A new look at q-exponential distributions via excess statistics

Jean-François Bercher, Christophe Vignat

A new look at q-exponential distributions via excess statistics

J.-F. Bercher [1]

Laboratoire des Signaux et Systèmes, CNRS-Univ Paris Sud-Supelec, 91192 Gif-sur-Yvette Cedex, France and Université Paris-Est, LabInfo-IGM, 5 bd Descartes, 77454 Marne-la-Vallée Cedex 2, France

C. Vignat

Université de Marne-la-Vallée, LabInfo-IGM, 5 bd Descartes, 77454 Marne-la-Vallée Cedex 2, France

Abstract

Q-exponential distributions play an important role in nonextensive statistics. They appear as the canonical distributions, i.e. the maximum generalized q-entropy distributions under mean constraint. Their relevance is also independently justified by their appearance in the theory of superstatistics introduced by Beck and Cohen. In this paper, we provide a third and independent rationale for these distributions. We indicate that q-exponentials are stable by a statistical normalization operation, and that Pickands' extreme values theorem plays the role of a CLT-like theorem in this context. This suggests that q-exponentials can arise in many contexts if the system at hand or the measurement device introduces some threshold. Moreover we give an asymptotic connection between excess distributions and maximum q-entropy. We also highlight the role of Generalized Pareto Distributions in many applications and present several methods for the practical estimation of q-exponentials parameters.

Key words:
q-exponentials, Generalized Pareto Distributions, excess distributions, maximum entropy principle, nonextensive statistics
PACS: 02.50.-r, 05.40.-a, 05.90.+m

1. Introduction

Since the pioneering paper by Tsallis (1988), nonextensive statistics have received an increasing interest in the statistical physics community. As a consequence, the canonical distributions that appear in this context have raised many studies: these distributions are the $q-$Gaussians under the assumption that the total energy is fixed, and the $q-$exponentials [1] when the mean value of the system is fixed.
In several cases and a wide variety of fields, experiments, numerical results and analytical derivations fairly agree with the formalism and the description by q-Gaussians or q-exponentials. This includes situations characterized by long-range interactions, long-range memory or space-time multifractal structure. Applications include fully developed turbulence, Levy anomalous diffusion, statistics of cosmic rays, econometry, and many others. Furthermore, recent studies have highlighted their excellent ability to approximate with high accuracy the exact distribution of complex systems [2].
Let us recall that $q-$exponential distributions are defined by

$$f_{q,\beta}(x) = \frac{1}{Z_{q,\beta}} e_q(-\beta x) \tag{1}$$

with

$$e_q(-\beta x) = (1 - \beta(1-q)x)^{\frac{1}{1-q}} \quad , x \in [0, A_q[\tag{2}$$

Email addresses: jf.bercher@esiee.fr (J.-F. Bercher), vignat@univ-mlv.fr (C. Vignat).
[1] On sabbatical leave from ESIEE-Paris, France.

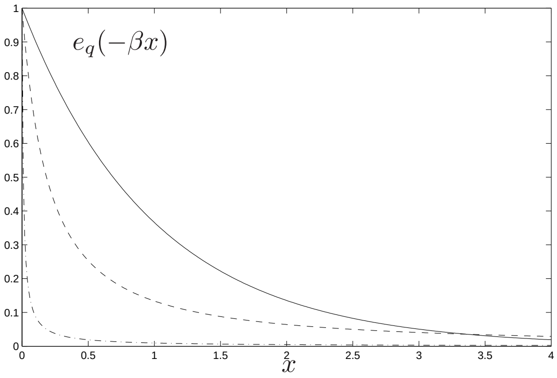

Figure 1. Infinite support $q-$exponential distributions with $\beta = \frac{1}{2-q}$ and $q = 1.01$ (solid), 1.99 (dash dot) and 1.8 (dash)

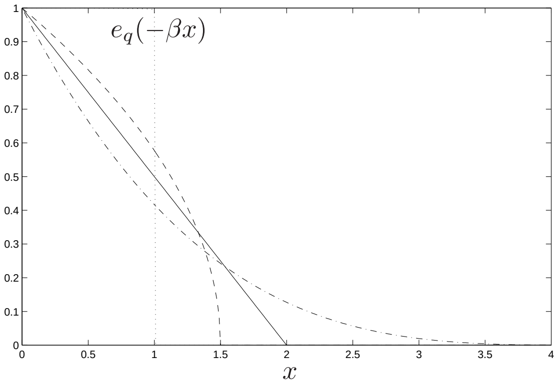

Figure 2. Finite support $q-$exponential distributions with $\beta = \frac{1}{2-q}$ and $q = 0$ (solid), $-\infty$ (dot), -1 (dash) and 0.7 (dashdot)

where A_q is the right endpoint, $A_q = \frac{1}{\beta(1-q)}$ if $q < 1$ and $A_q = +\infty$ if $1 < q < 2$, and where β is a positive scale parameter related to the mean m of $f_{q,\beta}$ as $m = \int x f_{q,\beta}(x)\, dx = \frac{1}{\beta(3-2q)}$, for $q < 3/2$. The partition function $Z_{q,\beta}$ is

$$Z_{q,\beta} = \frac{1}{\beta\,(2-q)},\ \ 0 < q < 2.$$

We will also denote $\bar{F}_{q,\beta}(x)$ the so-called survival function associated to the q-exponential distribution (1):

