Generalized-generalized entropies and limit distributions

Stefan Thurner[1,2] and Rudolf Hanel[1]
[1] *Complex Systems Research Group; HNO; Medical University of Vienna; Währinger Gürtel 18-20; A-1090; Austria*
[2] *Santa Fe Institute; 1399 Hyde Park Road; Santa Fe; NM 87501; USA*

(Received on 13 January, 2009)

Limit distributions are not limited to uncorrelated variables but can be constructively derived for a large class of correlated random variables, as was shown e.g. in the context of large deviation theory [1], and recently in a very general setting by Hilhorst and Schehr [2]. At the same time it has been conjectured, based on numerical evidence, that several limit distributions originating from specific correlated random processes follow q-Gaussians. It could be shown that this is not the case for some of these situations, and more complicated limit distributions are necessary. In this work we show the derivation of the analytical form of entropy which –under the maximum entropy principle, imposing ordinary constraints– provides exactly these limit distributions. This is a concrete example for the necessity of more general entropy functionals beyond q statistics.

Keywords: Limit distributions, Correlated random numbers, Entropy for arbitrary distributions

1. INTRODUCTION

The importance of non-exponential distributions arising from an impressive number of complex statistical systems is by now beyond any doubt. Clearly, for most of these systems the classical approach of statistical mechanics is highly unjustified and bound to fail. Since many complex systems are statistical in nature it should, in some cases, be possible to keep a generalized H-theorem or a 'maximum entropy principle' (MEP), however with a modified entropy functional. It has been known now for the past two decades that, for variables following asymptotic power-laws, an associated maximum entropy principle can be formulated in terms of generalized entropies where the usual logarithm of the Boltzmann H-function gets replaced by the so called q-logarithm [3] [1]. However, in many complex systems distribution functions are observed which do not fit into the framework of the class of such generalized entropies. This calls for a more general setting, which has been explored lately from several distinct directions. In [5] a more general 3-parameter logarithm was derived on the basis of the MEP, applicable to a special type of distribution functions. Also based on the MEP a very general form was derived in [6–8], suitable for practically all reasonable piecewise monotonous distribution functions. In [9, 10] information theoretic aspects have led to the same class of entropy functionals, as is the case for [11, 12], where a very similar entropic form was derived on convexity arguments. For these entropies we showed that for every reasonable class of distribution functions it is not only possible to derive a unique entropy functional, which guarantees not only the correct results under the maximum entropy principle [6], but it is also fully compatible with the first and the second law of thermodynamics [13]. The necessity of these entropy formulations can be demonstrated in several important examples, some of which are discussed in [14]. Since these entropies can be seen as a generalization of *previous generalizations* of Boltzmann-Gibbs entropies we therefore refer to them in the following as *generalized-generalized* entropies, S_{gg}.

Classical statistical mechanics is tightly related to the central limit theorem (CLT). From the later the Maxwell-Boltzmann distribution immediately follows if one assumes the velocity of gas particles being the result of $N \to \infty$ random collisions with other particles, given the velocity distribution being centered, stable and having a second moment. Complementary to this mathematical approach the same distribution can be derived from a *physical* principle (MEP), where Boltzmann's H-function is maximized under the constraint that average kinetic energy is proportional to a measurable constraint, the temperature k_BT. While the above is obvious for gases with short range interactions and for independent random numbers, as soon as correlations are considered things become more involved on both sides. Limit theorems for correlated random numbers have lately regained considerable interest [15–19]. Recently a general and transparent way of constructing limit distributions of such continuous random numbers was reported for the class of one dimensional exchangeable processes [2], i.e. all N-point distribution functions are totally symmetric. This is relevant for example if systems behave as if they were mean field. For many physical, biological, or sociological systems, generically only a small fraction of the relevant microscopic system variables and the interactions between them are known or accessible to measurement. Assume, e.g. a Hamiltonian system where only the momentum variables are observable. Interactions between the unknown state variables, which can be treated as mean field, will cause observable, totally symmetric correlations between the momentum variables. Similarly, different shapes of nails used on Galton's Board, a physical model usually demonstrating Gaussian limit distributions, introduce correlations of the binary process, i.e. bias the chances of a ball to bounce to the left or the right, producing non Gaussian limit distributions. It is interesting to note, that one of the recently proposed limit distributions leads to q-Gaussian [19, 20], while many others explicitly lead to more complicated distributions [2, 15, 18].

