A REFINEMENT OF JENSEN'S INEQUALITY WITH APPLICATIONS FOR f-DIVERGENCE MEASURES

S.S. DRAGOMIR

Abstract. A refinement of the discrete Jensen's inequality for convex functions defined on a convex subset in linear spaces is given. Application for f-divergence measures including the Kullback-Leibler and Jeffreys divergences are provided as well.

1. Introduction

The Jensen inequality for convex functions plays a crucial role in the Theory of Inequalities due to the fact that other inequalities such as that arithmetic mean-geometric mean inequality, Hölder and Minkowski inequalities, Ky Fan's inequality etc. can be obtained as particular cases of it.

Let C be a convex subset of the linear space X and f a convex set on C. If $\mathbf{p} = (p_1, \dots, p_n)$ is a probability sequence and $\mathbf{x} = (x_1, \dots, x_n) \in C^n$, then

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \le \sum_{i=1}^{n} p_i f(x_i), \tag{1.1}$$

is well known in the literature as Jensen's inequality.

In 1989, J. Pečarić and the author obtained the following refinement of (1.1):

$$\begin{aligned} f\left(\sum_{i=1}^{n} p_i x_i\right) &\le \sum_{i_1,\dots,i_{k+1}=1}^{n} p_{i_1}\dots p_{i_{k+1}} f\left(\frac{x_{i_1}+\dots+x_{i_{k+1}}}{k+1}\right) \\ &\le \sum_{i_1,\dots,i_k=1}^{n} p_{i_1}\dots p_{i_k} f\left(\frac{x_{i_1}+\dots+x_{i_k}}{k}\right) \\ &\le \dots \le \sum_{i=1}^{n} p_i f(x_i), \end{aligned} \tag{1.2}$$

for $k \ge 1$ and $\mathbf{p}$,$\mathbf{x}$ as above.

Date: 14 August, 2007.

2000 *Mathematics Subject Classification.* 26D15, 94A17.

Key words and phrases. Convex functions, Jensen's inequality, f-divergence measures, Kullback-Leibler divergence, Jeffreys divergence.

If $q_1, \ldots, q_k \geq 0$ with $\sum_{j=1}^{k} q_j = 1$, then the following refinement obtained in 1994 by the author [6] also holds:

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i_1,\ldots,i_k=1}^{n} p_{i_1} \ldots p_{i_k} f\left(\frac{x_{i_1}+\cdots+x_{i_k}}{k}\right) \leq \sum_{i_1,\ldots,i_k=1}^{n} p_{i_1} \ldots p_{i_k} f\left(q_1 x_{i_1}+\cdots+q_k x_{i_k}\right) \leq \sum_{i=1}^{n} p_i f(x_i), \tag{1.3}$$

where $1 \leq k \leq n$ and $\mathbf{p}$, $\mathbf{x}$ are as above.

For other refinements and applications related to Ky Fan's inequality, the arithmetic mean-geometric mean inequality, the generalised triangle inequality etc., see [3]-[8].

The main aim of the present paper is to establish a different refinement of the Jensen inequality for convex functions defined on linear spaces. Natural applications for the generalised triangle inequality in normed spaces and for the arithmetic mean-geometric mean inequality for positive numbers are given. Further applications for f-divergence measures of Csiszár with particular instances for the total variation distance, χ^2-divergence, Kullback-Leibler and Jeffreys divergences are provided as well.

2. General Results

The following result may be stated.

Theorem 1. *Let $f : C \to \mathbb{R}$ be a convex function on the convex subset C of the linear space X, $x_i \in C$, $p_i > 0$, $i \in \{1, \ldots, n\}$ with $\sum_{i=1}^{n} p_i = 1$. Then*

$$\begin{aligned} f\left(\sum_{j=1}^{n} p_j x_j\right) &\leq \min_{k\in\{1,\ldots,n\}} \left[(1-p_k) f\left(\frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k}\right) + p_k f(x_k)\right] \\ &\leq \frac{1}{n}\left[\sum_{k=1}^{n} (1-p_k) f\left(\frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k}\right) + \sum_{k=1}^{n} p_k f(x_k)\right] \\ &\leq \max_{k\in\{1,\ldots,n\}} \left[(1-p_k) f\left(\frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k}\right) + p_k f(x_k)\right] \\ &\leq \sum_{j=1}^{n} p_j f(x_j). \end{aligned} \tag{2.1}$$

