Trends Journal of Sciences Research (2015) 2(1):13-16 http://www.tjsr.org

Research Paper Open Access

Multiplicative Functions of Numbers Set and Logarithmic Identities. Shannon and factorial logarithmic Identities, Entropy and Coentropy

Yu. A. Kokotov

Agrophysical Institute, Sankt- Petersburg, 195220, Russia
*Correspondence: Yu. A. Kokotov(Kokotov.Yu@mail.ru)

Abstract The multiplicative functions characterizing the finite set of positive numbers are introduced in the work. With their help we find the logarithmic identities which connect logarithm of sum of the set numbers and logarithms of numbers themselves. One of them (contained in the work of Shannon) interconnects three information functions: information Hartley, entropy and coentropy. Shannon's identity allows better to understand the meaning and relationship of these collective characteristics of information (as the characteristics of finite sets and as probabilistic characteristics). The factorial multiplicative function and the logarithmic factorial identity are formed also from initial set numbers. That identity connects logarithms of factorials of integer numbers and logarithm of factorial of their sum.

Keywords: Mumber Sets, Multiplicative Functions, Logarithmic Identity, The Identity of Shannon, Information Entropy, Coentropy, Factorials

Citation: Yu. A. Kokotov. Multiplicative Functions of Numbers Set and Logarithmic Identities. Shannon and factorial logarithmic Identities, Entropy and Coentropy. Vol. 2, No. 1, 2015, pp.13-16

1. Introduction

The logarithmic scale is widely used in many areas of science, including information theory. In this theory the positive numbers logarithms are treated as Hartley information about a set divided into subsets.The widely used logarithmic characteristic is information entropy. The statistical entropy is also widely used in statistical physics (sometimes regarded from point of view of theory of information). In this regard, the following problem is interesting.

It is known that the logarithm of number considered as the sum is not divided into the sum of logarithms of her terms. Direct connection between these logarithms does not exist. However there are formulas that include logarithm of the sum and logarithms of all terms of sum. This suggests the existence of some causal relationship between them.

The work is devoted to analysis of this issue and its relationship with information theory.

Results and Discussion

Suppose we have a set of N positive numbers: m1 ... mN. This set may be characterized in different ways. So these numbers can simultaneously be viewed as terms the sum of which is equal to M, and as the factors forming the product, Π.

In information theory the logarithm of the sum, M, and the logarithms of its terms are considered as a measure of the information. These ones are respectively measure of information about the set as a whole (Hartley information, I0), and of information about each of the subsets (partial information, Ii). Therefore the relation between these values is important for the theory of information.

And the product and the sum can be considered as common integral characteristics of a set of numbers. It is evident that their ratio

$$\mu = \frac{\Pi}{M} \quad (1)$$

can also be seen as another its characteristic. It follows immediately from the above mentioned relations

$$\log M = \sum_{i=1}^{N} \log m_i - \log \mu \quad (2)$$

You can enter some other characteristics of set, which are convenient for taking the logarithm. Let's call them “multiplicative function of set”.

We may introduce the function formed from frequently used quantities: the geometric mean and the arithmetic mean:

$$\bar{\mu} = \frac{\sqrt[N]{\prod_{1}^{N} m_i}}{\frac{1}{N}\sum_{i=1}^{N} m_i} \quad (3)$$

This ratio is discussed usually in the statistics. Taking the logarithm, we obtain the identity:

$$\log M = \frac{1}{N}\sum_{i=1}^{N}\log m_i - \log\bar{\mu} + \log N; \quad (4)$$

The values of all N values m_i completely determine all quantities in these identities. Any set of terms in the sum of M can be transformed and represented in normalized form with weights of subsets, p_i:

$$p_i = \frac{m_i}{M}; \quad \sum_{i=1}^{N} p_i = 1; \quad (5)$$

The simplest characteristic of this set is another multiplicative function, product of weights:

$$\Psi = \prod_{1}^{N} p_i = \frac{\prod_{i=1}^{N} m_i}{M^N} \quad (6)$$

It follows another logarithmic identity, in fact, equivalent to eq. (4):

$$\log M = \frac{1}{N}\left(\sum_{i=1}^{N}\log m_i - \log\Psi\right); \quad (7)$$