Following recent work [8] we show that limit distributions and the MEP can be brought into a consistent framework also for *correlated* variables. By doing so we show how limit theorems for strongly correlated random variables – as recently introduced by Hilhorst and Schehr [2] – make a further level of generalization of entropy functionals necessary, which then

[1] Note that the traditional Tsallis entropy replaces the ordinary logarithm with the *dual* q-logarithm. For the definition of the dual, see [4].

is also able to explicitly take scaling relations in the distribution functions into account. We start by reviewing the derivation of the entropy functional necessary for arbitrary distribution functions following [8], and review limit distributions for sums of correlated random numbers, following [2]. Finally, we give the explicit form of the entropy leading to limit distributions used in [2].

2. GENERALIZED-GENERALIZED ENTROPY

Maybe the simplest way to derive the most general form of entropies, consistent with the maximum entropy condition, is by the following argument. Any (generalized or not) maximum entropy method, given the existence of some arbitrary stationary distribution function, $\tilde{P}(z_i)$, is formulated as

$$\frac{\delta G}{\delta P(z_i)} \mid_{P=\tilde{P}} = 0 \tag{1}$$

with

$$G \equiv S_{gg}[P] - \alpha \left\{ \sum_i P(z_i) - 1 \right\} - \beta \left\{ \sum_i f(z_i) P(z_i) - U \right\} \quad , \tag{2}$$

where α and β are Lagrange multipliers, z_i are state variables indexed by state i, and U denotes the expectation of function f, which depending on the problem may be a particular moment of z. Suppose that the distribution function is known (e.g. by experiment), and is parametrized as

$$P(z_i) = \frac{1}{\zeta} \mathcal{E}(-\alpha - \beta f(z_i)) \quad , \tag{3}$$

where $\mathcal{E}(x)$ is a *generalized exponential* [2].
The most general Ansatz for an entropy is that S_{gg} is some additive functional of the probability distributions

$$S_{gg} = -\sum_i L(P(z_i)) \quad . \tag{4}$$

Variation $\frac{\partial G}{\partial p_i} = 0$ directly leads to $-L'(P(z_i)) - \alpha - \beta f(z_i) = 0$, or $L(P(z_i)) = \int_0^{P(z_i)} dx \Lambda(\zeta x)$, meaning that the entropy has to be

$$S_{gg} = -\sum_i \int_0^{P(z_i)} dx \Lambda(\zeta x) \quad . \tag{5}$$

$P(z_i)$ is a normalized distribution function of some parameter set z [3], which could be e.g. energy or velocity. $\Lambda(x)$ is integrable in each interval $[0, P(z_i)]$. It can be seen as a *generalized logarithm* satisfying $\Lambda(x) < 0$ and $d\Lambda(x)/dx > 0$ for $0 < x < 1$, and $x\Lambda(x) \to 0$ $(x \to 0+)$, making $S_{gg}[P]$ non-negative and concave. Λ is the inverse function of $\mathcal{E}$: $\mathcal{E}(\Lambda(x)) = \Lambda(\mathcal{E}(x)) = x$. In other words, $\Lambda(x)$ needed for the entropy in Eq. (5) is chosen as the inverse function of the stationary (observed) distribution function. c is a constant, which ensures that $S_{gg}[P_0] = 0$ for a completely ordered state, i.e. $c = -\int_0^1 dx \Lambda(x)$. Through Ansatz (4) together with $S_{gg}[P_0] = 0$, S_{gg} is uniquely determined according to its exponential family $\mathcal{E}$, see also [21].

