Hindawi Publishing Corporation
Journal of Inequalities and Applications
Volume 2007, Article ID 37359, 8 pages
doi:10.1155/2007/37359

Research Article

On Logarithmic Convexity for Differences of Power Means

Slavko Simic

Received 20 July 2007; Accepted 17 October 2007

Recommended by László Losonczi

We proved a new and precise inequality between the differences of power means. As a consequence, an improvement of Jensen's inequality and a converse of Holder's inequality are obtained. Some applications in probability and information theory are also given.

1. Introduction

Let $\tilde{x}_n = \{x_i\}_1^n$, $\tilde{p}_n = \{p_i\}_1^n$ denote two sequences of positive real numbers with $\sum_1^n p_i = 1$. From Theory of Convex Means (cf. [1–3]), the well-known Jensen's inequality states that for $t < 0$ or $t > 1$,

$$\sum_1^n p_i x_i^t \geq \left(\sum_1^n p_i x_i\right)^t, \tag{1.1}$$

and vice versa for $0 < t < 1$. The equality sign in (1.1) occurs if and only if all members of $\tilde{x}_n$ are equal (cf. [1, page 15]). In this article, we will consider the difference

$$d_t = d_t^{(n)} = d_t^{(n)}(\tilde{x}_n, \tilde{p}_n) := \sum_1^n p_i x_i^t - \left(\sum_1^n p_i x_i\right)^t, \quad t \in \mathbb{R}/\{0,1\}. \tag{1.2}$$

By the above, d_t is identically zero if and only if all members of the sequence $\tilde{x}_n$ are equal; hence this trivial case will be excluded in the sequel. An interesting fact is that there exists an explicit constant $c_{s,t}$, independent of the sequences $\tilde{x}_n$ and $\tilde{p}_n$ such that

$$d_s d_t \geq c_{s,t}\left(d_{(s+t)/2}\right)^2 \tag{1.3}$$

for each $s,t \in \mathbb{R}/\{0,1\}$. More generally, we will prove the following inequality:

$$(\lambda_s)^{t-r} \le (\lambda_r)^{t-s}(\lambda_t)^{s-r}, \qquad -\infty < r < s < t < +\infty, \tag{1.4}$$

where

$$\lambda_t := \frac{d_t}{t(t-1)}, \qquad t \neq 0,1,$$
$$\lambda_0 := \log\left(\sum_1^n p_i x_i\right) - \sum_1^n p_i \log x_i; \qquad \lambda_1 := \sum_1^n p_i x_i \log x_i - \left(\sum_1^n p_i x_i\right) \log \sum_1^n p_i x_i. \tag{1.5}$$

This inequality is very precise. For example ($n = 2$),

$$\lambda_2\lambda_4 - (\lambda_3)^2 = \frac{1}{72}(p_1p_2)^2(1+p_1p_2)(x_1-x_2)^6. \tag{1.6}$$

Remark 1.1. Note that from (1.1) follows $\lambda_t > 0$, $t \neq 0,1$, assuming that not all members of $\tilde{x}_n$ are equal. The same is valid for λ_0 and λ_1. Corresponding integral inequalities will also be given. As a consequence of Theorem 2.2, a whole variety of applications arise. For instance, we obtain a substantial improvement of Jensen's inequality and a converse of Holder's inequality, as well. As an application to probability theory, we give a generalized form of Lyapunov-like inequality for moments of distributions with support on $(0,\infty)$. An inequality between the Kullback-Leibler divergence and Hellinger distance will also be derived.

2. Results

Our main result is contained in the following.

Theorem 2.1. *For $\tilde{p}_n$, $\tilde{x}_n$, d_t defined as above, then*

$$\lambda_t := \frac{d_t}{t(t-1)} \tag{2.1}$$

is log-convex for $t \in I := (-\infty,0)\cup(0,1)\cup(1,+\infty)$. As a consequence, the following general inequality is obtained.

