Physica A 463 (2016) 383–393

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Generalizations of Crooks and Lin's results on Jeffreys–Csiszár and Jensen–Csiszár f-divergences

Paweł Kluza, Marek Niezgoda *

Department of Applied Mathematics and Computer Science, University of Life Sciences in Lublin, Akademicka 13, 20-950 Lublin, Poland

HIGHLIGHTS

- We introduce and investigate new Jeffreys–Csiszár and Jensen–Csiszár f-divergences.
- We derive inequalities for such general divergences.
- We refine results by Crooks, Lin and Popescu et al. on Jeffreys and Jensen–Shannon divergences.

ARTICLE INFO

Article history:
Received 21 March 2016
Received in revised form 13 June 2016
Available online 28 July 2016

Keywords:
Shannon entropy
Tsallis relative entropy
Jeffreys–Csiszár divergence
Jensen–Csiszár divergence
Jensen functional

ABSTRACT

In this paper, Jeffreys–Csiszár and Jensen–Csiszár f-divergences are introduced and studied. Some bounds of Crooks and Lin types for such divergences are provided. To this end the concavity of the composition of monotone functions is discussed. Refinements of the original inequalities by Crooks and Popescu et al. are given.

1. Introduction

We begin with some notation.

Given a discrete probability distribution $\mathbf{p} = (p_1, \ldots, p_n)$ the *Shannon entropy* [1,2] is defined by

$$H_1(\mathbf{p}) = \sum_{i=1}^{n} p_i \ln \frac{1}{p_i} \tag{1}$$

with the convention $0 \ln \frac{1}{0} = 0$.

The *q-logarithm function* $\ln_q$ for $q \geq 0, q \neq 1$, is defined by (see Ref. [1, p. 388])

$$\ln_q x = \frac{x^{1-q} - 1}{1-q} \quad \text{for } x > 0.$$

The *q-exponential* $\exp_q$ is the inverse of the q-logarithm function $\ln_q$ given by (see Ref. [2, p. 281])

$$\exp_q t = [1 + (1-q)t]^{\frac{1}{1-q}} \quad \text{for } 1 + (1-q)t > 0.$$

* Corresponding author.
E-mail addresses: pawel.kluza@up.lublin.pl (P. Kluza), marek.niezgoda@up.lublin.pl (M. Niezgoda).

http://dx.doi.org/10.1016/j.physa.2016.07.062

For $q \geq 0$ and $q \neq 1$, the *Tsallis entropy* [3,1,2] is given by

$$H_q(\mathbf{p}) = \sum_{i=1}^{n} p_i \ln_q \frac{1}{p_i}. \tag{2}$$

Given two discrete probability distributions $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{r} = (r_1, \ldots, r_n)$, the *Tsallis relative entropy* [1,2] is defined by

$$D_q(\mathbf{p}|\mathbf{r}) = -\sum_{i=1}^{n} p_i \ln_q \left(\frac{r_i}{p_i}\right) \quad \text{for } q \geq 0 \text{ and } q \neq 1. \tag{3}$$

For $q = 1$,

$$D_1(\mathbf{p}|\mathbf{r}) = -\sum_{i=1}^{n} p_i \ln \left(\frac{r_i}{p_i}\right) \tag{4}$$

is the *Kullback–Leibler information (relative entropy)* [1].

The *Jeffreys divergence* [4,1] is defined by

$$J_1(\mathbf{p}|\mathbf{r}) = D_1(\mathbf{p}|\mathbf{r}) + D_1(\mathbf{r}|\mathbf{p}), \tag{5}$$

and the *Jensen–Shannon divergence* [5,1] is defined by

$$JS_1(\mathbf{p}|\mathbf{r}) = \frac{1}{2} D_1\left(\mathbf{p} \middle| \frac{\mathbf{p}+\mathbf{r}}{2}\right) + \frac{1}{2} D_1\left(\mathbf{r} \middle| \frac{\mathbf{p}+\mathbf{r}}{2}\right). \tag{6}$$

The *Jeffreys–Tsallis divergence* [1] is given by

$$J_q(\mathbf{p}|\mathbf{r}) = D_q(\mathbf{p}|\mathbf{r}) + D_q(\mathbf{r}|\mathbf{p}), \tag{7}$$

and the *Jensen–Shannon–Tsallis divergence* [1] is given by

$$JS_q(\mathbf{p}|\mathbf{r}) = \frac{1}{2} D_q\left(\mathbf{p} \middle| \frac{\mathbf{p}+\mathbf{r}}{2}\right) + \frac{1}{2} D_q\left(\mathbf{r} \middle| \frac{\mathbf{p}+\mathbf{r}}{2}\right). \tag{8}$$

Theorem A (*Crooks [6], Lin [7]*). *The following inequality holds:*

$$JS(\mathbf{p}|\mathbf{r}) \leq -\ln \frac{1 + \exp\left(-\frac{1}{2} J(\mathbf{p}|\mathbf{r})\right)}{2} \leq \frac{1}{4} J(\mathbf{p}|\mathbf{r}). \tag{9}$$

Theorem B (*Furuichi and Mitroi [1]*). *The following inequality holds:*

$$JS_r(\mathbf{p}|\mathbf{r}) \leq \min\left\{-\ln_r \frac{1 + \exp_q\left(-\frac{1}{2} J_q(\mathbf{p}|\mathbf{r})\right)}{2}, \frac{1}{4} J_{\frac{1+r}{2}}(\mathbf{p}|\mathbf{r})\right\} \tag{10}$$

for $0 \leq r \leq q$ *and* $1 < q$.

Theorem C (*Popescu et al. [2]*). *The following inequality holds:*

$$JS_r(\mathbf{p}|\mathbf{r}) \leq -\ln_r \frac{1 + \exp_q\left(-\frac{1}{2} J_q(\mathbf{p}|\mathbf{r})\right)}{2} \leq \frac{1}{4} J_q(\mathbf{p}|\mathbf{r}) \tag{11}$$

for $0 \leq r \leq q$ *and* $1 < q$.

Some further discussions on bounds for Jensen–Shannon divergence by using Jeffreys divergence can be found in Ref. [8].

Let $f : I_0 \to \mathbb{R}$ be a convex function on an interval $I_0 \subset \mathbb{R}$, and $\mathbf{p} = (p_1, \ldots, p_n)$ and $\mathbf{r} = (r_1, \ldots, r_n)$ be two discrete probability distributions. The *Csiszár f-divergence* is defined by

$$C_f(\mathbf{p}|\mathbf{r}) = \sum_{i=1}^{n} p_i f\left(\frac{r_i}{p_i}\right) \tag{12}$$

with the conventions $0f\left(\frac{0}{0}\right) = 0$ and $0f\left(\frac{c}{0}\right) = c \lim_{t\to\infty} \frac{f(t)}{t}$, $c > 0$ (see Ref. [9], cf. also Refs. [10,11]). Some operator versions of the Csiszár f-divergence can be found in Refs. [10,11].

In the present paper we develop the above symmetrization framework by replacing the logarithm based divergences by a family of Csiszár divergences. Therefore we now introduce the following notions.