Rend. Istit. Mat. Univ. Trieste
Volume 46 (2014), 231–235

On an inequality from Information Theory

Horst Alzer

Abstract. *We prove that the inequalities*

$$\sum_{j=1}^{n} \frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{\alpha}M_j^{1-\alpha}} \leq \sum_{j=1}^{n} p_j \log\frac{p_j}{q_j} \leq \sum_{j=1}^{n} \frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{\beta}M_j^{1-\beta}} \quad (\alpha,\beta\in\mathbb{R}),$$

where

$$m_j=\min(p_j^2,q_j^2) \quad \textit{and} \quad M_j=\max(p_j^2,q_j^2) \quad (j=1,...,n),$$

hold for all positive real numbers p_j, q_j $(j=1,...,n; n\geq 2)$ *with* $\sum_{j=1}^{n} p_j = \sum_{j=1}^{n} q_j$ *if and only if* $\alpha \leq 1/3$ *and* $\beta \geq 2/3$. *This refines a result of Halliwell and Mercer, who showed that the inequalities are valid with* $\alpha = 0$ *and* $\beta = 1$.

Keywords: Gibbs' inequality, Kullback-Leibler divergence, information theory, log-function.
MS Classification 2010: 26D15, 94A15.

1. Introduction

If p_j and q_j $(j=1,...,n)$ are positive real numbers with $\sum_{j=1}^{n} p_j = \sum_{j=1}^{n} q_j$, then

$$0 \leq \sum_{j=1}^{n} p_j \log\frac{p_j}{q_j}. \tag{1}$$

The sign of equality holds in (1) if and only if $p_j = q_j$ $(j=1,...,n)$. This inequality is known in the literature as Gibbs' inequality, named after the American scientist Josiah Willard Gibbs (1839-1903). A proof of (1) can be found, for instance, in [5, p. 382].

The expression on the right-hand side of (1) is called the Kullback-Leibler divergence. It is a measure of the difference between the probability distributions $P=\{p_1,...,p_n\}$ and $Q=\{q_1,...,q_n\}$. Gibbs' inequality has many applications in information theory and also in mathematical statistics. It attracted the attention of numerous researchers, who discovered remarkable extensions, improvements and related results. For details we refer to [1, 2, 4] and the references therein.

The work on this note has been inspired by an interesting paper published by Halliwell and Mercer [3] in 2004. They presented the following elegant refinement and converse of (1).

PROPOSITION 1.1. *Let p_j, q_j $(j = 1, ..., n)$ be positive real numbers satisfying $\sum_{j=1}^n p_j = \sum_{j=1}^n q_j$. Then,*

$$\sum_{j=1}^n \frac{q_j(q_j - p_j)^2}{q_j^2 + M_j} \le \sum_{j=1}^n p_j \log \frac{p_j}{q_j} \le \sum_{j=1}^n \frac{q_j(q_j - p_j)^2}{q_j^2 + m_j}, \tag{2}$$

where

$$m_j = \min(p_j^2, q_j^2) \quad \textit{and} \quad M_j = \max(p_j^2, q_j^2) \quad (j = 1, ..., n).$$

Double-inequality (2) can be written as

$$\sum_{j=1}^n \frac{q_j(q_j - p_j)^2}{q_j^2 + m_j^\alpha M_j^{1-\alpha}} \le \sum_{j=1}^n p_j \log \frac{p_j}{q_j} \le \sum_{j=1}^n \frac{q_j(q_j - p_j)^2}{q_j^2 + m_j^\beta M_j^{1-\beta}} \tag{3}$$

with $\alpha = 0$ and $\beta = 1$. With regard to this result it is natural to ask for all real parameters α and β such that (3) holds. In the next section, we establish that (3) is valid if and only if $\alpha \le 1/3$ and $\beta \ge 2/3$. In particular, setting $\alpha = 1/3$ and $\beta = 2/3$ leads to an improvement of both sides of (2).

2. Result

We need certain upper and lower bounds for the log-function.

LEMMA 2.1. *(i) If $0 < x \le 1$, then*

$$x - 1 - \frac{(x-1)^2}{x + x^{1/3}} \le \log x \le x - 1 - \frac{(x-1)^2}{x+1} \tag{4}$$

with equality if and only if $x = 1$.