For the discussion below note that the scaling factor ζ in the argument of Λ generates distributions of the form of Eq. (3), where ζ and α can be seen as two alternative normalization parameters, tightly related to each other. The condition $\zeta = 1$ leads to a normalization of the distribution as typically discussed in [6, 7, 9–12]. It has been shown explicitly that the first and second laws of thermodynamics are robust under this generalization of entropy for $\zeta = 1$ in [13]. Condition $\alpha = 0$ leads to a normalization of P where ζ plays the role of the *partition function*. This case is of importance e.g. for [8].

3. LIMIT DISTRIBUTIONS OF CORRELATED RANDOM NUMBERS

Hilhorst and Schehr [2] consider a totally symmetric correlated Gaussian N-point process

$$P_N(\mathbf{z}) = \frac{e^{-\frac{1}{2}\left(z, M^{-1} z\right)}}{\sqrt{(2\pi)^N \det(M)}} \quad , \tag{6}$$

with $\mathbf{z} = (z_1, \ldots, z_N)$ and M the covariance matrix. This stochastic process is used as a reference process for some other totally symmetric N-point distribution $\tilde{P}_N(u)$ which is related to $P_N(z)$ by a simple transformation of variables $u_i = h(z_i)$, for all $i = 1, \ldots, N$. Total symmetry dictates the form of the covariance matrix

$$M_{ij} = \delta_{ij} + \rho(1 - \delta_{ij}) \quad , \tag{7}$$

for $\rho \in (0, 1]$ with the inverse, $M^{-1}_{ij} = a\delta_{ij} - b(1 - \delta_{ij})$, where $a = \frac{1+(N-2)\rho}{(1-\rho)(1+(N-1)\rho)}$, and $b = \frac{\rho}{(1-\rho)(1+(N-1)\rho)}$. Note that the symmetry requirement condenses all information on correlation into the single parameter ρ. A short calculation yields that the marginal probability is a unit variance Gaussian,

$$P_1^{\text{Gauss}}(z_1) = \int dz_2 \ldots dz_N \, P_N(z_1, z_2, \ldots, z_N) \quad , \tag{8}$$

[2] For the exponential function both parameters α and ζ which can be used to normalize the distribution P can be combined into a single normalization constant $Z = \zeta \exp(\alpha)$, the partition function. For generalized exponentials $\mathcal{E}$ it is not *a priori* clear how normalization should be performed. Usually normalization of non-exponential distributions – with the q-exponential distributions as probably the most prominent example – is performed by means of the Lagrange multiplier α. This choice is compatible with thermodynamic reasoning and produces no pathologic results. The situation gets much more difficult when ζ is used for normalization and is not kept constant. This then raises questions including thermodynamic stability, potential phase-transitions in the parameter-space and the physical interpretation of quantities. A detailed discussion is beyond the scope of this paper and will be presented elsewhere. However, here we use the maximum entropy principle to describe distribution functions of averages of continuous correlated random variables. With foresight of Eq. (16) we see that *multiplicative* normalization is required when studying sets of correlated continuous random variables under scaling transformations. Inverse temperature β then basically is the parameter fixing the scale of the system. The reason for introducing ζ and its implicit β dependence is therefore not arbitrary but a strict consequence of the stochastic model.

