Honam Mathematical J. **44** (2022), No. 4, pp. 513–520
https://doi.org/10.5831/HMJ.2022.44.4.513

AN EXTENSION OF JENSEN-MERCER INEQUALITY WITH APPLICATIONS TO ENTROPY

Yamin Sayyari

Abstract. The Jensen and Mercer inequalities are very important inequalities in information theory. The article provides the generalization of Mercer's inequality for convex functions on the line segments. This result contains Mercer's inequality as a particular case. Also, we investigate bounds for Shannon's entropy and give some new applications in zeta function and analysis.

1. Introduction

Let $I := [a, b]$ be an interval, $\mathbf{x} := \{x_i\}_{i=1}^n \subseteq I$ and $\mathbf{p} := \{p_i\}_1^n \subseteq [0, 1]$ with $\sum_{i=1}^n p_i = 1$. The following inequality is well known in the literature as Jensen's inequality.

Theorem 1.1. [7] *(Jensen's inequality) If f is a convex function on an interval I, $\mathbf{x} := \{x_i\}_{i=1}^n \subseteq I$ and $\sum_{i=1}^n p_i = 1$, then*

$$0 \leq \sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right) := J_f(\mathbf{p}, \mathbf{x}).$$

A variant of Jensen's inequality is obtained by Mercer [5]. Mercer [5] proved that if f is a convex function on $[a, b]$, then

Theorem 1.2. [5] *If f is a convex function on an interval $I := [a, b]$, $x_i \in I$, $1 \leq i \leq n$ and $\sum_{i=1}^n p_i = 1$, then*

$$I_f(\mathbf{p}, \mathbf{x}) := f\left(a + b - \sum_{i=1}^n p_i x_i\right) + \sum_{i=1}^n p_i f(x_i) \leq f(a) + f(b). \tag{1.1}$$

Lemma 1.3. *([5], Lemma 1.3) Let $0 < a \leq y \leq b$. For f convex we have:*

$$f(a + b - y) \leq f(a) + f(b) - f(y).$$

The following lemma and Lemma 1.3 are equivalent.

Received December 11, 2021. Revised February 3, 2022. Accepted July 28, 2022.
2020 Mathematics Subject Classification. 26B25, 11Y40, 26D15, 94A17.
Key words and phrases. Shannon's entropy, Zeta function, Jensen's inequality, Mercer's inequality, convex function.

Lemma 1.4. *Let f be a convex function on $[a,b]$, $x_1, x_2 \in [a,b]$ and $x_1 + x_2 = a + b$. Then*

$$f(x_1) + f(x_2) \leq f(a) + f(b). \tag{1.2}$$

Proof. It follows from Lemma 1.3 with $x_1 = a + b - y$ and $x_2 = y$. □

2. An extension of Mercer inequality

In this section, we extend the Mercer inequality (1.2) for convex functions.

Theorem 2.1. *Let f be a convex function on $[a,b]$, $x_1, \ldots, x_n \in [a,b]$ and $\frac{x_1+\ldots+x_n}{n} = \frac{a+b}{2}$. Then*

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(x_1) + \ldots + f(x_n)}{n} \leq \frac{f(a)+f(b)}{2}.$$

Proof. Let $\frac{x_1+\ldots+x_n}{n} = \frac{a+b}{2}$. The first inequality is a direct consequence of Jensen's inequality, i.e.,

$$f\left(\frac{a+b}{2}\right) = f\left(\frac{x_1 + \ldots + x_n}{n}\right) \leq \frac{\sum_{i=1}^n f(x_i)}{n}.$$

The second inequality is verified as follows. Since $x_i \in [a,b]$, there is a sequence $\{\lambda_i\}_{i=1}^n$, $\lambda_i \in [0,1]$, such that $x_i = \lambda_i a + (1-\lambda_i)b$. On the other hand, since $\frac{x_1+\ldots+x_n}{n} = \frac{a+b}{2}$, we have

$$\frac{\sum_{i=1}^n \lambda_i}{n} = \sum_{i=1}^n \frac{1-\lambda_i}{n} = \frac{1}{2}.$$

