Effective temperature of self-similar time series

A.Olemskoi, S.Kokhan
Sumy State University, Ukraine

Received July 18, 2005, in final form October 17, 2005

Within slightly non-extensive statistics and the related numerical model, a picture is elaborated to treat self-similar time series as a thermodynamic system. Thermodynamic-type characteristics relevant to temperature, pressure, entropy, internal and free energies are introduced and tested. The statistics developed is shown to be governed by the effective temperature being exponential measure of the fractal dimension of the time series. Testing of the analytical consideration is based on the numerical scheme of non-extensive random walk. Effective temperature is found numerically to show that its value is reduced to averaged energy per one degree of freedom.

Key words: time series, non-extensive statistics, simulations

PACS: 05.90.+m, 05.45.Tp, 05.40.Fb

1. Introduction

Time series analysis allows one to elaborate and verify macroscopic models of complex system evolution based on the data analysis [1]. This analysis is known to be focused on numerical calculations of the correlation sum for delay vectors which make it possible to find principal characteristics of the time series. Being traditionally a branch of the theory of statistics, time series analysis is based on the class of harmonic oscillator models which are related to the simplest case of the Gaussian random process. However, a well known real time series is relevant to the Lévy stable processes, rather than to the Gaussian ones being a very special case [2]. Since the former processes are invariant with respect to dilatation transformation, the problem is reduced to considering the self-similar stochastic processes.

The simplest characteristic of a time series is known to be a set of the Lyapunov exponents whose largest positive value yields a predictability domain in the system behaviour. The degree of complexity in such a behaviour is determined using the Kolmogorov-Sinai entropy that is equal to the sum over positive magnitudes of the whole set of the Lyapunov exponents and can be reduced to the usual Shannon value in the information theory. Transferring to the nonlinear system, the probability $p_n$
of $n$-th scenario of the system behaviour is transformed into the power function $p_n^q$ determined by index $q \leq 1$, so that the Kolmogorov-Sinai entropy should be replaced by the Renyi entropy

$$K_q \equiv \frac{1}{1-q} \ln \sum_n p_n^q.$$  (1)

Correspondingly, the master equation $d p / d \epsilon = -\beta p^q$, $\beta = \text{const} > 0$ describes the probability variation with energy $\epsilon = \epsilon_n$ to derive the Tsallis distribution $p(\epsilon) \propto [1 - (1 - q)\beta \epsilon]^{1-q^{-1}}$ [3]. In the limit $q \to 1$, this distribution assumes the usual Boltzmann form $p(\epsilon) \propto \exp(-\beta \epsilon)$ decaying exponentially fast contrary to the power asymptotic of the Tsallis exponent. Physically, such a behaviour is caused by a self-similarity of the non-extensive system [4].

As a result, the problem appears to study the self-similar time series that present the processes corresponding to Zipf-Mandelbrot power-law distribution. This work is devoted to both analytical and numerical considerations of this type of time series as a thermodynamic system. It is worth stressing that our approach is principally different to the pseudo-thermodynamic formalism developed for considering the multifractal type objects [6]. Indeed, if within the latter formalism, the role of parameters of state is played by the related multifractal indices, we introduce a set of effective parameters of state, whose meaning is of true thermodynamic type (hence, an effective temperature is a measure of data scattering related to the fractal dimension of the time series).

The paper consists of two main sections 2 and 3, the first of which is devoted to the analytical consideration and the second one – to numerical study. We start the analytical consideration with the elaboration of a model which allows us to address a self-similar time series as slightly non-extensive thermodynamic system. Then, we calculate the entropy, internal energy and temperature of the time series. We show that a temperature governing the time series statistics is an exponential measure of a self-similarity index related to the fractal dimension. The testing of the analytical consideration is based on numerical scheme of non-extensive random walk [7] whose stochastic equation and its solutions are treated in section 3. As a result, we obtain non-trivial time series whose form is governed by the friction coefficient that fixes the related fractal dimension. We introduce a statistical scheme that allows us to consider the modelled time series as a grand canonical ensemble for which we calculate the entropy and the internal energy as functions of the particle number. We find the numerically effective temperature and show that its value is reduced to the averaged kinetic energy per one particle of the ideal gas.

