



**HAL**  
open science

# Some invariant probability and entropy as a measure of uncertainty

Qiuping A. Wang

► **To cite this version:**

Qiuping A. Wang. Some invariant probability and entropy as a measure of uncertainty. 2007. hal-00117452v2

**HAL Id: hal-00117452**

**<https://hal.science/hal-00117452v2>**

Preprint submitted on 6 Apr 2007 (v2), last revised 9 Jan 2008 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Some invariant probability and entropy as a measure of uncertainty

Qiuping A. Wang

Institut Supérieur des Matériaux et Mécaniques Avancées du Mans, 44 Av. Bartholdi,  
72000 Le Mans, France

### Abstract

The relationship between some probability distributions and their invariant property is discussed. A measure  $I$  of uncertainty (informational entropy) of the probability distribution  $p(x)$  is defined in a variational way by  $dI = \overline{d\bar{x}} - \overline{d\bar{x}}$  which makes it possible to derive three entropy forms directly from the distribution laws widely observed in nature without postulating a priori the entropy properties such as additivity and extensivity.

PACS numbers : 05.20.-y; 02.50.-r; 02.50.Tt

## 1) Introduction

It is well known that entropy and information can be considered as measures of uncertainty of probability distribution. However, the functional relationship between probability distribution and the associated entropy has since long been a question in statistical and informational science. There are many relationships established on the basis of the properties of entropy. In the conventional information theory and some of its extensions, these properties are postulated, such as the additivity and the extensivity in the Shannon information theory. The reader can refer to the references [1] to [9] to see several examples of entropies proposed on the basis of different postulated entropy properties. Among all these entropies, the most famous one is the Shannon informational entropy ( $S = -\sum_i p_i \ln p_i$ ) [2] which was almost the only one widely used in equilibrium thermodynamics (Boltzmann-Gibbs entropy) and in nonequilibrium dynamics (Kolmogorov-Sinai entropy for example). Recently, a nonextensive statistics [6][8][9] (NES) proposed to use other entropies for thermodynamics and stochastic dynamics of certain nonextensive systems. NES has given rise to a large number of papers in the last decade with very different viewpoints dealing with equilibrium and nonequilibrium systems, which have incited more and more debates [10][11] within the statistical physics community. In the debates, we can distinguish two kinds of questions. The first one is about whether or not NES can be used for some given systems. The discussions are rather related to technical aspects of numerical simulation and of the definition of physical quantities such as temperature and pressure<sup>1</sup>. The second one is more fundamental and about the physical foundation of the new entropies: whether or not it is necessary to replace Boltzmann-Gibbs-Shannon entropy with other ones in different physical situation? what should be the possible and useful forms of entropy? The present work is an investigation of these questions under a different angle without postulating mathematical properties of entropy as usual. We use entropy as a measure of uncertainty of given probability distribution and introduce a variational definition of entropy inspired by a thermodynamic relationship between entropy and energy. On this basis, three entropy functionals are derived for some well known invariant probability distributions.

One of the objectives of this work is to show that, Shannon entropy is not unique as measure of uncertainty. Other forms are possible if we change the rules of reasoning, modify

---

<sup>1</sup> We would like to indicate here that the criticisms raised by R. Balian, M. Nauenberg and R. Luzzi *et al* about the violation of thermodynamic laws by NES due to the absence of temperature and pressure within the NES formalisms and about its failure for systems having different  $q$  [10] are flawed and misleading. A review of the results on these points can be found in [12].

the postulated property of entropy or introduce new criterions as we have done in this work in introducing the variational definition of entropy. This definition underlies a new criterion of optimization of entropy for corresponding stable probability distribution. We will come back to this point in the conclusion of this paper. We notice that this is rather a conceptual work tackling in a general way the mathematical form of entropy without considering the detailed physics behind the well known and observed distribution laws considered in the calculations.