Hence,

$$\begin{aligned}\frac{f(x_1) + \ldots + f(x_n)}{n} &= \frac{\sum_{i=1}^n f(\lambda_i a + (1-\lambda_i)b)}{n} \\ &\leq \frac{\sum_{i=1}^n \lambda_i f(a) + \sum_{i=1}^n (1-\lambda_i) f(b)}{n} \\ &= \frac{f(a)+f(b)}{2}.\end{aligned}$$

□

Corollary 2.2. *With the notations in Theorem 2.1, if $n = 2$, then*

$$f(x_1) + f(x_2) \leq f(a) + f(b),$$

which is analogous to inequality (1.2).

Theorem 2.3. *Let $f : (0,\infty) \longrightarrow \mathbb{R}$ be a convex function. Then*

$$f\left(\frac{1+n}{2}\right) \leq \frac{f(1) + f(2) + \ldots + f(n)}{n} \leq \frac{f(1)+f(n)}{2}.$$

Proof. Since f is convex on $(0,\infty)$, f is convex on $[1,n]$ for all natural number n. Also, since $\frac{1+2+\ldots+n}{n}=\frac{1+n}{2}$, by the use of Theorem 2.1 we conclude that

$$f\left(\frac{1+n}{2}\right)\leq\frac{f(1)+f(2)+\ldots+f(n)}{n}\leq\frac{f(1)+f(n)}{2}.$$

□

Example 2.4. Let $a<b$, f be a convex function on $[a,b]$ and let $x\in[a,b]$, $x_1=\frac{a+x}{2}$, $x_2=\frac{b+x}{2}$ and $x_3=a+b-x$. Then $\frac{x_1+x_2+x_3}{3}=\frac{a+b}{2}$, hence by the use of Theorem 2.1 we have

$$f\left(\frac{a+b}{2}\right)\leq\frac{f\left(\frac{a+x}{2}\right)+f\left(\frac{b+x}{2}\right)+f(a+b-x)}{3}\leq\frac{f(a)+f(b)}{2}$$

for all $x\in[a,b]$. Integrating these inequalities over $[a,b]$ with respect to x yields

$$f\left(\frac{a+b}{2}\right)\leq\frac{1}{b-a}\int_a^b f(x)dx\leq\frac{f(b)+f(a)}{2}$$

which is the well-known Hermite–Hadamard inequality.

Theorem 2.5. *Let f be a convex function on $[a,b]$, $\sum_{k=1}^n p_k=1$, $[x_{ij}]_{n\times m}$ be a matrix with $a\leq x_{ij}\leq b$ for all $i(1\leq i\leq n)$, $j(1\leq j\leq m)$ and*

$$\frac{\sum_{i=1}^n x_{ij}}{n}=\frac{a+b}{2}$$

for all $j(1\leq j\leq m)$. Then

$$\frac{\sum_{j=1}^m\sum_{i=1}^{n-1}p_jf(x_{ij})+f\left[\frac{n}{2}(a+b)-\sum_{j=1}^m\sum_{i=1}^{n-1}p_jx_{ij}\right]}{n}$$
$$\leq\frac{f(a)+f(b)}{2}.$$

Proof. By the use of Theorem 2.1 we have

$$\begin{aligned}
&\frac{\sum_{j=1}^{m}\sum_{i=1}^{n-1} p_j f(x_{ij}) + f\left[\frac{n}{2}(a+b) - \sum_{j=1}^{m}\sum_{i=1}^{n-1} p_j x_{ij}\right]}{n} \\
&= \frac{\sum_{j=1}^{m}\sum_{i=1}^{n-1} p_j f(x_{ij}) + f\left[\sum_{j=1}^{m} p_j(\frac{n}{2}(a+b) - \sum_{i=1}^{n-1} x_{ij})\right]}{n} \\
&\leq \frac{\sum_{j=1}^{m}\sum_{i=1}^{n-1} p_j f(x_{ij}) + \sum_{j=1}^{m} p_j f\left[\frac{n}{2}(a+b) - \sum_{i=1}^{n-1} x_{ij}\right]}{n} \\
&= \sum_{j=1}^{m} p_j \left\{ \frac{\sum_{i=1}^{n-1} f(x_{ij}) + f\left[\frac{n}{2}(a+b) - \sum_{i=1}^{n-1} x_{ij}\right]}{n} \right\} \\
&\leq \sum_{j=1}^{m} p_j \left\{ \frac{f(a)+f(b)}{2} \right\} \\
&= \frac{f(a)+f(b)}{2},
\end{aligned}$$