We can introduce more complex multiplicative weight functions such as

$$\lambda = \prod_{i=1}^{N}\left(p_i\right)^{p_i} = \prod_{i=1}^{N}\left(\frac{m_i}{M}\right)^{\frac{m_i}{M}} = \frac{\prod_{i=1}^{N} m_i^{p_i}}{M}; \quad (8)$$

Logarithm of it is more complex logarithmic identity:

$$\log M = \sum_{i=1}^{N} p_i \log m_i - \log\lambda \quad (9)$$

This identity is equivalent to the equation:

$$\log M = \sum_{i=1}^{N} p_i \log m_i - \sum_{i=1}^{N} p_i \log p_i. \quad (10)$$

Eq. (9) and (10) can be written" in information form":

$$I_0(M) = J_N(M, N, p_1 ... p_{N-1}) + S_N(N, p_1 ... p_{N-1}) = \\ = \overline{I_i}(N, M, m_1 ... m_{N-1}) - \overline{\Lambda}_i(N, p_1 ... p_{N-1}) \quad (11)$$

where

$$I_0 = \log M \quad \text{(Hartley information)} \quad (12)$$

$$S_N = -\log\lambda = -\sum_{i=1}^{N} p_i \log p_i = \overline{\Lambda}$$

(Information entropy)(13)

$$J_N = \overline{I_i} = \sum_{i=1}^{N} p_i \log m_i \quad \text{(Coentropy)}^1 \quad (14)$$

This identity determines the information about the set as a whole as the difference between the two average values (arithmetic mean of the partial information about subsets and arithmetic mean of the logarithms of the weights of terms) or as sum of coentropy and entropy.

Entropy is usually considered in the theory of information as a common characteristic of set divided into subsets. The values *S* and *J* for a fixed value *M* are interrelated and change in the opposite direction.

Equation (10) was obtained on the basis of probability theory in the work of Shannon ([1], App. 2), as the mathematical foundation of intuitively introduced entropy. Information entropy is generally considered as closely related to the probability theory applied to a finite set of events. Until recently, eq. (10) was almost forgotten and was again got and used in the works Bianucci et al. [2] and Vyatkin [3]. It can be correct to name equation (10): Shannon identity.

It can be seen through the mathematical expectations of the corresponding quantities:

$$\log M = abs\left\langle \log p_i \right\rangle + \left\langle J_i \right\rangle; \quad (15)$$

In this case, the entropy becomes mathematical sense of the probability function of the random quantities ([4], Chapter 10).

However, the eq. (8) and functions *S* and *J* can be used for describing the properties of numerical sets and out of touch with the theory of probability. The function *J* named in [2] "coentropy" also proved important for the theory of information. Coentropy averages partial information about the subsets and therefore it is the proper measure of information about a particular set.

Entropy averages the information about weights, p_i , i.e. really only about relation of subsets in the given set. Both quantities are considered as the general characteristics of the set and as the collective characteristics of the information obtained from the experiment.

---

[1] This term is sometimes used in the literature in relation to different values, but not spread.
Proposed by Vyatkin [3] the term "negentropy" is not quite successful, because determined by it value differs from the actual "negative entropy" by a constant value.

In the probabilistic interpretation the weighting factors ("weights") are frequencies of events and are ttransformed to probabilities only in the limit case.

The values log M, log m_i and hence coentropy become meaningless at infinite M and m_i. Then the entropy is the only characteristic of a hypothetical infinite set of test results divided into N infinite subsets. All sets, finite or infinite, with the same relationship of terms, i.e. with the same structure, have the same entropy. Shannon identity determines the difference between information about all the set as a whole and information of the same set, divided into subsets. Obviously that the entropy is the difference of information between set as a whole and the same set divided into subsets [3]. In this way it determines the change of information in result of division of set into subsets. The usual interpretation of entropy as "the measure of uncertainty of results of the test having different probability" is much vaguer.

