Approximation of f-divergence measures by using two points Taylor's type representations with integral remainders

Bulletin of the *Transilvania* University of Braşov • Vol 12(61), No. 1 - 2019
Series III: Mathematics, Informatics, Physics, 21-40
https://doi.org/10.31926/but.mif.2019.12.61.1.3

APPROXIMATION OF f -DIVERGENCE MEASURES BY USING TWO POINTS TAYLOR'S TYPE REPRESENTATIONS WITH INTEGRAL REMAINDERS

Silvestru Sever DRAGOMIR[1]

Abstract

In this paper we establish some approximations of the f-divergence measures by the use of two points Taylor's type representations with integral remainders. Some inequalities for particular instances of interest are provided as well.

2000 *Mathematics Subject Classification:* 26D15; 26D10.
Key words: Taylor's formula, Power series, Logarithmic and exponential functions, f-divergence measures, Kullback-Leibler divergence, Hellinger discrimination, χ^2-divergence, Jeffrey's distance.

1 Introduction

One of the important issues in many applications of Probability Theory & Statistics is finding an appropriate measure of *distance* (*difference* or *discrimination*) between two probability distributions.

A number of *divergence measures* have been proposed and extensively studied by: Jeffreys 1946 [26], Kullback-Leibler 1951 [32], Rényi 1961 [39], Ali and Silvey 1966 [1], Csiszár 1967 [11], Havrda-Charvat 1967 [23], Sharma-Mittal 1977 [41], Rao 1982 [38], Burbea-Rao 1982 [8], Kapur 1984 [29], Vajda 1989 [48], Lin 1991 [33], Shioya and Da-te [42] and others, see [36]

These measures have been applied in a variety of fields such as: anthropology [38], genetics [36], finance, economics and political science [40], [45], [46], biology [37], the analysis of contingency tables [22], approximation of probability distributions [10], [30], signal processing [27], [28] and pattern recognition [7], [9].

Assume that a set Ω and the σ-finite measure μ are given. Consider the set of all probability densities on μ to be

$$\mathcal{P} := \left\{ p | p : \Omega \to \mathbb{R},\ p(x) \geq 0,\ \int_{\Omega} p(x)\, d\mu(x) = 1 \right\}.$$

[1]Mathematics, College of Engineering & Science, *Victoria University* Melbourne, Victoria, Australia, e-mail: sever.dragomir@vu.edu.au

The *Kullback-Leibler divergence* [32] is well known among the information divergences. It is defined for $p,\, q \in \mathcal{P}$ as follows:

$$D_{KL}(p,q) := \int_\Omega p(x) \ln \left[\frac{p(x)}{q(x)} \right] d\mu(x), \tag{1}$$

where ln is to base e.

In Information Theory and Statistics, various divergences are applied in addition to the Kullback-Leibler divergence. These are defined for $p,\, q \in \mathcal{P}$ as follows

$$D_v(p,q) := \int_\Omega |p(x) - q(x)|\, d\mu(x), \text{ \textit{variation distance}},$$

$$D_H(p,q) := \int_\Omega \left| \sqrt{p(x)} - \sqrt{q(x)} \right| d\mu(x), \text{ \textit{Hellinger distance} [24]},$$

$$D_{\chi^2}(p,q) := \int_\Omega p(x) \left[\left(\frac{q(x)}{p(x)} \right)^2 - 1 \right] d\mu(x), \ \chi^2\text{-\textit{divergence}},$$

$$D_\alpha(p,q) := \frac{4}{1-\alpha^2} \left[1 - \int_\Omega [p(x)]^{\frac{1-\alpha}{2}} [q(x)]^{\frac{1+\alpha}{2}} d\mu(x) \right], \ \alpha\text{-\textit{divergence}},$$

$$D_B(p,q) := \int_\Omega \sqrt{p(x) q(x)}\, d\mu(x), \text{ \textit{Bhattacharyya distance} [6]},$$

$$D_{Ha}(p,q) := \int_\Omega \frac{2p(x) q(x)}{p(x) + q(x)} d\mu(x), \text{ \textit{Harmonic distance}},$$

$$D_J(p,q) := \int_\Omega [p(x) - q(x)] \ln \left[\frac{p(x)}{q(x)} \right] d\mu(x), \text{ \textit{Jeffrey's distance} [26]},$$

and

$$D_\Delta(p,q) := \int_\Omega \frac{[p(x) - q(x)]^2}{p(x) + q(x)} d\mu(x), \text{ \textit{triangular discrimination} [44]}.$$

For other divergence measures, see the paper [29] by Kapur or the book on line [43] by Taneja.

In 1967, I. Csiszár [12] introduced the concept of f-*divergence* as follows

$$I_f(p,q) := \int_\Omega p(x) f\left[\frac{q(x)}{p(x)} \right] d\mu(x), \tag{2}$$

for $p,\, q \in \mathcal{P}$, where f *is convex on* $(0, \infty)$ and *normalised*, i.e. $f(1) = 0$.

Most of the above distances are particular instances of Csiszár f-divergence. There are also many others which are not in this class (see for example Taneja's book online [43]). For the basic properties of Csiszár f-divergence such as

$$I_f(p,q) \geq 0 \text{ for any } p,\ q \in \mathcal{P},$$

and

$$\mathcal{P} \times \mathcal{P} \ni (p,q) \mapsto I_f(p,q) \text{ is convex},$$

see [12], [13] and [48].

In the recent papers [14], [15] and [16] we obtained several reverses of Jensen's integral inequality. These applied to Csiszár f-divergence produce the following results:

