 APPROXIMATION OF  \( f \)-DIVERGENCE MEASURES BY USING TWO POINTS TAYLOR’S TYPE REPRESENTATIONS WITH INTEGRAL REMAINDEERS

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, \( \chi^2 \)-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 \( \Omega \) and the \( \sigma \)-finite measure \( \mu \) are given. Consider the set of all probability densities on \( \mu \) to be

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

\(^1\) 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:

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

(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

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

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

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

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

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

\[
D_H\alpha(p, q) := \int_{\Omega} \frac{2p(x)q(x)}{p(x) + q(x)} d\mu(x), \text{ Harmonic distance,}
\]

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

and

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

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

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

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

(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

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

Then we have the inequalities

$$0 \leq I_f (p, q) \leq \frac{(R - 1) (1 - r)}{R - r} \sup_{t \in (r,R)} \Psi_f (t; r, R) \quad (4)$$

$$\leq (R - 1) (1 - r) \frac{f'_- (R) - f'_+ (r)}{R - r}$$

$$\leq \frac{1}{4} (R - r) \left[ f'_- (R) - f'_+ (r) \right],$$

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

$$0 \leq I_f (p, q) \leq \frac{1}{4} (R - r) \frac{f (R) (1 - r) + f (r) (R - 1)}{(R - 1) (1 - r)} \quad (5)$$

$$\leq \frac{1}{4} (R - r) \left[ f'_- (R) - f'_+ (r) \right].$$

and the inequality

$$0 \leq I_f (p, q) \leq 2 \max \left\{ \frac{R - 1}{R - r}, \frac{1 - r}{R - r} \right\}$$

$$\times \left[ \frac{f (r) + f (R)}{2} - f \left( \frac{r + R}{2} \right) \right]$$

$$\leq \frac{1}{2} \max \{ R - 1, 1 - r \} \left[ f'_- (R) - f'_+ (r) \right]. \quad (6)$$

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

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

$$0 \leq I_f (p, q) \leq \frac{1}{2} \left[ f'_- (R) - f'_+ (r) \right] D_v (p, q) \quad (7)$$

$$\leq \frac{1}{2} \left[ f'_- (R) - f'_+ (r) \right] \left[ D_{\chi^2} (p, q) \right]^{1/2}$$

$$\leq \frac{1}{4} (R - r) \left[ f'_- (R) - f'_+ (r) \right].$$
and
\[
0 \leq I_f (p, q) \leq \frac{1}{2} ([1, R; f] - [r, 1; f]) D_v (p, q)
\]
\[
\leq \frac{1}{2} ([1, R; f] - [r, 1; f]) [D_x^2 (p, q)]^{1/2}
\]
\[
\leq \frac{1}{4} ([1, R; f] - [r, 1; f]) (R - r),
\]
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
\[
0 \leq I_f (p, q) \leq B_f (r, R)
\]
where
\[
B_f (r, R) := \left(\frac{R - 1}{R - r}\right) \int_r^1 |f' (t)| dt + \int_r^1 |f' (t)| dt.
\]

Moreover, we have the following bounds for \(B_f (r, R)\)
\[
B_f (r, R) \leq \left\{ \frac{1}{2} + \frac{1 - r + R}{R - r} \right\} \int_r^1 |f' (t)| dt
\]
\[
\leq \left\{ \frac{1}{2} \int_r^1 |f' (t)| dt + \int_r^1 |f' (t)| dt - \int_r^1 |f' (t)| dt \right\},
\]
and
\[
B_f (r, R) \leq \frac{(1 - r) (R - 1)}{R - r} \left[ \|f'\|_{1, R], \infty} + \|f'\|_{r, 1], \infty} \right]
\]
\[
\leq \frac{1}{2} (R - r) \|f'\|_{1, R], \infty} + \|f'\|_{r, 1], \infty} \leq \frac{1}{2} (R - r) \|f'\|_{r, R], \infty}
\]
and
\[
B_f (r, R) \leq \frac{1}{R - r} \left[ (1 - r) (R - 1)^{1/q} \|f'\|_{1, R], p} \right.
\]
\[
\left. + (R - 1) (1 - r)^{1/q} \|f'\|_{r, 1], p} \right]
\]
\[
\leq \frac{1}{R - r} \|f'\|_{r, R], p} [(1 - r)^q (R - 1) + (R - 1)^q (1 - r)]^{1/q},
\]