## 2) Invariance and probability distribution laws

In this section, by some trivial calculations one can find in many textbooks, we want to show the invariant properties of some probability distributions. In what follows, by invariance of a function  $f(x)$ , we mean that the dependence on  $x$  is invariant at transformation of  $x$  into  $x'$ , i.e.,  $f(x') \propto f(x)$ . We will consider three different invariances corresponding to three distribution laws widely observed in nature, i.e., exponential, power law and q-exponential distributions.

### a) Translation invariance and exponential law

Suppose that  $f(x)$  is invariant by a translation of  $x \rightarrow x+b$ , i.e.

$$f(x+b) = g(b)f(x) \quad (1)$$

where  $g(b)$  obviously depends on the form of  $f(x)$ . We have

$$\frac{df(x+b)}{db} = \frac{df(x+b)}{d(x+b)} = g'(b)f(x) \quad (2)$$

Let  $b=0$ , we get

$$\frac{df(x)}{dx} = g'(0)f(x) \quad \text{or} \quad \frac{df(x)}{f} = g'(0)dx \quad (3)$$

which means

$$\ln f(x) = g'(0)x + c \quad \text{or} \quad f(x) = ce^{g'(0)x} \quad (4)$$

If  $f(x)$  gives a probability such as  $p(x) = \frac{1}{Z}f(x)$  where  $Z = \sum_x f(x)$ , the normalization condition

$\sum_x p(x) = 1$  will make  $p(x)$  strictly invariant versus the transformation  $x \rightarrow x' = x+b$ , i.e.,

$p(x') = \frac{1}{Z}f(x') = \frac{1}{Z}f(x) = p(x)$  since  $Z = \sum_{x'} f(x') = \sum_x f(x+b) = g(b)\sum_x f(x) = Zg(b)$ . This calculation also

holds for the following two examples.

*b) Scale invariance and power law*

Now suppose that  $f(x)$  is scale invariant, we have

$$f(bx) = g(b)f(x) \quad (5)$$

Where  $b$  is the scale factor of the transformation. We make following calculation

$$\frac{df(bx)}{db} = \frac{df(bx)}{d(bx)} x = g'(b)f(x). \quad (6)$$

Let  $b=1$ , we get

$$\frac{df(x)}{dx} x = g'(1)f(x) \quad (7)$$

which means

$$f(x) = c x^{g'(1)}. \quad (8)$$

This kind of laws is widely observed in nature for different dynamical systems such as language systems[13] and scale free networks[14] among many others[15]. The well known Levy flight for large  $x$  is a good example of power law with  $g'(1)=-1-\alpha$  where  $0<\alpha<2$ .

*c) The q-exponential and its invariance properties*

Here we would like to mention a probability functional which has attracted a lot of attention in the last years:

$$f(x)=c[1+a\beta x]^{\frac{1}{a}}. \quad (9)$$

where  $a$  and  $\beta$  are some constants. The Zipf-Mandelbrot law  $f(x) = c[1+x]^{-\alpha}$  observed in textual systems and other evolutionary systems[16] can be considered as a kind of  $q$ -exponential law. Another example of this law is the equilibrium thermodynamic distribution for finite systems in contact with a finite heat bath, where  $a$  can be related to the number of elements  $N$  of the heat bath and tends to zero if  $N$  is very large[17], which implies

$$f(x)=c[1+a\beta x]^{\frac{1}{a}} \xrightarrow{a \rightarrow 0} c e^{\beta x}.$$

This distribution is not a power law in the sense of Eq.(8). It has neither the scale invariance nor the translation invariance mentioned above. The operator on  $x$  that keeps  $f(x)$

invariant is a generalized addition  $x+_ab=x+b+a\beta xb$ , i.e.  $f(x+_ab)=c[1+a\beta(x+b+a\beta xb)]^{\frac{1}{a}}$   
 $=[1+a\beta b]^{\frac{1}{a}}c[1+a\beta x]^{\frac{1}{a}}=g(b)f(x)$  where  $g(b)=[1+a\beta b]^{\frac{1}{a}}$ .