2. Analytical study of self-similar time series as a slightly non-extensive ideal gas

The principal peculiarity of time series is that it evolves during a large enough but finite time interval $T < \infty$. On the other hand, the self-similarity means that a
Effective temperature of self-similar time series

series is related to a fractal manifold characterized by the fractal dimension $D$. We show hereinafter that the both properties pointed out are taken into account in a natural way if we use non-extensive statistical mechanics where a simple combination of quantities $T$ and $D$ fixes the non-extensivity exponent $q$.

We start with considering the $d$-dimensional time series $x(t)$ related to the set \{\$x_i\}$ of consequent values $x_i \equiv x(t_i)$ of the principle variable $x(t)$ taken at a discrete time instant $t_i \equiv i\tau$ that we obtain as a result of dividing the whole time series length $T \equiv N_0\tau$ by $N_0$ equal intervals $\tau$. Following the ergodic hypothesis we shall imitate the time series $x(t)$ by a set of generalized coordinates \{\$x_n\}$ supplemented by a conjugated set \{\$v_n\}$ of velocities that show the jumping rates of coordinates $x_i$ with the time variation. It is principally important to take into account that the mapping of the time series $x(t)$ into the manifold \{\$x_n\}$ is not one-to-one correspondence because different terms $x(t_i)$ and $x(t_j)$ may be equal (similarly, this occurs at mapping the velocity time series $v(t)$ into the related manifold \{\$v_n\}$). This peculiarity is displayed first of all in that the number $N_0$ of specimens of the time series is much greater than the total number $N$ of the terms of the related manifold. Hereinafter, just the latter number $N$ plays the role of the particle number of the statistical system. Obviously, this system should be considered as grand canonical ensemble with variable number $N_n$ of particles in a state $n$.

Along the line of the ergodic hypothesis, the paradigm of our approach is to address the time series as a physical system defined by an effective Hamiltonian $H = H\{x_n, v_n\}$ based on which the statistical characteristics of this series could be found. If one proposes that the series terms $x_n$ related to different $n$ are not associated, then the effective Hamiltonian is additive:

$$H = \sum_{n=1}^{N} \varepsilon_n N_n, \quad \varepsilon_n \equiv \varepsilon(x_n, v_n).$$ (2)

Physically, this means that the series under consideration is relevant to an ideal gas comprising $N$ identical particles with energy $\varepsilon_n$ and number $N_n$ in state $n$. Further, we suppose different terms of the time series to be statistically identical, so that the effective particle energy does not depend on the coordinate $x_n$: $\varepsilon(x_n, v_n) \rightarrow \varepsilon(v_n)$. Moreover, since this energy does not vary with the inversion of the coordinate jumps $x_n - x_{n-1}$, the function $\varepsilon(v_n)$ should be even. We use the simplest square form

$$\varepsilon_n = \frac{1}{2}v_n^2,$$ (3)

which is reduced to the usual kinetic energy for a particle with mass 1. In the simplest case of Markovian consequence, the velocity is defined as follows:

$$v_n \equiv \frac{x_n - x_{n-1}}{\tau}.$$ (4)

To study the behaviour of the time series as a whole one needs to fulfill the summation over a set of states given by the manifold \{\$x_n, v_n\}$ that is relevant to the
system phase space. In so doing, it is convenient to pass to the related integrations as follows:

$$\sum_{\{x_n,v_n\}} \Rightarrow \int \int \prod_{n=1}^{N} \frac{dx_n dv_n}{N!\Delta} = \mathcal{N}^{-1} \prod_{n=1}^{N} \int \int dy_n du_n .$$  \hspace{1cm} (5)

Here, the factorial takes into account the statistical identity of the time series terms, $\Delta$ is the effective Planck constant that determines a minimal volume of the phase space per a particle related to a term. The inverted factor

$$\mathcal{N} \equiv N! \left( \frac{X^2}{\tau \Delta} \right)^{-dN} \simeq \left[ \frac{eX^{2d}}{N(\tau \Delta)^d} \right]^{-N}$$  \hspace{1cm} (6)

is caused by the change of variables $y_n \equiv x_n/X$, $u_n \equiv \tau v_n/X$ rescaled with respect to macroscopic length $X$ chosen to guarantee the conditions $\int dy_n = 1$, $\int du_n = 1$.

