|f\left(\frac{\lambda p^{(1)}\,(x)+(1-\lambda)\,p^{(2)}\,(x)}{\lambda q^{(1)}\,(x)+(1-\lambda)\,q^{(2)}\,(x)}\right)-f\left(\frac{p^{(1)}\,(x)}{q^{(1)}\,(x)}\right)\right. \tag{2.13}$$

$$\left.-\sum_{k=1}^{n}\frac{1}{k!}\left(\frac{\lambda p^{(1)}\,(x)+(1-\lambda)\,p^{(2)}\,(x)}{\lambda q^{(1)}\,(x)+(1-\lambda)\,q^{(2)}\,(x)}-\frac{p^{(1)}\,(x)}{q^{(1)}\,(x)}\right)^{k}f^{(k)}\left(\frac{p^{(1)}\,(x)}{q^{(1)}\,(x)}\right)\right|$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{\infty}\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)}-\frac{p^{(1)}(x)}{q^{(1)}(x)}\right|^{n+1} \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}}\left\|f^{(n+1)}\right\|_{\alpha}\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)}-\frac{p^{(1)}(x)}{q^{(1)}(x)}\right|^{n+\frac{1}{\beta}} \\ \frac{1}{n!}\left\|f^{(n+1)}\right\|_{1}\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)}-\frac{p^{(1)}(x)}{q^{(1)}(x)}\right|^{n}, \end{cases}$$

and

$$\left|f\left(\frac{\lambda p^{(1)}\,(x)+(1-\lambda)\,p^{(2)}\,(x)}{\lambda q^{(1)}\,(x)+(1-\lambda)\,q^{(2)}\,(x)}\right)-f\left(\frac{p^{(2)}\,(x)}{q^{(2)}\,(x)}\right)\right. \tag{2.14}$$

$$\left.-\sum_{k=1}^{n}\frac{1}{k!}\left(\frac{\lambda p^{(1)}\,(x)+(1-\lambda)\,p^{(2)}\,(x)}{\lambda q^{(1)}\,(x)+(1-\lambda)\,q^{(2)}\,(x)}-\frac{p^{(2)}\,(x)}{q^{(2)}\,(x)}\right)^{k}f^{(k)}\left(\frac{p^{(2)}\,(x)}{q^{(2)}\,(x)}\right)\right|$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_\infty \left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(2)}(x)}{q^{(2)}(x)}\right|^{n+1} \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}}\left\|f^{(n+1)}\right\|_\alpha \left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(2)}(x)}{q^{(2)}(x)}\right|^{n+\frac{1}{\beta}} \\ \frac{1}{n!}\left\|f^{(n+1)}\right\|_1 \left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(2)}(x)}{q^{(2)}(x)}\right|^{n}. \end{cases}$$

If we multiply (2.13) by $\lambda q^{(1)}(x)$ and (2.14) by $(1-\lambda)q^{(2)}(x)$, add them and apply the triangle inequality, we end up with

$$\begin{aligned} &\Bigg|\left(\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right) f\left(\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)}\right) \\ &-\lambda q^{(1)}(x) f\left(\frac{p^{(1)}(x)}{q^{(1)}(x)}\right) - (1-\lambda) q^{(2)}(x) f\left(\frac{p^{(2)}(x)}{q^{(2)}(x)}\right) \\ &-\lambda q^{(1)}(x)\sum_{k=1}^{n}\frac{1}{k!}\left(\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(1)}(x)}{q^{(1)}(x)}\right)^k f^{(k)}\left(\frac{p^{(1)}(x)}{q^{(1)}(x)}\right) \\ &-(1-\lambda)\sum_{k=1}^{n}\frac{1}{k!}\left(\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(2)}(x)}{q^{(2)}(x)}\right)^k f^{(k)}\left(\frac{p^{(2)}(x)}{q^{(2)}(x)}\right)\Bigg| \end{aligned} \tag{2.15}$$