In particular,

$$\begin{aligned} f\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right) &\leq \frac{1}{n} \min_{k\in\{1,\ldots,n\}} \left[(n-1) f\left(\frac{\sum_{j=1}^{n} x_j - x_k}{n-1}\right) + f(x_k)\right] \\ &\leq \frac{1}{n^2}\left[(n-1)\sum_{k=1}^{n} f\left(\frac{\sum_{j=1}^{n} x_j - x_k}{n-1}\right) + \sum_{k=1}^{n} f(x_k)\right] \end{aligned} \tag{2.2}$$

$$\leq \frac{1}{n} \max_{k\in\{1,\dots,n\}} \left[(n-1) f\left(\frac{\sum_{j=1}^{n} x_j - x_k}{n-1} \right) + f(x_k) \right]$$
$$\leq \frac{1}{n} \sum_{j=1}^{n} f(x_j).$$

Proof. For any $k \in \{1, \dots, n\}$, we have

$$\sum_{j=1}^{n} p_j x_j - p_k x_k = \sum_{\substack{j=1\\ j\neq k}}^{n} p_j x_j = \frac{\sum\limits_{\substack{j=1\\ j\neq k}}^{n} p_j}{\sum\limits_{\substack{j=1\\ j\neq k}}^{n} p_j} \sum_{\substack{j=1\\ j\neq k}}^{n} p_j x_j = (1-p_k) \cdot \frac{1}{\sum\limits_{\substack{j=1\\ j\neq k}}^{n} p_j} \sum_{\substack{j=1\\ j\neq k}}^{n} p_j x_j$$

which implies that

$$\frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k} = \frac{1}{\sum\limits_{\substack{j=1\\ j\neq k}}^{n} p_j} \sum_{\substack{j=1\\ j\neq k}}^{n} p_j x_j \in C \tag{2.3}$$

for each $k \in \{1, \dots, n\}$, since the right side of (2.3) is a convex combination of the elements $x_j \in C$, $j \in \{1, \dots, n\} \setminus \{k\}$.

Taking the function f on (2.3) and applying the Jensen inequality, we get successively

$$f\left(\frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k} \right) = f\left(\frac{1}{\sum\limits_{\substack{j=1\\ j\neq k}}^{n} p_j} \sum_{\substack{j=1\\ j\neq k}}^{n} p_j x_j \right) \leq \frac{1}{\sum\limits_{\substack{j=1\\ j\neq k}}^{n} p_j} \sum_{\substack{j=1\\ j\neq k}}^{n} p_j f(x_j)$$
$$= \frac{1}{1-p_k} \left[\sum_{j=1}^{n} p_j f(x_j) - p_k f(x_k) \right]$$

for any $k \in \{1, \dots, n\}$, which implies

$$(1-p_k) f\left(\frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k} \right) + p_k f(x_k) \leq \sum_{j=1}^{n} p_j f(x_j) \tag{2.4}$$

for each $k \in \{1, \dots, n\}$.

Utilising the convexity of f, we also have

$$(1-p_k) f\left(\frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k} \right) + p_k f(x_k) \tag{2.5}$$
$$\geq f\left[(1-p_k) \cdot \frac{\sum_{j=1}^{n} p_j x_j - p_k x_k}{1-p_k} + p_k x_k \right] = f\left(\sum_{j=1}^{n} p_j x_j \right)$$

for each $k \in \{1, \dots, n\}$.

Taking the minimum over k in (2.5), utilising the fact that

$$\min_{k\in\{1,\dots,n\}} \alpha_k \leq \frac{1}{n} \sum_{k=1}^{n} \alpha_k \leq \max_{k\in\{1,\dots,n\}} \alpha_k$$

and then taking the maximum in (2.4), we deduce the desired inequality (2.1). ∎

The following corollary may be stated as well.

