Entropies of deformed binomial distributions

H. Bergeron[a], E.M.F. Curado[b,c],

J.P. Gazeau[b,d], Ligia M.C.S. Rodrigues[b](*)

[a] *Univ Paris-Sud, ISMO, UMR 8214, 91405 Orsay, France*

[b] *Centro Brasileiro de Pesquisas Fisicas*

[c] *Instituto Nacional de Ciência e Tecnologia - Sistemas Complexos*

Rua Xavier Sigaud 150, 22290-180 - Rio de Janeiro, RJ, Brazil

[d] *APC, UMR 7164,*

Univ Paris Diderot, Sorbonne Paris Cité, 75205 Paris, France

December 2, 2014

Abstract

Asymptotic behavior (with respect to the number of trials) of symmetric generalizations of binomial distributions and their related entropies are studied through three examples. The first one derives from the q-exponential as a generating function. The second one involves the modified Abel polynomials, and the third one involves Hermite polynomials. The former and the latter have extensive Boltzmann-Gibbs whereas the second one (Abel) has extensive Rényi entropy. A probabilistic model is presented for this exceptional case.

1 Introduction

The content of our previous papers [1, 2, 3, 4] was devoted to a comprehensive study of discrete distributions generalizing the familiar Bernoulli-like (or binomial-like) distributions. The generalization consists in substituting the ordinary integers on which is based the binomial distribution with arbitrary sequences of positive numbers. They can be symmetrical or asymmetrical. The study concerned the positiveness of those formal distributions in order to view them as having a real probabilistic content. We have given many examples, which run from Delone sequences, q-sequences, sequences based on family of polynomials (modified Abel, Hermite...). A key point of our works was to display manageable generating functions. The existence of such functions allows to easily control positiveness and makes a series of computations easier. Hence,

---

*e-mail: herve.bergeron@u-psud.fr, evaldo@cbpf.br, gazeau@apc.univ-paris7.fr, ligia@cbpf.br

1

we have shown in the above references a palette of interesting properties. Nevertheless, except in a few cases, we did not give illustrating models of these new probabilities distributions, and we did not explore systematically their asymptotic behaviors, their associated entropies (Shannon or Boltzmann-Gibbs, Tsallis, Rényi ...), and related questions like extensiveness.

The aim of the present article is to examine comprehensively asymptotic behaviors and associated entropies in the three cases concerning symmetric deformations of the binomial distribution previously presented in [3, 4] and having a sound probabilistic content. Our interest is particularly concerned with the extensivity, asymptotic or not, of these three entropies. We recall that an entropy is extensive (resp. asymptotically extensive) if it is proportional (resp. asymptotically proportional) to the number n of events (resp. at large n). The three probability distributions mentioned above are denoted in this paper by

$$\mathcal{P} = \left(\mathfrak{p}_1^{(n)}, \mathfrak{p}_2^{(n)}, \ldots \mathfrak{p}_n^{(n)} \right) . \tag{1.1}$$

Due to symmetry, the multiplicity of states is the same as for the binomial distribution. In our evaluations of entropies, we adopt a "microscopic" point of view by ignoring the multiplicity.

The first entropy is the Boltzmann-Gibbs (BG) or Shannon [5, 6] entropy.

$$S_{\mathrm{BG}} = - \sum_{k=0}^{n} \mathfrak{p}_k^{(n)} \log \frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} . \tag{1.2}$$

The second one is the Tsallis entropy S_q [7], which is a deformation of (1.2), $S_q \to S_{\mathrm{BG}}$ as $q \to 1$,

$$S_q = \frac{1}{q-1} \left[1 - \sum_{k=0}^{n} \binom{n}{k} \left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} \right)^q \right] . \tag{1.3}$$

The third one is the Rényi entropy $S_{\mathrm{Re};q}$ [9], which is also a deformation of (1.2) $S_{\mathrm{Re};q} \to S_{\mathrm{BG}}$ as $q \to 1$,

$$S_{\mathrm{Re};q} = \frac{1}{1-q} \log \left[\sum_{k=0}^{n} \binom{n}{k} \left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} \right)^q \right] . \tag{1.4}$$

The organization of the paper is the following. In Section 2 the necessary background issued from [4] is rewieved. The first case, examined in Section 3, has the so-called q-exponential as a generating function. It gives rise to a nice probabilistic interpretation (e.g. Polya urns) and to an extensive Boltzmann-Gibbs entropy, as was already mentioned in [4]. We show that the Rényi entropy is also extensive for this distribution. The second case is related to modified Abel polynomials and has the exponential of the Lambert function as a generating function. This example forms the content of Section 4 and yields an unexpected nontrivial result. Indeed, the entropy which is asymptotically extensive in this case is not Boltzmann-Gibbs, nor the Tsallis q-entropy for any

q, but instead the Rényi one, and its asymptotic behavior does not depend on the Rényi parameter. Due to the importance of this result, we present in the same section a probabilistic model based on counting of words made with letters picked in several alphabets. This model is quite elaborate. Section 5 is devoted to our third example, involving Hermite polynomials. With this case, we return to the standard situation for which both Boltzmann-Gibbs and Rényi are extensive. Following our conclusions and comments in Section 6 are the first appendix where we present an historical survey of the concept of entropy, and the second one where we give the necessary technical details.

2 Symmetric deformations of binomial distributions

We remind in this section notations and main results of [4].

Let $\mathcal{X} = (x_n)_{n \in \mathbb{N}}$ be a sequence of positive numbers x_n for $n > 0$ and $x_0 = 0$. The "factorial" of x_n is defined as $x_n! = x_1 x_2 \cdots x_n, \quad x_0! \stackrel{\text{def}}{=} 1$, and from it we build the binomial coefficient

$$\binom{x_n}{x_k} := \frac{x_n!}{x_{n-k}! x_k!} \,.$$

We now associate to $\mathcal{X}$ the formal distribution

$$\mathfrak{p}_k^{(n)}(\eta) = \binom{x_n}{x_k} q_k(\eta) q_{n-k}(1-\eta) \,, \tag{2.1}$$

where the $q_k(\eta)$ are polynomials of degree k and the $\mathfrak{p}_k^{(n)}(\eta)$ are constrained by the *normalization* condition

$$\forall n \in \mathbb{N}, \quad \forall \eta \in [0,1], \quad \sum_{k=0}^{n} \mathfrak{p}_k^{(n)}(\eta) = 1, \tag{2.2}$$

and by the *non-negativeness* condition

$$\forall n, k \in \mathbb{N}, \quad \forall \eta \in [0,1], \quad \mathfrak{p}_k^{(n)}(\eta) \geq 0 \,. \tag{2.3}$$

The normalization implies

$$\forall \eta \in [0,1], \quad \mathfrak{p}_0^{(0)}(\eta) = q_0(\eta) q_0(1-\eta) = 1 \Rightarrow q_0(\eta) = \pm 1$$

From now on we keep the choice $q_0(\eta) = 1$. This implies

$$\forall n \in \mathbb{N}, \quad \forall \eta \in [0,1], \quad \mathfrak{p}_0^{(n)}(\eta) = q_n(1-\eta).$$

Therefore the non-negativeness condition is equivalent to the non-negativeness of the polynomials q_n on the interval $[0,1]$. The quantity $\mathfrak{p}_k^{(n)}(\eta)$ can be interpreted as the probability of having k wins and $n-k$ losses in a sequence of

correlated n trials. Besides, as we recover the invariance under $k \to n - k$ and $\eta \to 1 - \eta$ of the binomial distribution, no bias in the case $\eta = 1/2$ can exist favoring either win or loss.