Theorem 1 (Dragomir 2013, [15]). *Let $f:(0,\infty)\to\mathbb{R}$ be a convex function with the property that $f(1)=0$. Assume that $p,\ q\in\mathcal{P}$ and there exists the constants $0<r<1<R<\infty$ such that*

$$r\le\frac{q(x)}{p(x)}\le R \text{ for } \mu\text{-a.e. } x\in\Omega. \tag{3}$$

Then we have the inequalities

$$\begin{aligned}0\le I_f(p,q)&\le\frac{(R-1)(1-r)}{R-r}\sup_{t\in(r,R)}\Psi_f(t;r,R) \\ &\le(R-1)(1-r)\frac{f'_-(R)-f'_+(r)}{R-r} \\ &\le\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right],\end{aligned} \tag{4}$$

and $\Psi_f(\cdot;r,R):(r,R)\to\mathbb{R}$ is defined by

$$\Psi_f(t;r,R)=\frac{f(R)-f(t)}{R-t}-\frac{f(t)-f(r)}{t-r}.$$

We also have the inequality

$$\begin{aligned}0\le I_f(p,q)&\le\frac{1}{4}(R-r)\frac{f(R)(1-r)+f(r)(R-1)}{(R-1)(1-r)} \\ &\le\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right].\end{aligned} \tag{5}$$

and the inequality

$$\begin{aligned}0\le I_f(p,q)&\le 2\max\left\{\frac{R-1}{R-r},\frac{1-r}{R-r}\right\} \\ &\times\left[\frac{f(r)+f(R)}{2}-f\left(\frac{r+R}{2}\right)\right] \\ &\le\frac{1}{2}\max\{R-1,1-r\}\left[f'_-(R)-f'_+(r)\right].\end{aligned} \tag{6}$$

Some bounds in terms of the variation distance are as follows:

Theorem 2 (Dragomir 2016, [16]). *With the assumptions of Theorem 1 we have*

$$\begin{aligned}0\le I_f(p,q)&\le\frac{1}{2}\left[f'_-(R)-f'_+(r)\right]D_v(p,q) \\ &\le\frac{1}{2}\left[f'_-(R)-f'_+(r)\right]\left[D_{\chi^2}(p,q)\right]^{1/2} \\ &\le\frac{1}{4}(R-r)\left[f'_-(R)-f'_+(r)\right].\end{aligned} \tag{7}$$

and

$$0 \le I_f(p,q) \le \frac{1}{2}\left([1,R;f]-[r,1;f]\right) D_v(p,q) \tag{8}$$
$$\le \frac{1}{2}\left([1,R;f]-[r,1;f]\right)\left[D_{\chi^2}(p,q)\right]^{1/2}$$
$$\le \frac{1}{4}\left([1,R;f]-[r,1;f]\right)(R-r),$$

where $[a,b;f]$ *is the divided difference*

$$[a,b;f] := \frac{f(b)-f(a)}{b-a}.$$

Further bounds in terms of the Lebesgue norms of the derivative are embodied in the next theorem:

Theorem 3 (Dragomir 2013, [14])**.** *With the assumptions in Theorem 1 we have*

$$0 \le I_f(p,q) \le B_f(r,R) \tag{9}$$

where

$$B_f(r,R) := \frac{(R-1)\int_r^1 |f'(t)|\,dt + (1-r)\int_1^R |f'(t)|\,dt}{R-r}. \tag{10}$$

Moreover, we have the following bounds for $B_f(r,R)$

$$B_f(r,R) \tag{11}$$
$$\le \begin{cases} \left[\frac{1}{2}+\frac{\left|1-\frac{r+R}{2}\right|}{R-r}\right]\int_r^R |f'(t)|\,dt \\ \\ \frac{1}{2}\int_r^R |f'(t)|\,dt + \frac{1}{2}\left|\int_1^R |f'(t)|\,dt - \int_r^1 |f'(t)|\,dt\right|, \end{cases}$$

and

$$B_f(r,R) \le \frac{(1-r)(R-1)}{R-r}\left[\left\|f'\right\|_{[1,R],\infty} + \left\|f'\right\|_{[r,1],\infty}\right] \tag{12}$$
$$\le \frac{1}{2}(R-r)\frac{\|f'\|_{[1,R],\infty}+\|f'\|_{[r,1],\infty}}{2} \le \frac{1}{2}(R-r)\left\|f'\right\|_{[r,R],\infty}$$

and

$$B_f(r,R) \le \frac{1}{R-r}\left[(1-r)(R-1)^{1/q}\left\|f'\right\|_{[1,R],p}\right. \tag{13}$$
$$\left.+(R-1)(1-r)^{1/q}\left\|f'\right\|_{[r,1],p}\right]$$
$$\le \frac{1}{R-r}\left\|f'\right\|_{[r,R],p}\left[(1-r)^q(R-1)+(R-1)^q(1-r)\right]^{1/q},$$

Motivated by the above results, in this paper we establish some new inequalities for f-divergence measures by employing two points Taylor's type expansions that are presented below. Applications for particular instances of interest are provided as well.

2 Some Preliminary Identities

The following result is well known in the literature as Taylor's formula or Taylor's theorem with the integral remainder.

Lemma 1. *Let $I \subset \mathbb{R}$ be a closed interval, $c \in I$ and let n be a positive integer. If $f : I \longrightarrow \mathbb{C}$ is such that the n-derivative $f^{(n)}$ is absolutely continuous on I, then for each $y \in I$*

$$f(y) = T_n(f; c, y) + R_n(f; c, y), \tag{14}$$

where $T_n(f; c, y)$ is Taylor's polynomial, i.e.,

$$T_n(f; c, y) := \sum_{k=0}^{n} \frac{(y-c)^k}{k!} f^{(k)}(c). \tag{15}$$

Note that $f^{(0)} := f$ and $0! := 1$ and the remainder is given by

$$R_n(f; c, y) := \frac{1}{n!} \int_c^y (y-t)^n f^{(n+1)}(t)\, dt. \tag{16}$$

A simple proof of this lemma can be achieved by mathematical induction using the integration by parts formula in the Lebesgue integral.

For related results, see [2]-[5], [20]-[21], [28], [33]-[35] and [47].

The following identity can be stated:

Lemma 2. *Let $f : I \to \mathbb{C}$ be n-time differentiable function on the interior $\mathring{I}$ of the interval I and $f^{(n)}$, with $n \geq 1$, be locally absolutely continuous on $\mathring{I}$. Then for each distinct t, a, $b \in \mathring{I}$ and for any $\lambda \in \mathbb{R} \setminus \{0, 1\}$ we have the representation*

$$\begin{aligned} f(t) &= (1-\lambda) f(a) + \lambda f(b) \\ &+ \sum_{k=1}^{n} \frac{1}{k!} \left[(1-\lambda) f^{(k)}(a) (t-a)^k + (-1)^k \lambda f^{(k)}(b) (b-t)^k \right] \\ &+ S_{n,\lambda}(t, a, b), \end{aligned} \tag{17}$$

where the remainder $S_{n,\lambda}(t, a, b)$ is given by

$$\begin{aligned} &S_{n,\lambda}(t, a, b) \\ &:= \frac{1}{n!} \left[(1-\lambda)(t-a)^{n+1} \int_0^1 f^{(n+1)}((1-s)a + st)(1-s)^n ds \right. \\ &\left. + (-1)^{n+1} \lambda (b-t)^{n+1} \int_0^1 f^{(n+1)}((1-s)t + sb) s^n ds \right]. \end{aligned} \tag{18}$$

