BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **11**, No 3

Sofia • 2011

Some New Bounds of Weighted Entropy Measures

Amit Srivastava

Department of Mathematics, Jaypee Institute of Information Technology, NOIDA (Uttar Pradesh), INDIA
Email: raj_amit377@yahoo.co.in

Abstract: *In order to integrate the quantitative, objective and probabilistic concept of information with the qualitative, subjective and non – stochastic concept of utility, researchers over the past years have proposed several weighted information measures. These measures find applications in fields dealing with random events where it is necessary to take into account both the probabilities with which these events occur and some qualitative characteristic of these events. However it seems to the author that very little effort has been devoted by researchers in obtaining bounds on these weighted information Measures. In the present work, we have obtained bounds on these weighted information measures using the Lagrange's multiplier method and some well known inequalities. Without essential loss of insight we have restricted ourselves to discrete probability distributions.*

Keywords: *Weighted entropy, utility distribution, Arithmetic, Geometric & Harmonic mean inequality.*

1. Introduction

Let

(1) $\Gamma_n = \{(p_1, p_2, \dots, p_n): p_i \geq 0,\ i = 1, 2, \dots, n,\ \sum_{i=1}^{n} p_i = 1\},\ n = 2, 3, \dots$

denote the set of all finite discrete (n-ray) complete probability distributions. Shannon [7] introduced the following measure of information

(2) $$H_n(P) = -\sum_{i=1}^{n} p_i \log p_i$$

for all $P = (p_1, p_2, \dots, p_n) \in \Gamma_n$. The expression (2) is famous as Shannon's entropy or measure of uncertainty. The function $H_n(P)$ represents the expected value of uncertainty associated with the given probability scheme and it is uniquely determined by some rather natural postulates. Underlining the importance of Shannon's entropy, it is necessary to notice at the same time that this formula gives us a measure of information as a function of the probabilities with which various events occur, only. But there exists many fields dealing with random events where it becomes necessary to take into account both these probabilities and some qualitative characteristic of events. B e l i s and G u i a s u [1] raised this very important issue of integrating the quantitative, objective and probabilistic concept of information with the qualitative, subjective and non-stochastic concept of utility, and proposed the following weighted measure of information

(3) $$H(P,\ U) = -\sum_{i=1}^{n} u_i p_i \log p_i.$$

Here u_i denotes the weight of an elementary event with probability p_i. In general, u_i is a non-negative, finite, real number accounting for the relevance, significance or the utility of the occurrence of an event with probability p_i. Also, if one event is more relevant, more significant, and more useful (with respect to a given goal or from a given qualitative point of view) than another one, the weight of the first event will be greater than that of second one. L o n g o [5] studied the measure (3) in detail and raised objections about its applicability in various coding procedures. He further suggested that the concept of utility should be introduced in a different way in any information scheme and that a utility measure should exhibit a relative character rather than only a non – negative character. K a p u r [3, 4] further studied the measure (3) and asserted that (3) is not a measure of information since it depends on the units in which utility is measured and as such (3) can be expressed in dollar-bits or hour-bits and not in terms of bits only. Keeping in mind, K a p u r [4] considered the following probability distribution

$$\frac{u_i p_i}{\sum_{i=1}^{n} u_i p_i},\ i = 1, 2,\ \dots, n,$$

and used this distribution in obtaining the following weighted information measure

(4) $$\widetilde{H}(P,\ U) = -\sum_{i=1}^{n} \frac{u_i p_i}{\sum_{i=1}^{n} u_i p_i} \log \frac{u_i p_i}{\sum_{i=1}^{n} u_i p_i}.$$

M u n t e a n u and T a r n i c e r i u [8] considered a different set of axioms and proposed the following weighted information measure:

(5) $$\widetilde{H}^*(P,\ U) = -\sum_{i=1}^{n} p_i \log p_i + \sum_{i=1}^{n} u_i p_i.$$

In the present work, we have obtained bounds on the weighted information measures given by (2)-(5) using the Lagrange's multiplier method and some well known inequalities such as the weighted arithmetic, geometric & harmonic mean inequality and the Shannon inequality including its generalizations.