$$\leqslant \begin{cases} \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_\infty \Bigg[\lambda q^{(1)}(x)\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(1)}(x)}{q^{(1)}(x)}\right|^{n+1} \\ \qquad + (1-\lambda)q^{(2)}(x)\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(2)}(x)}{q^{(2)}(x)}\right|^{n+1}\Bigg] \\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}}\left\|f^{(n+1)}\right\|_\alpha \Bigg[\lambda q^{(1)}(x)\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(1)}(x)}{q^{(1)}(x)}\right|^{n+\frac{1}{\beta}} \\ \qquad + (1-\lambda)q^{(2)}(x)\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(2)}(x)}{q^{(2)}(x)}\right|^{n+\frac{1}{\beta}}\Bigg] \\ \frac{1}{n!}\left\|f^{(n+1)}\right\|_1 \Bigg[\lambda q^{(1)}(x)\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(1)}(x)}{q^{(1)}(x)}\right|^{n} \\ \qquad + (1-\lambda)q^{(2)}(x)\left|\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)} - \frac{p^{(2)}(x)}{q^{(2)}(x)}\right|^{n}\Bigg], \end{cases}$$

which is equivalent to

$$\left|\left(\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right)f\left(\frac{\lambda p^{(1)}(x)+(1-\lambda)p^{(2)}(x)}{\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)}\right)\right. \tag{2.16}$$

$$-\lambda q^{(1)}(x)f\left(\frac{p^{(1)}(x)}{q^{(1)}(x)}\right)-(1-\lambda)q^{(2)}(x)f\left(\frac{p^{(2)}(x)}{q^{(2)}(x)}\right)$$

$$-\lambda\sum_{k=1}^{n}\frac{1}{k!}\frac{(1-\lambda)^k(Det_{2,1})^k}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^k\left[q^{(1)}(x)\right]^{k-1}}f^{(k)}\left(\frac{p^{(1)}(x)}{q^{(1)}(x)}\right)$$

$$\left.-(1-\lambda)\sum_{k=1}^{n}\frac{1}{k!}\frac{\lambda^k(Det_{1,2})^k}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^k\left[q^{(2)}(x)\right]^{k-1}}f^{(k)}\left(\frac{p^{(2)}(x)}{q^{(2)}(x)}\right)\right|$$

$$\leqslant\begin{cases}\frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_\infty\left[\frac{\lambda(1-\lambda)^{n+1}\left|Det_{2,1}\right|^{n+1}}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^{n+1}\left[q^{(1)}(x)\right]^n}\right.\\ \qquad\left.+\frac{(1-\lambda)\lambda^{n+1}\left|Det_{2,1}\right|^{n+1}}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^{n+1}\left[q^{(2)}(x)\right]^n}\right];\\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}}\left\|f^{(n+1)}\right\|_\alpha\left[\frac{\lambda(1-\lambda)^{n+\frac{1}{\beta}}\left|Det_{2,1}\right|^{n+\frac{1}{\beta}}}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^{n+\frac{1}{\beta}}\left[q^{(1)}(x)\right]^{n+\frac{1}{\beta}-1}}\right.\\ \qquad\left.+\frac{(1-\lambda)\lambda^{n+\frac{1}{\beta}}\left|Det_{2,1}\right|^{n+\frac{1}{\beta}}}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^{n+\frac{1}{\beta}}\left[q^{(2)}(x)\right]^{n+\frac{1}{\beta}-1}}\right];\\ \frac{1}{n!}\left\|f^{(n+1)}\right\|_1\left[\frac{\lambda(1-\lambda)^n\left|Det_{2,1}\right|^n}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^n\left[q^{(1)}(x)\right]^{n-1}}\right.\\ \qquad\left.+\frac{(1-\lambda)\lambda^n\left|Det_{2,1}\right|^n}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^n\left[q^{(2)}(x)\right]^{n-1}}\right];\end{cases}$$