We now associate to the sequence $\mathcal{X}$ an "exponential" defined as the entire series

$$\mathcal{N}(t) = \sum_{n=0}^{\infty} \frac{t^n}{x_n!} \equiv \sum_{n=0}^{\infty} a_n t^n, \quad x_n = a_{n-1}/a_n, \tag{2.4}$$

which is supposed to have a non-vanishing radius of convergence. Hence $\mathcal{N}(\mathrm{t})$ is an element of Σ defined as the set of entire series $\sum_{n=0}^{\infty} a_n t^n$ possessing a non-vanishing radius of convergence and verifying $a_0 = 1$ and $\forall n \geq 1, a_n > 0$.

Starting from $\mathcal{N}(t) \in \Sigma$ and $\eta \in [0, 1]$, we consider the series $\mathcal{N}(t)^\eta$. It is easy to prove from $\mathcal{N}(t) = \mathcal{N}(t)^\eta \mathcal{N}(t)^{1-\eta}$ that it is a generating function for polynomials q_n obeying (2.1)-(2.2):

$$\forall \eta \in [0, 1], \quad \tilde{G}_{\mathcal{N}, \eta}(t) := \mathcal{N}(t)^\eta = \sum_{n=0}^{\infty} \frac{q_n(\eta)}{x_n!} t^n. \tag{2.5}$$

More precisely, the polynomials q_n issued from (2.5) have the following properties:

(a) $q_0(\eta) = 1$, $q_1(\eta) = \eta$ and more generally

$$\forall n \in \mathbb{N}, \, \forall \eta \in [0, 1], \, q_{n+1}(\eta) = \eta \frac{x_{n+1}}{n+1} \times$$
$$\times \sum_{k=0}^{n} \binom{x_n}{x_k} \frac{n-k+1}{x_{n-k+1}} q_k(\eta - 1). \tag{2.6}$$

(b) The q_n's are polynomials of degree n obeying

$$\forall n \in \mathbb{N}, \, q_n(1) = 1, \text{ and } \forall n \neq 0, \, q_n(0) = 0,$$

and they fulfill the normalization condition.

(c) The q_n's fulfill the functional relation

$$\forall z_1, z_2 \in \mathbb{C}, \, \forall n \in \mathbb{N}, \sum_{k=0}^{n} \binom{x_n}{x_k} q_k(z_1) q_{n-k}(z_2) = q_n(z_1 + z_2). \tag{2.7}$$

We note that these polynomials, suitably normalized, are of binomial type.

Since $\log(\mathcal{N}(t))$ is analytical around $t = 0$, $\mathcal{N}(t)^\eta = \exp(\eta \log(\mathcal{N}(t)))$ possesses a convergent series expansion around $t = 0$ (for all $\eta \in \mathbb{C}$).

Since we already know that $q_0(\eta) = 1$ and $\forall n \neq 0, q_n(0) = 0$, the non-negativeness condition is equivalent to specify that for any $\eta \in]0, 1]$, $q_n(\eta) > 0$ and then the function $t \mapsto \mathcal{N}(t)^\eta$ belongs to Σ. Defining Σ_0 as the set of entire series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ possessing a non-vanishing radius of convergence and

verifying the conditions $a_0 = 0$, $a_1 > 0$ and $\forall n \geq 2$, $a_n \geq 0$, it was proved in [4] that

$$\Sigma_+ := \{\mathcal{N} \in \Sigma \mid \forall \eta \in [0,1), \forall n \geq 0, q_n(\eta) > 0\} = \{e^F \mid F \in \Sigma_0\} \qquad (2.8)$$

is the set of deformed exponentials such that the generating functions $G_{\mathcal{N},\eta}(t)$ solve the non-negativeness problem.

3 Symmetric distribution from "q-exponential"

3.1 The probabilty distribution

We consider here the following family of functions belonging to Σ_+:

$$\mathcal{N}(t) = \left(1 - \frac{t}{\alpha}\right)^{-\alpha}, \quad \alpha > 0, \qquad (3.1)$$

that are q-exponentials in the sense that $e_{\mathsf{q}}(x) = [1 - (1 - \mathsf{q})x]^{(1/(1-\mathsf{q}))}$, where the parameter $\mathsf{q} = 1 + 1/\alpha$ with the notations in [8]. We first note that if $\alpha \to \infty$ then $\mathcal{N}(t) \to e^t$, i.e. we return to the ordinary binomial case. The corresponding sequence is bounded by α and given by

$$x_n = \frac{n\alpha}{n + \alpha - 1}, \quad \lim_{n \to \infty} x_n = \alpha. \qquad (3.2)$$

For the factorial we have:

$$x_n! = \alpha^n \frac{\Gamma(\alpha)n!}{\Gamma(n + \alpha)} = \frac{\alpha^n n!}{(\alpha)_n} = \frac{\alpha^n}{\binom{n+\alpha-1}{n}}, \qquad (3.3)$$

where $(z)_n = \Gamma(z + n)/\Gamma(z)$ is the Pochhammer symbol. The corresponding polynomials are given by

$$q_n(\eta) = \frac{\Gamma(\alpha)}{\Gamma(n + \alpha)} \frac{\Gamma(n + \alpha\eta)}{\Gamma(\alpha\eta)} = \frac{(\alpha\eta)_n}{(\alpha)_n}, \qquad (3.4)$$

and satisfy the recurrence relation

$$q_n(\eta) = \frac{n + \alpha\eta - 1}{n + \alpha - 1} q_{n-1}(\eta), \quad \text{with } q_0(\eta) = 1. \qquad (3.5)$$

In particular $q_1(\eta) = \eta$. The distribution $\mathfrak{p}_k^{(n)}(\eta)$ defined by these polynomials is given by

$$\mathfrak{p}_k^{(n)}(\eta) = \binom{n}{k} \frac{\Gamma(\alpha)}{\Gamma(\eta\alpha)\Gamma((1-\eta)\alpha)} \frac{\Gamma(\eta\alpha + k)\Gamma((1-\eta)\alpha + n - k)}{\Gamma(\alpha + n)} \qquad (3.6)$$

$$= \frac{\binom{\eta\alpha + k - 1}{k}\binom{(1-\eta)\alpha + n - k - 1}{n - k}}{\binom{\alpha + n - 1}{n}}. \qquad (3.7)$$

This is precisely the Pólya distribution [10], also called "Markov-Pólya" or "inverse hypergeometric" and more. It was considered by Pólya (1923) in the following urn scheme [11]. From a set of b black balls and r red balls contained in an urn one extracts one ball and return it to the urn together with c balls of the same color. The probability to have in the urn k black balls after the n-th trial is given by the ratio (3.7) with

$$\eta = \frac{b}{b+r} \,, \qquad \alpha = \frac{b+r}{c} \,, \tag{3.8}$$

which holds for rational parameters η and α. In this notation, the distribution (3.6) reads, in terms of Pochammer symbol,

$$\mathfrak{p}_k^{(n)}(b,c,r) = \binom{n}{k} \frac{\left(\dfrac{b}{c}\right)_k \left(\dfrac{r}{c}\right)_{n-k}}{\left(\dfrac{b+r}{c}\right)_n} \,. \tag{3.9}$$

We notice that if we take the medium value $\eta = 1/2$ and redefine the parameters according to $\alpha \to 2\nu$, $n \to N$ and $k \to n$ in the distribution given by Eq. (3.6) we recover the distribution r_n^N studied in reference [12], see Eqs. (4) and (10) therein, within the framework of the Laplace-de Finetti representation.