$$\bar{F}_{q,\beta}(x) = \int_x^{+\infty} f_{q,\beta}(z)dz = 1 - F_{q,\beta}(x) \tag{3}$$

$$= (1 - \beta(1-q)x)^{\frac{1}{1-q}+1} \tag{4}$$

where $F_{q,\beta}(x)$ is the probability distribution function.
The q-exponential distributions $f_{q,\beta}(x)$ are the solutions of the following maximum entropy problem:

$$\max_f H_q(f) \ \text{ with } \int_0^{+\infty} x f^q(x)\, dx = \theta \text{ and } \int_0^{+\infty} f(x)\, dx = 1 \tag{5}$$

where

$$H_q = \frac{1}{1-q}\left(\int_0^{+\infty} f^q(x)\, dx - 1\right) \tag{6}$$

is the Tsallis entropy. We note that the same solution is reached if Tsallis entropy is replaced by using Rényi entropy

$$S_q = \frac{1}{1-q}\log\int_0^{+\infty} f^q(x)\, dx,$$

since it is a simple monotone transform of the latter. In the limit case $q = 1$, Tsallis entropy coincides with Shannon entropy and the maximum entropy solution recovers the canonical exponential distribution

$$f_{1,\beta}(x) = \beta\exp(-\beta x),\ \ x \geq 0. \tag{7}$$

The shapes of several $q-$exponential distributions are shown in Fig.1 (infinite support case: $q > 1$ and $A_q = +\infty$) and in Fig.2 (finite support case: $q < 1$ and $A_q = \frac{1}{\beta(1-q)}$); in all cases, parameter β was chosen as $\beta = \frac{1}{2-q}$ so that $f_q(0) = 1$. These figures indicate the high versatility of q-exponentials for modeling probability distributions.
An important and alternate characterization of these distributions has been provided by the superstatistics theory as introduced in [3] (see also [4] for related results). In this approach, $q-$exponential distributions are recognized as scale mixtures of exponential distributions: more precisely, the integral representation

$$f_{q,\frac{1}{q-1}}(x) = \int_0^{+\infty} g_n(\beta) f_{1,\beta}(x) \, d\beta$$

where $g_n(\beta) \propto \beta^{\frac{n}{2}-1} e^{-\beta}$ is a χ_n^2 distribution with $n = \frac{2}{q-1} - 2$ degrees of freedom, allows to describe a $q-$exponential random variable as a (q=1) exponential variable with random scale parameter $\beta \sim \chi_n^2$. We note that this approach, in a physical setting, recovers similar results by Maguire (1952) [5] and Harris (1968) [6]. This superstatistical approach has found many applications in fluctuation theory: for example, in a recent paper, Briggs and Beck [7] have highlighted the relevance of the superstatistical approach in the modeling of train delays.
Hence, in the one hand, there are two constructions, the maximum q-entropy and the superstatistics one, that exhibit q-exponentials. In the other hand, many observations and experimental results assess the relevance of q-exponentials and their ubiquity. We propose in what follows a third and alternate statistical justification of q-exponential distributions, based on the theory of excess values.
In section 2, we introduce the "Focus on Excesses" operation, which will reveal as the fundamental operation in our setting. Then, in section 3 we present the Pickands-Balkema-de Haan result as a possible rationale for the ubiquity of q-exponentials in nature. Furthermore, we underline a stability property of q-exponentials, and show that these results form a Central Limit Theorem for the "Focus on Excesses" operation. In section 3.2, we draw attention to the Generalized Pareto Distribution (GPD), which is a q-exponential, and underline its applications outside the statistical physics field. Some examples are worked out in subsection 3.3. In section 4, we address the link between the maximum q-entropy and the distributions of excesses. We draw a possible connection and show that the distribution of excesses converges asymptotically to a maximum q-entropy distribution. Finally, we present and discuss in section 5 the estimation procedures for the parameters of q-exponentials-GPD.

2. The "Focus on Excesses" operation

Let us consider a one-dimensional positive random variable $X(\omega)$ that describes the state of a physical system; we denote as $\bar{F}_X$ its survival function

$$\bar{F}_X(x) = 1 - F_X(x) = \Pr\{X > x\}$$

with $F_X(x)$ the distribution function. Let us also consider a positive parameter $u > 0$ and assume that we have access only to the values of X that are larger than u : for example, we may have access to the system through a measurement device that, for a bandwidth or quantification reason, does not "see" any value $X < u$; or we may want to reject all values $X < u$ because we think that they are physically irrelevant. Since parameter u may be large, a wise decision is to shift all these measured values by a factor u : as a consequence, we will in fact record values of the excesses of X over the threshold u, that is values of the *conditional* random variable X_u defined as

$$X_u = X - u | X > u.$$

We denote $X_u = \mathcal{F}_u(X)$ the transform of X into X_u. By Bayes' theorem, the survival function of these excesses writes

$$\bar{F}_{X_u}(x) = \Pr(X > x + u | x > u) = \frac{\bar{F}_X(x+u)}{\bar{F}_X(u)}, \;\; x \geq 0. \tag{8}$$