Remarks: From now onwards, logarithms are taken to base 2.

2. Optimization of weighted entropy functionals

Consider the following function

(6) $$\varphi(p_1, p_2,\ \dots, p_n; u_1, u_2,\ \dots, u_n) = -\sum_{i=1}^{n} u_i p_i \log p_i + \lambda\left(\sum_{i=1}^{n} p_i - 1\right)$$

where λ is a real positive number (the lagrange multiplier).

The maxima of the function φ with respect to p_i for fixed u_i coincides with the maxima of the function $H(P,\ U)$. The necessary condition for the existence of extreme is given by the system

(7) $$\left\{\begin{matrix} \frac{\partial \varphi(p_1, p_2, \dots, p_n; u_1, u_2, \dots, u_n)}{\partial p_i} = 0,\ 1 \le i \le n \\ \sum_{i=1}^{n} p_i = 1 \end{matrix}\right\}.$$

Applying (7) on the function φ given by (6), we obtain

(8) $$p_i = \frac{1}{n},\ \ 1 \le i \le n.$$

The extreme given by (8) is a maximum because

$$\left\{\begin{matrix} \frac{\partial^2 \varphi}{\partial {p_i}^2} = -n u_i, & 1 \le i \le n \\ \frac{\partial^2 \varphi}{\partial p_i \partial p_j} = 0, & i \ne j \end{matrix}\right\}.$$

However the above maximum holds for fixed utility distributions. To obtain a more general result, we consider the weighted Arithmetic, Geometric & harmonic Mean inequality (weighted AGM inequality) which is as follows.

Let $x_1, x_2, \dots, x_n$ be positive real numbers and let $\alpha_1, \alpha_2, \dots, \alpha_n$ be non-negative real numbers whose sum is 1. Then it follows that for $n \ge 2$

(9) $$\frac{1}{\sum_{i=1}^{n} \frac{\alpha_i}{x_i}} \le \prod_{i=1}^{n} {x_i}^{\alpha_i} \le \sum_{i=1}^{n} \alpha_i x_i \quad \left(\sum_{i=1}^{n} \alpha_i = 1\right)$$

with equality if all ${x_i}'$s are equal.

Taking logarithm on both sides in (9), we obtain

(10) $$-\log \sum_{i=1}^{n} \frac{\alpha_i}{x_i} \le \sum_{i=1}^{n} \alpha_i \log x_i \le \log\left(\sum_{i=1}^{n} \alpha_i x_i\right).$$

Now replacing α_i by p_i and x_i by ${p_i}^{u_i}$ in the above inequality, we obtain

(11) $$-\log\left(\sum_{i=1}^{n} {p_i}^{u_i+1}\right) \le H(P, U) \le \log\left(\sum_{i=1}^{n} {p_i}^{1-u_i}\right).$$

The above inequality gives a lower bound and an upper bound for the weighted information measure given by (3). Further if we take $u_1 = u_2 = \dots = u_n = 1$ in (11), we obtain

$$-\log\left(\sum_{i=1}^{n} {p_i}^2\right) \le H_n(P) \le \log n$$

which gives a lower & an upper bound for the well known Shannon entropy given by (1).

Further if we take $u_1 = u_2 = \dots = u_n = u$ and $p_i = \frac{1}{n},\ i = 1, 2, \dots, n$ in (11), we obtain

$$\log n - (1-u) \log n \le H(P, U) \le -\log n + (1+u) \log n$$

The above inequality gives a lower bound & upper bound for weighted entropy in case of fixed weights and uniform probabilities. Again consider the following refinements [6] of the weighted AGM inequality given by

$$\frac{1}{A} \sum_{i=1}^{n} \frac{\alpha_i (x_i - A)^2}{x_i + \max(x_i, A)} \le \log(A) - \log(G) \le \frac{1}{A} \sum_{i=1}^{n} \frac{\alpha_i (x_i - A)^2}{x_i + \min(x_i, A)},$$