$$=\begin{cases}\frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_\infty\frac{\lambda(1-\lambda)\left|Det_{2,1}\right|^{n+1}}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^{n+1}}\left[\frac{(1-\lambda)^n}{\left[q^{(1)}(x)\right]^n}+\frac{\lambda^n}{\left[q^{(2)}(x)\right]^n}\right];\\ \frac{1}{n!(n\beta+1)^{\frac{1}{\beta}}}\left\|f^{(n+1)}\right\|_\alpha\frac{\lambda(1-\lambda)\left|Det_{2,1}\right|^{n+\frac{1}{\beta}}}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^{n+\frac{1}{\beta}}}\\ \qquad\times\left[\frac{(1-\lambda)^{n+\frac{1}{\beta}-1}}{\left[q^{(1)}(x)\right]^{n+\frac{1}{\beta}-1}}+\frac{\lambda^{n+\frac{1}{\beta}-1}}{\left[q^{(2)}(x)\right]^{n+\frac{1}{\beta}-1}}\right];\\ \frac{1}{n!}\left\|f^{(n+1)}\right\|_1\frac{\lambda(1-\lambda)\left|Det_{2,1}\right|^n}{\left[\lambda q^{(1)}(x)+(1-\lambda)q^{(2)}(x)\right]^n}\left[\frac{(1-\lambda)^{n-1}}{\left[q^{(1)}(x)\right]^{n-1}}+\frac{\lambda^{n-1}}{\left[q^{(2)}(x)\right]^{n-1}}\right],\end{cases}$$

where

$$Det_{y,z}=\det\begin{bmatrix}p^{(y)}(x) & p^{(z)}(x)\\ q^{(y)}(x) & q^{(z)}(x)\end{bmatrix},\ \text{where } (y,z)\in\{(1,2),(2,1)\}.$$

Integrating (2.16) over $x \in \Gamma$ and using the generalised triangle inequality, we deduce the desired inequality (2.12). □

In the recent paper [49], S. S. Dragomir proved the following perturbed Taylor's formula which is an improvement of a recent result due to Matić, Pečarić and Ujević from [50]. It is instructive to give the details here for the sake of completeness.

LEMMA 1. *Let $f : I \subseteq \mathbb{R} \to \mathbb{R}$ be such that $f^{(n)}$ is absolutely continuous and $f^{(n+1)} \in L_2(I)$. Then we have*

$$f(z) = \sum_{k=0}^{n} \frac{(z-a)^k}{k!} f^{(k)}(a) + \frac{(z-a)^{n+1}}{(n+1)!}\left[f^{(n)}; a, z\right] + G_n(f; a, z) \tag{2.17}$$

for all $z \in I$, where

$$\left[f^{(n)}; a, z\right] = \frac{f^{(n)}(z) - f^{(n)}(a)}{z-a}$$

and $G_n(f; a, z)$ satisfies the estimate

$$|G_n(f; a, z)| \leqslant \frac{n(z-a)^{n+1}}{(n+1)!\sqrt{2n+1}}\left[\frac{1}{z-a}\left\|f^{(n+1)}\right\|_2^2 - \left(\left[f^{(n)}; a, z\right]\right)^2\right]^{\frac{1}{2}} \tag{2.18}$$

for all $z \geqslant a$.

Proof. Recall Korkine's identity for the mapping h, g

$$\begin{aligned} &\frac{1}{z-a}\int_a^z h(t)g(t)\,dt - \frac{1}{(z-a)^2}\int_a^z h(t)\,dt \cdot \int_a^z g(t)\,dt \\ &= \frac{1}{2(z-a)}\int_a^z\int_a^z (h(t)-h(s))(g(t)-g(s))\,dtds. \end{aligned} \tag{2.19}$$