which completes the proof. □

Corollary 2.6. *With the notations in Theorem 2.5, if $n = 2$, then*

$$\sum_{j=1}^{m} p_j f(x_{1j}) + f(a+b-\sum_{j=1}^{m} p_j x_{1j}) \leq f(a)+f(b),$$

which is analogous to inequality (1.1).

3. Applications in information theory

Definition 3.1. The Shannon entropy of a positive probability distribution $P = (p_1, ..., p_n)$ is defined by $H(\mathbf{p}) := \sum_{i=1}^{n} p_i \log \frac{1}{p_i}$.

Let $\mathbf{x} = \{x_i\}_{i=1}^{n}$ be a positive real sequence and

$$A_n(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^{n} x_i \text{ and } G_n(\mathbf{x}) := \left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}}$$

denote the usual arithmetic and geometric means of $\{x_i\}$, respectively. From Theorem 2.1 we conclude the following result.

Proposition 3.2. *Let $a \geq 0$, $x_i \in [a,b]$ $i = 1, ..., n$ and $A_n(\mathbf{x}) = \frac{a+b}{2}$, then $\sqrt{ab} \leq G_n(\mathbf{x}) \leq \frac{a+b}{2}$.*

Proof. Setting $f(x) = -\log x$ in Theorem 2.1, we get

$$-\log\left(\frac{a+b}{2}\right) \leq \frac{-\log(x_1) - \ldots - \log(x_n)}{n} \leq \frac{-\log(a) - \log(b)}{2}.$$

Therefore,

$$\frac{\log(a)+\log(b)}{2} \leq \frac{\log(x_1)+\ldots+\log(x_n)}{n} \leq \log\left(\frac{a+b}{2}\right),$$

namely

$$\log\left(\sqrt{ab}\right) \leq \log(G_n(\mathbf{x})) \leq \log\left(\frac{a+b}{2}\right). \tag{3.1}$$

Since the function log is nondecreasing, the result follows from (3.1). □

Proposition 3.3. *Let $P = (p_1, ..., p_n)$ be a positive probability distribution and $0 < \max_{1\leq j\leq n} p_j \leq \frac{2}{n}$. Then*

1. $\log\left(\frac{n}{2}\right) \leq H(\mathbf{p}) \leq \log n$.
2. $0 \leq \log n - H(\mathbf{p}) \leq \log 2$.

Proof. 1. Applying Theorem 2.1 with

$$f(x) = \begin{cases} x\log(x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases},$$

$x_i = p_i$ for all $i = 1, \ldots, n$, $a = 0$ and $b = \frac{2}{n}$, we have

$$\frac{1}{n}\log\left(\frac{1}{n}\right) \leq \frac{p_1\log p_1 + \ldots + p_n \log p_n}{n} \leq \frac{1}{n}\log\left(\frac{2}{n}\right),$$

which completes the proof.

2. It follows from (i) that $\log\left(\frac{n}{2}\right) \leq H(\mathbf{p})$.

□

Proposition 3.4. *Let f be a convex function on $[a, b]$, $\sum_{k=1}^{n} p_k = 1$, $[x_{ij}]_{n\times m}$ be a matrix with $a \leq x_{ij} \leq b$ for all $i(1 \leq i \leq n)$, $j(1 \leq j \leq m)$ and*

$$\frac{\sum_{i=1}^{n} x_{ij}}{n} = \frac{a+b}{2}$$

for all $j(1 \leq j \leq m)$. Then

$$\frac{n}{2}(a+b) - \sum_{j=1}^{m}\left(p_j \sum_{i=1}^{n-1} x_{ij}\right) \geq \frac{\sqrt{a^n b^n}}{\prod_{j=1}^{m}\left(\prod_{i=1}^{n-1} x_{ij}\right)^{p_j}}.$$