(ii) If $x > 1$, then

$$x - 1 - \frac{(x-1)^2}{x+1} < \log x < x - 1 - \frac{(x-1)^2}{x + x^{1/3}}. \tag{5}$$

Proof. Let

$$f(x)=\log x-x+1+\frac{(x-1)^2}{x+1} \quad \text{and} \quad g(x)=-\log x+x-1-\frac{(x-1)^2}{x+x^{1/3}}.$$

Then,

$$f'(x)=\frac{(x-1)^2}{x(x+1)^2} \quad \text{and} \quad g'(x)=\frac{(t-1)^4(t^2+t+1)}{3t^4(t^2+1)^2} \quad (t=x^{1/3}).$$

It follows that f and g are strictly increasing on $(0,\infty)$. Since $f(1)=g(1)=0$, we conclude that (4) and (5) are valid. $\square$

We are now in a position to prove the following refinement of (2).

THEOREM 2.2. *Let $\alpha,\beta\in\mathbb{R}$. The inequalities* (3) *hold for all positive real numbers p_j,q_j $(j=1,...,n;n\geq 2)$ with $\sum_{j=1}^n p_j=\sum_{j=1}^n q_j$ if and only if $\alpha\leq 1/3$ and $\beta\geq 2/3$.*

Proof. First, we show that if $\alpha\leq 1/3$ and $\beta\geq 2/3$, then (3) is valid for all $p_j,q_j>0$ $(j=1,...,n)$ with $\sum_{j=1}^n p_j=\sum_{j=1}^n q_j$. Since the sums on the left-hand side and on the right-hand side of (3) are increasing with respect to α and β, respectively, it suffices to prove (3) for $\alpha=1/3$ and $\beta=2/3$.

First, let $q_j\leq p_j$. Applying (4) gives

$$\frac{q_j}{p_j}-1-\frac{(q_j/p_j-1)^2}{q_j/p_j+(q_j/p_j)^{1/3}}\leq\log\frac{q_j}{p_j}\leq\frac{q_j}{p_j}-1-\frac{(q_j/p_j-1)^2}{q_j/p_j+1}.$$

We multiply by p_j and sum up. This yields

$$\begin{aligned}
&\sum_{q_j\leq p_j} q_j-\sum_{q_j\leq p_j} p_j-\sum_{q_j\leq p_j}\frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{2/3}M_j^{1/3}}\\
&\quad=\sum_{q_j\leq p_j} q_j-\sum_{q_j\leq p_j} p_j-\sum_{q_j\leq p_j}\frac{p_j(q_j/p_j-1)^2}{q_j/p_j+(q_j/p_j)^{1/3}}\\
&\quad\leq\sum_{q_j\leq p_j} p_j\log\frac{q_j}{p_j}\\
&\quad\leq\sum_{q_j\leq p_j} q_j-\sum_{q_j\leq p_j} p_j-\sum_{q_j\leq p_j}\frac{p_j(q_j/p_j-1)^2}{q_j/p_j+1}\\
&\quad\leq\sum_{q_j\leq p_j} q_j-\sum_{q_j\leq p_j} p_j-\sum_{q_j\leq p_j}\frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{1/3}M_j^{2/3}}.
\end{aligned}\tag{6}$$

Next, let $q_j>p_j$. Using (5) leads to

$$\frac{q_j}{p_j}-1-\frac{(q_j/p_j-1)^2}{q_j/p_j+1}<\log\frac{q_j}{p_j}<\frac{q_j}{p_j}-1-\frac{(q_j/p_j-1)^2}{q_j/p_j+(q_j/p_j)^{1/3}}.$$

Again we multiply by p_j and sum up. Then we obtain

$$\begin{aligned}\sum_{q_j>p_j} q_j - \sum_{q_j>p_j} p_j - \sum_{q_j>p_j} \frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{2/3}M_j^{1/3}} \\ &< \sum_{q_j>p_j} q_j - \sum_{q_j>p_j} p_j - \sum_{q_j>p_j} \frac{p_j(q_j/p_j-1)^2}{q_j/p_j+1} \\ &< \sum_{q_j>p_j} p_j \log\frac{q_j}{p_j} \\ &< \sum_{q_j>p_j} q_j - \sum_{q_j>p_j} p_j - \sum_{q_j>p_j} \frac{p_j(q_j/p_j-1)^2}{q_j/p_j+(q_j/p_j)^{1/3}} \\ &= \sum_{q_j>p_j} q_j - \sum_{q_j>p_j} p_j - \sum_{q_j>p_j} \frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{1/3}M_j^{2/3}}.\end{aligned} \tag{7}$$