Corollary 1. *Let* $f : C \to \mathbb{R}$ *be a convex function on the convex subset* C, $0 \in C$, $y_j \in X$ *and* $q_j > 0$, $j \in \{1, \dots, n\}$ *with* $\sum_{j=1}^{n} q_j = 1$. *If* $y_j - \sum_{l=1}^{n} q_l y_l \in C$ *for any* $j \in \{1, \dots, n\}$, *then*

$$
\begin{aligned}
(2.6)\quad & f(0) \\
&\le \min_{k\in\{1,\dots,n\}} \left\{ (1-q_k) f\left[\frac{q_k}{1-q_k} \left(\sum_{l=1}^{n} q_l y_l - y_k \right) \right] + q_k f\left(y_k - \sum_{l=1}^{n} q_l y_l \right) \right\} \\
&\le \frac{1}{n} \left\{ \sum_{l=1}^{n} (1-q_k) f\left[\frac{q_k}{1-q_k} \left(\sum_{l=1}^{n} q_l y_l - y_k \right) \right] + \sum_{l=1}^{n} q_k f\left(y_k - \sum_{l=1}^{n} q_l y_l \right) \right\} \\
&\le \max_{k\in\{1,\dots,n\}} \left\{ (1-q_k) f\left[\frac{q_k}{1-q_k} \left(\sum_{l=1}^{n} q_l y_l - y_k \right) \right] + q_k f\left(y_k - \sum_{l=1}^{n} q_l y_l \right) \right\} \\
&\le \sum_{j=1}^{n} q_j f\left(y_j - \sum_{l=1}^{n} q_l y_l \right).
\end{aligned}
$$

In particular, if $y_j - \frac{1}{n}\sum_{l=1}^{n} y_l \in C$ *for any* $j \in \{1, \dots, n\}$, *then*

$$
\begin{aligned}
(2.7)\quad & f(0) \\
&\le \frac{1}{n} \min_{k\in\{1,\dots,n\}} \left\{ (n-1) f\left[\frac{1}{n-1} \left(\frac{1}{n}\sum_{l=1}^{n} y_l - y_k \right) \right] + f\left(y_k - \frac{1}{n}\sum_{l=1}^{n} y_l \right) \right\} \\
&\le \frac{1}{n^2} \left\{ (n-1) \sum_{k=1}^{n} f\left[\frac{1}{n-1} \left(\frac{1}{n}\sum_{l=1}^{n} y_l - y_k \right) \right] + \sum_{k=1}^{n} f\left(y_k - \frac{1}{n}\sum_{l=1}^{n} y_l \right) \right\} \\
&\le \frac{1}{n} \max_{k\in\{1,\dots,n\}} \left\{ (n-1) f\left[\frac{1}{n-1} \left(\frac{1}{n}\sum_{l=1}^{n} y_l - y_k \right) \right] + f\left(y_k - \frac{1}{n}\sum_{l=1}^{n} y_l \right) \right\} \\
&\le \frac{1}{n} \sum_{j=1}^{n} f\left(y_j - \frac{1}{n}\sum_{l=1}^{n} y_l \right).
\end{aligned}
$$

The above results can be applied for various convex functions related to celebrated inequalities as mentioned in the introduction.

Application 1. If $(X, \|\cdot\|)$ is a normed linear space and $p \ge 1$, then the function $f : X \to \mathbb{R}$, $f(x) = \|x\|^p$ is convex on X. Now, on applying Theorem 1 and

