Some Refinements of Entropy Inequality

J. Rooin

Department of Mathematics
Institute for Advanced Studies in Basic Sciences
Zanjan, Iran
rooin@iasbs.ac.ir

Abstract

In this note, using some refinements of Jensen's discrete inequality, we give some new refinements of Entropy inequality.

1 Introduction

Let $p_1, p_2, \cdots, p_n > 0$ with $\sum_{i=1}^n p_i = 1$, and

$$H(p_1, p_2, \cdots, p_n) = -\sum_{i=1}^n p_i \ln p_i. \tag{1}$$

Entropy inequality states that

$$H(p_1, p_2, \cdots, p_n) \leq \ln n, \tag{2}$$

see e.g. [1], or equivalently

$$\frac{1}{n} \leq \prod_{i=1}^n p_i^{p_i}. \tag{3}$$

In this note, we apply some refinements of Jensen's discrete inequality in the case of equal weights [2] in order to sharpen Eentropy inequality (2) and its equivalent (3).

2 Refinements

A quadratic real matrix with nonnegative entries is said to be a double stochastic matrix if the sum of each of its rows and columns is unit.
Let $B = [b_{ij}]$ and $C = [c_{ij}]$ be two $n \times n$ double stochastic matrices. Also, let X be a real linear space, $f : C \subseteq X \to \mathbb{R}$ a convex mapping on a convex subset C of X, and $x_1, \cdots, x_n$ a finite number of elements of C. Put

$$F(t) = \frac{\sum_{i=1}^{n} f\left(\sum_{j=1}^{n}[(1-t)b_{ij} + tc_{ij}]x_j\right)}{n} \qquad (0 \le t \le 1). \tag{4}$$

It is shown in [2] that, F is a bounded convex function on $[0,1]$ and so Riemann integrable,

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \le F(t) \le \frac{f(x_1) + \cdots + f(x_n)}{n} \qquad (0 \le t \le 1), \tag{5}$$

and

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \le \int_0^1 F(t)dt \le \frac{f(x_1) + \cdots + f(x_n)}{n}. \tag{6}$$

In the following theorem, we use these facts in order to get some refinements of Entropy inequality (2) and its equivalent (3).
Here, we denote by S_n the set of all permutations on $\{1, 2, \cdots, n\}$, and the Identric mean I is defined for each $a, b > 0$ by

$$I(a,b) = \begin{cases} a & \text{if } a = b, \\ \frac{1}{e}\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}} & \text{if } a \ne b. \end{cases}$$

Theorem 2.1. *With the above notations, if $p_1, p_2, \cdots, p_n > 0$, $\sum_{i=1}^{n} p_i = 1$ and $0 \le t \le 1$, then*

$$H(p_1, p_2, \cdots, p_n) \le H\left(\sum_{j=1}^{n}[(1-t)b_{1j} + tc_{1j}]p_j, \cdots, \sum_{j=1}^{n}[(1-t)b_{nj} + tc_{nj}]p_j\right) \le \ln n, \tag{7}$$

and

$$\frac{1}{n} \le \sqrt{\prod_{i=1}^{n} I\left(\left[\sum_{j=1}^{n} b_{ij}p_j\right]^2, \left[\sum_{j=1}^{n} c_{ij}p_j\right]^2\right)^{\sum_{j=1}^{n} \frac{b_{ij}+c_{ij}}{2} p_j}} \le \prod_{i=1}^{n} p_i^{p_i}. \tag{8}$$

In particular, we have

$$H(p_1, p_2, \cdots, p_n) \le H\left((1-t)p_1 + tp_n, \cdots, (1-t)p_n + tp_1\right) \le \ln n, \tag{9}$$

$$\frac{1}{n} \leq \sqrt{\prod_{i=1}^{n} I\left(p_i^2, p_{n+1-i}^2\right)^{\frac{p_i+p_{n+1-i}}{2}}} \leq \prod_{i=1}^{n} p_i^{p_i}, \tag{10}$$

$$\frac{1}{2} \leq \sqrt{I\left(p^2, q^2\right)} \leq p^p q^q \qquad (p, q > 0,\ p + q = 1). \tag{11}$$