Using (2.19), we have

$$\begin{aligned} &\int_a^z \frac{(z-t)^n}{n!} f^{(n+1)}(t)\,dt - \frac{1}{z-a}\int_a^z \frac{(z-t)^n}{n!}dt \cdot \int_a^z f^{(n+1)}(t)\,dt \\ &= \frac{1}{2(z-a)}\int_a^z\int_a^z \left(\frac{(z-t)^n-(z-s)^n}{n!}\right)\left(f^{(n+1)}(t) - f^{(n+1)}(s)\right)dtds. \end{aligned}$$

and then, using Taylor's representation (2.3) and the formula (2.17), we may conclude that

$$\begin{aligned} &G_n(f; a, z) \\ &= \frac{1}{2(z-a)}\int_a^z\int_a^z \left[\frac{(z-t)^n-(z-s)^n}{n!}\right]\left(f^{(n+1)}(t) - f^{(n+1)}(s)\right)dtds. \end{aligned} \tag{2.20}$$

Now, using the Cauchy-Buniakowski-Schwartz integral inequality for double integrals, we have

$$\begin{aligned} & \left|G_n(f;a,z)\right| \qquad (2.21)\\ \leqslant\ & \frac{1}{2(z-a)}\left[\int_a^z\int_a^z\left[\frac{(z-t)^n-(z-s)^n}{n!}\right]^2 dtds \right.\\ & \left.\times\int_a^z\int_a^z\left[f^{(n+1)}(t)-f^{(n+1)}(s)\right]^2 dtds\right]^{\frac{1}{2}}. \end{aligned}$$

Elementary calculations show that

$$\frac{1}{2(z-a)^2}\int_a^z\int_a^z\left[\frac{(z-t)^n-(z-s)^n}{n!}\right]^2 dtds=\frac{n^2(z-a)^{2n}}{[(n+1)!]^2(2n+1)}$$

and (see also [50])

$$\begin{aligned} & \frac{1}{2(z-a)^2}\int_a^z\int_a^z\left[f^{(n+1)}(t)-f^{(n+1)}(s)\right]^2 dtds\\ =\ & \frac{1}{z-a}\left\|f^{(n+1)}\right\|_2^2-\left(\left[f^{(n)};a,z\right]\right)^2, \end{aligned}$$

and so, by (2.21), we deduce (2.18). □

Now, by the Grüss inequality, we may state that

$$\begin{aligned} 0 &\leqslant \frac{1}{z-a}\int_a^z\left[f^{(n+1)}(t)\right]^2 dt-\left(\frac{1}{z-a}\int_a^z f^{(n+1)}(t)\,dt\right)^2\\ &\leqslant \frac{1}{4}\left(\Gamma(z)-\gamma(z)\right), \end{aligned}$$

where

$$\gamma(z)\leqslant f^{(n+1)}(t)\leqslant\Gamma(z)\quad\text{for all } t\in[a,z], \qquad (2.22)$$

then, by Lemma 1, we can obtain the result in [49], showing that (2.18) is an improvement on the pre-Grüss result obtained in [50].

COROLLARY 1. *Let $f: I\subseteq\mathbb{R}\to\mathbb{R}$ be such that $f^{(n)}$ is absolutely continuous and $f^{(n+1)}$ is bounded and satisfies (2.22). Then we have the representation (2.17) and the remainder $G_n(f;a,z)$ satisfies the estimate*

$$\left|G_n(f;a,z)\right|\leqslant\frac{n(z-a)^{n+1}}{2(n+1)!\sqrt{2n+1}}\left(\Gamma(z)-\gamma(z)\right) \qquad (2.23)$$

for all $z\geqslant a$.

If $z \leqslant a$, then a similar bound can be stated and so, in general, for any $a \in I$, we have the representation (2.17) and the bounds

$$
\begin{aligned}
& \left|G_n(f;a,z)\right| \qquad (2.24)\\
\leqslant\ & \frac{n\left|z-a\right|^{n+1}}{(n+1)!\sqrt{2n+1}}\left[\frac{\int_a^z \left[f^{(n+1)}(t)\right]^2 dt}{z-a}-\left(\left[f^{(n)};a,z\right]\right)^2\right]^{\frac{1}{2}}\\
\leqslant\ & \frac{n\left|z-a\right|^{n+1}}{2(n+1)!\sqrt{2n+1}}\left(\Gamma(z)-\gamma(z)\right),
\end{aligned}
$$

where

$$\Gamma := \sup_{z\in I} f^{(n+1)}(z) < \infty \ \text{ and } \ \gamma := \inf_{z\in I} f^{(n+1)}(z) > -\infty.$$

In what follows, we use the estimate (2.24).