[3] For continuous variables z replace $\sum_i \to \int dz$.

which allows to construct the set of variables u_i from z_i via the transformation of variables

$$\int_0^{u_i} du'\, P_1(u') = \int_0^{z_i} dz'\, P_1^{\text{Gauss}}(z') \quad . \tag{9}$$

P_1 is the one-point distribution of the u variables. Consequently, a function h, describing the variable transformation

$$u_i = h(z_i) \quad , \tag{10}$$

is defined through Eq. (9). The distribution of the average of the u variables,

$$\bar{u} = \frac{1}{N}\sum_{i=1}^{N} u_i \quad , \tag{11}$$

is thus found in terms of an integration over all z_i

$$\mathcal{P}(\bar{u}) = \int d\mathbf{z}\, P_N(\mathbf{z})\, \delta\left(\bar{u} - \frac{1}{N}\sum_{i=1}^{N} h(z_i)\right) \quad , \tag{12}$$

where $d\mathbf{z} = dz_1 \ldots dz_N$. After some calculation one arrives at the general result [2],

$$\mathcal{P}(\bar{u}) = \left(\frac{1-\rho}{2\pi\rho}\right)^{\frac{1}{2}} |k'(\nu_*(\bar{u}))|^{-1} \exp\left(-\frac{1-\rho}{2\rho}[\nu_*(\bar{u})]^2\right) \quad , \tag{13}$$

where ν_* is defined as the zero of the function

$$k(\nu) = \bar{u} - \frac{1}{\sqrt{2\pi}}\int dw\, e^{-\frac{w^2}{2}} h\left((w+\nu)\sqrt{1-\rho}\right) \quad , \tag{14}$$

and $k'(x) = d/dx\, k(x)$. For symmetric one-point distributions P_1, h and ν_* are both antisymmetric. Moreover it is seen that $\nu'_*(\bar{u}) = -k'(\nu_*(\bar{u}))^{-1} \geq 0$, so that

$$\mathcal{P}(\bar{u}) = \frac{1}{2}\frac{d}{d\bar{u}}\text{erf}\left(\sqrt{\frac{1-\rho}{2\rho}}\nu_*(\bar{u})\right) \quad . \tag{15}$$

4. ENTROPY FOR LIMIT DISTRIBUTIONS

All that remains to be done is to identify the distributions $\mathcal{P}$ and P from Eqs. (13) and (3). Since limit distributions in [2] are centered and symmetric, the first moment provides no information, and f from Eq. (2) is taken to be the second moment, $f(z) = z^2$. The one point distribution P_1 has fixed variance and so does $\mathcal{P}(\bar{u})$ in Eq. (13), whereas distributions obtained through the MEP, Eq. (3), scale with a function of the "inverse temperature", β. To take this into consideration for the identification of P and $\mathcal{P}$ a simple scale transformation $\bar{u} = \lambda z$ is needed, where $\lambda(\beta)$ depends explicitly on β. Consequently, $\mathcal{P}(\bar{u}) \to \lambda\mathcal{P}(\lambda z)$ and

$$\lambda\mathcal{P}(\lambda z) = \frac{1}{\zeta}\mathcal{E}\left(-\alpha - \beta z^2\right) \quad . \tag{16}$$

It can be shown that this particular identification and the independence of Lagrange multipliers requires the normalization condition $\alpha = const.$ for the limit distribution Eq. (3). Without loss of generality we choose $\alpha = 0$ [4]. To determine ζ and λ the two conditions valid for generalized exponentials functions are used, $\mathcal{E}(0) = 1$ and $\mathcal{E}'(0) = 1$. This leads to

$$\zeta^{-1} = \lambda\left(\frac{1-\rho}{2\pi\rho}\right)^{\frac{1}{2}} \nu'_*(0) \quad \text{and} \quad \lambda = \gamma\sqrt{\beta} \quad , \tag{17}$$

with $\gamma^2 \equiv \frac{2\rho\nu'_*(0)}{\nu'_*(0)^3(1-\rho)-\nu'''_*(0)\rho}$ [5]. The generalized exponential can now be identified as

$$\mathcal{E}(x) = \frac{\nu'_*(\gamma\sqrt{-x})}{\nu'_*(0)}\exp\left(-\frac{1-\rho}{2\rho}[\nu_*(\gamma\sqrt{-x})]^2\right) \quad , \tag{18}$$

where we have substituted $x = -\beta z^2$ in Eq. (16) . This uniquely defines $\mathcal{E}$ on the domain $(-\infty, 0]$. Finally, the generalized logarithm Λ is uniquely defined on the domain $(0,1]$ as the inverse function of $\mathcal{E}$ and can be given explicitly for specific examples.