Combining (6) and (7) gives

$$\begin{aligned}\sum_{j=1}^n q_j - \sum_{j=1}^n p_j - \sum_{j=1}^n \frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{2/3}M_j^{1/3}} \\ \le \sum_{j=1}^n p_j \log\frac{q_j}{p_j} \le \sum_{j=1}^n q_j - \sum_{j=1}^n p_j - \sum_{j=1}^n \frac{q_j(q_j-p_j)^2}{q_j^2+m_j^{1/3}M_j^{2/3}}.\end{aligned} \tag{8}$$

Since $\sum_{j=1}^n p_j = \sum_{j=1}^n q_j$, we conclude from (8) that (3) is valid with $\alpha = 1/3$ and $\beta = 2/3$.

It remains to prove that if (3) holds for all $p_j, q_j > 0$ $(j = 1, ..., n)$ with $\sum_{j=1}^n p_j = \sum_{j=1}^n q_j$, then $\alpha \le 1/3$ and $\beta \ge 2/3$.

Let $s, t \in \mathbb{R}$ with $1 < t < s+1$. We set

$$p_1 = \frac{s}{t}, \quad p_2 = \frac{1}{t}, \quad q_1 = \frac{s+1}{t} - 1, \quad q_2 = 1, \quad p_j = q_j \quad (j = 3, ..., n).$$

Then we have

$$\sum_{j=1}^n p_j = \sum_{j=1}^n q_j, \quad m_1 = q_1^2, \quad M_1 = p_1^2, \quad m_2 = p_2^2, \quad M_2 = q_2^2.$$

A short calculation reveals that (3) is equivalent to

$$F_\alpha(s,t) \le s \log\frac{s}{s+1-t} - \log t \le F_\beta(s,t),$$

where

$$F_c(s,t) = \frac{(t-1)^2}{s+1-t+s^{2(1-c)}(s+1-t)^{2c-1}} + \frac{(t-1)^2}{t+t^{1-2c}}.$$

We define

$$G_c(s,t) = s\log\frac{s}{s+1-t} - \log t - F_c(s,t).$$

Then,

$$G_c(s,1) = \frac{\partial}{\partial t}G_c(s,t)\Big|_{t=1} = \frac{\partial^2}{\partial t^2}G_c(s,t)\Big|_{t=1} = 0$$

and

$$\frac{s^2}{3(s^2+1)}\frac{\partial^3}{\partial t^3}G_c(s,t)\Big|_{t=1} = \frac{s^2+2}{3(s^2+1)} - c.$$

Since

$$\lim_{s\to 0}\frac{s^2+2}{3(s^2+1)} = \frac{2}{3} \quad \text{and} \quad \lim_{s\to\infty}\frac{s^2+2}{3(s^2+1)} = \frac{1}{3},$$

we conclude from $G_\alpha(s,t) \geq 0$ that $\alpha \leq 1/3$ and from $G_\beta(s,t) \leq 0$ that $\beta \geq 2/3$. □

REMARK 2.3. *The proof of the Theorem reveals that if $\alpha \leq 1/3$ and $\beta \geq 2/3$, then the sign of equality holds in each inequality of* (3) *if and only if $p_j = q_j$ $(j = 1, ..., n)$.*

Acknowledgements

I thank the referee for helpful comments.

References

[1] A. Clim and P. Cozma, *A new inequality of Shannon type*, Proc. Rom. Acad. Ser A **11** (2010), 294–302.
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, 2006.
[3] G. T. Halliwell and P. R. Mercer, *A refinement of an inequality from information theory*, J. Inequal. Pure Appl. Math. **5(1)** (2004), article 3, 3 pages.
[4] Pl. Kannappan and P. K. Sahoo, *On a generalization of the Shannon functional inequality*, J. Math. Anal. Appl. **140** (1989), 341–350.
[5] D. S. Mitrinović, *Analytic Inequalities*, Springer, New York, 1970.

Author's address:

Horst Alzer
Morsbacher Str. 10
51545 Waldbröl, Germany
E-mail: h.alzer@gmx.de

Received September 26, 2014
Revised October 14, 2014