Maximum entropy approach to central limit distributions of correlated variables

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*

(Dated: Version April 22, 2008)

Hilhorst and Schehr recently presented a straight forward computation of limit distributions of sufficiently correlated random numbers [1]. Here we present the analytical form of entropy which – under the maximum entropy principle (with ordinary constraints)– provides these limit distributions. These distributions are not q-Gaussians and can not be obtained with Tsallis entropy.

PACS numbers: 05.70.-a 05.90.+m 05.20.-y

INTRODUCTION

Classical statistical mechanics is tightly related to the central limit theorem (CLT). For example if one interprets the velocity of an ideal-gas particle as the result of N random collisions with other particles, the velocity distribution of particles corresponds to a N-fold convolution some distribution of momenta exchanges. For any such distribution, as long as it is centered, stable and has a second moment, the central limit theorem immediately guarantees Maxwell-Boltzmann distributions for $N \to \infty$. Alternative to this mathematical approach the same distribution can be derived from a *physical* principle, where Boltzmanns H-function (entropy) gets maximized under the constraint that average kinetic energy is proportional to temperature, $k_B T$. This principle is referred to as the maximum entropy principle (MEP). These results are trivial for the ideal-gas, or equivalently, for independent random numbers. However, as soon as correlations come into play things become more involved on both sides: CLTs for correlated random numbers have regained strong interest [2–4]. Recently a general and transparent method of obtaining limit distributions of correlated random numbers was reported for a wide class of processes [1]. In the context of the MEP there basically exist three methods to arrive at non-Boltzmann (non-Gaussian) distributions (with reasonably many constraints). The first is associated with Tsallis entropy which produces q-exponentials or q-Gaussians under entropy maximization [5], the second is based on the so-called κ-logarithm [6] leading to a three-parameter class of distribution functions. Recently a method was introduced to constructively design entropy functionals which – under maximization under ordinary constraints – produce *any* plausible distribution function [7]. This was later shown to be a thermodynamically relevant entropy [8]. This entropy is a generalization of previous generalizations of Boltzmann-Gibbs entropies therefore we call it *generalized-generalized* entropy, S_{gg}.

In this paper we show that CLTs and the MEP can be brought into a consistent framework for correlated variables. We start by reviewing the CLT for sums of correlated random numbers, following [1], and the MEP derivation for arbitrary distribution functions following, [7]. We then give the explicit form of the entropy leading to Hilhorst and Schehr's limit distributions.

Limit theorems of correlated random numbers

The idea in [1] is to consider a totally symmetric correlated Gaussian N-point process

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

with $\mathbf{z} = (z_1, \dots, z_N)$ 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 $i = 1, \dots, N$. Total symmetry dictates the form the covariance matrix

$$M_{ij} = \delta_{ij} + \rho(1 - \delta_{ij}) \quad , \tag{2}$$

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 $b = \frac{\rho}{(1-\rho)(1+(N-1)\rho)}$. Straight forward calculation yields that the marginal probability is a Gaussian with unit variance,

$$P_1^{\mathrm{Gauss}}(z_1) = \int dz_2 \dots dz_N \; P_N(z_1, z_2, \dots, z_N) \quad . \tag{3}$$

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

$$\int_0^{u_i} du' \; P_1(u') = \int_0^{z_i} dz' \; P_1^{\mathrm{Gauss}}(z') \quad , \tag{4}$$

P_1 being the one-point distribution of the u variables. Consequently, there is a unique function h, such that

$$u_i = h(z_i) \quad . \tag{5}$$

The distribution of the average of the u variables,

$$\bar{u} = \frac{1}{N}\sum_{i=1}^{N} u_i \quad , \tag{6}$$

is thus found in terms of an integration over all z_i

$$\mathcal{P}(\bar{u}) = \int d\mathbf{z}\; P_N(\mathbf{z})\; \delta\left(\bar{u} - \frac{1}{N}\sum_{i=1}^{N} h(z_i)\right) \quad , \tag{7}$$

where $d\mathbf{z} = dz_1 \ldots dz_N$. After some calculation one arrives at the general result [1],

$$\mathcal{P}(\bar{u}) = \left(\frac{1-\rho}{2\pi\rho}\right)^{\frac{1}{2}} |k'(\nu_*(\bar{u}))|^{-1} \exp\left(-\frac{1-\rho}{2\rho}[\nu_*(\bar{u})]^2\right), \tag{8}$$

where ν_* is defined as the zero of the function

$$k(\nu) = \bar{u} - \frac{1}{\sqrt{2\pi}}\int dw\, e^{-\frac{w^2}{2}} h\left((w+\nu)\sqrt{1-\rho}\right) \quad , \tag{9}$$

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 $\nu'_*(\bar{u}) = -k'(\nu_*(\bar{u}))^{-1} \geq 0$, such that

$$\mathcal{P}(\bar{u}) = \frac{1}{2}\frac{d}{d\bar{u}}\mathrm{erf}\left(\sqrt{\frac{1-\rho}{2\rho}}\nu_*(\bar{u})\right) \quad . \tag{10}$$