$$\frac{1}{A}\sum_{i=1}^{n}\frac{\alpha_i}{x_i}\frac{(x_i-H)^2}{H+\max(x_i,H)} \le \log(G)-\log(H) \le \frac{1}{A}\sum_{i=1}^{n}\frac{\alpha_i}{x_i}\frac{(x_i-H)^2}{H+\min(x_i,H)}$$

where $A=\sum_{i=1}^{n}\alpha_i x_i$, $G=\prod_{i=1}^{n}x_i^{\alpha_i}$, $H=\frac{1}{\sum_{i=1}^{n}\frac{\alpha_i}{x_i}}$ denote the weighted arithmetic mean, weighted geometric mean and weighted harmonic mean of the x_i's. Now replacing α_i by p_i and x_i by $p_i^{u_i}$ in the above inequalities, we obtain

$$\left(\frac{1}{\sum_{i=1}^{n}p_i^{u_i+1}}\right)\sum_{i=1}^{n}\frac{p_i(p_i^{u_i}-\sum_{i=1}^{n}p_i^{u_i+1})^2}{p_i^{u_i}+\max\left(p_i^{u_i},\sum_{i=1}^{n}p_i^{u_i+1}\right)} \le \log\left(\sum_{i=1}^{n}p_i^{u_i+1}\right)-H(P,U) \le$$

$$\le \left(\frac{1}{\sum_{i=1}^{n}p_i^{u_i+1}}\right)\sum_{i=1}^{n}\frac{p_i(p_i^{u_i}-\sum_{i=1}^{n}p_i^{u_i+1})^2}{p_i^{u_i}+\min\left(p_i^{u_i},\sum_{i=1}^{n}p_i^{u_i+1}\right)},$$

$$\sum_{i=1}^{n}(p_i^{1-u_i})\frac{\left(p_i^{u_i}-\frac{1}{\sum_{i=1}^{n}p_i^{1-u_i}}\right)^2}{\frac{1}{\sum_{i=1}^{n}p_i^{1-u_i}}+\max\left(p_i^{u_i},\frac{1}{\sum_{i=1}^{n}p_i^{1-u_i}}\right)} \le H(P,U)+\log\left(\sum_{i=1}^{n}p_i^{1-u_i}\right)$$

$$\le \sum_{i=1}^{n}(p_i^{1-u_i})\frac{\left(p_i^{u_i}-\frac{1}{\sum_{i=1}^{n}p_i^{1-u_i}}\right)^2}{\frac{1}{\sum_{i=1}^{n}p_i^{1-u_i}}+\min\left(p_i^{u_i},\frac{1}{\sum_{i=1}^{n}p_i^{1-u_i}}\right)}$$

which gives a refinement of the inequality given by (11).

If we consider the function

$$\varphi(p_1, p_2, \dots, p_n;\ u_1, u_2, \dots, u_n) =$$
$$= -\sum_{i=1}^{n}\frac{u_i p_i}{\sum_{i=1}^{n}u_i p_i}\log\frac{u_i p_i}{\sum_{i=1}^{n}u_i p_i}+\lambda\left(\sum_{i=1}^{n}p_i-1\right)$$

the maximum entropy discrete probability distribution is obtained as

(12) $$\frac{u_i p_i}{\sum_{i=1}^{n}u_i p_i}=\frac{1}{n},\ 1\le i\le n.$$

Again if we consider the function

$$\varphi(p_1, p_2, \dots, p_n;\ u_1, u_2, \dots, u_n) = -\sum_{i=1}^{n}p_i\log p_i+\sum_{i=1}^{n}u_i p_i+\lambda\left(\sum_{i=1}^{n}p_i-1\right)$$

the maximum entropy discrete probability distribution is obtained as

(13) $$p_i=\frac{2^{u_i}}{\sum_{i=1}^{n}2^{u_i}},\ 1\le i\le n.$$

In fact we will now derive some inequalities which gives an upper and a lower bound for the weighted entropy defined by (5) with equality at the point given by (13). We proceed as follows:

Let $p_i, q_i > 0$, where $1 \le i \le n$ and $\sum_{i=1}^{n} p_i = 1 = \sum_{i=1}^{n} q_i$. Then the following estimates hold.

$$\sum_{i=1}^{n} \frac{q_i(q_i - p_i)^2}{(q_i)^2 + (\max(q_i, p_i))^2} \le \tag{14}$$
$$\le \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \le \sum_{i=1}^{n} \frac{q_i(q_i - p_i)^2}{(q_i)^2 + (\min(q_i, p_i))^2}.$$

The above inequality is a refinement of well known Shannon inequality given by $\sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \ge 0$.