### 3) A definition of entropy as a measure of dynamical uncertainty

Suppose we have a random (discrete) variable  $x_i$  with a probability distribution  $p_i=\frac{1}{Z}f(x_i)$  where  $i$  is the state index. The average of  $x_i$  is given by  $\bar{x}=\sum_i x_i p_i$  and the normalization is  $\sum_i p_i=1$ . The uncertainty in this probability distribution of  $x$  can be measured by many quantities. For example, the standard deviation  $\sigma$  or the variance  $\sigma^2=\overline{x^2}-\bar{x}^2$  can surely be used if they exist. A disadvantage of  $\sigma^2$  is that it may not exist for many probability distributions. Here we propose another measure which seems much general. This is a variational definition of entropy as a measure of uncertainty given by following relationship

$$dI=d\bar{x}-d\bar{x}=\sum_i x_i d p_i. \quad (10)$$

This choice of uncertainty measure has been in a way inspired by the first and second laws of thermodynamics in equilibrium statistical thermodynamics. Considering the definition of internal energy  $\bar{E}=\sum_i p_i E_i$  where  $E_i$  is the energy of the state  $i$  with probability  $p_i$ , we can

write  $\delta\bar{E}=\sum_i \delta p_i E_i + \sum_i p_i \delta E_i = \sum_i \delta p_i E_i + \overline{\delta E_i}$ . It can be proved that

$\overline{\delta E_i} = \sum_i p_i \delta E_i = \sum_j \left( \sum_i p_i \frac{\partial E_i}{\partial q_j} \right) \delta q_j$  is the work done to the system by external forces

$F_j = \left( \sum_i p_i \frac{\partial E_i}{\partial q_j} \right)$  where  $q_j$  is extensive variables such as volume, distance or electrical

polarization. According to the first law of thermodynamics, the quantity  $\sum_i \delta p_i E_i = \delta\bar{E} - \overline{\delta E_i}$

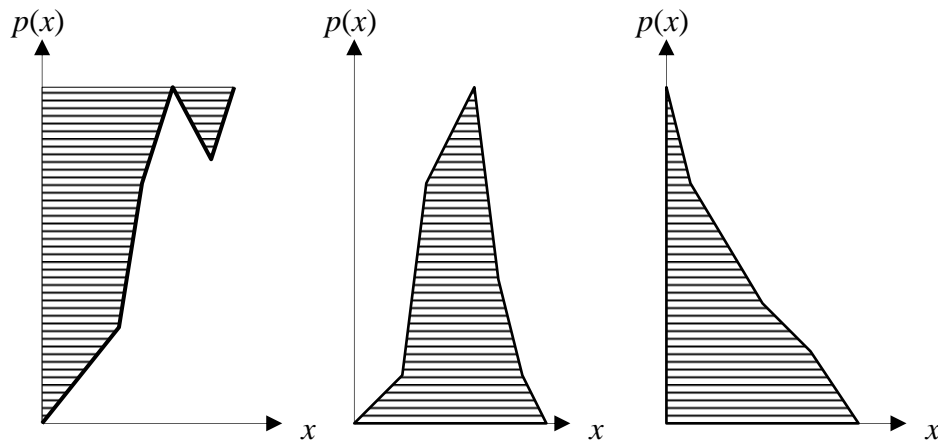
must be the heat change in the system, that is,  $\sum_i \delta p_i E_i = \delta Q = T \delta S$  for a reversible process,

where  $S$  is the thermodynamic entropy and  $T$  the absolute temperature. Hence the thermodynamic entropy must satisfy the following variational relation

$$\delta S = \frac{1}{T} (\delta\bar{E} - \overline{\delta E}). \quad (11)$$

This relationship is extended in Eq.(10) to arbitrary random variables  $x$ . By this definition, it is obvious that if the distribution is not exponential, the entropy functional may not be logarithmic.

The geometrical aspect of the uncertainty measure defined by Eq.(10) can be seen in the examples of Figure 1 which shows that  $dI$  and  $I$  are related to the width of the distributions on the one hand, and to the form of the distribution on the other.  $dI$  is not an increasing function of the distribution width. For example,  $dI=0$  for uniform distribution whatever the width of  $p(x)=\text{constant}$ . This means that  $I$  is a constant.