Note that identity Shannon can be written in a purely entropic form:

$$S_M = S_N + J_N = S_N + \sum_{i=1}^{N} p_i S_{m_i} = S_N + MO\left\langle S_{m_i} \right\rangle_N \quad (16)$$

where

$$S_M = \log M = -M\left(\frac{1}{M}\log\frac{1}{M}\right) \quad ;$$

$$S_{m_i} = I_{m_i} = \log m_i = -m_i\left(\frac{1}{m_i}\log\frac{1}{m_i}\right) \quad (17)$$

are entropy of sets M as whole (Hartley information) and entropy of subsets (partial Hartley information). Therefore, the entropy can be considered as general measure information of the whole set (before division into parts), of the same set but divided into parts and of each of set parts.

Information theory initially considered the integer set M divided into N integer subsets. If the relative placement of elements in the set and its subsets is not essential then the information is semantic.

Usually it is contested [1, 4] that the entropy of any set consisting of N parts (subsets) is maximum when all the weights, p_i and all parts of the set, m_i, are the same and equal to

$$S_N^{\max} = \log N \quad (18)$$

In this case coentropy is minimal and is equal to log (M/N). It is also believed that the minimal entropy set is zero.

However, the results are not accurate for finite integer sets. Indeed, the number of M can not be divided into N equal parts without integer remainder, R. Then the entropy of the integer set is maximized when the distribution of elements over the subset is the closest to the uniform. These sets are obtained by the distribution of R units of residue between any of N identical subsets consisting of Q units: It is clear that in this case there are different sets having the same maximum entropy:

$$S_{\max} = -R\frac{Q+1}{M}\log\frac{Q+1}{M} - (N-R)\frac{Q}{M}\log\frac{Q}{M}; \quad (19)$$

On the other hand, as shown in [3], the minimal entropy of integer set is attained at the most non-uniform distribution of elements across subsets (the all subsets except one contain only one element):

$$S_N^{min} = \log M - \frac{M-N+1}{M}\log\left(M-N+1\right); \quad (20)$$

In both cases, extreme value of coentropy is determined from the identity Shannon.

Note that the entropy can not be considered as characteristic of the individual set with a given structure. Entropy is ambiguous due his constituent, the function $y = p\log p$. It has the maximum at p =0.37. Contributions to the entropy of terms with the values p, greater and less than the 0.37 and, moreover, with a very large and very small values of p are equivalent, and in the latter case are small. The terms with p values in the neighborhood 0.37 give the largest contribution to the entropy. Therefore, very different distributions can have the same values of entropy. For the same reasons ambiguity is also characteristic for coentropy.

There may be introduced and other options of multiplicative functions, such as functions: $\prod_{i=1}^{N} q_i^{q_i}, \prod_{i=1}^{N} p_i^{q_i}, \prod_{i=1}^{N} q_i^{p_i}$, associated with values q = 1- p, sometimes used in probability theory. It follows logarithmic identities with components: $q_i \log q, \; q_i \log p_i, \; p_i \log q_i$.

The logarithmic function $q\log q$ is symmetric relative to functions $p\log p$ and has a maximum at the point p_{max} = 0.63, while the other logarithmic functions have no extreme. Function $p\log q$ tends to infinity at $p \to 1$ while the function $q\log p$ tends to infinity at $p \to 0$. These features are well reflected on a logarithmic scale, respectively, as an increase of reliability or unreliability of events.

Thus, there are many multiplicative functions and associated logarithm identities. The use of one or the other of them is related to the specific objectives of the work.

Logarithmic identities and information characteristics of finite sets can be used for the quantitative description of any number sets and without connect with theory of probability. Shannon's identity can be used, for example, to calculate the changes in the characteristics of information at combining the sets by various ways (by association, addition or intersection), as well as for the description of hierarchically organized systems. Really the partial Hartley information included into coentropy can be considered as proper information Hartley about sets of the following hierarchy level. This allows consistently applying the Shannon identity to sets of different levels of hierarchy, establishing links between their information characteristics.