THEOREM 3. *Let $f:(0,\infty)\to\mathbb{R}$ be n-time differentiable and such that $f^{(n)}$ is absolutely continuous on $[r,R]$, where $0<r\leqslant 1\leqslant R<\infty$. Assume that the probability distributions p,q satisfy the condition*

$$r \leqslant \frac{p(x)}{q(x)} \leqslant R \ \ a.e \text{ on } \Gamma. \qquad (2.25)$$

Then we have the inequalities

$$
\begin{aligned}
& \Bigg| I_f(p,q)-f(1)-\left[\sum_{k=1}^{n}\frac{f^{(k)}(1)}{k!}D_{\chi^k}(p,q)\right] \qquad (2.26)\\
& -\frac{1}{(n+1)!}I_{(\cdot-1)^k f^{(k)}(\cdot)}(p,q)+\frac{f^{(n)}(1)}{(n+1)!}D_{\chi^n}(p,q)\Bigg|\\
\leqslant\ & \frac{n}{(n+1)!\sqrt{2n+1}}B\left(p,q,f^{(n+1)}\right)\\
\leqslant\ & \frac{n(\Phi-\phi)}{(n+1)!\sqrt{2n+1}}D_{|\chi|^{n+1}}(p,q) \leqslant \frac{n(\Phi-\phi)}{(n+1)!\sqrt{2n+1}}(R-r)^{n+1},
\end{aligned}
$$

where

$$\Phi := \sup_{z\in[r,R]} f^{(n+1)}(z) < \infty \ \text{ and } \ \phi := \inf_{z\in[r,R]} f^{(n+1)}(z) > -\infty$$

and

$$B\left(p,q,f^{(n+1)}\right) := I_g(p,q)$$

where

$$g(z)=\left|z-1\right|^{n+1}\left[\frac{1}{z-1}\int_1^z\left[f^{(n+1)}(t)\right]^2 dt-\left(\frac{f^{(n)}(z)-f^{(n)}(1)}{z-1}\right)^2\right]^{\frac{1}{2}}.$$

Proof. Apply the inequality (2.24) for $a = 1$ and $z = \frac{p(x)}{q(x)}$ to obtain

$$\left| f\left(\frac{p(x)}{q(x)}\right) - f(1) - \sum_{k=1}^{n} \frac{f^{(k)}(1)}{k!}\left(\frac{p(x)}{q(x)} - 1\right)^{k} \right.$$
$$\left. - \frac{\left(\frac{p(x)}{q(x)} - 1\right)^{n}}{(n+1)!}\left[f^{(n)}\left(\frac{p(x)}{q(x)}\right) - f^{(n)}(1)\right] \right|$$
$$\leqslant \frac{n\left|\frac{p(x)}{q(x)} - 1\right|^{n}}{(n+1)!\sqrt{2n+1}} \frac{q(x)}{p(x) - q(x)} \int_{1}^{\frac{p(x)}{q(x)}} \left[f^{(n+1)}(t)\right]^{2} dt$$
$$\left[- \left(\frac{f^{(n)}\left(\frac{p(x)}{q(x)}\right) - f^{(n)}(1)}{\frac{p(x)}{q(x)} - 1}\right)^{2}\right]^{\frac{1}{2}} \leqslant \frac{n\left|\frac{p(x)}{q(x)} - 1\right|^{n+1}(\Phi - \phi)}{(n+1)!\sqrt{2n+1}}.$$