Proof. Using Taylor's representation with the integral remainder (14) we can write the following two identities

$$f(t) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a)(t-a)^k + \frac{1}{n!} \int_a^t f^{(n+1)}(\tau)(t-\tau)^n d\tau \tag{19}$$

and

$$f(t)=\sum_{k=0}^{n}\frac{(-1)^{k}}{k!}f^{(k)}(b)(b-t)^{k}+\frac{(-1)^{n+1}}{n!}\int_{t}^{b}f^{(n+1)}(\tau)(\tau-t)^{n}d\tau \tag{20}$$

for any t, a, $b\in \mathring{I}$.

For any integrable function h on an interval and any distinct numbers c, d in that interval, we have, by the change of variable $\tau=(1-s)c+sd$, $s\in[0,1]$ that

$$\int_{c}^{d}h(\tau)d\tau=(d-c)\int_{0}^{1}h((1-s)c+sd)ds.$$

Therefore,

$$\begin{aligned}&\int_{a}^{t}f^{(n+1)}(\tau)(t-\tau)^{n}d\tau\\&=(t-a)\int_{0}^{1}f^{(n+1)}((1-s)a+st)(t-(1-s)a-st)^{n}ds\\&=(t-a)^{n+1}\int_{0}^{1}f^{(n+1)}((1-s)a+st)(1-s)^{n}ds\end{aligned}$$

and

$$\begin{aligned}&\int_{t}^{b}f^{(n+1)}(\tau)(\tau-t)^{n}d\tau\\&=(b-t)\int_{0}^{1}f^{(n+1)}((1-s)t+sb)((1-s)t+sb-t)^{n}ds\\&=(b-t)^{n+1}\int_{0}^{1}f^{(n+1)}((1-s)t+sb)s^{n}ds.\end{aligned}$$

The identities (19) and (20) can then be written as

$$\begin{aligned}f(t)=&\sum_{k=0}^{n}\frac{1}{k!}f^{(k)}(a)(t-a)^{k}\\&+\frac{1}{n!}(t-a)^{n+1}\int_{0}^{1}f^{(n+1)}((1-s)a+st)(1-s)^{n}ds\end{aligned} \tag{21}$$

and

$$\begin{aligned}f(t)=&\sum_{k=0}^{n}\frac{(-1)^{k}}{k!}f^{(k)}(b)(b-t)^{k}\\&+(-1)^{n+1}\frac{(b-t)^{n+1}}{n!}\int_{0}^{1}f^{(n+1)}((1-s)t+sb)s^{n}ds.\end{aligned} \tag{22}$$

Now, if we multiply (21) with $1-\lambda$ and (22) with λ and add the resulting equalities, a simple calculation yields the desired identity (17). □

Remark 1. *If we take in (17) $t = \frac{a+b}{2}$, with $a, b \in \mathring{I}$, then we have for any $\lambda \in \mathbb{R}\setminus\{0,1\}$ that*

$$f\left(\frac{a+b}{2}\right) = (1-\lambda) f(a) + \lambda f(b) + \sum_{k=1}^{n} \frac{1}{2^k k!}\left[(1-\lambda) f^{(k)}(a) + (-1)^k \lambda f^{(k)}(b)\right](b-a)^k + \tilde{S}_{n,\lambda}(a,b), \tag{23}$$

where the remainder $\tilde{S}_{n,\lambda}(a,b)$ is given by

$$\tilde{S}_{n,\lambda}(a,b) := \frac{1}{2^{n+1} n!}(b-a)^{n+1}\left[(1-\lambda)\int_0^1 f^{(n+1)}\left((1-s)a + s\frac{a+b}{2}\right)(1-s)^n ds + (-1)^{n+1}\lambda \int_0^1 f^{(n+1)}\left((1-s)\frac{a+b}{2} + sb\right) s^n ds\right]. \tag{24}$$

In particular, for $\lambda = \frac{1}{2}$ we have

$$f\left(\frac{a+b}{2}\right) = \frac{f(a)+f(b)}{2} + \sum_{k=1}^{n} \frac{1}{2^{k+1} k!}\left[f^{(k)}(a) + (-1)^k f^{(k)}(b)\right](b-a)^k + \tilde{S}_n(a,b), \tag{25}$$

where the remainder $\tilde{S}_n(a,b)$ is given by

$$\tilde{S}_n(a,b) := \frac{1}{2^{n+2} n!}(b-a)^{n+1}\left[\int_0^1 f^{(n+1)}\left((1-s)a + s\frac{a+b}{2}\right)(1-s)^n ds + (-1)^{n+1}\int_0^1 f^{(n+1)}\left((1-s)\frac{a+b}{2} + sb\right) s^n ds\right]. \tag{26}$$

Remark 2. *The case $n = 0$, namely when the function f is locally absolutely continuous on $\mathring{I}$ with the derivative f' existing almost everywhere in $\mathring{I}$ is important and produces the following simple identities for each distinct t, a, $b \in \mathring{I}$ and $\lambda \in \mathbb{R}\setminus\{0,1\}$*

$$f(t) = (1-\lambda) f(a) + \lambda f(b) + S_\lambda(t,a,b), \tag{27}$$

where the remainder $S_\lambda(t,a,b)$ is given by

$$S_\lambda(t,a,b) := (1-\lambda)(t-a)\int_0^1 f'((1-s)a + st)\,ds - \lambda(b-t)\int_0^1 f'((1-s)t + sb)\,ds. \tag{28}$$

3 Two Points Estimates

Assume that $p,\ q\in\mathcal{P}$ and there exists the constants $0<r<1<R<\infty$ such that

$$r\leq\frac{q(x)}{p(x)}\leq R\text{ for }\mu\text{-a.e. }x\in\Omega. \tag{29}$$

We consider the following divergence measures

$$D_{\chi^k,r}(p,q):=\int_{\Omega.}\frac{(q(x)-rp(x))^k}{p^{k-1}(x)}d\mu(x)\geq 0\text{ for }k\in\mathbb{N}, \tag{30}$$

and

$$D_{R,\chi^k}(p,q):=\int_{\Omega.}\frac{(Rp(x)-q(x))^k}{p^{k-1}(x)}d\mu(x)\geq 0\text{ for }k\in\mathbb{N}. \tag{31}$$