Corollary 1 for $x_i \in X$, $p_i > 0$, $i \in \{1, \dots, n\}$ with $\sum_{i=1}^{n} p_i = 1$, we get:

$$\left\| \sum_{j=1}^{n} p_j x_j \right\|^p \le \min_{k \in \{1,\dots,n\}} \left[(1-p_k)^{1-p} \left\| \sum_{j=1}^{n} p_j x_j - p_k x_k \right\|^p + p_k \|x_k\|^p \right] \tag{2.8}$$
$$\le \frac{1}{n} \left[\sum_{k=1}^{n} (1-p_k)^{1-p} \left\| \sum_{j=1}^{n} p_j x_j - p_k x_k \right\|^p + \sum_{k=1}^{n} p_k \|x_k\|^p \right]$$
$$\le \max_{k \in \{1,\dots,n\}} \left[(1-p_k)^{1-p} \left\| \sum_{j=1}^{n} p_j x_j - p_k x_k \right\|^p + p_k \|x_k\|^p \right]$$
$$\le \sum_{j=1}^{n} p_j \|x_j\|^p$$

and

$$\max_{k \in \{1,\dots,n\}} \left\{ \left[(1-p_k)^{1-p} p_k^p + p_k \right] \left\| x_k - \sum_{l=1}^{n} p_l x_l \right\|^p \right\} \le \sum_{j=1}^{n} p_j \left\| x_k - \sum_{l=1}^{n} p_l x_l \right\|^p . \tag{2.9}$$

In particular, we have the inequality:

$$\left\| \frac{1}{n} \sum_{j=1}^{n} x_j \right\|^p \le \frac{1}{n} \min_{k \in \{1,\dots,n\}} \left[(n-1)^{1-p} \left\| \sum_{j=1}^{n} x_j - x_k \right\|^p + \|x_k\|^p \right] \tag{2.10}$$
$$\le \frac{1}{n^2} \left[(n-1)^{1-p} \sum_{k=1}^{n} \left\| \sum_{j=1}^{n} x_j - x_k \right\|^p + \sum_{k=1}^{n} \|x_k\|^p \right]$$
$$\le \frac{1}{n} \max_{k \in \{1,\dots,n\}} \left[(n-1)^{1-p} \left\| \sum_{j=1}^{n} x_j - x_k \right\|^p + \|x_k\|^p \right]$$
$$\le \frac{1}{n} \sum_{j=1}^{n} \|x_j\|^p$$

and

$$\left[(n-1)^{1-p} + 1 \right] \max_{k \in \{1,\dots,n\}} \left\| x_k - \frac{1}{n} \sum_{l=1}^{n} x_l \right\|^p \le \sum_{j=1}^{n} \left\| x_j - \frac{1}{n} \sum_{l=1}^{n} x_l \right\|^p . \tag{2.11}$$

If we consider the function $h_p(t) := (1-t)^{1-p} t^p + t$, $p \ge 1$, $t \in [0,1)$, then we observe that

$$h_p'(t) = 1 + pt^{p-1} (1-t)^{1-p} + (p-1) t^p (1-t)^{-p},$$

which shows that h_p is strictly increasing on $[0,1)$. Therefore,

$$\min_{k \in \{1,\dots,n\}} \left\{ (1-p_k)^{1-p} p_k^p + p_k \right\} = p_m + (1-p_m)^{1-p} p_m^p,$$

where $p_m := \min_{k\in\{1,\dots,n\}} p_k$. By (2.9), we then obtain the following inequality:

$$\left[p_m + (1-p_m)^{1-p} \cdot p_m^p\right] \max_{k\in\{1,\dots,n\}} \left\| x_k - \sum_{l=1}^n p_l x_l \right\|^p \leq \sum_{j=1}^n p_j \left\| x_j - \sum_{l=1}^n p_l x_l \right\|^p . \tag{2.12}$$

Application 2. Let $x_i, p_i > 0$, $i \in \{1,\dots,n\}$ with $\sum_{i=1}^n p_i = 1$. The following inequality is well known in the literature as the *arithmetic mean-geometric mean* inequality:

$$\sum_{j=1}^n p_j x_j \geq \prod_{j=1}^n x_j^{p_j}. \tag{2.13}$$

The equality case holds in (2.13) iff $x_1 = \cdots = x_n$.