3.2 Asymptotic behavior at large n

Let us now study the asymptotic behavior of (3.6) at large n. The probability distribution is given by:

$$\begin{aligned}
\mathfrak{p}_k^{(n)}(\eta) &= \binom{n}{k} \frac{\Gamma(\alpha)}{\Gamma(\eta\alpha)\Gamma((1-\eta)\alpha)} \frac{\Gamma(\eta\alpha + k)\Gamma((1-\eta)\alpha + n - k)}{\Gamma(\alpha + n)} \\
&= \binom{n}{k} \frac{B(\eta\alpha + k, (1-\eta)\alpha + n - k)}{B(\eta\alpha, (1-\eta)\alpha)} \,,
\end{aligned}$$

where $0 \le \eta \le 1$, $\alpha > 0$, and $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(q+p)$ is the beta function . We put $k = nx$, with $0 < x < 1$. Using the Stirling formula, $n! \sim \sqrt{2\pi}\, e^{-n} n^{n+1/2}$ or $\Gamma(z) \sim \sqrt{2\pi}\, e^{(z-1/2)\log z - z}$, we find

$$B(\eta\alpha + k, (1-\eta)\alpha + n - k) \sim \sqrt{\frac{2\pi}{n}}\, x^{\eta\alpha - 1/2} (1-x)^{(1-\eta)\alpha - 1/2}\, e^{-n\mathcal{C}(x)} \,,$$

where we introduced

$$\mathcal{C}(x) := -x \log x - (1-x)\log(1-x) \,,$$
$$\text{with} \quad \mathcal{C}'(x) = -\log \frac{x}{1-x} \,, \quad \mathcal{C}''(x) = -\frac{1}{x(1-x)} \,. \tag{3.10}$$

For $x \in (0,1)$ this function is nonnegative, concave and symmetric with respect to its maximum value $\log 2$ at $x = 1/2$. In fact, $\mathcal{C}(x)$ is the basic BG (or

Shannon) entropy in the case of two possibilities with probabilities x and $1-x$, and it appears in many places in the paper.

The asymptotic behavior of the binomial coefficient at large n is (see (B.2))

$$\binom{n}{k=nx} \sim \frac{1}{\sqrt{2\pi n x(1-x)}} \, e^{n\mathcal{C}(x)} \, .$$

Therefore, the limit distribution we find is the following:

$$\mathfrak{p}_{k=nx}^{(n)}(\eta) \sim \frac{1}{n} \frac{1}{B(\eta\alpha, (1-\eta)\alpha)} \, x^{\eta\alpha-1} \, (1-x)^{(1-\eta)\alpha-1} \, . \qquad (3.11)$$

We easily check that the probabilistic normalisation $\sum_{k=0}^{n} \mathfrak{p}_{k=nx}^{(n)} = 1$ remains valid at the limit $n \to \infty$. Indeed, replacing the sum $\sum_{k=0}^{n}$ by the integral $\int_0^1 n \, \mathrm{d}x$ leads to

$$\sum_{k=0}^{n} \mathfrak{p}_{k=nx}^{(n)} \sim \frac{1}{B(\eta\alpha, (1-\eta)\alpha)} \int_0^1 x^{\eta\alpha-1} \, (1-x)^{(1-\eta)\alpha-1} \, \mathrm{d}x = 1 \, . \qquad (3.12)$$

Moreover, our asymptotic formula (3.11), in the case $\eta = 1/2$ and after centering on the origin, becomes proportional to a Q-Gaussian [13] with $Q = (\alpha - 4)/(\alpha - 2)$. This result was recently obtained numerically by Ruiz and Tsallis [14].

3.3 Boltzmann-Gibbs entropy

We take a definition of the BG entropy for the distribution (3.6) which does not take into account the multiplicity of states, because as we have already mentioned in the introduction, we are adopting a microscopic point of view. Consequently we replace the random variable $-\log \mathfrak{p}_k^{(n)}$ by $-\log\left(\mathfrak{p}_k^{(n)}/\binom{n}{k}\right)$:

$$S_{\mathrm{BG}} = -\sum_{k=0}^{n} \mathfrak{p}_k^{(n)} \log \frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} \, . \qquad (3.13)$$

The division of the probability by the binomial coefficient in each logarithm in (3.13) means a counting of the degeneracy. As a preliminary numerical exploration, its extensive property is shown in Figure 1 where it is compared with the Tsallis entropy

$$S_q^{(n)} \overset{\mathrm{def}}{=} (1 - \sum_{k=0}^{n} \binom{n}{k} \left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}}\right)^q)/(q - 1) \, .$$

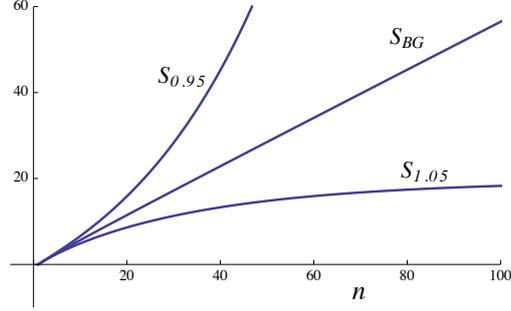Let us now establish the analytic formula for S_{BG} in the asymptotic limit as

Figure 1: n-dependence of entropies for the "q-exponential" case ($\mathcal{N}(t) = (1 - t/\alpha)^{-\alpha}$). The Boltzmann-Gibbs (BG) and Tsallis (S_q) entropies of the distribution are shown for $\eta = 1/2$ and $\alpha = 3$. Upper curve is for $q = 0.95$. Bottom curve is for $q = 1.05$.

$n \to \infty$. In the present case the latter behaves as

$$
S_{\text{BG}} \underset{\text{at large } n}{\sim} -\frac{1}{B(\eta\alpha, (1-\eta)\alpha)} \int_0^1 \mathrm{d}x\, x^{\eta\alpha-1} (1-x)^{(1-\eta)\alpha-1} \times
$$

$$
\times \log\left[\frac{1}{B(\eta\alpha, (1-\eta)\alpha)} \sqrt{\frac{2\pi}{n}}\, x^{\eta\alpha-1/2} (1-x)^{(1-\eta)\alpha-1/2}\, e^{-n\mathcal{C}(x)} \right]
$$

$$
= nI_1 + \frac{1}{2}\log n + I_2\,,
$$

with

$$
I_1 = \frac{1}{B(\eta\alpha, (1-\eta)\alpha)} \int_0^1 \mathrm{d}x\, x^{\eta\alpha-1} (1-x)^{(1-\eta)\alpha-1}\, \mathcal{C}(x)\,,
$$

$$
I_2 = \log\left(\frac{B(\eta\alpha, (1-\eta)\alpha)}{\sqrt{2\pi}} \right) - \frac{1}{B(\eta\alpha, (1-\eta)\alpha)} \int_0^1 \mathrm{d}x\, x^{\eta\alpha-1} (1-x)^{(1-\eta)\alpha-1} \times
$$

$$
\times \left[\left(\eta\alpha - \frac{1}{2} \right) \log x + \left((1-\eta)\alpha - \frac{1}{2} \right) \log(1-x) \right]\,.
$$

Since the term in n is dominant, the Boltzmann-Gibbs entropy is proved to be extensive in the present case. Let us calculate the integrals appearing in the above expressions. They are all of the type

$$
LB(p, q) := \int_0^1 \mathrm{d}x\, x^{p-1} \log x\, (1-x)^{q-1} = \int_0^1 \mathrm{d}x\, (1-x)^{p-1} \log(1-x)\, x^{q-1}
$$

$$
= \frac{\partial}{\partial p} B(p, q) = [\psi(p) - \psi(p+q)]\, B(p, q)\,, \quad \psi(t) = \frac{\Gamma'(t)}{\Gamma(t)}\,. \qquad (3.14)
$$

Finally, we find

$$S_{\text{BG}} \sim n[\psi(\alpha+1) - (\eta\psi(\eta\alpha+1) + (1-\eta)\psi((1-\eta)\alpha+1)] + \frac{1}{2}\log n +$$
$$+ \log\left(\frac{B(\eta\alpha, (1-\eta)\alpha)}{\sqrt{2\pi}}\right)$$
$$+ \alpha\psi(\alpha) - \left[\left(\eta\alpha - \frac{1}{2}\right)\psi(\eta\alpha) + \left((1-\eta)\alpha - \frac{1}{2}\right)\psi((1-\eta)\alpha)\right]. \quad (3.15)$$