Theorem 4. *Let I be an open interval with $[r,R]\subset I$ as above, $f:I\to\mathbb{C}$ be n-time differentiable function on I and $f^{(n)}$, with $n\geq 1$, be locally absolutely continuous on I. Then for any $p,\ q\in\mathcal{P}$ satisfying the condition (29) we have the representation*

$$\begin{aligned}I_f(p,q)&=(1-\lambda)f(r)+\lambda f(R)\\&+\sum_{k=1}^{n}\frac{1}{k!}\left[(1-\lambda)f^{(k)}(r)D_{\chi^k,r}(p,q)+(-1)^k\lambda f^{(k)}(R)D_{R,\chi^k}(p,q)\right]\\&+R_{f,n}(p,q;\lambda)\end{aligned} \tag{32}$$

and the reminder $R_{f,n}(p,q;\lambda)$ is given by

$$\begin{aligned}R_{f,n}(p,q;\lambda)=\frac{1}{n!}\Bigg[&(1-\lambda)\int_{\Omega}\frac{(q(x)-rp(x))^{n+1}}{p^n(x)}\\&\times\left(\int_0^1 f^{(n+1)}\left((1-s)r+s\frac{q(x)}{p(x)}\right)(1-s)^n ds\right)d\mu(x)\\&+(-1)^{n+1}\lambda\int_{\Omega}\frac{(Rp(x)-q(x))^{n+1}}{p^n(x)}\\&\times\left(\int_0^1 f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)}+sR\right)s^n ds\right)d\mu(x)\Bigg],\end{aligned} \tag{33}$$

where $\lambda\in[0,1]$.

In particular, for $\lambda=\frac{1}{2}$ we get

$$\begin{aligned}I_f(p,q)&=\frac{f(r)+f(R)}{2}\\&+\sum_{k=1}^{n}\frac{1}{k!}\left[\frac{f^{(k)}(r)D_{\chi^k,r}(p,q)+(-1)^k f^{(k)}(R)D_{R,\chi^k}(p,q)}{2}\right]\\&+R_{f,n}(p,q),\end{aligned} \tag{34}$$

where

$$R_{f,n}(p,q) = \frac{1}{2n!}\left[\int_\Omega \frac{(q(x)-rp(x))^{n+1}}{p^n(x)} \times \left(\int_0^1 f^{(n+1)}\left((1-s)r + s\frac{q(x)}{p(x)}\right)(1-s)^n ds\right) d\mu(x) + (-1)^{n+1}\int_\Omega \frac{(Rp(x)-q(x))^{n+1}}{p^n(x)} \times \left(\int_0^1 f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)} + sR\right) s^n ds\right) d\mu(x)\right]. \tag{35}$$

Proof. From Lemma 2 we have, by taking $t = \frac{q(x)}{p(x)}$, $a = r$ and $b = R$, that

$$f\left(\frac{q(x)}{p(x)}\right) = (1-\lambda) f(r) + \lambda f(R) + \sum_{k=1}^{n} \frac{1}{k!}\left[(1-\lambda) f^{(k)}(r)\left(\frac{q(x)}{p(x)} - r\right)^k + (-1)^k \lambda f^{(k)}(R)\left(R - \frac{q(x)}{p(x)}\right)^k\right] + S_{n,\lambda}\left(\frac{q(x)}{p(x)}, r, R\right), \tag{36}$$

where the remainder $S_{n,\lambda}\left(\frac{q(x)}{p(x)}, r, R\right)$ is given by

$$S_{n,\lambda}\left(\frac{q(x)}{p(x)}, r, R\right) = \frac{1}{n!}\left[(1-\lambda)\left(\frac{q(x)}{p(x)} - r\right)^{n+1}\int_0^1 f^{(n+1)}\left((1-s)r + s\frac{q(x)}{p(x)}\right)(1-s)^n ds + (-1)^{n+1}\lambda\left(R - \frac{q(x)}{p(x)}\right)^{n+1}\int_0^1 f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)} + sR\right)s^n ds\right], \tag{37}$$

where $x \in \Omega$.

If we multiply (36) by $p(x)$ and integrate on Ω we get

$$\int_\Omega p(x) f\left(\frac{q(x)}{p(x)}\right) d\mu(x) = [(1-\lambda) f(r) + \lambda f(R)]\int_\Omega p(x)\, d\mu(x) + \sum_{k=1}^{n}\frac{1}{k!}\left[(1-\lambda) f^{(k)}(r)\int_\Omega \frac{(q(x)-rp(x))^k}{p^{k-1}(x)} d\mu(x) + (-1)^k \lambda f^{(k)}(R)\int_\Omega \frac{(Rp(x)-q(x))^k}{p^{k-1}(x)} d\mu(x)\right] + R_{f,n}(p,q;\lambda), \tag{38}$$

where

$$\begin{aligned}R_{f,n}(p,q;\lambda) &= \int_\Omega p(x)\, S_{n,\lambda}\left(\frac{q(x)}{p(x)}, r, R\right) d\mu(x) \qquad (39)\\ &= \frac{1}{n!}\left[(1-\lambda)\int_\Omega p(x)\left(\frac{q(x)}{p(x)} - r\right)^{n+1}\right.\\ &\times \left(\int_0^1 f^{(n+1)}\left((1-s)\, r + s\frac{q(x)}{p(x)}\right)(1-s)^n\, ds\right) d\mu(x)\\ &+ (-1)^{n+1}\lambda \int_\Omega p(x)\left(R - \frac{q(x)}{p(x)}\right)^{n+1}\\ &\times \left.\left(\int_0^1 f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)} + sR\right) s^n ds\right) d\mu(x)\right],\end{aligned}$$

for $\lambda \in [0,1]$.