If X possesses a probability density f_X, then X_u has itself the density

$$f_{X_u}(x) = \frac{f_X(x+u)}{\bar{F}_X(u)}, \;\; x \geq 0. \tag{9}$$

The transformation of X into X_u is illustrated in Fig.3. The inset depicts the excess distribution, i.e. the normalized tail f_{X_u} (9) of the original distribution f_X given in the main part of the figure. For obvious reasons, we call "Focus on Excess" (FoE) operation the transformation of $\bar{F}_X$ into $\bar{F}_{X_u}$, or equivalently of f_X into f_{X_u}. With a slight abuse of notation we also denote $\mathcal{F}_u$ this transformation.
This approach is also known as the "Peaks over Threshold" technique and widely used in fields like climatology and hydrology. A rationale of this method can be found in [8].
The $q-$exponential distribution $f_{q,\beta}(x)$ (1) verifies a remarkable *stability* property with respect to this transformation as described in the following theorem (see [9]).

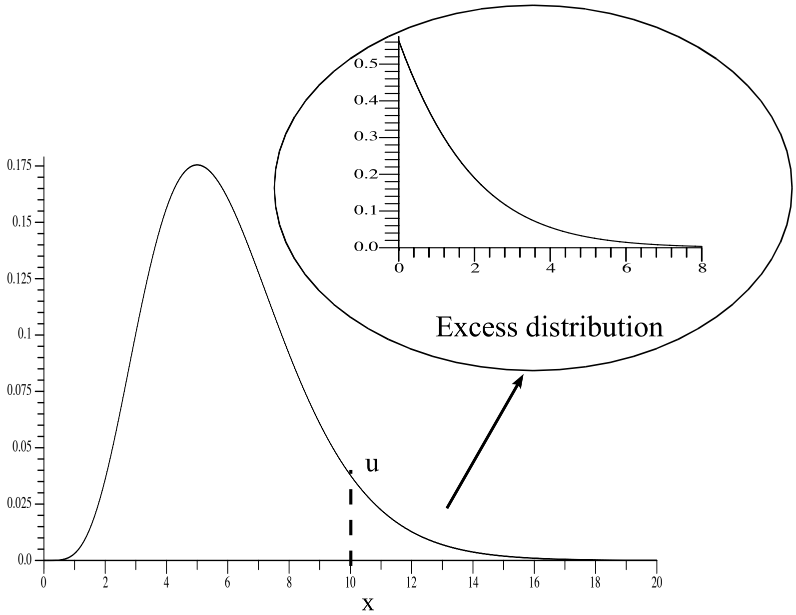

Figure 3. Illustration of the FoE transformation

Theorem 1 *For any $u < A_q$, the FoE of q-exponential remains q-exponential and leaves the tail index q unchanged:*

$$\mathcal{F}_u\left(f_{q,\beta}(x)\right) = f_{q,\beta'}(x)$$

with $\beta' = \frac{\beta}{1-\beta(1-q)u}$.
Proof. By equation (9), we know that

$$\mathcal{F}_u\left(e_q(-\beta x)\right) \propto (1-\beta(1-q)(x+u))^{\frac{1}{1-q}} \propto (1-\beta'(1-q)x)^{\frac{1}{1-q}}, \tag{10}$$

$$= e_q(-\beta' x) \text{ for } 0 \leq x \leq A_q - u \tag{11}$$

where the last term, with $\beta' = \beta/(1-\beta(1-q)u)$, is obtained by factoring $(1-\beta(1-q)u)$. ■
This highlights the particular status of the q-exponential in the FoE transformation. We can also remark that $\beta' = \beta$ in the special exponential $(q = 1)$ case: the normalized tail of an exponential distribution is exactly invariant by FoE transformation.

3. The Limit theorem for q-exponentials

3.1. *Pickands' theorem*

An important theorem in probability as well as in statistical physics is the Central Limit theorem; its extension to the nonextensive context has been the subject of several recent papers [10]. According to this theorem, if a probability density f belongs to the domain of attraction of the Gaussian distribution, and if X_i are centered and independently chosen distributed according to f, then the distribution of

$$Z_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \tag{12}$$

converges with n to the Gaussian distribution. Moreover, the Gaussian distribution appears as a fixed point in this context since if all X_i are independent and Gaussian in (12) then Z_n is *exactly* Gaussian for any value of n.
Let us denote by $\mathcal{T}_n$ the transform

$$\mathcal{T}_n : f \mapsto f_n$$

where f_n is the distribution of Z_n, i.e. the normalized $n-$fold convolution of f. Consider now a variable submitted to a serie of k successive thresholding operations, with thresholds $u_1, \ldots u_k$: the resulting excess variable, say Z_u, is

$$Z_u = \mathcal{F}_{u_k} \circ \mathcal{F}_{u_{k-1}} \circ \ldots \mathcal{F}_{u_1}(X)$$

where $\circ$ is the composition of functions and $\mathcal{F}_{u_i}$ as defined above. Of course, the excess variable can also be viewed as resulting of a single thresholding operation, with $u = \sum_{i=1}^{k} u_i$:

$$Z_u = \mathcal{F}_u(X).$$

An analogy appears between $\mathcal{T}_n$ and the FoE transform $\mathcal{F}_u$ which associates the survival functions:

$$\mathcal{F}_u : \bar{F} \mapsto \bar{F}_u,$$

where the real threshold parameter u plays the role of the integer parameter n. Thus, an important question at this point arises: does there exist a limit result, analogous to the CLT, in the context of the FoE transform? Let us first introduce the Fréchet (or heavy-tailed) family of distributions, that will play in extreme statistics the role of the Gaussian domain of attraction in the CLT.