As explained in the introduction, the self-similarity condition leads to an application of Tsallis statistics. The last one is characterized by distribution $p_n$ whose tail decays as a power-law function. The distribution obeys the condition

$$\sum_n p_n^{\beta} \equiv \langle 1 \rangle_q \neq 1$$  \hspace{1cm} (7)

so that the definition of the internal energy reads:

$$E = \sum_n \frac{\varepsilon_n p_n^{\beta}}{\langle 1 \rangle_q} .$$  \hspace{1cm} (8)

To take into account these constraints, an escort distribution was proposed to use [8]

$$\mathcal{P}_n \equiv \frac{p_n^{\beta}}{\langle 1 \rangle_q} .$$  \hspace{1cm} (9)

In the explicit form it reads as follows:

$$\mathcal{P}_q \{y_n, u_n\} = \begin{cases} \frac{1}{Z} \left[ 1 - (1 - q) \frac{\mathcal{H}\{y_n, u_n\} - E}{\langle 1 \rangle_q T_s} \right]^{\frac{1}{1-q}} \text{ at } (1 - q) \frac{\mathcal{H}\{y_n, u_n\} - E}{\langle 1 \rangle_q T_s} < 1, \\ 0 \text{ otherwise.} \end{cases}$$  \hspace{1cm} (10)

Here, the partition function is defined by the condition

$$Z \equiv \mathcal{N}^{-1} \prod_{n=1}^{N} \int \int \left[ 1 - (1 - q) \frac{\mathcal{H}\{y_n, u_n\} - E}{\langle 1 \rangle_q T_s} \right]^{\frac{1}{1-q}} dy_n du_n ,$$  \hspace{1cm} (11)

where $0 < q < 1$ is a parameter of non-extensivity, $T_s$ is energy scale. Internal energy $E$ is determined by the equality

$$E \equiv \mathcal{N}^{-1} \prod_{n=1}^{N} \int \int \mathcal{H}\{y_n, u_n\} \mathcal{P}_q \{y_n, u_n\} dy_n du_n ,$$  \hspace{1cm} (12)

764
and normalization parameter $\langle 1 \rangle_q = Z^{1-q}$ is expressed by the partition function (11) in accordance with normalization condition.

To check the statistical scheme proposed let us first address the trivial case of time series $x_n = \text{const}$. Here, the particle energy $\varepsilon$ is a constant as well, so that the Hamiltonian is $H = N\varepsilon$. The partition function $Z = N^{-1}$ and the normalization parameter $\langle 1 \rangle_q = N^{-(1-q)}$ are given by the inverted normalization factor (6), while the internal energy $E = N\varepsilon$ is reduced to the Hamiltonian. Then, the entropy $H = -a \ln N$, $a = 1/2 \cdot (1 - q)dN \neq 0$ is reduced to zero if only the normalization factor takes the value $N = 1$. As a result, we find the effective Planck constant:

$$\Delta = \left( \frac{e}{N} \right)^{\frac{1}{2}} \frac{X^2}{\tau}. \quad (13)$$

Our future consideration is based on the assumption that the volume $V \equiv X^d$ of $d$-dimensional domain of the $x_n$ coordinate variation depending on the particle number $N$ is governed by Lévy-type law

$$X^d = x^d N^{\frac{1}{z}}. \quad (14)$$

Here, $x$ is a microscopic scale and a dynamic exponent $z$ is reduced to the fractal dimension $D$ of self-similar manifold [5]

$$z = D. \quad (15)$$

Then, one obtains the following scaling relation for the phase space volume per a term of the time series:

$$\Delta^d = e \left( \frac{\varepsilon^2}{\tau} \right)^d N^{\frac{2}{z} - 1}. \quad (16)$$

In the case of Gaussian scattering, when $D = 2$, the minimal volume $\Delta^d$ of the phase space does not depend on the number $N$ of the particles. Such a condition approves of our choice of the relation (14) for the whole volume $V \equiv X^d$ as the function of the number $N$ of time series terms.