3.4 Rényi Entropy

We finally explore, for the present case, the Rényi entropy

$$S_{\text{Re};q} = \frac{1}{1-q}\log\left[\sum_{k=0}^{n}\binom{n}{k}\left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}}\right)^q\right], \quad (3.16)$$

which becomes S_{BG} as $q \to 1$. Using the asymptotic formula of Eqs. (B.2) and (3.11), the approximation $\sum_{k=0}^{n} \sim \int_0^1 n\mathrm{d}x$, and the Laplace formula (see (5.16)), we obtain the asymptotic expression for $q < 1$

$$\sum_{k=0}^{n}\binom{n}{k}\left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}}\right)^q \sim \frac{1}{2\sqrt{1-q}}\left[\frac{2^{5-2\alpha}\pi}{nB^2(\eta\alpha, (1-\eta)\alpha)}\right]^{q/2} e^{n(1-q)\log 2}. \quad (3.17)$$

By taking the logarithm of (3.17), we see that the dominant term is

$$S_{\text{Re};q} \sim n\log 2. \quad (3.18)$$

and the Rényi entropy is obviously extensive. A point to be noticed is that this asymptotic behavior is independent of the Rényi parameter q. Actually this remarkable feature is encountered in many distributions [15], including the next two cases considered in this paper. We will give a special attention to this fact in Section 6.

4 Symmetric distribution from modified Abel polynomials

4.1 Probability distribution

We take here the specific generating function $\mathcal{N}(t)$ given by

$$\mathcal{N}(t) = e^{-\alpha W(-t/\alpha)}, \quad \alpha > 0, \quad (4.1)$$

where W is the Lambert function [16], i.e. solving the functional equation $W(t)e^{W(t)} = t$. We first note that if $\alpha \to \infty$ then $\mathcal{N}(t) \to e^t$. The corresponding sequence is bounded by α/e and given by

$$x_n = \frac{n\alpha}{n+\alpha}\left(1 - \frac{1}{n+\alpha}\right)^{n-2}, \quad \lim_{n\to\infty} x_n = \alpha/e. \quad (4.2)$$

9

We also note that $x_n \to n$ as $\alpha \to \infty$. The corresponding factorial is

$$x_n! = n! \, \frac{\alpha^{n-1}}{(n+\alpha)^{n-1}} \,. \tag{4.3}$$

The polynomials q_n's read as

$$q_n(\eta) = \eta \frac{\left(\eta + \frac{n}{\alpha}\right)^{n-1}}{\left(1 + \frac{n}{\alpha}\right)^{n-1}} \,. \tag{4.4}$$

We verify that $q_0(\eta) = 1$ and $q_1(\eta) = \eta$. The polynomials above are a kind of modified Abel polynomials [17] which look like

$$P_n(x) = x(x+na)^{n-1} \,, a \in \mathbb{Q} \,, \tag{4.5}$$

with the difference of the presence of a normalization factor in the denominator of (5.3) and the relaxing of the rational condition.

The corresponding probability distribution is found to be:

$$\mathfrak{p}_k^{(n)}(\eta) = \binom{n}{k} \eta(1-\eta) \frac{(\eta + k/\alpha)^{k-1}(1 - \eta + (n-k)/\alpha)^{n-k-1}}{(1 + n/\alpha)^{n-1}} \,, \tag{4.6}$$

with $0 \le \eta \le 1$.

4.2 Regularization at the limit $n \to \infty$

Putting $k = nx$ in (4.6), with $0 < x < 1$, and using the Stirling formula, we find the limit distribution

$$\mathfrak{p}_{nx}^{(n)}(\eta) \underset{\text{large } n}{\sim} \frac{\alpha\eta(1-\eta)}{\sqrt{2\pi}} (nx(1-x))^{-3/2} \,. \tag{4.7}$$

The problem is that if one replaces the discrete sum $\sum_{k=0}^{n}$ by the integral $\int_0^1 n\,\mathrm{d}x$ this expression leads to a divergent integral. However, there is a simple way to give it a finite value through a sort of principal value. First, let us consider the finite convergent integral

$$B_\epsilon(a,a) := \int_\epsilon^{1-\epsilon} x^{a-1}(1-x)^{a-1}\mathrm{d}x \tag{4.8}$$

with a small $\epsilon > 0$ and arbitrary a. Due to the symmetry of the integrand under the interchange $x \to 1 - x$, we have

$$B_\epsilon(a,a) = 2\int_\epsilon^{\frac{1}{2}} x^{a-1}(1-x)^{a-1}\mathrm{d}x \,. \tag{4.9}$$

By expanding the binomial $(1-x)^{a-1}$ we easily find its expression in terms of Gauss hypergeometric function:

$$B_\epsilon(a,a) = \frac{2}{a} \left[2^{-a}{}_2F_1\left(a, 1-a; a+1; \frac{1}{2}\right) - \epsilon^a {}_2F_1\left(a, 1-a; a+1; \epsilon\right) \right] \,. \tag{4.10}$$

We now consider our particular case $a = -1/2$. By using the formula [19]

$$_2F_1\left(a, 1-a; b; \frac{1}{2}\right) = \sqrt{\pi} 2^{1-b}\Gamma(b)\left[\Gamma\left(\frac{a+b}{2}\right)\Gamma\left(\frac{1+b-a}{2}\right)\right]^{-1}, \quad (4.11)$$

and from $1/\Gamma(0) = 0$, $_2F_1(a, 1-a; a+1; \epsilon) \approx 1$ at small ϵ, we eventually find

$$B_\epsilon\left(-\frac{1}{2}, -\frac{1}{2}\right) \sim \frac{4}{\sqrt{\epsilon}}. \quad (4.12)$$

Now, from (4.7), we have

$$\sum_{k=0}^{n} \mathfrak{p}_{nx}^{(n)}(\eta) \underset{\text{large } n}{\sim} \frac{\alpha\eta(1-\eta)}{\sqrt{2\pi}} n^{-1/2} \lim_{\epsilon\to 0} \int_{\epsilon}^{1-\epsilon} x^{-\frac{3}{2}}(1-x)^{-\frac{3}{2}}\mathrm{d}x \quad (4.13)$$

$$\underset{\text{large } n}{\sim} \lim_{\epsilon\to 0} \frac{4\alpha\eta(1-\eta)}{\sqrt{2\pi}} \frac{1}{\sqrt{n\epsilon}}. \quad (4.14)$$

It is then legitimate to put $\epsilon = A/n$, where the arbitrarily constant A is consistently chosen as $A = 8(\alpha\eta(1-\eta))^2/\pi$, in such a way that the original expression remains equal to 1.