Theorem 2.2. *For $-\infty < r < s < t < +\infty$, then*

$$\lambda_s^{t-r} \le (\lambda_r)^{t-s}(\lambda_t)^{s-r}, \tag{2.2}$$

with

$$\lambda_0 := \log\left(\sum_1^n p_i x_i\right) - \sum_1^n p_i \log x_i,$$
$$\lambda_1 := \sum_1^n p_i x_i \log x_i - \left(\sum_1^n p_i x_i\right) \log\left(\sum_1^n p_i x_i\right). \tag{2.3}$$

Applying standard procedure (cf. [1, page 131]), we pass from finite sums to definite integrals and obtain the following theorem.

Theorem 2.3. *Let $f(x), p(x)$ be nonnegative and integrable functions for $x \in (a,b)$, with $\int_a^b p(x)dx = 1$.Denote*

$$D_s = D_s(a,b,f,p) := \int_a^b p(x)f^s(x)dx - \left(\int_a^b p(x)f(x)dx\right)^s. \tag{2.4}$$

For $0 < r < s < t$, $r,s,t \neq 1$, then

$$\left(\frac{D_s}{s(s-1)}\right)^{t-r} \leq \left(\frac{D_r}{r(r-1)}\right)^{t-s}\left(\frac{D_t}{t(t-1)}\right)^{s-r}. \tag{2.5}$$

3. Applications

Finally, we give some applications of our results in analysis, probability, and information theory. Also, since the involved constants are independent on n, we will write $\sum(\cdot)$ instead of $\sum_1^n(\cdot)$.

3.1. An improvement of Jensen's inequality.

By the inequality (2.2) various improvements of Jensen's inequality (1.1) can be established such as the following proposition.

Proposition 3.1. *There exist*

(i) *for $s > 3$,*

$$\sum p_i x_i^s \geq \left(\sum p_i x_i\right)^s + \binom{s}{2}\left(\frac{d_3}{3d_2}\right)^{s-2} d_2; \tag{3.1}$$

(ii) *for $0 < s < 1$,*

$$\sum p_i x_i^s \leq \left(\sum p_i x_i\right)^s - \frac{s(1-s)}{2}\left(\frac{3d_2}{d_3}\right)^{2-s} d_2, \tag{3.2}$$

where d_2 and d_3 are defined as above.

3.2. A converse of Holder's inequality.

The following converse statement holds.

Proposition 3.2. *Let $\{a_i\}$, $\{b_i\}$, $i = 1,2,\ldots$, be arbitrary sequences of positive real numbers and $1/p + 1/q = 1$, $p > 1$. Then*

$$\begin{aligned} pq&\left[\left(\sum a_i^p\right)^{1/p}\left(\sum b_i^q\right)^{1/q} - \sum a_i b_i\right] \\ &\leq \left(\sum a_i^p \log\frac{a_i^p}{b_i^q} - \left(\sum a_i^p\right)\log\frac{\sum a_i^p}{\sum b_i^q}\right)^{1/p}\left(\sum b_i^q \log\frac{b_i^q}{a_i^p} - \left(\sum b_i^q\right)\log\frac{\sum b_i^q}{\sum a_i^p}\right)^{1/q}. \end{aligned} \tag{3.3}$$

For $0 < p < 1$, the inequality (3.3) is reversed.

3.3. A new moments inequality. Apart from Jensen's inequality, in probability theory is very important Lyapunov moments inequality which asserts that for $0 < m < n < p$,

$$(\mathrm{E}X^n)^{p-m} \leq (\mathrm{E}X^m)^{p-n}(\mathrm{E}X^p)^{n-m}. \tag{3.4}$$

This inequality is valid for any probability law with support on $(0,+\infty)$. A consequence of Theorem 2.2 gives a similar but more precise moments inequality.

Proposition 3.3. *For $1 < m < n < p$ and for any probability distribution P with* $\operatorname{supp} P = (0,+\infty)$, *then*

$$(\mathrm{E}X^n - (\mathrm{E}X)^n)^{p-m} \leq C(m,n,p)(\mathrm{E}X^m - (\mathrm{E}X)^m)^{p-n}(\mathrm{E}X^p - (\mathrm{E}X)^p)^{n-m}, \tag{3.5}$$

where the constant $C(m,n,p)$ is given by

$$C(m,n,p) = \frac{\binom{n}{2}^{p-m}}{\binom{m}{2}^{p-n}\binom{p}{2}^{n-m}}. \tag{3.6}$$

There remains an interesting question: under what conditions on m, n, p is the inequality (3.5) valid for distributions with support on $(-\infty,+\infty)$?