**Figure 1**, Three geometrical representations of the variation  $dI$  defined in Eq.(10) for some distributions. The hatched areas represent the value of the entropy variation  $dI = \overline{dx} - \overline{dx} = \sum_i x_i d p_i$  for each case. It is easily seen that  $dI=0$  if the width (uncertainty) of the distribution is zero or if the distribution is uniform.

#### 4) Entropy functional and probability distribution

In this section, on the basis of the uncertainty measure defined in Eq.(10), we will derive the entropy functionals for the three probability laws discussed in section 2.

##### a) Translation invariant probability and Shannon entropy

The following calculation is trivial. From Eq.(10), for exponential distribution  $p_i = \frac{1}{Z} e^{-x_i}$ , we have

$$dI = -\sum_i \ln(p_i Z) d p_i = -\sum_i \ln p_i d p_i - \ln Z \sum_i d p_i = -\sum_i \ln p_i d p_i = -d \sum_i p_i \ln p_i$$

We get

$$I = -\sum_i p_i \ln p_i$$

This is Shannon information or, within the conventional statistical mechanics, the Gibbs formula for Clausius entropy. Remember that the maximization of this entropy using lagrange multiplier associated with expectation of  $x$  yields exponential distribution law.

### *b) Scale invariant probability and entropy functional*

We have in this case power law probability distribution  $p_i$  :

$$p_i = \frac{1}{Z} x_i^{-a} \quad (12)$$

Put it into Eq.(10) to get

$$dI = \sum_i (p_i Z)^{-1/a} d p_i = Z^{-1/a} \frac{1}{1-1/a} d \sum_i p_i^{1-1/a} = -Z^{-1/a} d[\sum_i p_i^{1-1/a} / (1-1/a) + c] \quad (13)$$

where  $c$  is an arbitrary constant. We finally obtain

$$I \propto \sum_i p_i^{1-1/a} / (1-1/a) + c \quad (14)$$

In order to determine  $c$ , we imagine a system with two states  $i=1$  and 2 with  $p_1 = 0$  and  $p_2 = 1$ . In this case,  $I=0$  so that

$$\left( \frac{0+1}{1-1/a} \right) + c = 0 \quad (15)$$

i.e.,

$$c = -\frac{1}{1-1/a} \quad (16)$$

We finally get

$$I = \frac{\sum_i p_i^{1-1/a} - 1}{1-1/a} \quad (17)$$



Let  $q = \frac{1}{a}$ , we can write

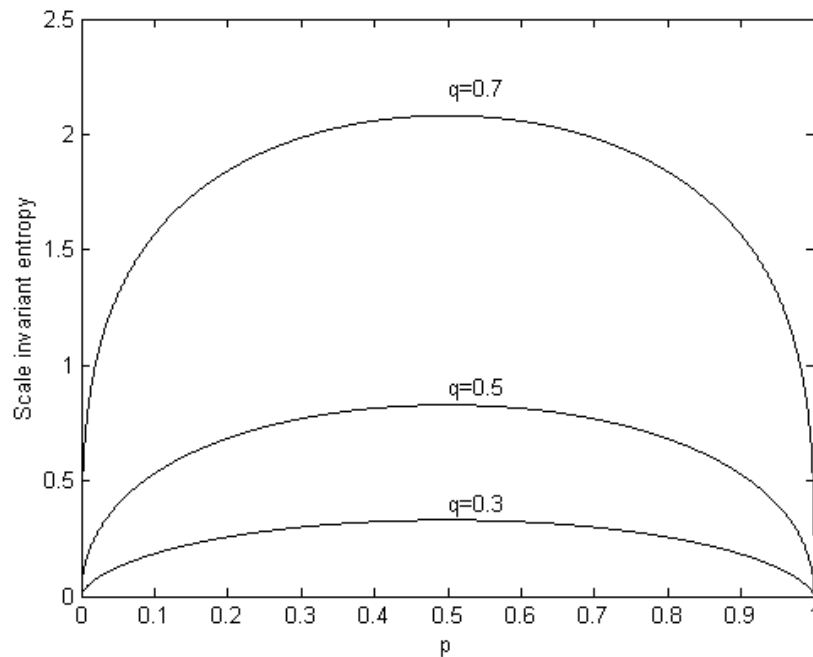
$$I = -\frac{1 - \sum_i p_i^{1-q}}{1-q} = -\sum_i \frac{p_i - p_i^{1-q}}{1-q} \quad (18)$$