Now replacing q_i by $\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}$ in the above inequality, we obtain

$$\sum_{i=1}^{n} \frac{\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} \left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} - p_i\right)^2}{\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}\right)^2 + \left(\max\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}, p_i\right)\right)^2} \le \log\left(\sum_{i=1}^{n} 2^{u_i}\right) - \widetilde{H}^*(P, U) \le$$
$$\le \sum_{i=1}^{n} \frac{\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} \left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} - p_i\right)^2}{\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}\right)^2 + \left(\min\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}, p_i\right)\right)^2}$$

which gives

$$\log\left(\sum_{i=1}^{n} 2^{u_i}\right) - \sum_{i=1}^{n} \frac{\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} \left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} - p_i\right)^2}{\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}\right)^2 + \left(\min\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}, p_i\right)\right)^2} \le \widetilde{H}^*(P, U) \le$$
$$\le \log\left(\sum_{i=1}^{n} 2^{u_i}\right) - \sum_{i=1}^{n} \frac{\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} \left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}} - p_i\right)^2}{\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}\right)^2 + \left(\max\left(\frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}, p_i\right)\right)^2}.$$

Equality holds in the above inequality if $p_i = \frac{2^{u_i}}{\sum_{i=1}^{n} 2^{u_i}}$ for each *i*.

3. Discussion

In the present work, we have obtained bounds on the weighted information measures given by (3), (4) and (5) using the Lagrange's multiplier method and some well known inequalities. The optimum obtained for the weighted entropy given by (8) coincides with the optimum for the Shannon entropy given by (1) if the weights are fixed. However we have obtained a more general result using the weighted AGM inequality when the weights associated with events are not fixed. .

Further the maxima given by (12) are obtained for the weighted entropy given by (4). But it seems very difficult to find a utility (weight) & a probability distribution which satisfy (12). The effect of weights seems more significant in a maximum given by (13) which is obtained for weighted entropy given by (5). Work on one parametric generalization of these results is in progress and will be reported elsewhere.

References

1. B e l i s, M., S. G u i a s u. A Quantitative – Qualitative Measure of Information in Cybernetic Systems. – IEEE Trans. on Inf. Theory, Vol. **14**, 1968, 593-594.
2. H a l l i w e l l, G. T., P. R. M e r c e r. A Refinement of an Inequality from Information Theory. – Journal of Inequalities in Pure and Applied Mathematics, Vol. **5**, 2003, No 1, Issue 3, 29-31.
3. K a p u r, J. N. New Qualitative – Quantitative Measures of Information. – Nat. Acad. Sci. Lett., Vol. **9**, 1986, No 2, 51-54.
4. K a p u r, J. N. On the Concept of Useful Information. – J. Org. Behav. Stat., Vol. **2**, 1985, No 3-4, 147-162.
5. L o n g o, G. Qualitative – Quantitative Measures of Information. New York, Springer-Verlag, 1972.
6. M e r c e r, P. R. Refined Arithmetic, Geometric and Harmonic Mean Inequalities. – Rocky Mountain Journal of Mathematics, Vol. **33**, No 4, 1459-1464.
7. S h a n n o n, C. E. A Mathematical Theory of Communication. – Bell Syst. Tech. Jour., Vol. **2**, 1948, 623-659.
8. M u n t e a n u, V., D. T a r n i c e r i u. Optimal Coding for Qualitative Sources on Noiseless Channels. The Annals of “Dunarea De Jos” University of Galati Fascicle II (Electrotechnics, Electronics, Automatic Control, Informatics), 2006 **http://www.ann.ugal.ro/eeai/index.html**