Moreover,

$$H(p_1, p_2, \cdots, p_n) \leq \frac{1}{n!} \sum_{\sigma \in S_n} H\left((1-t)p_{\sigma(1)} + tp_{\sigma(n)}, \cdots, (1-t)p_{\sigma(n)} + tp_{\sigma(1)}\right) \leq \ln n, \tag{12}$$

$$\frac{1}{n} \leq \sqrt{\frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i=1}^{n} I\left(p_{\sigma(i)}^2, p_{\sigma(n+1-i)}^2\right)^{\frac{p_{\sigma(i)}+p_{\sigma(n+1-i)}}{2}}} \leq \prod_{i=1}^{n} p_i^{p_i}, \tag{13}$$

which are independent from ordering of p_i's.
Proof. Take the convex function $f(t) = t \ln t$ $(t > 0)$ and consider $F(t)$ as in (4) with $(0, \infty)$ instead of C and p_i's instead of x_i's . We have

$$F(t) = -\frac{1}{n} H\left(\sum_{j=1}^{n} [(1-t)b_{1j} + tc_{1j}]p_j, \cdots, \sum_{j=1}^{n} [(1-t)b_{nj} + tc_{nj}]p_j\right), \tag{14}$$

and since $\int_a^b f(t)dt = \frac{b^2-a^2}{4} \ln I(a^2, b^2)$,

$$\int_0^1 F(t)dt = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{n}(c_{ij} - b_{ij})p_j} \int_{\sum_{j=1}^{n} b_{ij}p_j}^{\sum_{j=1}^{n} c_{ij}p_j} f(t)dt \tag{15}$$

$$= \frac{1}{2n} \ln \prod_{i=1}^{n} I\left(\left[\sum_{j=1}^{n} b_{ij}p_j\right]^2, \left[\sum_{j=1}^{n} c_{ij}p_j\right]^2\right)^{\sum_{j=1}^{n} \frac{b_{ij}+c_{ij}}{2} p_j}.$$

Now (7) and (8) follow by substituting (14) and (15) in (5) and (6) respectively, and taking into account that

$$f\left(\frac{p_1 + \cdots + p_n}{n}\right) = f(\frac{1}{n}) = \frac{1}{n} \ln(\frac{1}{n})$$

and

$$\frac{f(p_1) + \cdots + f(p_n)}{n} = -\frac{1}{n} H(p_1, \cdots, p_n) = \frac{1}{n} \ln(\prod_{i=1}^{n} p_i^{p_i}).$$

In particular, (9) and (10) follow from (7) and (8) respectively by taking $b_{ij} = \delta_{ij}$ and $c_{ij} = \delta_{i,n+1-j}$ $(i, j = 1, 2, \cdots, n)$, where δ_{ij} is the Kronecker delta.
(11) is an special case of (10), taking $n = 2$, $p_1 = p$ and $p_2 = q$.
Finally, (12) and (13) follow by applying (9) and (10) for arbitrary permutations $p_{\sigma(1)}, p_{\sigma(2)}, \cdots, p_{\sigma(n)}$
$(\sigma \in S_n)$ of $p_1, p_2, \cdots, p_n$, and summing all of them with some calculations.

REFERENCES

1. Gareth A. Jones and J. Mary Jones, *Information and Coding Theory*, Springer-Verlag, 2000.

2. J. Rooin, Some aspects of convex functions and their applications, *Jipam*, Vol 2, Issue 1, Article 4 (2001).

Institute for Advanced Studies in Basic Sciences,
P.O. Box 45195-159,
Zanjan, IRAN.
e-mail: Rooin@iasbs.ac.ir