Definition 1 *A distribution F belongs the Fréchet domain if it is a heavy tail distribution, that is if its survival distribution $\bar{F}$ writes*

$$\bar{F}(x) = x^{-\frac{1}{\gamma}} l(x), \ \forall x \geq 0$$

where $l(x)$is a slowly varying function:

$$\lim_{x \to +\infty} \frac{l(xt)}{l(x)} = 1 \ \forall t > 0$$

and parameter γ is called the tail index of F.

Pickands' theorem, which can be viewed as the analogue of the CLT in excess statistics, is the following.

Theorem 2 *A necessary and sufficient condition for F to belong to the Fréchet domain is*

$$\lim_{u \to A_q} \|\bar{F}_{X_u}(x) - \bar{F}_{q,\beta(u)}(x)\|_\infty = 0$$

for some function $\beta(u)$ and with

$$q = \frac{2\gamma + 1}{\gamma + 1}. \tag{13}$$

This result means, roughly speaking, that if F is in the domain of attraction of the Fréchet distribution, i.e. has heavy tails with tail index γ, then its FoE transform converges (in the infinite norm sense) to a limit distribution which is nothing but a q-exponential with q given by (13). To the accumulation of random variates in the classical CLT, with $n \to \infty$, corresponds here an increase in the threshold u in the FoE. In the classical CLT, we know that the limit distribution (the Gaussian) once reached remains stable by convolution of Gaussian pdf (addition of Gaussian variables). Accordingly, theorem 1 shows that q-exponentials are stable by the FoE operation. This means that once the FoE operations have converged to the limit distribution, then it remains stable by further applications of FoE. In other words, q-exponentials are a fixed point of the FoE operation.

3.2. *Generalized Pareto Distributions*

In the original Pickands' formulation [11] (see also the similar work by Balkema and De Haan [12]), the name coined for the limit distribution of excesses is Generalized Pareto Distribution (GPD). Its survival function is given by

$$\bar{F}_X(x) = P(X > x) = \left(1 + \frac{\gamma}{\sigma}x\right)_+^{-\frac{1}{\gamma}} \text{ for } x \geq 0, \tag{14}$$

where $x_+ = \max(x, 0)$. Its pdf is

$$f_X(x) = \frac{1}{\sigma}\left(1 + \frac{\gamma}{\sigma}x\right)_+^{-\frac{1}{\gamma}-1} \text{ for } x \geq 0, \tag{15}$$

where σ and γ are respectively the scale and shape parameters. A location parameter μ can also be specified according to

$$f_X(x) = \frac{1}{\sigma}\left(1 + \frac{\gamma}{\sigma}(x - \mu)\right)_+^{-\frac{1}{\gamma}-1} \text{ for } x \geq \mu.$$

Obviously, for $\mu = 0$, the GPD coincides with the q-exponential, with

$$\gamma = -\left(\frac{1-q}{2-q}\right) \text{ and } \sigma = \frac{1}{\beta(2-q)}$$

The interesting point is that the GPD is employed outside the statistical physics field for modeling heavy tailed distributions. Examples of applications are numerous, ranging from reliability theory [13], traffic in networks [14], hydrology [15,16], climatology [17,18], geophysics [19], materials science [20,21], radar imaging [22] or actuarial sciences [23]. These uses are related to the POT (Peaks over Threshold) method [15,8,24], and the underlying rationale is of course the fact that the Balkema-de Haan-Pickands theorem asserts that the distribution of excesses over a threshold often follows approximately a GPD for large values of the threshold.

3.3. *Some examples*

Let us now detail Pickands' theorem for classical distributions that belong to the Fréchet domain, namely: the alpha-stable, Student and Cauchy distributions.

– **The alpha-stable case**

  In the case of an alpha-stable distribution with parameter $0 < \alpha < 1$, we know from [25, Th.1 p. 448] that the tail behavior of the survival distribution is

$$\bar{F}_Z(z) \sim \frac{1}{\Gamma(1-\alpha)} z^{-\alpha}$$

  so that the tail index is $\gamma = \frac{1}{\alpha}$ and

$$\bar{F}_{Z_u}(z) \sim \left(1 + \frac{z}{u}\right)^{-\alpha}.$$

  We note that the Cauchy distribution ($\alpha = 1$) can not be deduced from this result.