This proves the representations (32) and (33). □

Corollary 1. *With the assumptions of Theorem 4 and if $f^{(n+1)} \in L_\infty[r,R]$, then we have the following bounds for the reminder*

$$\begin{aligned}&|R_{f,n}(p,q;\lambda)| \qquad (40)\\ &\le \frac{1}{(n+1)!}\left[(1-\lambda)\int_\Omega \frac{(q(x) - rp(x))^{n+1}}{p^n(x)}\left\|f^{(n+1)}\right\|_{\left[r,\frac{q(x)}{p(x)}\right],\infty} d\mu(x)\right.\\ &\left.+\lambda\int_\Omega \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)}\left\|f^{(n+1)}\right\|_{\left[\frac{q(x)}{p(x)},R\right],\infty} d\mu(x)\right]\\ &\le \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}\left[(1-\lambda)\, D_{\chi^{n+1},r}(p,q) + \lambda D_{R,\chi^{n+1}}(p,q)\right]\\ &\le \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}(R-r)^{n+1}\end{aligned}$$

for any $\lambda \in [0,1]$, and, in particular, for $\lambda = \frac{1}{2}$

$$\begin{aligned}|R_{f,n}(p,q)| &\le \frac{1}{2(n+1)!}\left[\int_\Omega \frac{(q(x) - rp(x))^{n+1}}{p^n(x)}\left\|f^{(n+1)}\right\|_{\left[r,\frac{q(x)}{p(x)}\right],\infty} d\mu(x)\right. \qquad (41)\\ &\left.+ \int_\Omega \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)}\left\|f^{(n+1)}\right\|_{\left[\frac{q(x)}{p(x)},R\right],\infty} d\mu(x)\right]\\ &\le \frac{1}{2(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}\left[D_{\chi^{n+1},r}(p,q) + D_{R,\chi^{n+1}}(p,q)\right]\\ &\le \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}(R-r)^{n+1}.\end{aligned}$$

Proof. From (33) we have

$$\begin{aligned}
|R_{f,n}(p,q;\lambda)| &\le \frac{1}{n!}\left[(1-\lambda)\int_\Omega \frac{(q(x)-rp(x))^{n+1}}{p^n(x)}\right. \qquad (42)\\
&\times\left|\int_0^1 f^{(n+1)}\left((1-s)r+s\frac{q(x)}{p(x)}\right)(1-s)^n ds\right| d\mu(x)\\
&+\lambda\int_\Omega \frac{(Rp(x)-q(x))^{n+1}}{p^n(x)}\\
&\times\left.\left|\int_0^1 f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)}+sR\right)s^n ds\right| d\mu(x)\right]\\
&\le \frac{1}{n!}\left[(1-\lambda)\int_\Omega \frac{(q(x)-rp(x))^{n+1}}{p^n(x)}\right.\\
&\times\left(\int_0^1 \left|f^{(n+1)}\left((1-s)r+s\frac{q(x)}{p(x)}\right)\right|(1-s)^n ds\right) d\mu(x)\\
&+\lambda\int_\Omega \frac{(Rp(x)-q(x))^{n+1}}{p^n(x)}\\
&\times\left.\int_0^1 \left|f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)}+sR\right)\right| s^n ds d\mu(x)\right]\\
&= K_n(p,q;\lambda)
\end{aligned}$$

for any $\lambda\in[0,1]$.

We have

$$\begin{aligned}
&\int_0^1 \left|f^{(n+1)}\left((1-s)r+s\frac{q(x)}{p(x)}\right)\right|(1-s)^n ds\\
&\qquad\le \operatorname{essup}_{s\in[0,1]}\left|f^{(n+1)}\left((1-s)r+s\frac{q(x)}{p(x)}\right)\right|\int_0^1(1-s)^n ds\\
&\qquad\qquad= \frac{1}{n+1}\left\|f^{(n+1)}\right\|_{\left[r,\frac{q(x)}{p(x)}\right],\infty}\le\frac{1}{n+1}\left\|f^{(n+1)}\right\|_{[r,R],\infty}
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^1 \left|f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)}+sR\right)\right| s^n ds\\
&\qquad\le \operatorname{essup}_{s\in[0,1]}\left|f^{(n+1)}\left((1-s)\frac{q(x)}{p(x)}+sR\right)\right|\int_0^1 s^n ds\\
&\qquad\qquad= \frac{1}{n+1}\left\|f^{(n+1)}\right\|_{\left[\frac{q(x)}{p(x)},R\right],\infty}\le\frac{1}{n+1}\left\|f^{(n+1)}\right\|_{[r,R],\infty}
\end{aligned}$$

for $x\in\Omega$.

Therefore

$$K_n(p,q;\lambda) \le \frac{1}{(n+1)!}\left[(1-\lambda)\int_\Omega \frac{(q(x)-rp(x))^{n+1}}{p^n(x)}\left\|f^{(n+1)}\right\|_{\left[r,\frac{q(x)}{p(x)}\right],\infty} d\mu(x)\right.$$
$$\left.+\lambda\int_\Omega \frac{(Rp(x)-q(x))^{n+1}}{p^n(x)}\left\|f^{(n+1)}\right\|_{\left[\frac{q(x)}{p(x)},R\right],\infty} d\mu(x)\right]$$
$$\le \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}\left[(1-\lambda)\int_\Omega \frac{(q(x)-rp(x))^{n+1}}{p^n(x)}d\mu(x)\right.$$
$$\left.+\lambda\int_\Omega \frac{(Rp(x)-q(x))^{n+1}}{p^n(x)}d\mu(x)\right]$$
$$= \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}\left[(1-\lambda)\int_\Omega p(x)\left(\frac{q(x)}{p(x)}-r\right)^{n+1} d\mu(x)\right.$$
$$\left.+\lambda\int_\Omega p(x)\left(R-\frac{q(x)}{p(x)}\right)^{n+1} d\mu(x)\right]$$
$$\le \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}(R-r)^{n+1}$$
$$\times\left[(1-\lambda)\int_\Omega p(x)\,d\mu(x)+\lambda\int_\Omega p(x)\,d\mu(x)\right]$$
$$= \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],\infty}(R-r)^{n+1},$$

and from (42) we get (40). □

We consider the divergence measures

$$D_{\chi^{n+1+1/s},r}(p,q) := \int_{\Omega.} \frac{(q(x)-rp(x))^{n+1+1/s}}{p^{n+1/s}(x)}d\mu(x) \ge 0 \text{ for } n\in\mathbb{N},\ s>1 \quad (43)$$

and

$$D_{R,\chi^{n+1+1/s}}(p,q) \quad (44)$$
$$:= \int_{\Omega.} \frac{(Rp(x)-q(x))^{n+1+1/s}}{p^{n+1/s}(x)}d\mu(x) \ge 0 \text{ for } n\in\mathbb{N},\ s>1.$$