Example: In [2] the special case of a block function $P_1(u_j) = 1$ for $-\frac{1}{2} \leq u_j \leq \frac{1}{2}$ was discussed, implying

$$\begin{aligned} k'(\nu) &= -\frac{\kappa e^{-\kappa^2\nu^2}}{\sqrt{\pi}} \quad , \quad \kappa \equiv \left(\frac{1-\rho}{2(2-\rho)}\right)^{\frac{1}{2}} \\ k(\nu) &= \bar{u} - \frac{1}{2}\text{erf}(\kappa\nu) \quad , \quad \nu_* = \kappa^{-1}\text{erf}^{-1}(2\bar{u}). \end{aligned} \tag{19}$$

Thus the limit distribution Eq. (13) becomes

$$\mathcal{P}(\bar{u}) = \left(\frac{2-\rho}{\rho}\right)^{\frac{1}{2}} \exp\left(-\frac{2(1-\rho)}{\rho}[\text{erf}^{-1}(2\bar{u})]^2\right) \quad . \tag{20}$$

The block function has been used earlier [18] where it was conjectured on numerical evidence that the limiting distribution would be a q-Gaussian. This is obviously ruled out by Eq. (20), however, actual discrepancy is small, see Fig. 1. For this example Eq. (18) becomes

$$\mathcal{E}(x) \equiv \exp\left(-(\pi\gamma^2)^{-1}\left[\text{erf}^{-1}(2\gamma\sqrt{-x})\right]^2\right) \quad , \tag{21}$$

where $\gamma = \sqrt{\rho/(2\pi(1-\rho))}$. The associated generalized logarithm is given on the domain $(0,1]$ by

$$\Lambda(x) = -\left[(2\gamma)^{-1}\text{erf}\left(\gamma\sqrt{-\pi\ln x}\right)\right]^2 \quad . \tag{22}$$

The difference when compared to q-logarithms in Fig. 1 is small but visible.

---

[4] Choosing $\alpha \neq 0$ has the consequence that – due to the form invariance $\mathcal{P}(z) \to \lambda\mathcal{P}(\lambda z)$ under scaling transformation $z \to \lambda z - \alpha$ effectively has to be subtracted again and renders no consequence.

[5] Here we explicitly see how Eq. (16) requires the introduction of the multiplicative normalization term ζ and how its β dependency follows as a consequence. The introduction of ζ therefore is indeed a consequence of the model and *not* an ad hoc invention.

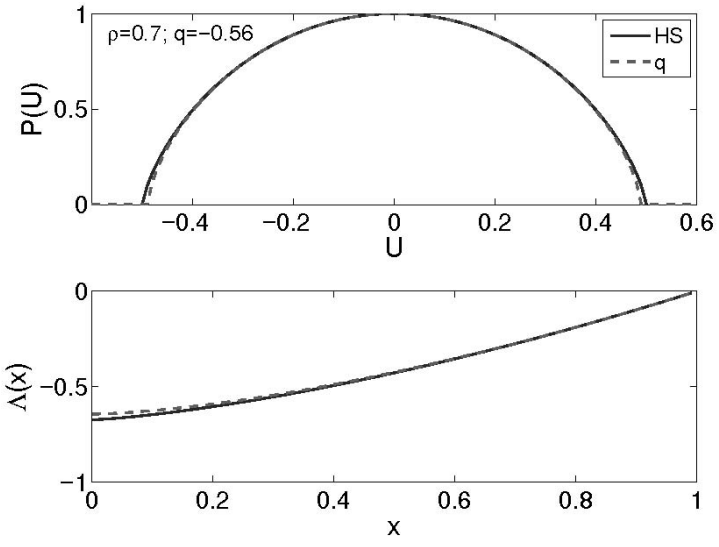

FIG. 1: Top: Limit distributions *al la* Hilhorst-Schehr for $\rho = 0.7$ (lines). Broken lines are q-Gaussians with a q from the reported best-fit value of $q = \frac{1-\frac{5}{3}\rho}{1-\rho}$, [2]. Bottom: Generalized logarithms needed to derive Hilhorst-Schehr distribution functions under the maximum entropy principle (lines). Broken lines are q-logarithms $\ln_q(x) = (x^{1-q} - 1)/(1-q)$ for the same q values.