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 $f : I \rightarrow \mathbb{C}$ is such that the $n$-derivative $f^{(n)}$ is absolutely continuous on $I$, then for each $y \in I$

$$f (y) = T_n (f; c, y) + R_n (f; c, y),$$

(14)

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

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

(15)

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

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

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

$$f (t) = (1 - \lambda) f (a) + \lambda f (b)$$

(17)

$$+ \sum_{k=1}^{n} \frac{1}{k!} \left[ (1 - \lambda) f^{(k)} (a) (t-a)^k + (-1)^k \lambda f^{(k)} (b) (b-t)^k \right]$$

$$+ S_{n,\lambda} (t, a, b),$$

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

$$S_{n,\lambda} (t, a, b)$$

(18)

$$:= \frac{1}{n!} \left[ \left(1 - \lambda\right) (t-a)^{n+1} \int_{0}^{1} f^{(n+1)} ((1-s)a + st) (1-s)^n ds 

+ (-1)^{n+1} \lambda (b-t)^{n+1} \int_{0}^{1} f^{(n+1)} ((1-s)b + sb) s^n ds \right].$$

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

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

(19)
and
\[ f(t) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} f^{(k)}(b) (b-t)^k + \frac{(-1)^{n+1}}{n!} \int_t^b f^{(n+1)}(\tau) (\tau - t)^n \, d\tau \quad (20) \]
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,
\[ \int_a^t f^{(n+1)}(\tau) (t-\tau)^n \, d\tau = (t-a) \int_0^1 f^{(n+1)}((1-s)a + st) (t-(1-s)a-st)^n \, ds \]
\[ = (t-a)^{n+1} \int_0^1 f^{(n+1)}((1-s)a + st) (1-s)^n \, ds \]
and
\[ \int_t^b f^{(n+1)}(\tau) (\tau-t)^n \, d\tau = (b-t) \int_0^1 f^{(n+1)}((1-s)t + sb) ((1-s)t + sb-t)^n \, ds \]
\[ = (b-t)^{n+1} \int_0^1 f^{(n+1)}((1-s)t + sb) s^n \, ds. \]

The identities (19) and (20) can then be written as
\[ f(t) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a) (t-a)^k \quad (21) \]
\[ + \frac{1}{n!} (t-a)^{n+1} \int_0^1 f^{(n+1)}((1-s)a + st) (1-s)^n \, ds \]
and
\[ f(t) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} f^{(k)}(b) (b-t)^k \quad (22) \]
\[ + (-1)^{n+1} \frac{(b-t)^{n+1}}{n!} \int_0^1 f^{(n+1)}((1-s)t + sb) s^n \, ds. \]

Now, if we multiply (21) with \( 1 - \lambda \) and (22) with \( \lambda \) 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 \( a, b \in \hat{I} \), then we have for any \( \lambda \in \mathbb{R} \setminus \{0,1\} \) that

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

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

\[
\begin{align*}
\tilde{S}_{n,\lambda}(a,b) &:= \frac{1}{2^{n+1}n!} (b-a)^{n+1} \left[ (1-\lambda) \int_{0}^{1} f^{(n+1)} \left( (1-s) a + s \frac{a+b}{2} \right) (1-s)^n ds 
+ (-1)^{n+1} \lambda \int_{0}^{1} f^{(n+1)} \left( (1-s) \frac{a+b}{2} + sb \right) s^n ds \right].
\end{align*}
\]