If we multiply by $p(x) \geqslant 0$, integrate over $x \in \Gamma$ and use the generalised triangle inequality, we deduce

$$\left| I_{f}(p,q) - f(1) - \sum_{k=1}^{n} \frac{f^{(k)}(1)}{k!} D_{\chi^{k}}(p,q) \right.$$
$$\left. - \frac{1}{(n+1)!}\int_{\Gamma} q(x)\left(\frac{p(x)}{q(x)} - 1\right)^{k} f^{(k)}\left(\frac{p(x)}{q(x)}\right) d\mu(x) + \frac{f^{(n)}(1)}{(n+1)!} D_{\chi^{n}}(p,q) \right|$$

$$\leqslant \frac{n(\Phi - \phi)}{(n+1)!\sqrt{2n+1}} \int_{\Gamma} q(x)\left|\frac{p(x)}{q(x)} - 1\right|^{n+1} \left[\frac{q(x)}{p(x) - q(x)} \int_{1}^{\frac{p(x)}{q(x)}} \left[f^{(n+1)}(t)\right]^{2} dt \right.$$
$$\left. - q^{2}(x) \frac{\left|f^{(n)}\left(\frac{p(x)}{q(x)}\right) - f^{(n)}(1)\right|^{2}}{(p(x) - q(x))^{2}}\right]^{\frac{1}{2}} d\mu(x)$$
$$\leqslant \frac{n(\Phi - \phi)}{2(n+1)!\sqrt{2n+1}} D_{|\chi|^{n+1}}(p,q) \leqslant \frac{n(\Phi - \phi)}{2(n+1)!\sqrt{2n+1}} (R - r)^{n+1}$$

and the theorem is proved. □

3. Some particular inequalities

The following proposition holds.

PROPOSITION 1. *Let p,q be two probability distributions satisfying the condition*

$$0 < r \leqslant \frac{p(x)}{q(x)} \leqslant R < \infty \text{ a.e on } \Gamma. \tag{3.1}$$

Then, for $n \geqslant 1$, we have the inequality

$$\left| D_{KL}(q,p) - \sum_{k=1}^{n} \frac{(-1)^k}{k!} D_{\chi^k}(p,q) \right| \tag{3.2}$$

$$\leqslant \begin{cases} \dfrac{1}{(n+1)\,r^{n+1}} D_{|\chi|^{n+1}}(p,q); \\[2ex] \dfrac{1}{(n\beta+1)^{\frac{1}{\beta}}\left[(n+1)\alpha-1\right]^{\frac{1}{\alpha}}} \left[\dfrac{R^{(n+1)\alpha-1}-r^{(n+1)\alpha-1}}{R^{(n+1)\alpha-1}r^{(n+1)\alpha-1}}\right]^{\frac{1}{\alpha}} D_{|\chi|^{n+\frac{1}{\beta}}}(p,q), \\ \qquad\qquad\qquad \alpha > 1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\[2ex] \dfrac{1}{n}\cdot\dfrac{R^n-r^n}{R^n r^n} D_{|\chi|^n}(p,q); \end{cases}$$

$$\leqslant \begin{cases} \dfrac{1}{(n+1)\,r^{n+1}} (R-r)^{n+1}; \\[2ex] \dfrac{1}{(n\beta+1)^{\frac{1}{\beta}}\left[(n+1)\alpha-1\right]^{\frac{1}{\alpha}}} \left[\dfrac{R^{(n+1)\alpha-1}-r^{(n+1)\alpha-1}}{R^{(n+1)\alpha-1}r^{(n+1)\alpha-1}}\right]^{\frac{1}{\alpha}} (R-r)^{n+\frac{1}{\beta}}, \\ \qquad\qquad\qquad \alpha > 1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\[2ex] \dfrac{1}{n}\cdot\dfrac{R^n-r^n}{R^n r^n} (R-r)^n; \end{cases} \tag{3.3}$$