4.3   Boltzmann-Gibbs Entropy

We first examine the Boltzmann-Gibbs entropy for the limit distribution (4.7). From (4.7) and (B.2) we find the asymptotic behavior of each term in the defining sum of the BG entropy:

$$\mathfrak{p}_k^{(n)} \log \frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} \sim \frac{\alpha\eta(1-\eta)}{\sqrt{2\pi}}(nx(1-x))^{-3/2}\times$$

$$\times \log\left[\alpha\eta(1-\eta)(nx(1-x))^{-1}e^{-n\mathcal{C}(x)}\right]. \quad (4.15)$$

After replacing the sum $\sum_{k=0}^{n}$ by the integral $\int_0^1 n\mathrm{d}x$ in

$$S_{\text{BG}} = -\sum_{k=0}^{n} \mathfrak{p}_k^{(n)} \log \frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}}, \quad (4.16)$$

the BG entropy behaves as the sum of three terms

$$S_{\text{BG}} \underset{\text{large } n}{\sim} \frac{1}{\sqrt{n}}\left(\lim_{\epsilon_1\to 0} R_{1;\epsilon_1} + \lim_{\epsilon_2\to 0} R_{2;\epsilon_2}\right) + \sqrt{n}R_3, \quad (4.17)$$

where

$$R_{1;\epsilon_1} = -\frac{\alpha\eta(1-\eta)}{\sqrt{2\pi}}\log\left(\frac{\alpha\eta(1-\eta)}{n}\right)B_{\epsilon_1}\left(-\frac{1}{2}, -\frac{1}{2}\right)$$

$$R_{2;\epsilon_2} = 2\frac{\alpha\eta(1-\eta)}{\sqrt{2\pi}}LB_{\epsilon_2}\left(-\frac{1}{2}, -\frac{1}{2}\right)$$

$$R_3 = -2\frac{\alpha\eta(1-\eta)}{\sqrt{2\pi}}\int_0^1 x^{-1/2}(\log x)(1-x)^{-3/2}\mathrm{d}x.$$

Here we have introduced the notation

$$LB_\epsilon(a,a) := \int_\epsilon^{1-\epsilon} x^{a-1}(\log x)(1-x)^{a-1}\mathrm{d}x$$
$$= \int_\epsilon^{1-\epsilon} x^{a-1}(\log(1-x))(1-x)^{a-1}\mathrm{d}x, \qquad (4.18)$$

and made use of symmetries $x \mapsto 1-x$ in the integrals. From the general expression

$$LB_\epsilon(a,a) = \frac{1}{2}\frac{\partial}{\partial a}B_\epsilon(a,a) = -\frac{1}{2a}B_\epsilon(a,a) - \frac{\sqrt{\pi}}{a}2^{1-2a}\log 2\frac{\Gamma(a+1)}{\Gamma\left(a+\frac{1}{2}\right)} +$$
$$+ \frac{\sqrt{\pi}}{a}2^{-2a}\frac{\Gamma(a+1)}{\Gamma\left(a+\frac{1}{2}\right)}\left[\psi(a+1) - \psi\left(a+\frac{1}{2}\right)\right] +$$
$$- \frac{1}{a}\log\epsilon\,\epsilon^a\,{}_2F_1(a,1-a;a+1;\epsilon) - \frac{1}{a}\epsilon^a\,O(\epsilon),$$

we find

$$LB_\epsilon\left(-\frac{1}{2}, -\frac{1}{2}\right) \sim \frac{4}{\sqrt{\epsilon}} + \frac{2\log\epsilon}{\sqrt{\epsilon}} \sim \frac{2\log\epsilon}{\sqrt{\epsilon}}.$$

In (4.17) choosing $\epsilon_1 \propto 1/n$ and $\epsilon_2(\log\epsilon_2)^2 \propto 1/n$ we get rid of both the integral divergences in $R_{1;\epsilon}$ and $R_{2;\epsilon}$ respectively. The value of R_3 is easily found from [18]:

$$LI_p := \int_0^1 x^{p-1}(\log x)(1-x)^{-p-1}\,\mathrm{d}x = -\frac{\pi}{p}\csc p\pi, \quad 0 < p < 1. \qquad (4.19)$$

In the present case, $p = 1/2$ and so $LI_p = -2\pi$. Hence, we see that the dominant term in (4.17) is

$$S_{\mathrm{BG}} \sim 2\sqrt{2\pi}\,\alpha\,\eta\,(1-\eta)\,\sqrt{n}. \qquad (4.20)$$

Hence, we conclude that the Boltzmann-Gibbs entropy is not extensive for this type of deformation of the binomial distribution and with the chosen regularization of integrals. It behaves as $\sqrt{n}$, as is also shown in Figure 2 obtained from numerical computations of (4.16).

4.4 Tsallis Entropy

Let us now explore, for the present case, the Tsallis entropy defined as

$$S_q = \frac{1}{q-1}\left[1 - \sum_{k=0}^n \binom{n}{k}\left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}}\right)^q\right], \qquad (4.21)$$

which becomes S_{BG} as $q \to 1$. We first estimate the general term in the sum:

$$\binom{n}{nx}\left(\frac{\mathfrak{p}_{nx}^{(n)}}{\binom{n}{nx}}\right)^q \sim \frac{(\alpha\eta(1-\eta))^q}{\sqrt{2\pi}}n^{-q-1/2}(x(1-x))^{-q-1/2}e^{-n(q-1)\mathcal{C}(x)}. \qquad (4.22)$$
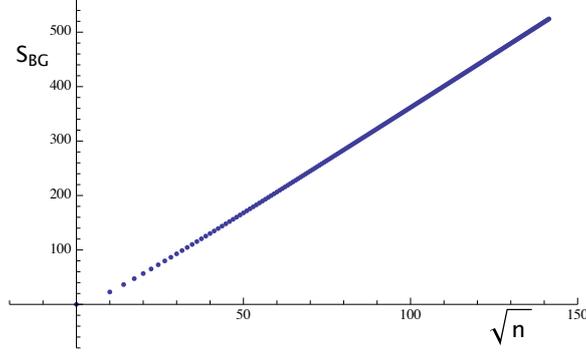
Figure 2: Numerical behavior of the Boltzmann-Gibbs entropy (4.16) versus $\sqrt{n}$ for the symmetric distribution from modified Abel polynomials, with $\eta = 0.8$, $\alpha = 5$ and n up to 20 000.

We now replace the sum $\sum_{k=0}^{n}$ by the integral $\int_0^1 n\,\mathrm{d}x$ and obtain

$$\sum_{k=0}^{n} \binom{n}{k} \left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} \right)^q \sim \frac{(\alpha\eta(1-\eta))^q}{\sqrt{2\pi}} \, n^{1/2-q} \int_0^1 (x(1-x))^{-q-1/2} \, e^{-n(q-1)\mathcal{C}(x)} \, \mathrm{d}x \,, \tag{4.23}$$

and this would impose a convergence condition $q < 1/2$ if we were not in the very large n regime. With the properties (3.10) of $\mathcal{C}(x)$, the use of the Laplace approximation method with the condition $q < 1$ yields

$$\sum_{k=0}^{n} \binom{n}{k} \left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} \right)^q \sim \frac{1}{\sqrt{|q-1|}} \, 2^{2q}(\alpha\eta(1-\eta))^q \, n^{-q} \, e^{n(1-q)\log 2} \,, \tag{4.24}$$

and the Tsallis entropy becomes

$$S_q \sim \frac{1}{(q-1)\sqrt{|q-1|}} \left[\sqrt{|q-1|} - 2^{2q}(\alpha\eta(1-\eta))^q \, n^{-q} \, e^{n(1-q)\log 2} \right] \,. \tag{4.25}$$

We see that the Tsallis entropy is not extensive for any value of $q < 1$. However, we should be aware that our derivation prevents us to consider the asymptotic form of S_{BG} in (4.20) as the limit at $q \to 1$ of (4.25), since the Laplace approximation method in (4.23) loses its validity for $q = 1$.

4.5   Rényi Entropy

We finally explore the Rényi entropy

$$S_{\mathrm{Re};q} = \frac{1}{1-q} \log \left[\sum_{k=0}^{n} \binom{n}{k} \left(\frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}} \right)^q \right] \,. \tag{4.26}$$

13

From (4.24) we derive immediately

$$\log \left[\sum_{k=0}^{n} \binom{n}{nx} \left(\frac{\mathfrak{p}_{nx}^{(n)}}{\binom{n}{nx}} \right)^q \right] \sim \log \left(\frac{1}{\sqrt{|q-1|}} 2^{2q} (\alpha \eta (1-\eta))^q \right) + $$
$$ - q \log n + n(1-q) \log 2 \,. \qquad (4.27)$$

Therefore, the Rényi entropy is extensive for $q < 1$:

$$S_{\mathrm{Re};q} \sim n \log 2 \,. \qquad (4.28)$$

We recover the asymptotic q-independence already noticed in the case of the previous example.