Corollary 2. *With the assumptions of Theorem 4 and if* $f^{(n+1)} \in L_s[r,R]$, *with*

$s,\ q>1$, and $\frac{1}{s}+\frac{1}{q}=1$, then we have the following bounds for the reminder

$$\begin{aligned}
&\left|R_{f,n}\left(p,q;\lambda\right)\right| \qquad (45)\\
&\leq \frac{1}{(n+1)!}\left[(1-\lambda)\int_{\Omega}\frac{\left(q\left(x\right)-rp\left(x\right)\right)^{n+1+1/s}}{p^{n+1/s}\left(x\right)}\left\|f^{(n+1)}\right\|_{\left[r,\frac{q(x)}{p(x)}\right],s}d\mu\left(x\right)\right.\\
&\left.+\lambda\int_{\Omega}\frac{\left(Rp\left(x\right)-q\left(x\right)\right)^{n+1+1/s}}{p^{n+1/s}\left(x\right)}\left\|f^{(n+1)}\right\|_{\left[\frac{q(x)}{p(x)},R\right],s}d\mu\left(x\right)\right]\\
&\leq \frac{1}{\left(qn+1\right)^{1/q}(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],s}\\
&\times\left[(1-\lambda)D_{\chi^{n+1+1/s},r}\left(p,q\right)+\lambda D_{R,\chi^{n+1+1/s}}\left(p,q\right)\right]\\
&\leq \frac{1}{\left(qn+1\right)^{1/q}(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],s}\left(R-r\right)^{n+1+1/s}
\end{aligned}$$

for any $\lambda\in[0,1]$, and, in particular, for $\lambda=\frac{1}{2}$

$$\begin{aligned}
&\left|R_{f,n}\left(p,q\right)\right| \qquad (46)\\
&\leq \frac{1}{2(n+1)!}\left[\int_{\Omega}\frac{\left(q\left(x\right)-rp\left(x\right)\right)^{n+1+1/s}}{p^{n+1/s}\left(x\right)}\left\|f^{(n+1)}\right\|_{\left[r,\frac{q(x)}{p(x)}\right],s}d\mu\left(x\right)\right.\\
&\left.+\int_{\Omega}\frac{\left(Rp\left(x\right)-q\left(x\right)\right)^{n+1+1/s}}{p^{n+1/s}\left(x\right)}\left\|f^{(n+1)}\right\|_{\left[\frac{q(x)}{p(x)},R\right],s}d\mu\left(x\right)\right]\\
&\leq \frac{1}{2\left(qn+1\right)^{1/q}(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],s}\\
&\times\left[D_{\chi^{n+1+1/s},r}\left(p,q\right)+D_{R,\chi^{n+1+1/s}}\left(p,q\right)\right]\\
&\leq \frac{1}{\left(qn+1\right)^{1/q}(n+1)!}\left\|f^{(n+1)}\right\|_{[r,R],s}\left(R-r\right)^{n+1+1/s}.
\end{aligned}$$

Proof. Using Hölder's integral inequality for $s,\ q>1$ and $\frac{1}{s}+\frac{1}{q}=1$, we have

$$\begin{aligned}
&\int_0^1\left|f^{(n+1)}\left((1-\tau)r+\tau\frac{q\left(x\right)}{p\left(x\right)}\right)\right|(1-\tau)^n d\tau\\
&\leq\left(\int_0^1\left|f^{(n+1)}\left((1-\tau)r+\tau\frac{q\left(x\right)}{p\left(x\right)}\right)\right|^s ds\right)^{1/s}\left(\int_0^1(1-\tau)^{qn}d\tau\right)^{1/q}\\
&=\left(\left(\frac{q\left(x\right)}{p\left(x\right)}-r\right)\int_r^{\frac{q(x)}{p(x)}}\left|f^{(n+1)}\left(u\right)\right|^s du\right)^{1/s}\left(\frac{1}{qn+1}\right)^{1/q}\\
&=\frac{1}{\left(qn+1\right)^{1/q}}\left(\frac{q\left(x\right)}{p\left(x\right)}-r\right)^{1/s}\left\|f^{(n+1)}\right\|_{\left[r,\frac{q(x)}{p(x)}\right],s}\\
&\leq\frac{1}{\left(qn+1\right)^{1/q}}\left(\frac{q\left(x\right)}{p\left(x\right)}-r\right)^{1/s}\left\|f^{(n+1)}\right\|_{[r,R],s}
\end{aligned}$$

and, similarly

$$\int_0^1 \left| f^{(n+1)} \left((1-\tau) \frac{q(x)}{p(x)} + \tau R \right) \right| \tau^n d\tau \leq \frac{1}{(qn+1)^{1/q}} \left(R - \frac{q(x)}{p(x)} \right)^{1/s} \left\| f^{(n+1)} \right\|_{\left[\frac{q(x)}{p(x)}, R\right], s} \leq \frac{1}{(qn+1)^{1/q}} \left(R - \frac{q(x)}{p(x)} \right)^{1/s} \left\| f^{(n+1)} \right\|_{[r,R], s}$$

for $x \in \Omega$.

Therefore,

$$\begin{aligned} &K_n(p,q;\lambda) \\ &\leq \frac{1}{(qn+1)^{1/q}(n+1)!} \\ &\times \left[(1-\lambda) \int_\Omega \frac{(q(x) - rp(x))^{n+1+1/s}}{p^{n+1/s}(x)} \left\| f^{(n+1)} \right\|_{\left[r, \frac{q(x)}{p(x)}\right], s} d\mu(x) \right. \\ &\left. + \lambda \int_\Omega \frac{(Rp(x) - q(x))^{n+1+1/s}}{p^{n+1/s}(x)} \left\| f^{(n+1)} \right\|_{\left[\frac{q(x)}{p(x)}, R\right], s} d\mu(x) \right] \\ &\leq \frac{1}{(qn+1)^{1/q}(n+1)!} \left\| f^{(n+1)} \right\|_{[r,R], s} \\ &\times \left[(1-\lambda) \int_\Omega \frac{(q(x) - rp(x))^{n+1+1/s}}{p^{n+1/s}(x)} d\mu(x) + \lambda \int_\Omega \frac{(Rp(x) - q(x))^{n+1+1/s}}{p^{n+1/s}(x)} d\mu(x) \right] \\ &\leq \frac{1}{(qn+1)^{1/q}(n+1)!} \left\| f^{(n+1)} \right\|_{[r,R], s} \left[(1-\lambda)(R-r)^{n+1+1/s} + \lambda (R-r)^{n+1+1/s} \right] \\ &= \frac{1}{(qn+1)^{1/q}(n+1)!} \left\| f^{(n+1)} \right\|_{[r,R], s} (R-r)^{n+1+1/s}, \end{aligned}$$

which, by (42), produces the desired result (45). □

4 Application for Kullback-Leibler Divergence

Consider the logarithmic function $f(t) = -\ln t$, $t > 0$. Then

$$I_f(p,q) = -\int_\Omega p(x) \ln \left[\frac{q(x)}{p(x)} \right] d\mu(x) = D_{KL}(p,q)$$

for $p, q \in \mathcal{P}$.