Proof. Consider the mapping $f(t) = \ln t$. We have

$$\begin{aligned} I_f(p,q) &= \int_\Gamma q(x) f\left(\frac{p(x)}{q(x)}\right) d\mu(x) = \int_\Gamma q(x) \ln\left(\frac{p(x)}{q(x)}\right) d\mu(x) \\ &= -\int_\Gamma q(x) \ln\left(\frac{q(x)}{p(x)}\right) d\mu(x) = -D_{KL}(q,p), \end{aligned}$$

$$f^{(k)}(t) = \frac{(-1)^{k-1}(k-1)!}{t^k}, \quad k \in \mathbb{N}, k \geqslant 1$$

for $\alpha > 1$ and

$$\begin{aligned}\left\|f^{(n+1)}\right\|_{\infty} &:= \sup_{t\in[r,R]}\left|f^{(n+1)}(t)\right| = n!\sup_{t\in[r,R]}\left\{\frac{1}{t^{n+1}}\right\} = \frac{n!}{r^{n+1}};\\ \left\|f^{(n+1)}\right\|_{\alpha} &:= \left(\int_r^R\left|f^{(n+1)}(t)\right|^{\alpha}dt\right)^{\frac{1}{\alpha}} = n!\left[\int_r^R\frac{dt}{t^{(n+1)\alpha}}\right]^{\frac{1}{\alpha}}\\ &= n!\left[\frac{t^{-(n+1)\alpha+1}}{-(n+1)\alpha+1}dt\Big|_r^R\right]^{\frac{1}{\alpha}}\\ &= n!\left[\frac{R^{(n+1)\alpha-1}-r^{(n+1)\alpha-1}}{[(n+1)\alpha-1]R^{(n+1)\alpha-1}r^{(n+1)\alpha-1}}\right]^{\frac{1}{\alpha}}.\end{aligned}$$

Applying Theorem 1 and using the above assumptions, we deduce the desired inequality (3.2). □

The following proposition also holds.

PROPOSITION 2. *Let p,q be as in the above Proposition 1. Then we have the inequality*

$$\left|D_{KL}(q,p)-\sum_{k=2}^{n}\frac{(-1)^k}{(k-1)k}D_{\chi^k}(p,q)\right| \tag{3.4}$$

$$\leqslant \begin{cases}\dfrac{1}{n(n+1)r^{n+1}}D_{|\chi|^{n+1}}(p,q);\\[2ex] \dfrac{1}{n(n\beta+1)^{\frac{1}{\beta}}(n\alpha-1)^{\frac{1}{\alpha}}}\left[\dfrac{R^{n\alpha-1}-r^{n\alpha-1}}{R^{n\alpha-1}r^{n\alpha-1}}\right]^{\frac{1}{\alpha}}D_{|\chi|^{n+\frac{1}{\beta}}}(p,q),\\ \qquad\alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1;\\[2ex] \dfrac{1}{(n-1)n}\cdot\dfrac{R^{n-1}-r^{n-1}}{R^{n-1}r^{n-1}}D_{|\chi|^n}(p,q);\end{cases}$$

$$\leqslant \begin{cases}\dfrac{1}{n(n+1)r^{n+1}}(R-r)^{n+1};\\[2ex] \dfrac{1}{n(n\beta+1)^{\frac{1}{\beta}}(n\alpha-1)^{\frac{1}{\alpha}}}\left[\dfrac{R^{n\alpha-1}-r^{n\alpha-1}}{R^{n\alpha-1}r^{n\alpha-1}}\right]^{\frac{1}{\alpha}}(R-r)^{n+\frac{1}{\beta}},\\ \qquad\alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1;\\[2ex] \dfrac{1}{(n-1)n}\cdot\dfrac{R^{n-1}-r^{n-1}}{R^{n-1}r^{n-1}}(R-r)^n,\end{cases}$$

for $x \geqslant a$.