4.6   Probabilistic interpretation

Choosing the parameters α and η in the expression (4.6) as

$$\alpha = \frac{p+q}{c} \quad \text{and} \quad \eta = \frac{p}{p+q} \,, \qquad (4.29)$$

where p, q and c are three positive integers, we obtain

$$\mathfrak{p}_k^{(n)} = \binom{n}{k} \frac{p(p+kc)^{k-1} q(q+(n-k)c)^{n-k-1}}{(p+q)(p+q+nc)^{n-1}} \,. \qquad (4.30)$$

¿From the sum of probabilities, we deduce the finite expansion formula

$$\sum_{k=0}^{n} \binom{n}{k} p(p+kc)^{k-1} q(q+(n-k)c)^{n-k-1} = (p+q)(p+q+nc)^{n-1} \,. \qquad (4.31)$$

We now present counting interpretation of this expansion and its resulting urn model. We define a finite set for which the numbers $\binom{n}{k} p(p+kc)^{k-1} q(q+(n-k)c)^{n-k-1}$ for $k = 0, 1 \ldots$ correspond to counting of partitions. As our main interest is to present at least one sound probabilistic model, for the sake of simplicity we consider the case $c = 1$.

4.6.1   The model

Let $\mathcal{A}(2n, p, q) = \mathcal{A}_{C\ell}(2n) \cup \mathcal{A}_\ell(p) \cup \mathcal{A}_C(q)$ be an alphabet of $2n + p + q$ letters viewed as the union of three sub-alphabets:

- $\mathcal{A}_\ell(p)$, $p \geq 1$, is a set $\{b_1, b_2, \ldots, b_p\}$ of p letters which are only lowercase, by convention $\mathcal{A}_\ell(0) = \emptyset$,

- $\mathcal{A}_C(q)$, $q \geq 1$, is a set $\{C_1, C_2, \ldots, C_q\}$ of q letters which are solely capital, by convention $\mathcal{A}_C(0) = \emptyset$,

- The family $\{\mathcal{A}_{C\ell}(2n)\}_{n=1}^{\infty}$ where $\mathcal{A}_{C\ell}(2n) = \bigcup_{i=1}^{n}\{a_i, A_i\}$ made of $2n$ mixed letters, built from a possible infinite sequence of pairs

$$(a_1, A_1), \ldots (a_i, A_i), \ldots$$

  Each pair (a_i, A_i) is made from the same letter in both sizes (lowercase and capital), and the letters are assumed to be different in different pairs, independently of their size. The inclusion $\mathcal{A}_{C\ell}(2n) \subset \mathcal{A}_{C\ell}(2m)$ holds for any $n \leq m$.
  In the following we introduce also the lowercase part of $\mathcal{A}_{C\ell}(2n)$ as $\mathcal{A}_{C\ell}^{\ell}(n) = \bigcup_{i=1}^{n}\{a_i\}$ and the capital part of $\mathcal{A}_{C\ell}(2n)$ as $\mathcal{A}_{C\ell}^{C}(n) = \bigcup_{i=1}^{n}\{A_i\}$.

- All letters, independently of their size, are assumed to be different: in $\mathcal{A}(2n, p, q)$ we have $n + p$ different lowercase letters and $n + q$ different capital letters.

We consider the set of words $\mathcal{W}_n$ with n letters picked from $\mathcal{A}(2n, p, q)$, built as $\mathcal{W}_n = \bigcup_{k=0}^{n}\mathcal{W}_k^n$ where the subsets $\mathcal{W}_k^n$ contain the words with n letters, k of them being lowercase and $n - k$ capital. The words are built with the following rules.

  (i) Different orderings of letters are assumed to give different words,

 (ii) In a word in $\mathcal{W}_k^n$, starting from the left, the first lowercase letter encountered (if $k \neq 0$) belongs to $\mathcal{A}_\ell(p)$, and the first capital letter encountered (if $k \neq n$) belongs to $\mathcal{A}_C(q)$.

(iii) In a word in $\mathcal{W}_k^n$, all the lowercase letters ($k \neq 0$) belong to $\mathcal{A}_\ell(p) \cup \mathcal{A}_{C\ell}^{\ell}(k)$, and all the capital letters ($k \neq n$) belong to $\mathcal{A}_C(q) \cup \mathcal{A}_{C\ell}^{C}(n - k)$.

Now let us evaluate the number of words $\mathcal{N}_k^n$ in $\mathcal{W}_k^n$.

- If $k = 0$ the words contain exactly n capital letters. The first one (from the left) belongs to $\mathcal{A}_C(q)$ and the $n-1$ remaining ones belong to $\mathcal{A}_C(q) \cup \mathcal{A}_{C\ell}^{C}(n)$. This gives

$$\mathcal{N}_0^n = q(q + n)^{n-1} . \tag{4.32}$$

- If $k = 1$, the words contain a unique lowercase letter that belongs to $\mathcal{A}_\ell(p)$, and $n-1$ capital letters. The first capital letter belongs to $\mathcal{A}_C(q)$, the $n-2$ remaining (capital letters) belong to $\mathcal{A}_C(q) \cup \mathcal{A}_{C\ell}^{C}(n - 1)$. Since there is $n = \binom{n}{1}$ ways to locate the lowercase letter in the word, we have

$$\mathcal{N}_1^n = \binom{n}{1}pq(q + n - 1)^{n-2} \tag{4.33}$$

- For $2 \leq k \leq n - 2$, we first choose the k positions of the lowercase letters in the word, there are $\binom{n}{k}$ possibilities. The first lowercase letter belongs to $\mathcal{A}_\ell(p)$, the following $k - 1$ ones belong to $\mathcal{A}_\ell(p) \cup \mathcal{A}_{C\ell}^{\ell}(k)$, then for each choice of the k positions, we have $p(p+k)^{k-1}$ possibilities for the lowercase

15

letters. For the capital letters, we obtain similarly $q(q + n - k)^{n-k-1}$ possibilities. We deduce

$$\mathcal{N}_k^n = \binom{n}{k} p(p + k)^{k-1} q(q + n - k)^{n-k-1} \tag{4.34}$$

- The cases $k = n - 1$ and $k = n$ are analyzed following the same rules, leading to

$$\mathcal{N}_{n-1}^n = \binom{n}{n-1} pq(p + n - 1)^{n-2} \quad \text{and} \quad \mathcal{N}_n^n = p(p + n)^{n-1}. \tag{4.35}$$

We conclude that the formula of Eq.(4.34) is valid for $k = 0, 1 \ldots, n - 1, n$. Using Eq.(4.31) we deduce that the total number $\mathcal{N}_n$ of words of $\mathcal{W}_n$ is $\mathcal{N}_n = (p + q)(p + q + n)^{n-1}$.

Remark The value of $\mathcal{N}_n$ can be easily understood. A generic word of $\mathcal{W}_n$ contains:

- One letter that belongs either to $\mathcal{A}_\ell(p)$ or to $\mathcal{A}_C(q)$: this gives $p + q$ possibilities,
- Each remaining letter is either lowercase belonging to $\mathcal{A}_\ell(p) \cup \mathcal{A}_{C\ell}^\ell(k)$, or capital belonging to $\mathcal{A}_C(q) \cup \mathcal{A}_{C\ell}^C(n - k)$ for some k. This gives $(p + k) + (q + n - k) = p + q + n$ possibilities for each $n - 1$ letters.

Therefore $\mathcal{N}_n = (p + q)(p + q + n)^{n-1}$.

Conclusion The probabilities $\mathfrak{p}_k^{(n)}$ of Eq.(4.30) are the probabilities to extract a word with k lowercase letters after a draw at random from the "urn" $\mathcal{W}_n$.