In the statistical approach to the information initial integer set and its integer subsets are displayed by a set of permutations of their elements. The specific permutation of elements of original set and subsets are the elements of the new factorial information sets. The permutations are associated with the ordered arrangement in relation to each other pre-numbered specific elements of the original set and its subsets. The statistical information is no longer determined on Hartley but as the Boltzmann information, through permutation and factorials:

$$I^{0}_{(st)} = \log M!; \quad I_{i(st)} = \log m_i! \qquad (21)$$

and acquires syntactic character.

The usual formula for the information entropy can be written as

$$S_N = -\log \frac{\prod_{1}^{N} m_i^{\frac{m_i}{M}}}{M}; \qquad (22)$$

In the case of integer sets it should be compared with statistical definition of information entropy [5] in terms of factorials of M and m_i:

$$S_{st} = -\log \frac{\prod_{1}^{N} m_i!}{M!} = -\log \Phi \qquad (23)$$

The ratio

$$\Phi = \frac{\prod_{1}^{N} m_i!}{M!} \qquad (24)$$

can be regarded as factorial multiplicative function of initial set. It is the factorial analog of function μ. It is also completely determined by the initial separation of initial set into subsets and indirectly reflects this division.

Statistical entropy $S = -\log \Phi$ is a "factorial" analogue of function -log μ in the identity (3).

It follows immediately and logarithmic "factorial" identity:

$$\log M! = \sum_{i=1}^{N} \log m_i! - \log \Phi = J_{st} + S_{st}; \quad (25)$$

This identity is a "factorial" analogue of identity (2). It establishes a relationship between the logarithm of the factorial of the sum of M and the logarithms of factorials of its terms, m_i, formed from initial integer sum of the integer terms. This ratio is determined entirely by the division of the original sum to the terms. Logarithmic identities are not trivial as it may be seemed at first glance. All of them are related to the multiplicative functions of the original set of numbers and entirely determined by its properties. They allow you to specify the meaning of the basic concepts of information theory and can find specific application in the description of complex systems, including those ones with factorials.

Conclusions

1. The "multiplicative functions" introduced in this work reflect in different ways the properties of set numbers simultaneously considered as the terms of sum and as factors of product.
2. The different logarithmic identities arising from the multiplicative functions describe the indirect relationships between the logarithm of the sum of the numbers and the logarithm of its terms. Identities are interesting to information theory and other areas of science using a logarithmic scale.
3. The little known "probabilistic" Shannon equation applied to a finite set is one of the logarithmic identities. It combines three functions of information: information Hartley, entropy and "coentropy". Identity allows better to understand the meaning and properties of the collective characteristics of information and can be used in the description of complex systems.
4. Integer set divided into integer subsets can be displayed with a plurality of permutations of their numbered elements derived from the initial set. Numbers of elements of such a "statistical information set" are factorials of elements numbers of the original set. Logarithms of factorials are considered as a measure of statistical (and simultaneously syntactic) information.

Various "factorial" multiplicative function of the new set can be formed of factorials. Logarithm of the simplest of these ones is the statistical entropy. The "factorial" logarithmic identity includes the logarithms of factorials and the logarithm of the factorial terms of their sum. The connection between the logarithms of factorial established through statistical entropy and therefore through the factorial multiplicative functions.

References

[1] Shannon C.E. Mathematical Theory of Communications. Bell System Technical Journal 1948 v. 23.

[2] Bianucci D., Cattaneo G., Ciucci D. Information Entropy and Co-entropy of Crispand Fussy Granulation in “Application of Fuzzy Sets Theory”. Camogli, 2007.

[3] Vijatkin V.B. Synergetic information theory. Information technology: 2009, 12; Scientific Journal of Kuban Agronomic University: 2008, v.44, p. 10; 2009, v. 45, p. 10; 2009, v.47, p. 3.

[4] Ventzel E.S. Probability Theory. Science, Moscow. M. 1964.

[5] Tribus M. Thermostatic and Thermodynamics. Van Nostrand. N.Y. 1961.