Notice that this functional is different from the Tsallis entropy in  $q$ -logarithm (see below). It does not yield Shannon entropy when  $q \rightarrow 1$ . As a matter of fact,  $q$  must be positive and smaller than unity.  $I$  is negative if  $q$  is greater than unity or smaller than zero, which does not make sense. For large  $x$  Lévy flight for example,  $1 < a < 3$ , so  $\frac{1}{3} < q < 1$ .

It can be calculated that

$$I = -\sum_i \frac{p_i - p_i^{1-q}}{1-q} = -\sum_i p_i \frac{1 - p_i^{-q}}{1-q} = -\sum_i p_i \frac{1 - \frac{1}{Z^{-q}x}}{1-q} = -\frac{1 - Z^q \bar{x}}{1-q} \quad (19)$$

Its variation with probability value is shown in Figure 2. Its maximization with a Lagrange multiplier  $\beta$  such as  $\delta(I - \beta \bar{x}) = 0$  yields the power law of Eq.(12) with  $\beta = Z^q = Z^{1/a}$ .



**Figure 2**, The variation of the scale invariant entropy  $I = -\sum_{i=1}^2 \frac{p_i - p_i^{1-q}}{1-q}$  with  $p_1 = p$  and  $p_2 = 1-p$  for different  $q$  values. It can be shown that if  $q \rightarrow 0$ ,  $S = -\frac{q}{1-q} \sum p \ln p \rightarrow 0$ , and if  $q \rightarrow 1$ ,  $S = \frac{\sum 1 - 1}{1-q} + \sum \ln p \rightarrow \infty$ .

### c) The entropy for $q$ -exponential probability

We have seen above that the probability  $p_i = c[1 - a\beta x_i]_a^{\frac{1}{a}}$  had a special invariant property. Let us express  $x$  as a function of  $p_i$  and put it into Eq.(10) to get

$$dI = \sum_i \frac{1 - (p_i/c)^a}{a\beta} d p_i = -\frac{1}{a\beta c^a} \sum_i p_i^a d p_i = -\frac{1}{a\beta c^a (1+a)} d(\sum_i p_i^{1+a} + c) \quad (20)$$

By the same tricks for determining  $c$  in the above section, we get  $c=-1$ . So we can write

$$I = -\frac{\sum_i p_i^{1+a} - 1}{a} = -\sum_i \frac{p_i - p_i^q}{1-q} \quad (21)$$

Where  $q=1+a$  and we have used the normalization  $\sum_i p_i = 1$ . This is the Tsallis entropy which

tends to the Shannon entropy in the case of  $q \rightarrow 1$  or  $a \rightarrow 0$  where  $p_i = c[1 - a\beta x_i]_a^{\frac{1}{a}}$  tends to an exponential distribution.

## 5) Concluding remarks

We have derived three entropy functionals directly from three known probability distributions. This was done on the basis of a variational definition of uncertainty measure, or entropy without postulating entropy property (such as additivity) as in the usual information theory. The variational definition  $dI = d\bar{x} - \overline{dx}$  is valid for any probability distributions of  $x$  as long as it has finite expectation value. According to the results, the exponential probability

has Shannon entropy, the power law distribution has an entropy like  $I = -\frac{1 - \sum_i p_i^{1-q}}{1-q}$  where

$0 < q < 1$ , and the  $q$ -exponential distribution has Tsallis entropy  $I = -\sum_i \frac{p_i - p_i^q}{1-q}$  where  $q$  is positive.