Remark Other interesting probabilities emerge from this urn model. For example let us call $P(\{l_1, l_2, \ldots\})$ the probability that a word of $\mathcal{W}_n$ contains at least one of the letters of the family $\{l_1, l_2, \ldots\}$. We have the following results

$$\begin{cases} P(\mathcal{A}_\ell(p)) = 1 - \mathfrak{p}_0^{(n)} \\ P(\mathcal{A}_C(q)) = 1 - \mathfrak{p}_n^{(n)} \\ \forall k \geq 2, \; P(\mathcal{A}_\ell(p) \cup \mathcal{A}_{C\ell}^\ell(k)) = \sum_{i=k}^n \mathfrak{p}_i^{(n)} \\ P(\mathcal{A}_\ell(p) \cup \mathcal{A}_{C\ell}^\ell(1)) = \frac{1}{2} P(\mathcal{A}_\ell(p) \cup \mathcal{A}_{C\ell}^\ell(2)) \end{cases} \tag{4.36}$$

4.6.2 An example

Let us illustrate the above counting with the manageable although not trivial case $n = 3$, $p = q = 1$ and the alphabet

$$\mathcal{A} = \{a, b, c, d, A, B, C, D\} \equiv \mathcal{A}_{C\ell}(6) \cup \mathcal{A}_\ell(1) \cup \mathcal{A}_C(1),$$

$$\mathcal{A}_{C\ell}(2) = \{a, A\}, \; \mathcal{A}_{C\ell}(4) = \{a, A\} \cup \{b, B\}, \; \mathcal{A}_{C\ell}(6) = \{a, A\} \cup \{b, B\} \cup \{c, C\},$$

$$\mathcal{A}_{C\ell}^{\ell}(1) = \{a\}, \ \mathcal{A}_{C\ell}^{\ell}(2) = \{a, b\}, \ \mathcal{A}_{C\ell}^{\ell}(3) = \{a, b, c\},$$
$$\mathcal{A}_{C\ell}^{C}(1) = \{A\}, \ \mathcal{A}_{C\ell}^{C}(2) = \{A, B\}, \ \mathcal{A}_{C\ell}^{C}(3) = \{A, B, C\},$$
$$\mathcal{A}_{\ell}(1) = \{d\}, \quad \mathcal{A}_{C}(1) = \{D\}.$$

The total number of possible words of $\mathcal{W}_3$ is $\mathcal{N}_3 = 50$. The set of allowed words with 3 letters built from the above rules is described as follows.

- The subset of words $\mathcal{W}_0^3$ is

$$\begin{pmatrix} DAA & DAB & DAC & DAD \\ DBA & DBB & DBC & DBD \\ DCA & DCB & DCC & DCD \\ DDA & DDB & DDC & DDD \end{pmatrix},$$

  corresponding to $\mathcal{N}_0^3 = 16$ words.

- The subset of words $\mathcal{W}_1^3$ is

$$\begin{pmatrix} dDA & dDB & dDD \\ DdA & DdB & DdD \\ DAd & DBd & DDd \end{pmatrix}.$$

  corresponding to $\mathcal{N}_1^3 = 9$ words.

- The subset of words $\mathcal{W}_2^3$ is

$$\begin{pmatrix} dDa & dDb & dDd \\ daD & dbD & ddD \\ Dda & Ddb & Ddd \end{pmatrix}.$$

  corresponding to $\mathcal{N}_2^3 = 9$ words.

- The subset of words $\mathcal{W}_3^3$ is

$$\begin{pmatrix} daa & dab & dac & dad \\ dba & dbb & dbc & dbd \\ dca & dcb & dcc & dcd \\ dda & ddb & ddc & ddd \end{pmatrix},$$

  corresponding to $\mathcal{N}_3^3 = 16$ words.

The total number of words is $2 \times 16 + 2 \times 9 = 50$. Finally, the probabilities $\mathfrak{p}_k^{(3)}$ corresponding to these 4 situations are given in Table 1.

5 Symmetric distribution from Hermite polynomials

Here the function $\mathcal{N}(t)$ is chosen as

$$\mathcal{N}(t) = e^{t + \frac{a}{2} t^2}, \quad 0 < a < 1. \tag{5.1}$$

17

Table 1: values of $\mathfrak{p}_k^{(n)}$ for $n = 3$, $\alpha = 2 = p + q$, $\eta = p/(p+q) = 1/2$, $p = q = 1$

k	$\mathfrak{p}_k^{(n)}$
0	8/25
1	9/50
2	9/50
3	8/25

The corresponding sequence x_n has the following factorial form:

$$x_n! = \left[\frac{i^n \left(\frac{a}{2}\right)^{n/2}}{n!} H_n\left(\frac{-i}{\sqrt{2a}}\right) \right]^{-1} = \left[\sum_{m=0}^{\lfloor \frac{n}{2}\rfloor} \frac{(a/2)^m}{m!(n-2m)!} \right]^{-1} := \frac{1}{\varphi_n(a)}. \quad (5.2)$$

In particular, $x_1! = x_1 = 1$, $x_2! = 2/(a+1)$. Also, $x_n = \varphi_{n-1}(a)/\varphi_n(a)$, and we know from [2] that $x_n \approx \sqrt{n/a}$ as $n \to \infty$. The corresponding polynomials and probability distributions are respectively given by

$$q_n(\eta) = \frac{x_n!}{n!} \left(i\sqrt{\frac{a\eta}{2}} \right)^n H_n\left(-i\sqrt{\frac{\eta}{2a}}\right) \quad (5.3)$$

and

$$\mathfrak{p}_k^{(n)}(\eta) = \eta^k (1-\eta)^{n-k} \frac{\varphi_k(a/\eta)\varphi_{n-k}(a/(1-\eta))}{\varphi_n(a)}. \quad (5.4)$$

5.1 Asymptotic behavior at large n

Let us evaluate the asymptotic behavior of the probability distribution (5.4). For that, let us rewrite it in terms of Hermite polynomials:

$$\mathfrak{p}_k^{(n)}(\eta) = \binom{n}{k} \eta^{\frac{k}{2}} (1-\eta)^{\frac{n-k}{2}} \frac{H_k\left(-i\sqrt{\frac{\eta}{2a}}\right) H_{n-k}\left(-i\sqrt{\frac{1-\eta}{2a}}\right)}{H_n\left(-i\sqrt{\frac{1}{2a}}\right)}. \quad (5.5)$$

Putting $k = nx$, with $0 < x < 1$, using the Stirling formula (B.1), and the asymptotic behavior of Hermite polynomials versus their respective degree when argument is not real [19][1],

$$|H_n(t)| \sim \frac{n!}{2\,\Gamma\left(\frac{n}{2}+1\right)} e^{\sqrt{2n}|\mathrm{Im}(t)|}, \quad (5.6)$$

we find

$$\mathfrak{p}_{k=nx}^{(n)} \sim \frac{1}{2} \binom{\frac{n}{2}}{\frac{nx}{2}} \eta^{\frac{k}{2}} (1-\eta)^{\frac{n-k}{2}} \exp\left[\sqrt{\frac{n}{a}} \left(\sqrt{x\eta} + \sqrt{(1-x)(1-\eta)} - 1\right) \right]$$

$$\sim \frac{1}{2} \frac{1}{\sqrt{n\pi x(1-x)}} e^{nA(x)}, \quad (5.7)$$

---

[1] Page 255. Actually, a factor 2 in front of $|H_n(t)|$ is missing there.

where

$$A(x) = \frac{x}{2} \log \frac{\eta}{x} + \frac{(1-x)}{2} \log \frac{1-\eta}{1-x} + \frac{1}{\sqrt{na}} (\sqrt{x\eta} + \sqrt{(1-x)(1-\eta)} - 1). \quad (5.8)$$