3.4. An inequality on symmetrized divergence. Define probability distributions P and Q of a discrete random variable by

$$P(X = i) = p_i, \quad Q(X = i) = q_i, \quad i = 1,2,\ldots, \quad \sum p_i = \sum q_i = 1. \tag{3.7}$$

Among the other quantities, of importance in information theory, are Kullback-Leibler divergence $D_{\mathrm{KL}}(P\|Q)$ and Hellinger distance $H(P,Q)$, defined to be

$$\begin{aligned} D_{\mathrm{KL}}(P\|Q) &:= \sum p_i \log\frac{p_i}{q_i}, \\ H(P,Q) &:= \sqrt{\sum\left(\sqrt{p_i}-\sqrt{q_i}\right)^2}. \end{aligned} \tag{3.8}$$

The distribution P represents here data, observations, while Q typically represents a model or an approximation of P. Gibbs' inequality states that $D_{\mathrm{KL}}(P\|Q) \geq 0$ and $D_{\mathrm{KL}}(P\|Q) = 0$ if and only if $P = Q$. It is also well known that

$$D_{\mathrm{KL}}(P\|Q) \geq H^2(P,Q). \tag{3.9}$$

Since Kullback and Leibler themselves (see [4]) defined the divergence as

$$D_{\mathrm{KL}}(P\|Q) + D_{\mathrm{KL}}(Q\|P), \tag{3.10}$$

we will give a new inequality for this symmetrized divergence form.

Proposition 3.4. *Let*

$$D_{\mathrm{KL}}(P\|Q) + D_{\mathrm{KL}}(Q\|P) \geq 4H^2(P,Q). \tag{3.11}$$

4. Proofs

Before we proceed with proofs of the above assertions, we give some preliminaries which will be used in the sequel.

Definition 4.1. It is said that a positive function $f(s)$ is log-convex on some open interval I if

$$f(s)f(t) \geq f^2\left(\frac{s+t}{2}\right) \tag{4.1}$$

for each $s,t \in I$.

We quote here a useful lemma from log-convexity theory (cf. [5], [6, pages 284–286].

Lemma 4.2. *A positive function f is log-convex on I if and only if the relation*

$$f(s)u^2 + 2f\left(\frac{s+t}{2}\right)uw + f(t)w^2 \geq 0 \tag{4.2}$$

holds for each real u, w, and $s,t \in I$. This result is nothing more than the discriminant test for the nonnegativity of second-order polynomials. Another well-known assertions are the following (cf. [1, pages 74, 97-98]).

Lemma 4.3. *If $g(x)$ is twice differentiable and $g''(x) \geq 0$ on I, then $g(x)$ is convex on I and*

$$\sum p_i g(x_i) \geq g\left(\sum p_i x_i\right) \tag{4.3}$$

for each $x_i \in I$, $i = 1,2,\ldots$, and any positive weight sequence $\{p_i\}$, $\sum p_i = 1$.

Lemma 4.4. *If $\phi(s)$ is continuous and convex for all s_1, s_2, s_3 of an open interval I for which $s_1 < s_2 < s_3$, then*

$$\phi(s_1)(s_3 - s_2) + \phi(s_2)(s_1 - s_3) + \phi(s_3)(s_2 - s_1) \geq 0. \tag{4.4}$$