We have $f^{(k)}(t) = \frac{(-1)^k (k-1)!}{t^k}$, $k \in \mathbb{N}$, $k \geq 1$ and for $[a,b] \subset (0,\infty)$,

$$\left\| f^{(n+1)} \right\|_{[a,b],\infty} := \sup_{t \in [a,b]} \left| f^{(n+1)}(t) \right| = n! \sup_{t \in [a,b]} \left\{ \frac{1}{t^{n+1}} \right\} = \frac{n!}{a^{n+1}};$$

and for $\alpha \geq 1$

$$\left\| f^{(n+1)} \right\|_{[a,b],\alpha} := \left(\int_a^b \left| f^{(n+1)}(t) \right|^{\alpha} dt \right)^{\frac{1}{\alpha}} = n! \left[\int_a^b \frac{dt}{t^{(n+1)\alpha}} \right]^{\frac{1}{\alpha}}$$
$$= n! \left[\frac{b^{(n+1)\alpha-1} - a^{(n+1)\alpha-1}}{[(n+1)\alpha - 1] b^{(n+1)\alpha-1} a^{(n+1)\alpha-1}} \right]^{\frac{1}{\alpha}}.$$

Assume that $p, q \in \mathcal{P}$ and there exists the constants $0 < r < 1 < R < \infty$ such that

$$r \leq \frac{q(x)}{p(x)} \leq R \text{ for } \mu\text{-a.e. } x \in \Omega.$$

By using Theorem 4 we have

$$\begin{aligned} & D_{KL}(p,q) \\ & = \ln\left[r^{-(1-\lambda)} R^{-\lambda} \right] \\ & + \sum_{k=1}^{n} \frac{1}{k} \left[\frac{(-1)^k (1-\lambda)}{r^k} D_{\chi^k, r}(p,q) + \frac{\lambda}{R^k} D_{R,\chi^k}(p,q) \right] + D_{f,n}(p,q;\lambda) \end{aligned} \tag{47}$$

and the reminder $D_n(p,q;\lambda)$ is given by

$$\begin{aligned} D_n(p,q;\lambda) = (1-\lambda)(-1)^{n+1} & \int_{\Omega} \frac{(q(x) - rp(x))^{n+1}}{p^n(x)} \\ & \times \left(\int_0^1 \frac{(1-s)^n ds}{\left((1-s) r + s\frac{q(x)}{p(x)} \right)^{n+1}} \right) d\mu(x) \\ & + \lambda \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)} \\ & \times \left(\int_0^1 \frac{s^n ds}{\left((1-s)\frac{q(x)}{p(x)} + sR \right)^{n+1}} \right) d\mu(x), \end{aligned} \tag{48}$$

where $\lambda \in [0,1]$.

In particular, for $\lambda = \frac{1}{2}$ we get

$$\begin{aligned} D_{KL}(p,q) = \ln & \left[r^{-1/2} R^{-1/2} \right] \\ & + \frac{1}{2} \sum_{k=1}^{n} \frac{1}{k} \left[\frac{(-1)^k}{r^k} D_{\chi^k, r}(p,q) + \frac{1}{R^k} D_{R,\chi^k}(p,q) \right] + D_{f,n}(p,q) \end{aligned} \tag{49}$$

and the reminder $D_n(p,q)$ is given by

$$\begin{aligned} & D_n(p,q) \quad (50) \\ & = \frac{1}{2}(-1)^{n+1}\int_\Omega \frac{(q(x)-rp(x))^{n+1}}{p^n(x)}\left(\int_0^1 \frac{(1-s)^n\,ds}{\left((1-s)r+s\frac{q(x)}{p(x)}\right)^{n+1}}\right)d\mu(x) \\ & + \frac{1}{2}\int_\Omega \frac{(Rp(x)-q(x))^{n+1}}{p^n(x)}\left(\int_0^1 \frac{s^n ds}{\left((1-s)\frac{q(x)}{p(x)}+sR\right)^{n+1}}\right)d\mu(x). \end{aligned}$$

By Corollary 1 we have

$$\begin{aligned} |D_n(p,q;\lambda)| & \leq \frac{1}{(n+1)r^{n+1}}\left[(1-\lambda)D_{\chi^{n+1},r}(p,q)+\lambda D_{R,\chi^{n+1}}(p,q)\right] \quad (51) \\ & \leq \frac{1}{(n+1)}\left(\frac{R}{r}-1\right)^{n+1} \end{aligned}$$

for any $\lambda\in[0,1]$, and, in particular, for $\lambda=\frac{1}{2}$

$$\begin{aligned} |D_n(p,q)| & \leq \frac{1}{2(n+1)r^{n+1}}\left[D_{\chi^{n+1},r}(p,q)+D_{R,\chi^{n+1}}(p,q)\right] \quad (52) \\ & \leq \frac{1}{(n+1)}\left(\frac{R}{r}-1\right)^{n+1}. \end{aligned}$$