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

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

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

\[
\begin{align*}
\tilde{S}_{n}(a,b) &:= \frac{1}{2^{n+2}n!} (b-a)^{n+1} \left[ \int_{0}^{1} f^{(n+1)} \left( (1-s) a + s \frac{a+b}{2} \right) (1-s)^n ds 
+ (-1)^{n+1} \int_{0}^{1} f^{(n+1)} \left( (1-s) \frac{a+b}{2} + sb \right) s^n ds \right].
\end{align*}
\]

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

\[
f(t) = (1-\lambda) f(a) + \lambda f(b) + S_{\lambda}(t,a,b),
\]

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

\[
S_{\lambda}(t,a,b) := (1-\lambda) (t-a) \int_{0}^{1} f'((1-s)a+st) ds 
- \lambda (b-t) \int_{0}^{1} f'((1-s)t+sb) ds.
\]
3 Two Points Estimates

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

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

We consider the following divergence measures

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

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

**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

\[
I_f (p, q) = (1 - \lambda) f(r) + \lambda f(R)
\]

\[
+ \sum_{k=1}^{n} \frac{1}{k!} \left[ (1 - \lambda) f^{(k)}(r) D_{\chi^k, r} (p, q) + (-1)^k \lambda f^{(k)}(R) D_{R, \chi^k} (p, q) \right]
\]

\[
+ R_{f,n} (p, q; \lambda)
\]

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

\[
R_{f,n} (p, q; \lambda) = \frac{1}{n!} \left[ (1 - \lambda) \int_{\Omega} \frac{(q(x) - rp(x))^{n+1}}{p^n(x)} d\mu(x) \right]
\]

\[
\times \left( \int_0^1 f^{(n+1)} \left( (1 - s) r + s \frac{q(x)}{p(x)} \right) (1 - s)^n ds \right) d\mu(x)
\]

\[
+ (-1)^{n+1} \lambda \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)} d\mu(x)
\]

\[
\times \left( \int_0^1 f^{(n+1)} \left( (1 - s) \frac{q(x)}{p(x)} + sR \right) s^n ds \right) d\mu(x),
\]

where \( \lambda \in [0, 1] \).

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

\[
I_f (p, q) = \frac{f(r) + f(R)}{2}
\]

\[
+ \sum_{k=1}^{n} \frac{1}{k!} \left[ f^{(k)}(r) D_{\chi^k, r} (p, q) + (-1)^k \frac{f^{(k)}(R)}{2} D_{R, \chi^k} (p, q) \right]
\]

\[
+ R_{f,n} (p, q),
\]
Approximation of $f$-divergence measures

where

$$R_{f,n}(p,q) = \frac{1}{2n!} \left[ \int_{\Omega} (q(x) - rp(x))^{n+1} \right]$$

$$\times \left( \int_{0}^{1} f^{(n+1)} \left( (1-s) r + s \frac{q(x)}{p(x)} \right) (1-s)^n ds \right) d\mu(x)$$

$$+ (-1)^{n+1} \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)}$$

$$\times \left( \int_{0}^{1} f^{(n+1)} \left( (1-s) \frac{q(x)}{p(x)} + R \right) s^n ds \right) d\mu(x) \right].$$

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

$$f \left( \frac{q(x)}{p(x)} \right)$$

$$= (1 - \lambda) f(r) + \lambda f(R)$$

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

$$+ S_{n\lambda} \left( \frac{q(x)}{p(x)}, r, R \right),$$

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

$$S_{n\lambda} \left( \frac{q(x)}{p(x)}, r, R \right)$$

$$= \frac{1}{n!} \left[ (1 - \lambda) \left( \frac{q(x)}{p(x)} - r \right)^{n+1} \int_{0}^{1} f^{(n+1)} \left( (1-s) r + s \frac{q(x)}{p(x)} \right) (1-s)^n ds \right.$$

$$\left. + (-1)^{n+1} \lambda \left( R - \frac{q(x)}{p(x)} \right)^{n+1} \int_{0}^{1} f^{(n+1)} \left( (1-s) \frac{q(x)}{p(x)} + R \right) s^n ds \right],$$

where $x \in \Omega$.