Now we consider the main thermodynamic quantities of a time series represented as an ideal gas. In so doing, we are based on the expressions for the partition function (11) and the internal energy (12), whose explicit form has been obtained in [8–10]. The principal point of our approach is the entropy definition

$$H \equiv aK_q = a \ln Z = \frac{dN}{2} \ln \langle 1 \rangle_q, \quad a \equiv \frac{1}{2} (1 - q)dN, \quad (17)$$

that is chosen to guarantee the main thermodynamic relations. Within the limits

$$1 - q \ll 2/d, \quad N \gg 1 \quad (18)$$

one has

$$H \simeq \frac{Na}{2(1-a)} \ln \left[ e^{2+d} \left( \frac{2\pi T_s}{\Delta^2} \right)^d \left( \frac{X^d}{N} \right)^2 \right]. \quad (19)$$
Taking into consideration the scaling relation (14), this expression takes the usual form

\[ H = N \frac{D - 1}{D} \ln \left( \frac{G}{N} \right), \quad G \equiv (2\pi e T_s) \frac{aD}{\tau} \left( \frac{x}{\tau} \right)^{-dD} \]  

(20)

if the dynamic exponent is determined as

\[ z \equiv D = (1 - a)^{-1}. \]  

(21)

Respectively, the internal energy and the normalization parameter read:

\[ E = \frac{dN}{2} \left( \frac{G}{N} \right)^{\frac{2a}{D}} T_s, \quad \langle 1 \rangle_q = \left( \frac{G}{N} \right)^{\frac{2a}{D}}, \quad a \equiv \frac{D - 1}{D}. \]  

(22)

The temperature is defined as follows [9]

\[ T \equiv \langle 1 \rangle_q T_s = \left( \frac{G}{N} \right)^{\frac{2a}{D}} T_s, \]  

(23)

where the last equality takes into account the second of the relations (22). This definition guarantees the equipartition law

\[ E = CT, \quad C \equiv cN, \quad c \equiv d/2, \]  

(24)

where the quantity \( C = \frac{\partial E}{\partial T} \) is the specific heat. It is easy to prove that equations (20)–(23) arrive at standard thermodynamic relation \( \frac{\partial H}{\partial E} \equiv T^{-1} \).

Hereinabove we have considered the simplest model which makes it possible to examine analytically a self-similar time series in a standard statistical manner. A peculiarity of the related equalities is a scale invariance with respect to the variation of the non-extensivity parameter \( 1 - q \) which is contained everywhere through the parameter \( a \), given the last equation (17), that is related to the dynamic exponent \( z \) and the fractal dimension \( D \) of the time series according to equation (21). This invariance is clear to have been caused by self-similarity of the system under consideration.

The main progress in our consideration is that the time series statistics is governed completely by the temperature (23). Taking into consideration the second of the equalities (20) and rescaling the temperature unit \( T_s \) into \( T_{sc} \equiv (2\pi e)^a T_s \) we arrive at the expression

\[ \frac{T}{T_{sc}} = \left[ \left( \frac{\tau}{X} \right)^2 T_{sc} \right]^{\frac{a}{1-a}}. \]  

(25)

Being independent of the number of terms \( N \), the time series temperature shows the exponential dependence on the index \( a \) related to the fractal dimension according to equation (21). To establish the character of the power dependence on the ratio of the range \( X \) of the principle variable to the time interval \( \tau \), it is natural to choose
the measure units of the temperature in the following manner: $T_{sc} \equiv e^{(X/\tau)^2}$, $T_s \equiv e^{1-a(2\pi)^a(X/\tau)^2}$. Then, the expression (25) for the time series temperature takes the simplest form

$$T = \left(\frac{X}{\tau}\right)^2 e^D, \quad D \equiv \frac{1}{1-a}$$

(26)

according to which the value $T$ is the exponential measure of the fractal dimension $D$ of the self-similar time series.

3. Numerical study of time series

The purpose of this section is to numerically verify the definitions based on expressions (4)–(3), (10)–(12), (20)–(26). As the above analysis has shown, one of the peculiarities of the statistical system under consideration is that it is slightly non-extensive due to condition (18). Thus, we can put $q = 1$ in the following numerical consideration.