From Corollary 2 we have for s, $q>1$ with $\frac{1}{s}+\frac{1}{q}=1$, that

$$\begin{aligned} & |D_n(p,q;\lambda)| \quad (53) \\ & \leq \frac{1}{(qn+1)^{1/q}(n+1)}\left[\frac{R^{(n+1)s-1}-r^{(n+1)s-1}}{[(n+1)s-1]R^{(n+1)s-1}r^{(n+1)s-1}}\right]^{\frac{1}{s}} \\ & \times\left[(1-\lambda)D_{\chi^{n+1+1/s},r}(p,q)+\lambda D_{R,\chi^{n+1+1/s}}(p,q)\right] \\ & \leq \frac{1}{(qn+1)^{1/q}(n+1)}\left[\frac{R^{(n+1)s-1}-r^{(n+1)s-1}}{[(n+1)s-1]R^{(n+1)s-1}r^{(n+1)s-1}}\right]^{\frac{1}{s}} \\ & \times(R-r)^{n+1+1/s} \end{aligned}$$

for any $\lambda \in [0,1]$, and, in particular, for $\lambda = \frac{1}{2}$

$$\begin{aligned}
&|D_n(p,q)| \qquad (54)\\
&\le \frac{1}{2(qn+1)^{1/q}(n+1)}\left[\frac{R^{(n+1)s-1}-r^{(n+1)s-1}}{[(n+1)s-1]R^{(n+1)s-1}r^{(n+1)s-1}}\right]^{\frac{1}{s}}\\
&\times\left[D_{\chi^{n+1+1/s},r}(p,q)+D_{R,\chi^{n+1+1/s}}(p,q)\right]\\
&\le \frac{1}{(qn+1)^{1/q}(n+1)}\left[\frac{R^{(n+1)s-1}-r^{(n+1)s-1}}{[(n+1)s-1]R^{(n+1)s-1}r^{(n+1)s-1}}\right]^{\frac{1}{s}}\\
&\times(R-r)^{n+1+1/s}.
\end{aligned}$$

References

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

[2] Akkouchi, M., *Improvements of some integral inequalities of H. Gauchman involving Taylor's remainder.* Divulg. Mat. **11** (2003), no. 2, 115–120.

[3] Anastassiou, G. A., *Taylor-Widder representation formulae and Ostrowski, Grüss, integral means and Csiszár type inequalities.* Comput. Math. Appl. **54** (2007), no. 1, 9–23.

[4] Anastassiou, G. A., *Ostrowski type inequalities over balls and shells via a Taylor-Widder formula.* J. Inequal. Pure Appl. Math. **8** (2007), no. 4, Article 106, 13 pp.

[5] Barnett, N. S., Cerone, P., Dragomir, S. S. and Sofo, A., *Approximating Csiszár f-divergence by the use of Taylor's formula with integral remainder*, Math. Ineq. & Appl., **5**(3) (2002), 417-434.

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

[7] Beth Bassat, M., *f-entropies, probability of error and feature selection,* Inform. Control, **39** (1978), 227-242.

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

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

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

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

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

[13] Csiszár, I. and Körner, J., *Information Theory: Coding Theorem for Discrete Memoryless Systems,* Academic Press, New York, 1981.

[14] Dragomir, S. S., *Reverses of the Jensen inequality in terms of the first derivative and applications,* Acta Math. Vietnam. **38** (2013), no. 3, 429–446. Preprint, RGMIA Res. Rep. Coll. **14** (2011), Art. 71, Online: http://rgmia.org/v14.php.

[15] Dragomir, S. S., *Some reverses of the Jensen inequality with applications*, Bull. Aust. Math. Soc. **87** (2013), no. 2, 177–194. Preprint, RGMIA Res. Rep. Coll. **14** (2011), Art. 72, Online: http://rgmia.org/v14.php.

[16] Dragomir, S. S., *A refinement and a divided difference reverse of Jensen's inequality with applications*, Rev. Colombiana Mat. **50** (2016), no. 1, 17–39. Preprint, *RGMIA Res. Rep. Coll.* **14** (2011), Art. 74, Online: http://rgmia.org/v14.php.

[17] Dragomir, S. S., *New estimation of the remainder in Taylor's formula using Grüss' type inequalities and applications.* Math. Inequal. Appl. **2** (1999), no. 2, 183–193.

[18] Dragomir, S. S. and Thompson, H. B., *A two points Taylor's formula for the generalised Riemann integral.* Demonstratio Math. **43** (2010), no. 4, 827–840.

[19] Dragomir, S. S. and Ionescu, N. M., *Some converse of Jensen's inequality and applications.* Rev. Anal. Numér. Théor. Approx. **23** (1994), no. 1, 71–78.

[20] Gauchman, H., *Some integral inequalities involving Taylor's remainder. I.* J. Inequal. Pure Appl. Math. **3** (2002), no. 2, Article 26, 9 pp. (electronic).

[21] Gauchman, H. *Some integral inequalities involving Taylor's remainder. II.* J. Inequal. Pure Appl. Math. **4** (2003), no. 1, Article 1, 5 pp. (electronic).

[22] Gokhale, D. V.and Kullback, S., *Information in Contingency Tables*, New York, Marcel Decker, 1978.

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

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

[25] Hwang, D.-Y., *Improvements of some integral inequalities involving Taylor's remainder.* J. Appl. Math. Comput. **16** (2004), no. 1-2, 151–163.

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

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

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

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

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

[31] Kechriniotis, A. I. and Assimakis, N. D., *Generalizations of the trapezoid inequalities based on a new mean value theorem for the remainder in Taylor's formula.* J. Inequal. Pure Appl. Math. **7** (2006), no. 3, Article 90, 13 pp. (electronic).

[32] Kullback, S. andLeibler, R. A., *On information and sufficiency*, Annals Math. Statist., **22** (1951), 79-86.

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

[34] Liu, Z., *Note on inequalities involving integral Taylor's remainder.* J. Inequal. Pure Appl. Math. **6** (2005), no. 3, Article 72, 6 pp. (electronic).

[35] Liu, W. and Zhang, Q., *Some new error inequalities for a Taylor-like formula.* J. Comput. Anal. Appl. **15** (2013), no. 6, 1158–1164.

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

[37] Pielou, E. C., *Ecological Diversity,* Wiley, New York, 1975.

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

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

[40] Sen, A., *On Economic Inequality,* Oxford University Press, London 1973.

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

[42] Shioya, H. and Da-Te, T., *A generalisation of Lin divergence and the derivative of a new information divergence*, Elec. and Comm. in Japan, **78** (7) (1995), 37-40.

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

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

[45] Theil, H., *Economics and Information Theory*, North-Holland, Amsterdam, 1967.

[46] Theil, H., *Statistical Decomposition Analysis*, North-Holland, Amsterdam, 1972.

[47] Ujević, N., *Error inequalities for a Taylor-like formula.* Cubo **10** (2008), no. 1, 11–18.

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