Applying the inequality (2.1) for the convex function $f : (0,\infty) \to \mathbb{R}$, $f(x) = -\ln x$ and performing the necessary computations, we derive the following refinement of (2.13):

$$\begin{aligned} \sum_{i=1}^n p_i x_i &\geq \max_{k\in\{1,\dots,n\}} \left\{ \left(\frac{\sum_{j=1}^n p_j x_j - p_k x_k}{1-p_k} \right)^{1-p_k} \cdot x_k^{p_k} \right\} \\ &\geq \prod_{k=1}^n \left[\left(\frac{\sum_{j=1}^n p_j x_j - p_k x_k}{1-p_k} \right)^{1-p_k} \cdot x_k^{p_k} \right]^{\frac{1}{n}} \\ &\geq \min_{k\in\{1,\dots,n\}} \left\{ \left(\frac{\sum_{j=1}^n p_j x_j - p_k x_k}{1-p_k} \right)^{1-p_k} \cdot x_k^{p_k} \right\} \geq \prod_{i=1}^n x_i^{p_i}. \end{aligned} \tag{2.14}$$

In particular, we have the inequality:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n x_i &\geq \max_{k\in\{1,\dots,n\}} \left\{ \left(\frac{\sum_{j=1}^n x_j - x_k}{n-1} \right)^{\frac{n-1}{n}} \cdot x_k^{\frac{1}{n}} \right\} \\ &\geq \prod_{k=1}^n \left[\left(\frac{\sum_{j=1}^n x_j - x_k}{n-1} \right)^{\frac{n-1}{n}} \cdot x_k^{\frac{1}{n}} \right]^{\frac{1}{n}} \\ &\geq \min_{k\in\{1,\dots,n\}} \left\{ \left(\frac{\sum_{j=1}^n x_j - x_k}{n-1} \right)^{\frac{n-1}{n}} \cdot x_k^{\frac{1}{n}} \right\} \geq \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}. \end{aligned}$$

3. Applications for f-Divergences

Given a convex function $f : [0,\infty) \to \mathbb{R}$, the *$f$-divergence functional*

$$I_f(\mathbf{p}, \mathbf{q}) := \sum_{i=1}^n q_i f\left(\frac{p_i}{q_i}\right), \tag{3.1}$$

where $\mathbf{p} = (p_1, \dots, p_n)$, $\mathbf{q} = (q_1, \dots, q_n)$ are positive sequences was introduces by Csiszár in [1], as a generalised measure of information, a "distance function" on the set of probability distributions $\mathbb{P}^n$. As in [1], we interpret undefined expressions by

$$f(0) = \lim_{t \to 0+} f(t), \qquad 0f\left(\frac{0}{0}\right) = 0,$$
$$0f\left(\frac{a}{0}\right) = \lim_{q \to 0+} f\left(\frac{a}{q}\right) = a \lim_{t \to \infty} \frac{f(t)}{t}, \quad a > 0.$$

The following results were essentially given by Csiszár and Körner [2]:

(i) *If f is convex, then $I_f(\mathbf{p}, \mathbf{q})$ is jointly convex in p and q;*
(ii) *For every $p, q \in R_+^n$, we have*

$$I_f(\mathbf{p}, \mathbf{q}) \geq \sum_{j=1}^{n} q_j f\left(\frac{\sum_{j=1}^n p_j}{\sum_{j=1}^n q_j}\right). \tag{3.2}$$

If f is strictly convex, equality holds in (3.2) iff

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} = \dots = \frac{p_n}{q_n}.$$

If f is normalised, i.e., $f(1) = 0$, then for every $p, q \in R_+^n$ with $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i$, we have the inequality

$$I_f(\mathbf{p}, \mathbf{q}) \geq 0. \tag{3.3}$$

In particular, if $\mathbf{p}, \mathbf{q} \in \mathbb{P}^n$, then (3.3) holds. This is the well-known positive property of the f-divergence.

The following refinement of (3.3) may be stated.