Proof. Applying Theorem 2.5 with $f(x) = -\log x$, we get

$$\frac{-\sum_{j=1}^{m}\sum_{i=1}^{n-1} p_j \log(x_{ij}) - \log\left[\frac{n}{2}(a+b) - \sum_{j=1}^{m}\sum_{i=1}^{n-1} p_j x_{ij}\right]}{n} \leq \frac{-\log(a)-\log(b)}{2}.$$

After some calculations the desired assertion follows.. □

Corollary 3.5. *With the notations in Proposition 3.4, if $n = 2$ and $x_j := x_{1j}$ for all $j = 1, ..., m$, then we have*

$$a + b - \sum_{j=1}^{m} p_j x_j \geq \frac{ab}{\prod_{j=1}^{m} x_j^{p_j}},$$

for all $\{x_j\} \in [a, b]$.

4. Applications in zeta function

We now provide some applications of Theorem 2.3 in number theory. The Riemann zeta function is defined as follows:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \; s \in \mathbb{C}.$$

Consequently, ζ converges for all complex numbers s such that $Re(s) > 1$.

Proposition 4.1. *Let $s > 0$ and n be a natural number, then*

$$\frac{n2^s}{(1+n)^s} \leq 1 + \frac{1}{2^s} + \frac{1}{3^s} + \ldots + \frac{1}{n^s} \leq \frac{n + n^{1-s}}{2}.$$

Proof. Let $s > 0$ and $f(x) = x^{-s}$. Since $f''(x) = s(s+1)x^{-s-2} > 0$, applying Theorem 2.3 with $f(x) = x^{-s}$, we obtain

$$\left(\frac{1+n}{2}\right)^{-s} \leq \frac{1^{-s} + 2^{-s} + \ldots + n^{-s}}{n} \leq \frac{1 + n^{-s}}{2}.$$

Therefore,

$$\frac{n2^s}{(1+n)^s} \leq 1 + \frac{1}{2^s} + \frac{1}{3^s} + \ldots + \frac{1}{n^s} \leq \frac{n + n^{1-s}}{2}.$$

□

Example 4.2. Let n be a natural number and $s = 3$. Then Proposition 4.1 yields (see figure 1)

$$\frac{8n}{(1+n)^3} \leq 1 + \frac{1}{2^3} + \frac{1}{3^3} + \ldots + \frac{1}{n^3} \leq \frac{n^3 + 1}{2n^2}.$$

Proposition 4.3. *Let n be a natural number. Then*

1. $n^n \leq (n!)^2 \leq \left(\frac{1+n}{2}\right)^{2n}$.
2. $\left(\frac{1+n}{2}\right)^{n+n^2} \leq 1^2 \times 2^4 \times \ldots \times n^{2n} \leq (n)^{n^2}$.
3. $n^{2n} \leq (n!)^4 \leq (n)^{n^2}$.

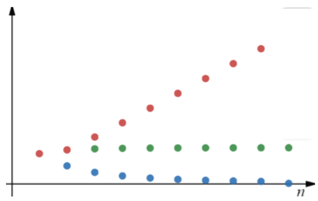

FIGURE 1. $\frac{8n}{(1+n)^3}$, $1+\frac{1}{2^3}+\frac{1}{3^3}+\ldots+\frac{1}{n^3}$ and $\frac{n^3+1}{2n^2}$

Proof. 1. On applying Theorem 2.3 for the function $f(x) = -\ln x$, we get

$$-\ln\left(\frac{1+n}{2}\right) \leq \frac{-\sum_{i=1}^{n}\ln i}{n} \leq \frac{-\ln n}{2}.$$

Thus,

$$2n\ln\left(\frac{2}{1+n}\right) \leq 2\ln\left(\prod_{i=1}^{n}\frac{1}{i}\right) \leq n\ln\frac{1}{n}.$$

Hence, $n^n \leq (n!)^2 \leq \left(\frac{1+n}{2}\right)^{2n}$.