We shall follow the numerical scheme of non-extensive random walk based on the discrete stochastic equation [7]

$$x_{i+1} = \sqrt{\tau} \zeta_i + [(1 - \gamma \tau) + \sqrt{\tau} \xi_i] x_i.$$  

(27)

Here, discrete time $t_i = i \tau$ is fixed by integers $i = 0, 1, \ldots, N_0$ and minimal interval $\tau$; $\zeta_i$ and $\xi_i$ are additive and multiplicative stochastic sources normed with white-noise conditions $\langle \zeta_i \zeta_j \rangle = \langle \xi_i \xi_j \rangle = \delta_{ij}$; friction coefficient $\gamma$ determines the parameter $\nu = \gamma/(1+\gamma)$ which fixes, in accordance with stationary distribution, the non-extensivity parameter $q \equiv (2-\nu)^{-1} = (1+\gamma)/(2+\gamma)$. Making use of iteration procedure (27) we arrive at stochastic time series, which are explicitly relevant to the Lévy flights at friction coefficient $-1 < \gamma$, when the parameter $\nu$ is negative. Transferring to positive parameters $\gamma$, $\nu$, their growth arrives at gradual transformation of superdiffusion process into Brownian diffusion that is related to the magnitudes $\gamma = \infty$, $\nu = 1$. As $\nu$ increases, the relevant fractal dimension $D$ grows as well. Making use of the origin time series $x(t)$, it is easy to obtain the velocity time dependencies $v(t)$, in accordance with definition (4).

To study the statistical properties of the above time series we keep to the ergodic hypothesis that assumes the identity of averaging over both velocity time series $v(t_i)$, $i = 1, 2, \ldots, N_0$ and statistical ensemble $\{v_n\}$, $n = 1, 2, \ldots, N$, which describes the velocity scattering in the phase space. We determine such an ensemble dividing maximum interval of the velocity variation domain into $N \gg 1$ zones $n = 1, 2, \ldots, N$, within which the velocities have the mean value $v_n$ and small variation $\delta v_n$. Then, the probability $p_n$ to hit the interval $[v_n - \delta v_n/2, v_n + \delta v_n/2]$ and the relevant probability density function $\pi(v_n)$ are determined as follows:

$$p_n \equiv \frac{v_n}{N_0}, \quad \pi(v_n) \equiv N_0^{-1} \frac{v_n}{\delta v_n},$$  

(28)
Figure 1. The probability density function of the velocity distribution at different fractal dimensions (curves 1 – 5 correspond to parameters \( \nu = 0.85; 0.50; 0.35; 0.99; -0.10 \); respectively, fractal dimensions are \( \nu = 1.81; 1.58; 1.45; 2.57; 1.45 \)).

where \( \nu_n \) is a number of the time series specimens with the mean value \( v_n \). According to figure 1 the probability density function \( \pi(v) \) has the usual bell-shaped form centered at the velocity \( v = 0 \).

Since the distribution (28) corresponds to the grand canonical ensemble, the related thermodynamic functions are determined by the \( n \)-state number

\[
N_n \equiv Np_n = \frac{N}{N_0} \nu_n
\]  

and energy (3). Thus, the total energy of the ideal gas reads:

\[
E \equiv \sum_{n=1}^{N} \varepsilon_n N_n = \frac{N}{N_0} \sum_{n=1}^{N} \nu_n^2 \nu_n.
\]  

The plots of the dependencies \( E(N) \) of energy (30) accompanied with the related dependencies \( H(N) \) of entropy (17) are shown in figures 2 for different values of fractal dimensions \( D \). It is principally important, that the energy \( E \) should be directly proportional to the particle number \( N \), while the entropy \( H \) should increase much more slowly with \( N \). As for the velocity distribution \( \pi(v) \), non-monotonic variations of both energy and entropy as the functions of the fractal dimension take place, but the main tendency is the decrease of \( E \) and the increase of \( H \) with the growth of \( D \).