Theorem 2. *For any $\mathbf{p}, \mathbf{q} \in \mathbb{P}^n$, we have the inequalities*

$$\begin{aligned} I_f(\mathbf{p}, \mathbf{q}) &\geq \max_{k \in \{1, \dots, n\}} \left[(1 - q_k) f\left(\frac{1 - p_k}{1 - q_k}\right) + q_k f\left(\frac{p_k}{q_k}\right)\right] \\ &\geq \frac{1}{n}\left[\sum_{k=1}^{n} (1 - q_k) f\left(\frac{1 - p_k}{1 - q_k}\right) + \sum_{k=1}^{n} q_k f\left(\frac{p_k}{q_k}\right)\right] \\ &\geq \min_{k \in \{1, \dots, n\}} \left[(1 - q_k) f\left(\frac{1 - p_k}{1 - q_k}\right) + q_k f\left(\frac{p_k}{q_k}\right)\right] \geq 0, \end{aligned} \tag{3.4}$$

provided $f : [0, \infty) \to \mathbb{R}$ is convex and normalised on $[0, \infty)$.

The proof is obvious by Theorem 1 applied for the convex function $f : [0, \infty) \to \mathbb{R}$ and for the choice $x_i = \frac{p_i}{q_i}$, $i \in \{1, \dots, n\}$ and the probabilities q_i, $i \in \{1, \dots, n\}$.

If we consider a new divergence measure $R_f(\mathbf{p}, \mathbf{q})$ defined by

$$R_f(\mathbf{p}, \mathbf{q}) := \frac{1}{n - 1} \sum_{k=1}^{n} (1 - q_k) f\left(\frac{1 - p_k}{1 - q_k}\right) \tag{3.5}$$

and call it the *reverse $f-$divergence*, we observe that

$$R_f(\mathbf{p}, \mathbf{q}) = I_f(\mathbf{r}, \mathbf{t}) \tag{3.6}$$

with

$$\mathbf{r} = \left(\frac{1 - p_1}{n - 1}, \dots, \frac{1 - p_n}{n - 1}\right), \quad \mathbf{t} = \left(\frac{1 - q_1}{n - 1}, \dots, \frac{1 - q_n}{n - 1}\right) \qquad (n \geq 2).$$

With this notation, we can state the following corollary of the above proposition.

Corollary 2. *For any* $\mathbf{p},\mathbf{q}\in\mathbb{P}^n$, *we have*

$$I_f(\mathbf{p},\mathbf{q})\geq R_f(\mathbf{p},\mathbf{q})\geq 0. \tag{3.7}$$

The proof is obvious by the second inequality in (3.4) and the details are omitted.

In what follows, we point out some particular inequalities for various instances of divergence measures such as: the *total variation distance*, χ^2-*divergence*, *Kullback-Leibler divergence*, *Jeffreys divergence*.

The *total variation distance* is defined by the convex function $f(t)=|t-1|$, $t\in\mathbb{R}$ and given in:

$$V(p,q):=\sum_{j=1}^{n}q_j\left|\frac{p_j}{q_j}-1\right|=\sum_{j=1}^{n}|p_j-q_j|. \tag{3.8}$$

The following improvement of the positivity inequality for the total variation distance can be stated as follows.

Proposition 1. *For any* $\mathbf{p},\mathbf{q}\in\mathbb{P}^n$, *we have the inequality:*

$$V(p,q)\geq 2\max_{k\in\{1,\dots,n\}}|p_k-q_k|\quad(\geq 0). \tag{3.9}$$

The proof follows by the first inequality in (3.7) for $f(t)=|t-1|$, $t\in\mathbb{R}$.

The K. Pearson χ^2-*divergence* is obtained for the convex function $f(t)=(1-t)^2$, $t\in\mathbb{R}$ and given by

$$\chi^2(p,q):=\sum_{j=1}^{n}q_j\left(\frac{p_j}{q_j}-1\right)^2=\sum_{j=1}^{n}\frac{(p_j-q_j)^2}{q_j}. \tag{3.10}$$

Proposition 2. *For any* $\mathbf{p},\mathbf{q}\in\mathbb{P}^n$,

$$\chi^2(p,q)\geq\max_{k\in\{1,\dots,n\}}\left\{\frac{(p_k-q_k)^2}{q_k(1-q_k)}\right\}\geq 4\max_{k\in\{1,\dots,n\}}(p_k-q_k)^2\quad(\geq 0). \tag{3.11}$$