– **The Student-t case**

  A Student -t distribution with $m = 2\nu - 1$ degrees of freedom writes

$$f_X(x) = A_\nu \left(1 + x^2\right)^{-\nu}, \ x \in \mathbb{R}$$

  with $A_\nu = \frac{\Gamma(\nu)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\nu - \frac{1}{2}\right)}$. Thus an equivalent of the Student survival distribution is

$$\bar{F}_Z(z) = \frac{A_\nu}{2\nu - 1} z^{-2\nu+1},$$

  the tail index is $\gamma = (2\nu - 1)^{-1}$ and the excess distribution behaves as

$$\bar{F}_{Z_u}(z) \sim \left(1 + \frac{z}{u}\right)^{-2\nu+1}.$$

– **The Cauchy case**

  As a special case, the Cauchy distribution corresponds to $m = 1$ degrees of freedom and

$$\bar{F}_{Z_u}(z) \sim \left(1 + \frac{z}{u}\right)^{-1}. \tag{16}$$

In nonextensive statistics, the Student-t distributions described above are well known under the name of $q-$Gaussian distributions ($q > 1$). What is shown here is that the excess distribution of a $q-$Gaussian distribution with nonextensivity parameter q converges to a $q-$exponential distribution with nonextensivity parameter

$$q' = \frac{1+q}{2}.$$

We remark that in limit case where $q \to 1$, the $q-$Gaussian distribution converges to the classical Gaussian distribution. In this case we also obtain that $q' \to 1$, which means that the excess distribution of a Gaussian converges to the exponential distribution. Indeed, Pickands' theorem extends to the Gumbel family of distributions characterized by an exponential tail: the associated excess pdf is simply the distribution given by (7), which is the limit case of the $q-$exponential (1) for $q \to 1$.

3.4. *Numerical Illustration*

As a numerical illustration, we present the Cauchy case

$$f_X(x) = \frac{2}{\pi(1+x^2)}, \ x \geq 0.$$

The survival function of the excess variable X_u is

$$\bar{F}_{X_u}(x) = \frac{\bar{F}_X(x+u)}{\bar{F}_X(u)} = \frac{1 - \frac{2}{\pi}\arctan(x+u)}{1 - \frac{2}{\pi}\arctan(u)}.$$

Using the fact that $\arctan(t) \approx \pi/2 - 1/t$, for $t \gg 1$, we have

$$\bar{F}_{X_u}(x) \approx \frac{1}{u+x} / \frac{1}{u} = \left(1 + \frac{x}{u}\right)^{-1}$$

which is (16) again. It is the survival function of a q-exponential with entropic index $q = 3/2$.

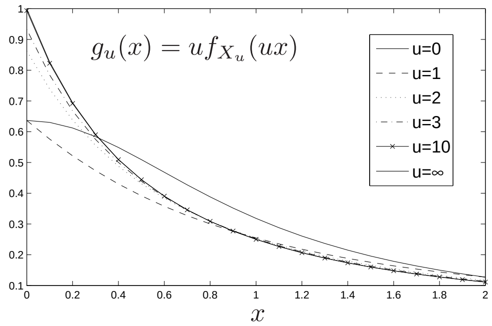

Figure 4. Illustration of Pickands' theorem in the Cauchy case. The figure presents the normalized excess pdfs $g_u(x)$, for several values of the threshold u. It shows that the normalized pdf quickly converges, as u increases, to the limit distribution $1/(1+x)^2$.

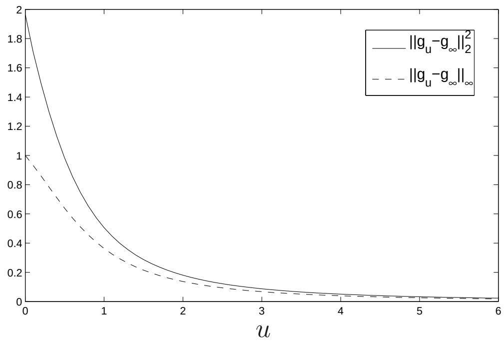

Figure 5. Rate of convergence to the limit distribution. The rate of convergence is illustrated by the quadratic and infinite norms of the approximation error $g_u - g_\infty$, as a function of the threshold u.

The excess density can be explicited as

$$f_{X_u}(x)\,(x) = \frac{2}{\pi}\frac{1}{\left(1-\frac{2}{\pi}\arctan(u)\right)\left(1+(x+u)^2\right)}$$

The convergence of the survival distribution to a Generalized Pareto Distribution with $\gamma = 1$ as in (16) is illustrated in Fig. 4. In this figure, we have plotted the normalized pdfs $x \mapsto u f_{X_u}(ux) = g_u(x)$, corresponding to the normalized random variable $Y_u = X_u/u$, for several values of the threshold u. The figure shows the convergence to the GPD limit $g_\infty(x) = \frac{1}{(1+x)^2}$. We note that Pickands' theorem only ensures convergence of the distribution functions, but the Cauchy case is smooth enough to observe a convergence of the normalized pdf as well. We remark on Fig. 4 that small values of parameter u (typically $u = 3$) ensure an already accurate approximation of f_{X_∞}. Rate of convergence to the limit distribution are examined in Fig. 5, where we report the quadratic and infinite norms of the approximation error, as a function of u:

$$\|g_u - g_\infty\|_\infty \text{ and } \|g_u - g_\infty\|_2^2$$

Note that the rate of convergence of the GPD approximation in Pickands' theorem is rigorously quantified in [26].

4. A connection with Maximum q-entropy

Since there exists an entropic proof of the classical central limit theorem, it is appealing to look for relationships between Pickands' result and the maximum entropy principle. We show here that a distribution of the excesses over a threshold converges, as this threshold goes to infinity, to a maximum q-entropy solution. Details of similar derivations are given in [27].