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

$$\int_{\Omega} p(x) f \left( \frac{q(x)}{p(x)} \right) d\mu(x)$$

$$= [(1 - \lambda) f(r) + \lambda f(R)] \int_{\Omega} p(x) d\mu(x)$$

$$+ \sum_{k=1}^{n} \frac{1}{k!} \left[ (1 - \lambda) f^{(k)}(r) \int_{\Omega} \frac{(q(x) - rp(x))^k}{p^{k-1}(x)} d\mu(x) \right.$$}

$$\left. + (-1)^k \lambda f^{(k)}(R) \int_{\Omega} \frac{(Rp(x) - q(x))^k}{p^{k-1}(x)} d\mu(x) \right] + R_{f,n}(p,q;\lambda),$$
where

$$R_{f,n} (p,q;\lambda) = \int_{\Omega} p(x) S_{n,\lambda} \left( \frac{q(x)}{p(x)}, r, R \right) d\mu(x)$$

(39)

$$\leq \frac{1}{(n+1)!} \left[ (1-\lambda) \int_{\Omega} \left( q(x) - rp(x) \right)^{n+1} \left\| f^{(n+1)} \right\|_{[r, q(x), p(x)]}, \infty} d\mu(x) \\
+ \lambda \int_{\Omega} \left( Rp(x) - q(x) \right)^{n+1} \left\| f^{(n+1)} \right\|_{[q(x), p(x), R], \infty} d\mu(x) \\
\leq \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_{[r, R], \infty} \left( 1 - \lambda \right) D_{\chi^{n+1}, r} (p, q) + \lambda D_{R, \chi^{n+1}} (p, q) \\
\leq \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_{[r, R], \infty} (R - r)^{n+1}$$

(40)

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

$$\left| R_{f,n} (p,q) \right| \leq \frac{1}{2(n+1)!} \left[ \int_{\Omega} \left( q(x) - rp(x) \right)^{n+1} \left\| f^{(n+1)} \right\|_{[r, q(x), p(x)]}, \infty} d\mu(x) \\
+ \int_{\Omega} \left( Rp(x) - q(x) \right)^{n+1} \left\| f^{(n+1)} \right\|_{[q(x), p(x), R], \infty} d\mu(x) \\
\leq \frac{1}{2(n+1)!} \left\| f^{(n+1)} \right\|_{[r, R], \infty} \left( D_{\chi^{n+1}, r} (p, q) + D_{R, \chi^{n+1}} (p, q) \right) \\
\leq \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_{[r, R], \infty} (R - r)^{n+1}.$$
Approximation of $f$-divergence measures

Proof. From (33) we have

\[
|R_{f,n}(p, q; \lambda)| \leq \frac{1}{n!} \left[ (1 - \lambda) \int_{\Omega} \frac{(q(x) - rp(x))^{n+1}}{p^n(x)} \right]
\times \left| \int_{0}^{1} f^{(n+1)} \left( (1 - s) r + s \frac{q(x)}{p(x)} \right) (1 - s)^n ds \right| d\mu(x)
+ \lambda \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)} \times \left| \int_{0}^{1} f^{(n+1)} \left( (1 - s) \frac{q(x)}{p(x)} + sR \right) s^n ds \right| d\mu(x)
\leq \frac{1}{n!} \left[ (1 - \lambda) \int_{\Omega} \frac{(q(x) - rp(x))^{n+1}}{p^n(x)} \right]
\times \left( \int_{0}^{1} \left| f^{(n+1)} \left( (1 - s) r + s \frac{q(x)}{p(x)} \right) (1 - s)^n ds \right| d\mu(x) \right)
+ \lambda \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)} \times \left| \int_{0}^{1} f^{(n+1)} \left( (1 - s) \frac{q(x)}{p(x)} + sR \right) s^n ds \right| d\mu(x)
= K_n(p, q; \lambda)
\]