Proof. On applying the first inequality in Theorem 2 for the function $f(t)=(1-t)^2$, $t\in\mathbb{R}$, we get

$$\begin{aligned}\chi^2(p,q)&\geq\max_{k\in\{1,\dots,n\}}\left\{(1-q_k)\left(\frac{1-p_k}{1-q_k}-1\right)^2+q_k\left(\frac{p_k}{q_k}-1\right)^2\right\}\\&=\max_{k\in\{1,\dots,n\}}\left\{\frac{(p_k-q_k)^2}{q_k(1-q_k)}\right\}.\end{aligned}$$

Since

$$q_k(1-q_k)\leq\frac{1}{4}[q_k+(1-q_k)]^2=\frac{1}{4},$$

then

$$\frac{(p_k-q_k)^2}{q_k(1-q_k)}\geq 4(p_k-q_k)^2$$

for each $k\in\{1,\dots,n\}$, which proves the last part of (3.11). ∎

The *Kullback-Leibler divergence* can be obtained for the convex function $f:(0,\infty)\to\mathbb{R}$, $f(t)=t\ln t$ and is defined by

$$KL(p,q):=\sum_{j=1}^{n}q_j\cdot\frac{p_j}{q_j}\ln\left(\frac{p_j}{q_j}\right)=\sum_{j=1}^{n}p_j\ln\left(\frac{p_j}{q_j}\right). \tag{3.12}$$

Proposition 3. *For any* $\mathbf{p},\mathbf{q}\in\mathbb{P}^n$, *we have:*

$$KL(p,q)\geq \ln\left[\max_{k\in\{1,\dots,n\}}\left\{\left(\frac{1-p_k}{1-q_k}\right)^{1-p_k}\cdot\left(\frac{p_k}{q_k}\right)^{p_k}\right\}\right] \tag{3.13}$$
$$\geq \ln\left[\max_{k\in\{1,\dots,n\}}\left\{\frac{(1-q_k)\,q_k}{q_k\,(1-p_k)^2+(1-q_k)\,p_k^2}\right\}\right].$$

Proof. The first inequality is obvious by Theorem 2. Utilising the inequality between the *geometric mean and the harmonic mean,*

$$x^{\alpha}y^{1-\alpha}\geq\frac{1}{\frac{\alpha}{x}+\frac{1-\alpha}{y}},\qquad x,y>0,\ \alpha\in[0,1]$$

we have

$$\left(\frac{1-p_k}{1-q_k}\right)^{1-p_k}\cdot\left(\frac{p_k}{q_k}\right)^{p_k}\geq\frac{1}{(1-p_k)\cdot\left(\frac{1-p_k}{1-q_k}\right)+p_k\cdot\frac{p_k}{q_k}}$$
$$=\frac{(1-q_k)\,q_k}{q_k\,(1-p_k)^2+(1-q_k)\,p_k^2},$$

for any $k\in\{1,\dots,n\}$, which implies the second part of (3.13). ∎

Another divergence measure that is of importance in Information Theory is the *Jeffreys divergence*

$$J(p,q):=\sum_{j=1}^{n}q_j\cdot\left(\frac{p_j}{q_j}-1\right)\ln\left(\frac{p_j}{q_j}\right)=\sum_{j=1}^{n}(p_j-q_j)\ln\left(\frac{p_j}{q_j}\right), \tag{3.14}$$

which is an f-divergence for $f(t)=(t-1)\ln t$, $t>0$.