Proposition 3 gives a simplified form of a q-exponential distribution function as the solution of a maximum q-norm problem. Then, we state in theorem 4 that the 1-norm, q-norm, and q-expectation of a suitably normalized version of the excess variable (in the Fréchet domain) converges asymptotically to constant values. Finally, with an appropriate choice of q, this shows that the distribution of excesses is necessarily, asymptotically, the q-exponential solution of the maximum q-norm problem (or equivalently of Tsallis' q-entropy maximization).

Proposition 3 *With $\sqrt{2} > q > 1$, consider the class of functions*

$$\mathcal{F} = \left\{G : \mathbb{R}^+ \to \mathbb{R};\ \int_0^{+\infty} zG(z)\,dz < \infty\right\}.$$

The function

$$G_*(z) = (1+z)^{\frac{1}{1-q}+1} \tag{17}$$

is the unique solution of the following maximum norm problem:

$$\begin{aligned} &\max_{G\in\mathcal{F}} \|G\|_q^q = \max_{G\in\mathcal{F}} \int_0^{+\infty} G(z)^q\, dz \\ &\textit{such that} \\ &\int_0^{+\infty} zG(z)^q\, dz = \frac{(q-1)^2}{(q^2-q-1)(q^2-2)} \textit{ and } \int_0^{+\infty} G(z)\, dz = \frac{1-q}{2q-3}. \end{aligned} \tag{18}$$

Moreover, the corresponding maximum norm is

$$\|G_*\|_q^q = \int_0^{+\infty} G_*(z)^q\, dz = \frac{1-q}{q^2-q-1}.$$

Let us now consider a random variable X that belongs to the Fréchet domain with tail index $\gamma < \frac{1}{2}$, and a normalized version $Y_u = X_u/g(u)$ of its excesses, where function $g(u) \sim u$, as $u \to +\infty$. Then, we have the following results.

Theorem 4 *With $\gamma < q < 1$, as $u \to +\infty$, the following asymptotics hold*

$$\|\bar{F}_{Y_u}\|_q^q \sim \frac{1}{\frac{q}{\gamma}-1}$$

with asymptotical q-expectation

$$\int_0^{+\infty} z\bar{F}_{Y_u}(z)^q\, dz \sim \frac{1}{\left(1-\frac{q}{\gamma}\right)\left(2-\frac{q}{\gamma}\right)}$$

and asymptotical $1-$norm

$$\int_0^{+\infty} \bar{F}_{Y_u}(z)\, dz = \frac{1}{\frac{1}{\gamma}-1}.$$

As a consequence, with the choice of

$$q = \frac{1+2\gamma}{1+\gamma} \text{ or } \gamma = \frac{1-q}{q-2}$$

we obtain that the excess distribution of any suitably normalized random variable $Y = X/g(u)$, where X belongs to the Fréchet domain and $g(u) \sim u$, has asymptotically same $1-$norm, $q-$norm and q-expectation as the q-exponential distribution (17). Since the maximum norm problem (18) has a unique solution, we deduce that the maximum q-norm distribution G_* and the distribution of excesses F coincide asymptotically as $u \to +\infty$. The physical significance behind this result is that it enables to connect the distribution of the excesses over a threshold and a maximum q-entropy construction, and thus add a new motivation to the use of Tsallis entropy: q-exponentials, as the Tsallis entropy maximizers, can arise in many contexts in which the system at hand or the measurement device introduces some threshold.

Moreover, function G_* (17) can now be identified as the survival function associated to the q-exponential density

$$g_*(x) \propto (1+x)^{\frac{1}{1-q}},$$

which is itself the solution of a maximum q-norm problem (5) similar to (18). Hence, the asymptotic density of excesses (if it exists) is a q-exponential density. The addition of a scale parameter $\beta(q-1)$ simply enables to adjust the mean or variance of the distribution.

5. Parameters estimation

The statistical analysis of heavy-tailed data requires some caution [28]. A recent reference for this topic is [29].

An important issue for applications of q-exponential-GPD and the assessment of their role in real situations is the estimation of their parameters. As far as the tail exponent $\alpha = 1/\gamma$ is concerned, several methods exist. The corresponding, and relevant problem, in the nonextensive framework is the estimation of parameter q.

The Hill estimator - The Hill estimator [30] is given by

$$\hat{\gamma}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\log X_{(n-i+1)} - \log X_{(n-k)}\right)$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ are the order statistics of the sample of size n, and k is a smoothing parameter. This estimator, which can be derived by several rationale (including maximum likelihood) is valid for $\gamma > 0$, i.e. $q > 1$.

Pickands' estimator - Another, more general estimator, is the Pickands' estimator, which is valid for all γ and given by

$$\hat{\gamma}(k) = \frac{1}{\log 2} \log \frac{X_{(n-k+1)} - X_{(n-2k+1)}}{X_{(n-2k+1)} - X_{(n-4k+1)}}$$

Both Hill and Pickands estimators are consistent and asymptotically normal [31,29].