Proof. Consider the mapping $f(t) = t\ln(t)$. We have

$$\begin{aligned}
I_f(p,q) &= \int_\Gamma q(x) f\left(\frac{p(x)}{q(x)}\right) d\mu(x) = \int_\Gamma q(x)\frac{p(x)}{q(x)}\ln\left(\frac{p(x)}{q(x)}\right) d\mu(x) \\
&= \int_\Gamma p(x)\ln\left(\frac{p(x)}{q(x)}\right) d\mu(x) = D_{KL}(p,q), \\
f^{(1)}(t) &= \ln t + 1, \\
f^{(k)}(t) &= \frac{(-1)^k (k-2)!}{t^{k-1}}, \quad k \geqslant 2 \\
\left\| f^{(n+1)} \right\|_\infty &= \frac{(n-1)!}{r^n}, \\
\left\| f^{(n+1)} \right\|_\alpha &= (n-1)!\left[\frac{R^{n\alpha-1} - r^{n\alpha-1}}{R^{n\alpha-1} r^{n\alpha-1}}\right]^{\frac{1}{\alpha}}, \ \alpha > 1
\end{aligned}$$

and

$$\left\| f^{(n+1)} \right\|_1 = (n-2)!\frac{R^{n-1} - r^{n-1}}{R^{n-1} r^{n-1}}.$$

Applying Theorem 1 for the mapping $f(t) = t\ln t$, we have

$$\left| D_{KL}(p,q) - f^{(1)}(1) D_\chi(p,q) - \sum_{k=2}^{n} \frac{(-1)^k (k-2)!}{k!} D_{\chi^k}(p,q) \right|$$

$$\leqslant \begin{cases}
\frac{1}{(n+1)!}\frac{(n-1)!}{r^n} D_{|\chi|^{n+1}}(p,q); \\[2ex]
\frac{1}{n(n\beta+1)^{\frac{1}{\beta}}} \cdot \frac{(n-1)!}{(n\alpha-1)^{\frac{1}{\alpha}}} \left[\frac{R^{n\alpha-1} - r^{n\alpha-1}}{R^{n\alpha-1} r^{n\alpha-1}}\right]^{\frac{1}{\alpha}} D_{|\chi|^{n+\frac{1}{\beta}}}(p,q); \\[2ex]
\frac{1}{n!}(n-2)!\frac{R^{n-1} - r^{n-1}}{R^{n-1} r^{n-1}} D_{|\chi|^n}(p,q).
\end{cases}$$

That is,

$$\left| D_{KL}(q,p) - \sum_{k=2}^{n} \frac{(-1)^k}{(k-1)k} D_{\chi^k}(p,q) \right|$$

$$\leqslant \begin{cases}
\frac{1}{n(n+1) r^{n+1}} D_{|\chi|^{n+1}}(p,q); \\[2ex]
\frac{1}{n(n\beta+1)^{\frac{1}{\beta}} (n\alpha-1)^{\frac{1}{\alpha}}} \left[\frac{R^{n\alpha-1} - r^{n\alpha-1}}{R^{n\alpha-1} r^{n\alpha-1}}\right]^{\frac{1}{\alpha}} D_{|\chi|^{n+\frac{1}{\beta}}}(p,q); \\[2ex]
\frac{1}{(n-1)n}\frac{R^{n-1} - r^{n-1}}{R^{n-1} r^{n-1}} D_{|\chi|^n}(p,q)
\end{cases}$$

and the first inequality in (3.4) is proved.

The second inequality is obvious and we omit the details. □

REMARK 1. Similar results can be obtained if we apply Theorem 1 for other particular mappings f , generating the Hellinger, Jeffrey's, Bhattacharyya, or other divergence measures as considered in the introduction.

REMARK 2. Perturbed inequalities, such as those stated in Theorem 3, may also be considered. We omit the details.

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 8001,
Victoria, Australia.

e-mail: neil@matilda.vu.edu.au
pc@matilda.vu.edu.au
sofo@matilda.vu.edu.au
sever.dragomir@vu.edu.au
URL: http://rgmia.vu.edu.au/SSDRagomirWeb.html