Proposition 4. *For any* $\mathbf{p},\mathbf{q}\in\mathbb{P}^n$, *we have:*

$$J(p,q)\geq\max_{k\in\{1,\dots,n\}}\left\{(q_k-p_k)\ln\left[\frac{(1-p_k)\,q_k}{(1-q_k)\,p_k}\right]\right\} \tag{3.15}$$
$$\geq\max_{k\in\{1,\dots,n\}}\left[\frac{(q_k-p_k)^2}{p_k+q_k-2p_kq_k}\right]\geq 0.$$

Proof. Writing the first inequality in Theorem 2 for $f(t)=(t-1)\ln t$, we have

$$J(p,q)\geq\max_{k\in\{1,\dots,n\}}\left\{(1-q_k)\left[\left(\frac{1-p_k}{1-q_k}-1\right)\ln\left(\frac{1-p_k}{1-q_k}\right)\right]+q_k\left(\frac{p_k}{q_k}-1\right)\ln\left(\frac{p_k}{q_k}\right)\right\}$$
$$=\max_{k\in\{1,\dots,n\}}\left\{(q_k-p_k)\ln\left(\frac{1-p_k}{1-q_k}\right)-(q_k-p_k)\ln\left(\frac{p_k}{q_k}\right)\right\}$$
$$=\max_{k\in\{1,\dots,n\}}\left\{(q_k-p_k)\ln\left[\frac{(1-p_k)\,q_k}{(1-q_k)\,p_k}\right]\right\},$$

proving the first inequality in (3.15).

Utilising the elementary inequality for positive numbers,

$$\frac{\ln b-\ln a}{b-a}\geq\frac{2}{a+b},\qquad a,b>0$$

we have

$$\begin{aligned}
&(q_k - p_k)\left[\ln\left(\frac{1-p_k}{1-q_k}\right) - \ln\left(\frac{p_k}{q_k}\right)\right] \\
&= (q_k - p_k)\cdot\frac{\ln\left(\frac{1-p_k}{1-q_k}\right) - \ln\left(\frac{p_k}{q_k}\right)}{\frac{1-p_k}{1-q_k} - \frac{p_k}{q_k}}\cdot\left[\frac{1-p_k}{1-q_k} - \frac{p_k}{q_k}\right] \\
&= \frac{(q_k - p_k)^2}{q_k(1-q_k)}\cdot\frac{\ln\left(\frac{1-p_k}{1-q_k}\right) - \ln\left(\frac{p_k}{q_k}\right)}{\frac{1-p_k}{1-q_k} - \frac{p_k}{q_k}} \\
&\geq \frac{(q_k - p_k)^2}{q_k(1-q_k)}\cdot\frac{2}{\frac{1-p_k}{1-q_k} - \frac{p_k}{q_k}} = \frac{2(q_k - p_k)^2}{p_k + q_k - 2p_kq_k} \geq 0,
\end{aligned}$$

for each $k \in \{1, \dots, n\}$, giving the second inequality in (3.15). ∎

References

[1] I. CSISZÁR, Information-type measures of differences of probability distributions and indirect observations, *Studia Sci. Math. Hung.*, **2** (1967), 299-318.

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

[3] S.S. DRAGOMIR, An improvement of Jensen's inequality, *Bull. Math. Soc. Sci. Math. Roumanie*, **34**(**82**) (1990), No. 4, 291-296.

[4] S.S. DRAGOMIR, Some refinements of Ky Fan's inequality, *J. Math. Anal. Appl.*, **163**(2) (1992), 317-321.

[5] S.S. DRAGOMIR, Some refinements of Jensen's inequality, *J. Math. Anal. Appl.*, **168**(2) (1992), 518-522.

[6] S.S. DRAGOMIR, A further improvement of Jensen's inequality, *Tamkang J. Math.*, **25**(1) (1994), 29-36.

[7] S.S. DRAGOMIR, A new improvement of Jensen's inequality, *Indian J. Pure and Appl. Math.*, **26**(10) (1995), 959-968.

[8] S.S. DRAGOMIR, J. PEČARIĆ and L.E. PERSSON, Properties of some functionals related to Jensen's inequality, *Acta Math. Hung.*, **70**(1-2) (1996), 129-143.

[9] J. PEČARIĆ and S.S. DRAGOMIR, A refinements of Jensen inequality and applications, *Studia Univ. Babeş-Bolyai, Mathematica,* **24**(1) (1989), 15-19.

School of Computer Science and Mathematics, Victoria University, PO Box 14428, Melbourne City Mail Centre, VIC, 8001, Australia.

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

URL: `http://rgmia.vu.edu.au/dragomir`