Maximum Likelihood - The log-likelihood for a set of n independent and identically distributed observations X_i, distributed according to a GPD(γ, σ) given by (15), is:

$$\log p_n(X_1, X_2, \ldots, X_n) = -n \log \sigma - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^{n} \log\left(1 + \frac{\gamma}{\sigma} X_i\right).$$

Then, the Maximum Likelihood procedure gives two equations for the unknown parameters. These relations are implicit but can be solved numerically. We obtain

$$\sigma = \frac{1+\gamma}{n} \sum_{i=1}^{n} \frac{X_i}{1 + \frac{\gamma}{\sigma} X_i}$$

and

$$\gamma = \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + \frac{\gamma}{\sigma} X_i\right)$$

More details on Maximum Likelihood estimation of GPD parameters can be found in [32,33]. The Maximum Likelihood for q-exponentials has also be discussed recently in [34]. It shall be mentioned that the ML does not exist for $\gamma < -1$, which corresponds to the case $q > 2$, where $e_q(x)$ is not normalizable.

Method of moments - When both the mean m_X and variance s_X^2 exist, that is for $\gamma < 0.5$ ($q > 0.5$), one can apply a *method of moments* (MOM) since there is a simple relationship between these moments and the parameters γ and σ, namely

$$\sigma = 0.5 m_X \left(1 + (m_X/s_X)^2\right) \tag{19}$$

and

$$\gamma = 0.5 \left(1 - (m_X/s_X)^2\right) \tag{20}$$

Then it suffices to use some estimates of the mean and standard deviation in (19,20) to obtain a (rough) estimate of the GPD parameters.

Conditional mean exceedance - Variations on the method of moments include the so-called "conditional mean exceedance method" which relies on the analysis of the plot of the mean of excesses over u, which is linear according to

$$E[X - u | X > u] = \frac{\sigma + \gamma u}{1 - \gamma},$$

with slope $\gamma/(1-\gamma)$ and intercept $\sigma/(1-\gamma)$. For several values of the threshold u, and therefore estimates of the mean of excesses, a least-squares procedure can be used in order to identify the parameters of the linear model.

Probability Weighted Moments - An alternative method to the MOM, known as "Probability Weighted Moments" (PWM) was introduced by [9,35]. The paper [9] also provides and interesting comparison of ML, MOM and PWM estimates. This procedure relies on the definition of a "weighted moment":

$$m_p = E\left[X \bar{F}_X(X)^p\right]$$

where $\bar{F}_X$ is the survival function of the GPD given in (14). For $\gamma < 1$ ($q > 0$), we readily obtain that

$$m_p = \frac{\sigma}{(p+1)(p+1-\gamma)}.$$

Then, two weighted moments are sufficient to exhibit the values of the parameters γ and σ; e.g. with m_1 and m_0, one has

$$\gamma = \frac{4m_1 - m_0}{2m_1 - m_0} \text{ and } \beta = \frac{2m_1 m_0}{m_0 - 2m_1}$$

With more moments, estimated classically, one can look for a least-squares solution. PWM is reported to compete with Maximum Likelihood.

Synthesis and Software - Another method is the Elemental Percentile Method, discussed in [36]. A synthesis of estimation methods of GPD parameters can be found in [33, chapter 20, pages 614-620] and in [37].

It shall be mentioned that statistical softwares include estimation methods for GPD-q-exponentials parameters. An example is the "Vector Generalized Additive Models" package [38,39]. Another example is the statistical toolbox of the highly employed Matlab™software, which includes fitting tools for the GPD since its 5.1 version (2006).

6. Conclusion

In this paper, we have proposed a possible rationale for understanding the ubiquity of q-exponential distributions in nature. The point is that, given an heavy-tailed distributed system, q-exponentials can occur as soon as the measurement device, or the system at hand, involves some threshold. This is the essence of Pickands' result.
Furthermore, the q-exponentials are stable by thresholding. This shows that q-exponentials are the limit distributions in a Central Limit like theorem, where the underlying operation is the "Focus on Excesses" operation we have introduced.
We have also underlined that q-exponentials can also be recognized as Generalized Pareto Distributions which are of importance in many applications outside the statistical physics field.
We have drawn a connection between excess distributions and the maximum q-entropy, by showing that a distribution of excesses converges to a maximum q-entropy solution.
Finally, we have presented some procedures for the crucial problem of estimation of q-exponential parameters. We have also provided a large number of bibliographic entries on these different topics.
In our view, the rationale proposed here is complementary to the classical maximum q-entropy approach, and to the Superstatistics approach. We hope that it will prompt some new viewpoints in physical applications.