The peculiarity of the self-similar system under consideration is that it is finite and is characterized by the following effective temperatures:

\[
\Theta^{-1} \equiv \left. \frac{\partial H}{\partial E} \right|_D, \quad T^{-1} \equiv \left. N \frac{\partial H}{\partial E} \right|_N.
\]  

768
Effective temperature of self-similar time series

Figure 2. Dependencies of energy $E$ and entropy $H$ on the number $N$ of effective particles (curves 1 – 4 correspond to fractal dimensions $D = 1.50; 1.45; 1.58; 1.81$).

Being determined at a constant value of the fractal dimension $D$, the first of these is a function of the particle number $N$, whereas the second one depends on the magnitude $D$ to be determined at a fixed value $N$. According to the calculations, the above temperatures are associated by the following relation:

$$T = \frac{\Theta}{N} \left[1 - \frac{1}{(1-q)D^2} \left( \frac{H}{D-1} + \frac{D-1}{D^2}N\ln N \right)^{-1} \right]^{-1}. \quad (32)$$

Taking into account the definition (20) where $T_s$ is rescaled by factor $N$, one obtains

$$T = \frac{\Theta}{N} \left\{ 1 - \left[ D\ln \left( \frac{NT}{T_s} \right) + \frac{2(D-1)^2}{D} \ln N \right]^{-1} \right\}^{-1}. \quad (33)$$

Making use of the data presented in figure 2 we receive the dependence $\Theta(N)$ for the first temperature (31) shown in the left panel of figure 3. This temperature takes very large values which increase directly proportionally to the $N$ number growth. Such a behaviour displays an intermediate character of the quantity $\Theta$ that plays the role of an effective energy complementary to the energy (30). On the other hand, using the relation (33) we arrive at the dependencies of the temperature $T(N)$ shown in the right panel of figure 3. It is seen that with the increase of the fractal dimension to the values $D$ which are close to $D = 2$, the dependencies $T(N)$ approach the constant magnitudes. These dimensions relate to large values of the parameter $\nu \eq 1$ which is connected with non-extensivity index $q$ according to the equality $q = (2-\nu)^{-1}$. Thus, we find that the temperature $T$ takes the values which are non-dependent on the particle number $N$ in the region $q \approx 1$ that is relevant to a slightly non-extensive limit where the above analytical consideration is valid. Due to such a behaviour we can conclude that the temperature $T$ has a usual physical sense.
Figure 3. Dependencies of effective $\Theta$ and physical $T$ temperatures on the number $N$ of effective particles (curves 1 – 4 correspond to fractal dimensions $D = 1.50; 1.45; 1.58; 1.81$).

To confirm this conclusion we examine the equipartition law (24) rewritten using the averaged kinetic energy (3):

$$T = \langle \varepsilon \rangle \equiv \frac{\langle v^2 \rangle}{2}. \quad (34)$$

To this end we compare the values $\langle \varepsilon \rangle$ and $T$ at different magnitudes of the fractal dimension $D$ in figure 4. We can see that the difference between the above values does not exceed 10% if one does not take into account the points related to the three smallest magnitudes of $D$ where our analysis is not applicable.

Figure 4. Dependencies of the averaged kinetic energy $\langle \varepsilon \rangle$ and the physical temperature $T$ on the fractal dimension $D$. 
Effective temperature of self-similar time series

References

Ефективна температура самоподібних часових наборів

О.Олємськой, С.Кохан
Сумський державний університет, Україна


В рамках слабо неекстенсивної статистики та пов’язаної числової моделі, розробляється картина розгляду самоподібних часових наборів як термодинамічної системи. Вводяться та тестуються характеристики термодинамічного типу, що відповідають температурі, тиску, ентропії, внутрішній та вільній енергіям. Показано, що розвинута статистика підкоряється ефективній температурі, будучи експонентною мірою фрактальної розмірності часових наборів. Тестування аналітичного розгляду базується на числовій схемі неекстенсивних випадкових блукань. Чисельно знаходиться ефективна температура з метою показати, що її значення приводить до усередненої енергії на один ступінь вільності.

Ключові слова: часові набори, неекстенсивна статистика, симуляції

PACS: 05.90.+m, 05.45.Tp, 05.40.Fb