5. DISCUSSION

Correlations between observable sets of microscopic state variables of physical systems can have considerable effect on the distribution of associate macroscopic variables, i.e. averages. Such correlations may enter physical processes e.g. in biased subsequent decision making (e.g. Galton's Board with biased nails) or through disregarded, unobservable, or unknown subspaces of microscopic state variables of a system affecting the observable state variables (e.g. spacial mean field interactions affecting momentum distributions). Chronologically, based on numerical evidence [18] it was conjectured that several [15, 18] limit distributions of sums of correlated random numbers, such as in Eq. (8), are q-Gaussians, which would have established a direct connection to Tsallis entropy. In [2] it was shown for these examples that this is not the case. As an explicit example for the necessity of generalized entropies extending beyond traditional q-statistics, in this work we have treated the question that if Tsallis entropy does not lead to the exact limit distributions under the MEP, which entropy functional does? The importance of this question gains momentum by the fact that different choices of the function h in Eq. (10) may lead to distributions that substantially deviate from q-Gaussians, much more drastically than for the case of the bloc function example shown in Fig. 1. We constructively answered the question by building on a recently proposed generalization of q-generalized entropy [6], which allows to treat scaling relations within generalized families of exponential functions [8].

Acknowledgements

We thank J. Naudts for clarifying discussions. Supported by Austrian Science Fund FWF Projects P17621 and P19132.

[1] E. Bertin, M. Clusel, J. Phys. A: Math. Gen. **39**, 7607-7619 (2006).
[2] H.J. Hilhorst, G. Schehr, J. Stat. Mech., P06003 (2007).
[3] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[4] J. Naudts, Physica A **316**, 323 - 334 (2002).
[5] G. Kaniadakis, Phys. Rev. E **66**, 056125 (2002).
[6] R. Hanel, S. Thurner, Physica A **380**, 109-114 (2007).
[7] R. Hanel, S. Thurner, in Complexity, Metastability, and Nonextensivity, AIP **965** 68-75 (2007).
[8] S. Thurner, R. Hanel, arXiv:0804.3477, cond-mat.stat-mech
[9] J. Naudts, Rev. Math. Phys. **16** 809 (2004).
[10] J. Naudts, Journal of Inequalities in Pure and Applied Mathematics **5** 102 (2004).
[11] S. Abe, J. Phys. A: Mathematical and General **36** 8733-8738 (2003).
[12] P.H. Chavanis, Eur. Phys. J. B **62**, 179208 (2008).
[13] S. Abe, S. Thurner, Europhys. Lett. **81**, 10004 (2008).
[14] J. Naudts, Contribution to Sigma Phi, Crete 2008.
[15] L.G. Moyano, C. Tsallis, M. Gell-Mann, Europhys. Lett. **73**, 813-819 (2006).
[16] S. Umarov, C. Tsallis and S. Steinberg, Milan J. Math. **76** (2008).
[17] F. Baldovin, A.L. Stella, Phys. Rev. E **75**, 020101(R) (2007).
[18] W. Thistleton, J.A. Marsh, K. Nelson and C. Tsallis, unpublished.
[19] A. Rodriguez, V. Schwämmle, C. Tsallis, JSTAT P09006 (2008).
[20] R. Hanel, S. Thurner, C. Tsallis, in preparation (2009).
[21] J. Naudts, Entropy **10**, 131-149 (2008).