Generalized-generalized entropy

The presently most general form of entropy that is consistent with the maximum entropy condition reads in dimensionless notation [7],

$$S_{gg}[P] = -\sum_i P(z_i)\Lambda(P(z_i)) + \eta[P] \tag{11}$$

with

$$\eta[P] = \sum_i \int_0^{P(z_i)} dx\, x\frac{d\Lambda(x)}{dx} + c \quad , \tag{12}$$

which can be rewritten as

$$S_{gg}[P] = -\sum_i \int_0^{P(z_i)} dx\; \Lambda(x) + c \quad . \tag{13}$$

In these equations, $P(z_i)$ is a normalized distribution function of some parameter set z [11] In physics this 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. c is a constant, which ensures that $S_{gg}[P_0] = 0$ for a completely ordered state, i.e. $c = -\int_0^1 dx\Lambda(x)$.

A generalized maximum entropy method, given the existence of some arbitrary stationary distribution function, $\tilde{P}(z_i)$, is formulated as

$$\frac{\delta G}{\delta P(z_i)}\;|_{P=\tilde{P}} = 0 \tag{14}$$

with

$$G \equiv S_{gg}[P] - \alpha\left\{\sum_i P(z_i) - 1\right\} - \beta\left\{\sum_i f(z_i)P(z_i) - U\right\}, \tag{15}$$

where α and β are Lagrange multipliers, and U denotes the expectation of function f, which depending on the problem may be a particular moment of z. The stationary solution to this problem is given by

$$\tilde{P}(z_i) = \mathcal{E}\left(-\alpha - \beta f(z_i)\right) \quad , \tag{16}$$

where $\mathcal{E}(x)$ is a generalized exponential, which is the inverse function of $\Lambda(x)$: $\mathcal{E}(\Lambda(x)) = \Lambda(\mathcal{E}(x)) = x$. In other words, $\Lambda(x)$ from Eq. (11) is chosen as the inverse function of the stationary (observed) distribution function.

The form of entropy in Eq. (13) is enforced by the maximum entropy principle. Under variation the first term on the right hand side of Eq. (11) yields $\frac{d}{dx}P\Lambda(P) = \Lambda + P\frac{d\Lambda}{dx}$. The term $P\frac{d\Lambda}{dx}$ can neither get absorbed into the logarithmic terms, nor by the constants. As long as this term is present, the only solution to the MEP is $\Lambda = \ln$. The idea of the generalized-generalized entropy is to introduce η (which modifies entropy), such that the term $P\frac{d\Lambda}{dx}$ cancels out exactly under the variation, for details see [7]. It has been shown explicitly that the first and second laws of thermodynamics are robust under this generalization of entropy [8]. For the discussion below note that introducing a scaling factor ζ in the argument,

$$S_{gg}[P] = -\sum_i \int_0^{P(z_i)} dx\; \Lambda(\zeta x) \quad , \tag{17}$$

corresponds to distributions of the form

$$P(z) = \frac{1}{\zeta}\mathcal{E}(-\alpha - \beta f(z)) \quad . \tag{18}$$

ζ and α can be seen as two alternative normalization parameters. For instance the condition $\zeta = 1$ leads to a normalization of the distribution as discussed in [7, 8], while the condition $\alpha = 0$ leads to a normalization of P where ζ plays the role of the *partition function.*

ENTROPY FOR LIMIT DISTRIBUTIONS

Let us now properly identify the distributions P and $\mathcal{P}$ from Eqs. (8) and (17). First, note that the limit distributions in [1] are centered and symmetric thus the first

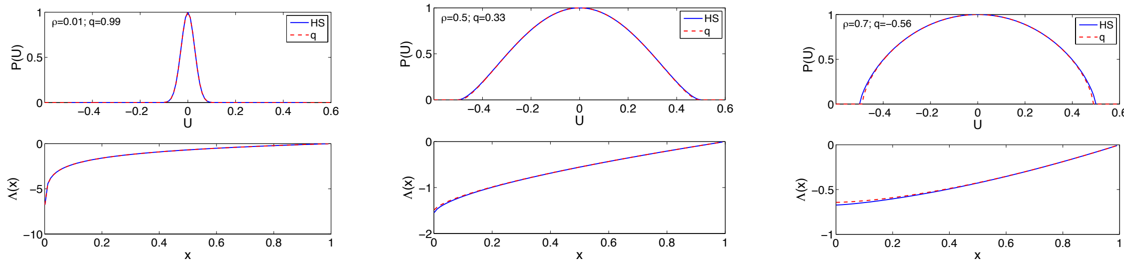

FIG. 1: Top: Limit distributions *al la* Hilhorst-Schehr for $\rho = 0.001$ (left) 0.5 (middle), and 0.7 (right) (lines). Broken lines are q-Gaussians with a q from the reported best-fit value of $q = \frac{1-\frac{5}{3}\rho}{1-\rho}$, [1]. Bottom: Generalized logarithms necessary to provide Hilhorst-Schehr 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.