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

We have

\[
\int_{0}^{1} \left| f^{(n+1)} \left( (1 - s) r + s \frac{q(x)}{p(x)} \right) (1 - s)^n ds \right|
\leq \sup_{s \in [0,1]} \left| f^{(n+1)} \left( (1 - s) r + s \frac{q(x)}{p(x)} \right) \right| \int_{0}^{1} (1 - s)^n ds
= \frac{1}{n+1} \left\| f^{(n+1)} \right\|_{[r, q(x) / p(x)], \infty} \leq \frac{1}{n+1} \left\| f^{(n+1)} \right\|_{[r, R], \infty}
\]

and

\[
\int_{0}^{1} \left| f^{(n+1)} \left( (1 - s) \frac{q(x)}{p(x)} + sR \right) s^n ds \right|
\leq \sup_{s \in [0,1]} \left| f^{(n+1)} \left( (1 - s) \frac{q(x)}{p(x)} + sR \right) \right| \int_{0}^{1} s^n ds
= \frac{1}{n+1} \left\| f^{(n+1)} \right\|_{[q(x) / p(x), R], \infty} \leq \frac{1}{n+1} \left\| f^{(n+1)} \right\|_{[r, R], \infty}
\]

for $x \in \Omega$. 
Therefore

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

and from (42) we get (40). \(\square\)

We consider the divergence measures

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

and

\[ D_{R, X^{n+1+1/s}} (p, q) \quad (44) \]

\[ := \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1+1/s}}{p^{n+1/s}(x)} \, d\mu(x) \geq 0 \text{ for } n \in \mathbb{N}, s > 1. \]

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

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

$$|R_{f,n}(p,q;\lambda)|$$

(45)

$$\leq \frac{1}{(n+1)!} \left[ (1-\lambda) \int_{\Omega} \frac{(q(x) - r\mu(x))^{n+1/s}}{p^{n+1/s}(x)} \|f^{(n+1)}\|_{[r,\frac{q(x)}{p(x)}]_s} d\mu(x)\right.$$  

$$+ \lambda \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1/s}}{p^{n+1/s}(x)} \|f^{(n+1)}\|_{[\frac{q(x)}{p(x)},R]_s} d\mu(x)$$

$$\leq \frac{1}{(qn+1)^{1/q}(n+1)!} \|f^{(n+1)}\|_{[r,R],s} \times [(1-\lambda) D_{\chi_{n+1/s},r}(p,q) + \lambda D_{R,\chi^{n+1/s}}(p,q)]$$

(46)

$$\leq \frac{1}{(qn+1)^{1/q}(n+1)!} \|f^{(n+1)}\|_{[r,R],s} (R-r)^{n+1/s}.$$  

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

$$|R_{f,n}(p,q)|$$


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

$$\int_0^1 \left| f^{(n+1)} \left( (1-\tau) r + \frac{\tau q(x)}{p(x)} \right) \right|^s (1-\tau)^n d\tau$$

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

$$= \left( \frac{(q(x)}{p(x)} - r \right) \int_r^{\frac{q(x)}{p(x)}} \left| f^{(n+1)}(u) \right|^s du \left( \frac{1}{qn+1} \right)^{1/q}$$

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

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

\[ \int_0^1 \left| f^{(n+1)} \left( (1 - \tau) \frac{q(x)}{p(x)} + \tau R \right) \right| \tau^n d\tau \]

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

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

for \( x \in \Omega \).