It is worth mentioning again that the present definition of entropy as a measure of uncertainty offers the possibility of introducing the maximum entropy principle in a natural way with Lagrange multipliers associated with expectation of the random variables. It is easy to verify, with the above three entropies, that the maximum entropy calculus yields the original probability distributions. At first glance this seems an ordinary and fortuitous mutual

invertibility. But it is not, since the probability and the entropy are not reciprocal functions and the maximum entropy calculus is not a usual mathematical operation. As a matter of fact, this invertibility between entropy and probability resides in the variational definition  $dI = d\bar{x} - \overline{dx}$ . As discussed in the section 3,  $\overline{dx}$  can be considered as an extended work whatever the nature of  $x$ . So to get the “equilibrium state”, we can put  $\overline{dx}=0$  just as in the mechanical equilibrium condition where the vector sum of all forces acting on an object should be 0. We straightforwardly get  $dI - \beta \overline{dx} = 0$  or  $d(I + \alpha \sum_i p_i + \beta \overline{x}) = 0$  if we add the normalization condition. This is the usual maximum entropy principle using Lagrange multipliers  $\alpha$  and  $\beta$ . That is the entropy  $I$  defined by  $dI = d\bar{x} - \overline{dx}$  goes naturally to conditioned maximum for its corresponding probability distribution. This formalism of variational method has been applied to stochastic dynamics to develop an extended least action principle where we have postulated  $\overline{\delta A} = 0$  for Lagrange action  $A$  on the random paths[18].

## Acknowledgement

We thank Professor F. Tsohnang and Professor S. Abe for very useful discussion and comments.

## References

- [1] M.D. Esteban, *Kybernetika*, **31**(1995)337
- [2] C.E. Shannon, *The Bell System Technical Journal*, **27**(1948)379
- [3] A. Rényi, *Calcul de probabilit e*,(Paris, Dunod, 1966)P522  
A. Wehrl, *Rev. Mod. Phys.*, **50**(1978)221
- [4] J. Harvda and F. Charvat, *Kybernetika*, **3**(1967)30
- [5] Z. Daroczy, *Information and control*, **16**(1970)36
- [6] C. Tsallis, *J. Statis. Phys.*, **52**(1988)479
- [7] S. Abe, *Phys. Lett. A*, **271**(2000)74
- [8] G. Kaniadakis, *Phys. Rev. E*, **296** (2000)405
- [9] Q.A. Wang, *Chaos, Solitons & Fractals*, **12**(2001)1431;  
Q.A. Wang, *Entropy*, **5**(2) (2003)
- [10] R. Balian and M. Nauenberg, *Europhysics news*, **37/2** (2006)9;  
R. Luzzy, A.R. Vasconcellos and J.G. Ramos, *Europhysics news*, **37/2** (2006)11
- [11] See also the special issue of *Comptes Rendus Physique*, (2006) and the special issue of *Europhysics news*, **36/6** (2005)
- [12] Qiuping A. Wang, Laurent Nivanen, Alain Le Mehaute, Michel Pezeril, Temperature and pressure in nonextensive thermostatics, *Europhys. Lett.*,**65**(2004)606;  
Qiuping A. Wang, Laurent Nivanen, Alain Le Mehaute, A composition of different  $q$  nonextensive systems with the normalized expectation based on escort probability, (2006), cond-mat/0601255
- [13] W. Li, [Random texts exhibit Zipf's-law-like word frequency distribution](#), *IEEE Transactions on Information Theory*, **38**(6), (1992)1842-1845
- [14] M. E. J. Newman, [The structure and function of complex networks](#), *SIAM Review*, **45**, (2003) 167-256
- [15] M. E. J. Newman, [Power laws, Pareto distributions and Zipf's law](#), *Contemporary Physics* **46**, (2005) 323-351
- [16] Z. K. Silagadze, [Citations and the Zipf-Mandelbrot's law](#), physics/9901035

- [17] Y.P. Terletsii, Statistical physics, North-Holland Publishing Company, Amsterdam, 1971
- [18] Q.A. Wang, Non quantum uncertainty relations of stochastic dynamics, Chaos, Solitons and Fractals 26 (2005) 1045