References

[1] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, Journal of Statistical Physics 52 (1988) 479–487.
[2] H. Hilhorst, G. Schehr, A note on q-gaussians and non-gaussians in statistical mechanics, Journal of Statistical Mechanics: Theory and Experiment (P06003).
[3] C. Beck, E. G. D. Cohen, Superstatistics, Physica A 322 (2003) 267.
[4] G. Wilk, Z. Wlodarczyk, Interpretation of the nonextensitivity parameter q in some applications of Tsallis statistics and Lévy distributions, Phys. Rev. Lett 84 (2000) 2770.
[5] B. A. Maguire, E. S. Pearson, H. A. Wynn, The time intervals between industrial accidents, Biometrika 39 (1952) 168–180.
[6] C. M. Harris, The Pareto distribution as a queue service discipline, Operations Research 16 (1968) 307–313.
[7] K. Briggs, C. Beck, Modelling train delays with q-exponential functions, Physica A 378 (2007) 498.
[8] R. Leadbetter, On a basis for 'Peaks over Threshold' modeling, Statistics and Probability Letters 12 (1991) 357–362.
[9] J. R. M. Hosking, J. R. Wallis, Parameter and quantile estimation for the Generalized Pareto Distribution, Technometrics 29 (1987) 339–349.
[10] S. Umarov, C. Tsallis, On multivariate generalizations of the q-central limit theorem consistent with nonextensive statistical mechanics, in: S. Abe, H. Herrmann, P. Quarati, A. Rapisarda, C. Tsallis (Eds.), Complexity, Metastability and Nonextensivity, Vol. 965, American Institute of Physics Conference Proceedings, New York, 2007, p. 34.
[11] J. Pickands, Statistical inference using extreme order statistics, The Annals of Statistics 3 (1975) 119–131.
[12] A. A. Balkema, L. de Haan, Residual life time at great age, The Annals of Probability 2 (1974) 792–804.
[13] M. Asadi, I. Bayramoglu, The mean residual life function of a k-out-of-n structure at the system level, IEEE Transactions on Reliability 55 (2006) 314–318.
[14] P.-H. Hsieh, An exploratory first step in teletraffic data modeling: evaluation of long-run performance of parameter estimators, Computational Statistics & Data Analysis 40 (2002) 263–283.
[15] M. A. J. Van Montfort, J. V. Witter, Generalized Pareto Distribution applied to rainfall depths, Hydrological Sciences Journal 31 (1986) 151–162.
[16] S. Coles, L. R. Pericchi, S. Sisson, A fully probabilistic approach to extreme rainfall modeling, Journal of Hydrology 273 (2003) 35–50.
[17] S. Coles, E. Casson, Extreme value modelling of hurricane wind speeds, Structural Safety 20 (1998) 283–296.
[18] I. Harris, Generalised Pareto methods for wind extremes. Useful tool or mathematical mirage?, Journal of Wind Engineering and Industrial Aerodynamics 93 (2005) 341–360.
[19] V. F. Pisarenko, D. Sornette, Characterization of the frequency of extreme events by the Generalized Pareto Distribution, Pure and Applied Geophysics 160 (2003) 2343.
[20] G. Shi, H. V. Atkinson, C. M. Sellars, C. W. Anderson, Application of the Generalized Pareto Distribution to the estimation of the size of the maximum inclusion in clean steels, Acta Materialia 47 (1999) 1455–1468.
[21] D. Fullwood, B. Adams, S. Kalidindi, Generalized Pareto front methods applied to second-order material property closures, Computational Materials Science 38 (2007) 788–799.
[22] B. LaCour, Statistical characterization of active sonar reverberation using extreme value theory, IEEE Journal of Oceanic Engineering 29 (2004) 310–316.
[23] A. C. Cebrian, M. Denuit, P. Lambert, Generalized Pareto fit to the society of actuaries' large claims database, North American Actuarial Journal 7 (2003) 18–36.

[24] J. A. Lechner, E. Simiu, N. A. Heckert, Assessment of ‘Peaks over Threshold’ methods for estimating extreme value distribution tails, Structural Safety 12 (1993) 305–314.
[25] W. Feller, An introduction to Probability theory and its applications, Vol. 2, John Wiley and Sons, 1971.
[26] R. Worms, Vitesse de convergence de l’approximation de Pareto gÃ©nÃ©ralisee de la loi des excÃ¨s - Rate of convergence for the generalized Pareto approximation of the excesses, Comptes Rendus de l’Academie des Sciences - Series I - Mathematics 333 (2001) 65–70.
[27] J. F. Bercher, C. Vignat, An entropic view of Pickands’ theorem, in: International Symposium on Information Theory ISIT2008, Toronto, 2008. URL http://arxiv.org/abs/0802.3110
[28] A. Clauset, C. R. Shalizi, M. E. J. Newman, Power-law distributions in empirical data, 0706.1062.
[29] N. Markovich, Nonparametric Analysis of Univariate Heavy-Tailed Data, Series in Probability and Statistics, Wiley, 2007.
[30] B. Hill, A simple general approach to inference about the tail of a distribution, Annals of Statistics 3 (1975) 1163–1174.
[31] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events for Insurance and Finance, 1st Edition, Springer, 2008.
[32] S. D. Grimshaw, Computing maximum likelihood estimates for the Generalized Pareto Distribution, Technometrics 35 (1993) 185–191.
[33] N. L. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, 2nd Edition, Wiley-Interscience, 1994.
[34] C. R. Shalizi, Maximum likelihood estimation for q-exponential (Tsallis) distributions, math/0701854.
[35] J. Hosking, Some theory and practical uses of trimmed l-moments, Journal of Statistical Planning and Inference 137 (2007) 3024–3039.
[36] E. Castillo, A. S. Hadi, Fitting the Generalized Pareto Distribution to data, Journal of the. Americal Statistical Association 92 (1997) 1609–1620.
[37] J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels, Statistics of Extremes: Theory and Applications, 1st Edition, Wiley, 2004.
[38] T. W. Yee, C. J. Wild, Vector Generalized Additive Models, Journal of the Royal Statistical Society, Series B, Methodological 58 (1996) 481–493.
[39] T. W. Yee, The VGAM package (2007).
URL http://www.stat.auckland.ac.nz/~yee/VGAM/