Therefore,

\[ K_n (p, q; \lambda) \]

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

\[ \times \left[ 1 - \lambda \right] \int_{\Omega} \left( \frac{q(x) - rp(x)}{p^{n+1/s}(x)} \right)^{n+1+1/s} d\mu(x) + \lambda \int_{\Omega} \left( \frac{Rp(x) - q(x)}{p^{n+1/s}(x)} \right)^{n+1+1/s} d\mu(x) \]

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

\[ = \frac{1}{(qn + 1)^{1/q}} \left( R - r \right)^{n+1+1/s}, \]

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

\[ \square \]

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 \( [a, b] \subset (0, \infty) \),

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

and for $\alpha \geq 1$

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

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

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

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

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

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

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

$$
D_{KL} (p, q) = \ln \left[ r^{-1/2} R^{-1/2} \right]
+ \frac{1}{2} \sum_{k=1}^n \frac{1}{k} \left[ \frac{(-1)^k}{r^k} D_{\chi^k, r} (p, q) + \frac{1}{R^k} D_{R, \chi^k} (p, q) \right] + D_{f,n} (p, q)
$$
and the reminder \( D_n (p, q) \) is given by

\[
D_n (p, q) = \frac{1}{2} (-1)^{n+1} \int_{\Omega} \frac{(q(x) - rp(x))^{n+1}}{p^n(x)} \left( \int_0^1 \frac{(1 - s)^n ds}{(1 - s) r + s \frac{q(x)}{p(x)}} \right)^{n+1} d\mu(x)
\]

\[
+ \frac{1}{2} \int_{\Omega} \frac{(Rp(x) - q(x))^{n+1}}{p^n(x)} \left( \int_0^1 \frac{s^n ds}{(1 - s) \frac{q(x)}{p(x)} + s R} \right)^{n+1} d\mu(x).
\]

By Corollary 1 we have

\[
|D_n (p, q; \lambda)| \leq \frac{1}{(n+1) R^{n+1}} \left[ (1 - \lambda) D_{\chi^{n+1, r}} (p, q) + \lambda D_{R, \chi^{n+1}} (p, q) \right]
\]

(51)

\[
\leq \frac{1}{(n+1)} \left( \frac{R}{r} - 1 \right)^{n+1}
\]

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

\[
|D_n (p, q)| \leq \frac{1}{2 (n+1) R^{n+1}} \left[ D_{\chi^{n+1, r}} (p, q) + D_{R, \chi^{n+1}} (p, q) \right]
\]

(52)

\[
\leq \frac{1}{(n+1)} \left( \frac{R}{r} - 1 \right)^{n+1}.
\]

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

\[
|D_n (p, q; \lambda)| \leq \frac{1}{(gn + 1)^{1/q} (n+1)} \left[ \frac{R^{(n+1)s-1} - r^{(n+1)s-1}}{[(n + 1) s - 1] R^{(n+1)s-1} r^{(n+1)s-1}} \right]^{1/2}
\]

\[
\times \left[ (1 - \lambda) D_{\chi^{n+1, r+1/s}} (p, q) + \lambda D_{R, \chi^{n+1+1/s}} (p, q) \right]
\]

(53)

\[
\leq \frac{1}{(gn + 1)^{1/q} (n+1)} \left[ \frac{R^{(n+1)s-1} - r^{(n+1)s-1}}{[(n + 1) s - 1] R^{(n+1)s-1} r^{(n+1)s-1}} \right]^{1/2}
\]

\[
\times (R - r)^{n+1+1/s}
\]
Approximation of $f$-divergence measures

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

\[
|D_n (p, q)| \leq \frac{1}{2 (qn + 1)^{1/q} (n + 1)} \left[ \frac{R^{(n+1)s-1} - r^{(n+1)s-1}}{[(n + 1)s - 1] R^{(n+1)s-1}r^{(n+1)s-1}} \right]^{1/2} \\
\times \left[ D_{\chi^{n+1/s},r} (p, q) + D_{R,\chi^{n+1/s}} (p, q) \right] \\
\leq \frac{1}{(qn + 1)^{1/q} (n + 1)} \left[ \frac{R^{(n+1)s-1} - r^{(n+1)s-1}}{[(n + 1)s - 1] R^{(n+1)s-1}r^{(n+1)s-1}} \right]^{1/2} \\
\times (R - r)^{n+1/s}.
\]

References