Proof of Theorem 2.1. Consider the function $f(x,u,w,r,s,t)$ given by

$$f(x,u,w,r,s,t) := f(x) = u^2\frac{x^s}{s(s-1)} + 2uw\frac{x^r}{r(r-1)} + w^2\frac{x^t}{t(t-1)}, \tag{4.5}$$

where $r := (s+t)/2$ and u, w, r, s, t are real parameters with $r,s,t \notin \{0,1\}$. Since

$$f''(x) = u^2x^{s-2} + 2uwx^{r-2} + w^2x^{t-2} = \left(ux^{s/2-1} + wx^{t/2-1}\right)^2 \geq 0, \qquad x > 0, \tag{4.6}$$

by Lemma 4.3, we conclude that $f(x)$ is convex for $x > 0$. Hence, by Lemma 4.3 again,

$$u^2\frac{\sum p_i x_i^s}{s(s-1)} + 2uw\frac{\sum p_i x_i^r}{r(r-1)} + w^2\frac{\sum p_i x_i^t}{t(t-1)} \geq u^2\frac{\left(\sum p_i x_i\right)^s}{s(s-1)} + 2uw\frac{\left(\sum p_i x_i\right)^r}{r(r-1)} + w^2\frac{\left(\sum p_i x_i\right)^t}{t(t-1)}, \tag{4.7}$$

that is,

$$u^2\lambda_s + 2uw\lambda_r + w^2\lambda_t \geq 0 \tag{4.8}$$

holds for each $u, w \in \mathbb{R}$. By Lemma 4.2 this is possible only if

$$\lambda_s \lambda_t \geq \lambda_r^2 = \lambda_{(s+t)/2}^2, \tag{4.9}$$

and the proof is done. □

Proof of Theorem 2.2. Note that the function λ_s is continuous at the points $s = 0$ and $s = 1$ since

$$\begin{aligned} \lambda_0 &:= \lim_{s\to 0} \lambda_s = \log\left(\sum_1^n p_i x_i\right) - \sum_1^n p_i \log x_i, \\ \lambda_1 &:= \lim_{s\to 1} \lambda_s = \sum_1^n p_i x_i \log x_i - \left(\sum_1^n p_i x_i\right) \log\left(\sum_1^n p_i x_i\right). \end{aligned} \tag{4.10}$$

Therefore, $\log \lambda_s$ is a continuous and convex function for $s \in \mathbb{R}$. Applying Lemma 4.4 for $-\infty < r < s < t < +\infty$, we get

$$(t-r)\log\lambda_s \leq (t-s)\log\lambda_r + (s-r)\log\lambda_t, \tag{4.11}$$

which is equivalent to the assertion of Theorem 2.2. □

Remark 4.5. The method of proof we just exposed can be easily generalized. This is left to the reader.

Proof of Theorem 2.3 can be produced by standard means (cf. [1, pages 131–134]) and therefore is omitted.

Proof of Proposition 3.1. Applying Theorem 2.2 with $2 < 3 < s$, we get

$$\lambda_2^{s-3} \lambda_s \geq \lambda_3^{s-2}, \tag{4.12}$$

that is,

$$\lambda_s = \frac{\sum p_i x_i^s - \left(\sum p_i x_i\right)^s}{s(s-1)} \geq \left(\frac{\lambda_3}{\lambda_2}\right)^{s-2} \lambda_2, \tag{4.13}$$

and the proof of Proposition 3.1, part (i), follows. Taking $0 < s < 1 < 2 < 3$ in Theorem 2.2 and proceeding as before, we obtain the proof of the part (ii). Note that in this case

$$\lambda_s = \frac{\left(\sum p_i x_i\right)^s - \sum p_i x_i^s}{s(1-s)}. \tag{4.14}$$

□

Proof of Proposition 3.2. From Theorem 2.2, for $r = 0$, $s = s$, $t = 1$, we get

$$\lambda_s \leq \lambda_0^{1-s} \lambda_1^s, \tag{4.15}$$

that is,

$$\begin{aligned} &\frac{\left(\sum p_i x_i\right)^s - \sum p_i x_i^s}{s(1-s)} \\ &\quad \leq \left(\log \sum p_i x_i - \sum p_i \log x_i\right)^{1-s} \left(\sum p_i x_i \log x_i - \left(\sum p_i x_i\right) \log \sum p_i x_i\right)^s. \end{aligned} \tag{4.16}$$