2. Since $f(x) = x\log x$ is convex on $(0,\infty)$, Theorem 2.3 imply that

$$\frac{1+n}{2}\ln\left(\frac{1+n}{2}\right) \leq \frac{1\ln 1 + 2\ln 2 + \ldots + n\ln n}{n} \leq \frac{n\ln n}{2}.$$

Therefore,

$$\ln\left(\frac{1+n}{2}\right)^{n+n^2} \leq 2\ln\left(\prod_{i=1}^{n} i^i\right) \leq n^2\ln n.$$

It follows that

$$\left(\frac{1+n}{2}\right)^{n+n^2} \leq 1^2\times 2^4\times\ldots\times n^{2n} \leq (n)^{n^2}.$$

3. Since $1^2\times 2^4\times\ldots\times n^{2n} \geq (n!)^4$, this result follows from (1) and (2). □

5. Conclusion

The Mercer's inequality and its derivative theorems play an important role in mathematical analysis. In this paper, we introduced an extension of Mercer's inequality for convex functions. Based on these results, we have given some bounds for zeta function and Shannon's entropy. We anticipate that the present results will find some applications in p-series as well as other related disciplines.

References

[1] I. Budimir, S. S. Dragomir, and J. Pecaric, *Further reverse results for Jensen's discrete inequality and applications in information theory*, J. Inequal. Pure Appl. Math. **2** (2001), no. 1, Art. 5.

[2] S. S. Dragomir, *A converse result for Jensen's discrete inequality via Grüss inequality and applications in information theory*, An. Univ. Oradea. Fasc. Mat. **7** (1999-2000), 178–189.

[3] S.S. Dragomir and C.J. Goh, *Some bounds on entropy measures in Information Theory*, Appl. Math. Lett. **10** (1997), no. 3, 23–28.

[4] JLWV Jensen, *Om konvekse Funktioner og Uligheder mellem Middelværdier*, Nyt Tidsskr Math. B **16** (1905), 49–68 (in Danish).

[5] A. McD. Mercer, *A variant of Jensen's inequality*, JIPAM. J. Inequal. Pure Appl. Math. **4** (2003), no. 4, Article ID 73.

[6] D. S. Mitrinovic, *Analytic Inequalities,* Springer, New York, 1970.

[7] D. S. Mitrinovic, J. E. Pecaric, and A. M. Fink, *Classical and new inequalities in analysis*, Kluwer Academic Publishers, Dordrecht, 1993.

[8] Y. Sayyari, *An improvement of the upper bound on the entropy of information sources*, J. Math. Ext. **15** (2021), no. 5, 1–12.

[9] Y. Sayyari, *New bounds for entropy of information sources*, Wavelet and Linear Algebra, **7** (2020), no. 2, 1–9.

[10] Y. Sayyari, *New entropy bounds via uniformly convex functions*, Chaos, Solitons and Fractals **141** (2020), 110360.

[11] Y Sayyari, *New refinements of Shannon's entropy upper bounds*, Journal of Information and Optimization Sciences **42** (2021), no . 8, 1869–1883.

[12] Y. Sayyari, M. R. Molaei, and S. M. Moghayer, *Entropy of continuous maps on quasi-metric spaces,* Journal of Advanced Research in Dynamical and Control Systems **7** (2015), no. 4, 1–10.

[13] S. Simic, *On a global bound for Jensen's inequality,* J. Math. Anal. Appl. **343** (2008), 414–419.

[14] S. Simic, *Jensen's inequality and new entropy bounds,* Appl. Math. Lett. **22** (2009), no. 8, 1262–1265.

[15] P. Walters, *An Introduction to Ergodic Theory*, Springer Verlag, New York, 2000.

Yamin Sayyari
Department of Mathematics,
Sirjan University Of Technology,
Sirjan, Iran
E-mail: ysayyari@gmail.com, y.sayyari@sirjantech.ac.ir