moment does not provide any information. We therefore choose f from Eq. (15) to be $f(z) = z^2$. Second, the one point distribution P_1 has fixed variance and so does $\mathcal{P}(\bar{u})$ in Eq. (8), whereas distributions obtained through the MEP Eq. (18) scale with a function of the "inverse temperature", β. To take this into consideration for the identification of P and $\mathcal{P}$ we need a simple scale transformation $\bar{u} = \lambda z$, where $\lambda(\beta)$ depends explicitly on β. Consequently, $\mathcal{P}(\bar{u}) \to \lambda\mathcal{P}(\lambda z)$ and

$$\lambda\mathcal{P}(\lambda z) = \frac{1}{\zeta}\mathcal{E}\left(-\alpha - \beta z^2\right) \quad . \tag{19}$$

This particular identification and the independence of Lagrange parameters requires the normalization condition $\alpha = 0$ for the limit distribution Eq. (18). Further, to determine ζ and λ we use two conditions usually valid for generalized exponential functions, $\mathcal{E}(0) = 1$ and $\mathcal{E}'(0) = 1$. This leads to

$$\zeta^{-1} = \lambda\left(\frac{1-\rho}{2\pi\rho}\right)^{\frac{1}{2}}\nu_*'(0) \quad \text{and} \quad \lambda = \gamma\sqrt{\beta} \quad , \tag{20}$$

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

$$\mathcal{E}\left(x\right) = \frac{\nu_*'(\gamma\sqrt{-x})}{\nu_*'(0)}\exp\left(-\frac{1-\rho}{2\rho}[\nu_*(\gamma\sqrt{-x})]^2\right) \quad , \tag{21}$$

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.

An example

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 in [1]. This choice implies

$$\begin{aligned} k'(\nu) &= -\frac{\kappa e^{-\kappa^2\nu^2}}{\sqrt{\pi}} \quad , \quad \kappa \equiv \left(\frac{1-\rho}{2(2-\rho)}\right)^{\frac{1}{2}} \\ k(\nu) &= \bar{u} - \frac{1}{2}\mathrm{erf}(\kappa\nu) \quad , \quad \nu_* = \kappa^{-1}\mathrm{erf}^{-1}(2\bar{u}), \end{aligned} \tag{22}$$

and the limit distribution Eq. (8) becomes

$$\mathcal{P}(\bar{u}) = \left(\frac{2-\rho}{\rho}\right)^{\frac{1}{2}}\exp\left(-\frac{2(1-\rho)}{\rho}[\mathrm{erf}^{-1}(2\bar{u})]^2\right) \quad . \tag{23}$$

This block function has been used earlier [9] where it was conjectured on numerical evidence that the limiting distribution would be a q-Gaussian. This is obviously ruled out by Eq. (23), however, actual discrepancy is small, see Fig. 1. For this example Eq. (21) becomes

$$\mathcal{E}(x) \equiv \exp\left(-(\pi\gamma^2)^{-1}\left[\mathrm{erf}^{-1}(2\gamma\sqrt{-x})\right]^2\right) \quad , \tag{24}$$

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

$$\Lambda(x) = -\left[(2\gamma)^{-1}\mathrm{erf}\left(\gamma\sqrt{-\pi\ln x}\right)\right]^2 \quad . \tag{25}$$

It is compared to q-logarithms in Fig. 1; the discrepancy is small but visible.

DISCUSSION

Historically it was hypothesized on numerical evidence [9] that limit distributions of sums of correlated random numbers generated along the lines of Eq. (3) might be deeply related to q-Gaussians, thus lending fundamental support for q-entropy. In [1] it was shown that this is not the case. If q-entropy does not lead to the exact limit distributions under the MEP, which form of entropy does? Here we constructively answer this question by using a recently proposed generalization of q-generalized entropy [7], which is thermodynamically relevant [8]. Interestingly, a similar program has been carried out some time ago for Levy stable distributions, which result from a generalized CLT where higher momenta do not exist. There it was shown that the corresponding entropy functional is Tsallis entropy [10], however under q-constraints.

Supported by Austrian Science Fund FWF Projects P17621 and P19132.

[1] H.J. Hilhorst, G. Schehr, J. Stat. Mech., P06003 (2007).
[2] F. Baldovin, A.L. Stella, Phys. Rev. E **75**, 020101(R) (2007).
[3] L.G. Moyano, C. Tsallis, M. Gell-Mann, Europhys. Lett. **73**, 813-819 (2006).
[4] S. Umarov, C. Tsallis, M. Gell-Mann, S. Steinberg, arXiv:cond-mat/0606038; cond-mat/0606040.
[5] C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
[6] G. Kaniadakis, Phys. Rev. E **66**, 056125 (2002).
[7] R. Hanel, S. Thurner, Physica A **380**, 109-114 (2007).
[8] S. Abe, S. Thurner, Europhys. Lett. **81**, 10004 (2008).
[9] W. Thistleton, J.A. Marsh, K. Nelson, C. Tsallis, arXiv:cond-mat/0605570.
[10] S. Abe, A.K. Rajagopal, Europhys. Lett. **52**, 610-614 (2000).
[11] For continuous variables z replace $\sum_i \to \int dz$.