Putting

$$s = \frac{1}{p}, \qquad 1 - s = \frac{1}{q}; \qquad p_i = \frac{b_i^q}{\sum b_j^q}, \quad x_i = \frac{a_i^p}{b_i^q}, \quad i = 1, 2, \dots, \tag{4.17}$$

after some calculations, we obtain the inequality (3.3). In the case $0 < p < 1$, put $r = 0$, $s = 1$, $t = s$ and proceed as above. □

Proof of Proposition 3.3. For a probability distribution P of a discrete variable X, defined by

$$P(X = x_i) = p_i, \quad i = 1, 2, \dots; \quad \sum p_i = 1, \tag{4.18}$$

its expectance EX and moments EX^r of rth-order (if exist) are defined by

$$\mathrm{EX} := \sum p_i x_i; \qquad \mathrm{EX}^r := \sum p_i x_i^r. \tag{4.19}$$

Since supp $P = (0, \infty)$, for $1 < m < n < p$, the inequality (2.2) reads

$$\left(\frac{\mathrm{EX}^n - (\mathrm{EX})^n}{n(n-1)}\right)^{p-m} \le \left(\frac{\mathrm{EX}^m - (\mathrm{EX})^m}{m(m-1)}\right)^{p-n} \left(\frac{\mathrm{EX}^p - (\mathrm{EX})^p}{p(p-1)}\right)^{n-m}, \tag{4.20}$$

which is equivalent with (3.5). If P is a distribution with a continuous variable, then, by Theorem 2.3, the same inequality holds for

$$\mathrm{EX} := \int_0^\infty t\,dP(t); \qquad \mathrm{EX}^r := \int_0^\infty t^r\,dP(t) < \infty. \tag{4.21}$$

□

Proof of Proposition 3.4. Putting $s = 1/2$ in (4.16), we get

$$\begin{aligned} &\left(\log \sum p_i x_i - \sum p_i \log x_i\right)^{1/2} \left(\sum p_i x_i \log x_i - \left(\sum p_i x_i\right) \log \sum p_i x_i\right)^{1/2} \\ &\quad \ge 4\left(\left(\sum p_i x_i\right)^{1/2} - \sum p_i x_i^{1/2}\right). \end{aligned} \tag{4.22}$$

Now, for $x_i = q_i/p_i$, $i = 1, 2, \dots$, and taking in account that $\sum p_i = \sum q_i = 1$, we obtain

$$\sqrt{D_{\mathrm{KL}}(P\|Q) D_{\mathrm{KL}}(Q\|P)} \ge 4\left(1 - \sum \sqrt{p_i q_i}\right) = 2 \sum \left(p_i + q_i - 2\sqrt{p_i q_i}\right) = 2H^2(P, Q). \tag{4.23}$$

Therefore,

$$D_{\mathrm{KL}}(P\|Q) + D_{\mathrm{KL}}(Q\|P) \ge 2\sqrt{D_{\mathrm{KL}}(P\|Q) D_{\mathrm{KL}}(Q\|P)} \ge 4H^2(P, Q). \tag{4.24}$$

□

References

[1] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, UK, 1978.

[2] D. S. Mitrinović, *Analytic Inequalities*, Die Grundlehren der mathematischen Wisenschaften, Band 1965, Springer, Berlin, Germany, 1970.

[3] H.-J. Rossberg, B. Jesiak, and G. Siegel, *Analytic Methods of Probability Theory*, vol. 67 of *Mathematische Monographien*, Akademie, Berlin, Germany, 1985.

[4] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[5] P. A. MacMahon, *Combinatory Analysis*, Chelsea, New York, NY, USA, 1960.

[6] J. E. Pečarić, F. Proschan, and Y. L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*, vol. 187 of *Mathematics in Science and Engineering*, Academic Press, Boston, Mass, USA, 1992.

Slavko Simic: Mathematical Institute, Serbian Academy of Sciences and Arts (SANU),
Kneza Mihaila 35, 11000 Belgrade, Serbia
Email address: ssimic@turing.mi.sanu.ac.yu