Let us check if the asymptotic distribution (5.7), continuous with respect to the measure $n\mathrm{d}x$, is correctly normalized,

$$\frac{1}{2} \sqrt{\frac{n}{\pi}} \int_0^1 [x(1-x)]^{-1/2} e^{nA(x)} \, \mathrm{d}x = 1 ? \quad (5.9)$$

For showing this, we use Laplace's method. The two first derivatives of the function $A(x)$ are given by

$$A'(x) = \frac{1}{2} \left(\log \frac{\eta}{x} - \log \frac{1-\eta}{1-x} \right) + \frac{1}{2\sqrt{na}} \left(\sqrt{\frac{\eta}{x}} - \sqrt{\frac{1-\eta}{1-x}} \right),$$

$$A''(x) = -\frac{1}{2} \left(\frac{1}{x} + \frac{1}{1-x} \right) - \frac{1}{4\sqrt{na}} \left(\sqrt{\frac{\eta}{x^3}} + \sqrt{\frac{1-\eta}{(1-x)^3}} \right).$$

We see that in the integration interval $A''(x) < 0$, $A'(x) = 0$ for $x = \eta$ (unique root), and that the values assumed by $A(x)$ and $A''(x)$ at this value are respectively

$$A(\eta) = 0, \quad A''(\eta) \sim -\frac{1}{2\eta(1-\eta)}. \quad (5.10)$$

Then let us apply the Laplace approximation formula (with suitable conditions on the functions involved)

$$\int_a^b h(x) \, e^{nA(x)} \, \mathrm{d}x \sim \sqrt{\frac{2\pi}{n|A''(x_0)|}} \, h(x_0) \, e^{nA(x_0)} \quad \text{as } n \to \infty, \quad (5.11)$$

where $A'(x_0) = 0$ for $x_0 \in [a,b]$, $A''(x_0) < 0$ and h is positive. We get in our case,

$$\frac{1}{2} \int_0^1 \mathfrak{p}_k^{(n)} \, n\mathrm{d}x \sim \sqrt{\frac{n}{\pi}} \int_0^1 [x(1-x)]^{-1/2} e^{nA(x)} \, \mathrm{d}x \sim 1. \quad (5.12)$$

$\square$

5.2 Boltzmann-Gibbs entropy

¿From the asymptotic behavior (5.7) and (B.2) we infer the following behavior

$$\mathfrak{p}_{nx}^{(n)} \log \frac{\mathfrak{p}_{nx}^{(n)}}{\binom{n}{nx}} \sim \frac{1}{2\sqrt{n\pi}} \, h(x) \, e^{nA(x)}, \quad (5.13)$$

where $A(x)$ is given by (5.8) and the function $h(x)$ is given by

$$h(x) = \sqrt{x(1-x)} \left[-\frac{1}{2}\log 2 + \frac{n}{2}[x\log(x\eta) + (1-x)\log((1-x)(1-\eta))] + \right.$$
$$\left. + \sqrt{\frac{n}{a}}[\sqrt{x\eta} + \sqrt{(1-x-(1-\eta))} - 1] \right] . \tag{5.14}$$

After the usual replacement $\sum_{k=0}^{n} \mapsto \int_0^1 n\,dx$, we get for the BG entropy,

$$S_{\mathrm{BG}} = -\sum_{k=0}^{n} \mathfrak{p}_k^{(n)} \log \frac{\mathfrak{p}_k^{(n)}}{\binom{n}{k}}$$
$$\sim -\frac{1}{2}\sqrt{\frac{n}{\pi}} \int_0^1 h(x)\, e^{nA(x)}\, dx\,. \tag{5.15}$$

Applying the Laplace approximation method

$$S_{\mathrm{BG}} = -\frac{1}{2}\sqrt{\frac{n}{\pi}}\sqrt{\frac{2\pi}{n|A''(\eta)|}}\, h(\eta)\, e^{nA(\eta)}$$
$$\sim \frac{1}{2}\log 2 - n[\eta\log\eta + (1-\eta)\log(1-\eta)]\,. \tag{5.16}$$

So we can conclude that S_{BG} is extensive in this model.

5.3 Tsallis and Rényi entropy

To estimate the asymptotic behavior of both entropies, we first use the approximation resulting from (5.7) and (B.2)

$$\left[\binom{n}{k=nx}\right]^{q-1}\left(\mathfrak{p}_{k=nx}^{(n)}\right)^q \sim \frac{1}{\sqrt{2^{q+1}n\pi x(1-x)}}e^{nB(x)}\,, \tag{5.17}$$

with

$$B(x) = qA(x) - (q-1)\mathcal{C}(x) = \frac{q}{2}[x\log\eta + (1-x)\log(1-\eta)] +$$
$$+ \left(\frac{q}{2}-1\right)[x\log x + (1-x)\log(1-x)] + \tag{5.18}$$
$$+ \frac{q}{\sqrt{na}}[\sqrt{x\eta} + \sqrt{(1-x)(1-\eta)} - 1]\,.$$

Next we transform the sum into an integral, as usual,

$$\sum_{k=nx}\left[\binom{n}{k=nx}\right]^{q-1}\left(\mathfrak{p}_{k=nx}^{(n)}\right)^q \sim \sqrt{\frac{n}{2^{q+1}\pi}}\int_0^1 dx\,(x(1-x))^{-1/2}\,e^{nB(x)}\,. \tag{5.19}$$

20

In order to implement the Laplace method, we calculate B' and B''.

$$B'(x) = \frac{q}{2}[\log \eta - \log(1 - \eta)] + \left(\frac{q}{2} - 1\right)[\log x - (1 - x)\log(1 - x)] +$$
$$+ \frac{q}{2\sqrt{na}}\left[\sqrt{\frac{\eta}{x}} + \sqrt{\frac{1 - \eta}{1 - x}}\right], \tag{5.20}$$

$$B''(x) = \left(\frac{q}{2} - 1\right)\frac{1}{x(1 - x)} - \frac{q}{4\sqrt{na}}\left[\sqrt{\frac{\eta}{x^3}} + \sqrt{\frac{1 - \eta}{(1 - x)^3}}\right]. \tag{5.21}$$

We see that for $q < 2$ we have $B''(x) < 0$ for all $x \in (0, 1)$. Hence, if $q < 2$ and if we find one and only one $x_0 \in (0, 1)$ such that $B'(x_0) = 0$, the Laplace approximation method is valid, and we obtain the behavior of the sum at large n:

$$\sum_k \left[\binom{n}{k}\right]^{q-1}\left(\mathfrak{p}_k^{(n)}\right)^q \sim \sqrt{\frac{1}{2^q|B''(x_0)|}}\,(x_0(1 - x_0))^{-1/2}\,e^{nB(x_0)}. \tag{5.22}$$

Now, for the median value $\eta = 1/2$, we find immediately the unique solution $x_0 = 1/2$. Then, $B''(1/2) = 2(q - 2) - q/\sqrt{na}$, $B(1/2) = (1 - q)\log 2$, and so

$$\sum_k \left[\binom{n}{k}\right]^{q-1}\left(\mathfrak{p}_k^{(n)}\right)^q \underset{\text{at large } n}{\sim} 2^{(3-q)/2}(q - 2)^{-1/2}\,e^{n(1-q)\log 2}. \tag{5.23}$$

Therefore, for $\eta = 1/2$, while the Tsallis entropy is not extensive, the Rényi entropy is extensive,

$$S_{\text{Re};q} \sim n \log 2. \tag{5.24}$$

One can easily show that with $\eta = 1/2 + \delta$, $|\delta| \ll 1/2$, the value of the root x_0 is $x_0 = \frac{q}{2-q}\delta + O(\delta^2)$ and that the behavior (5.24) holds too. We have checked numerically that it holds for all $\eta \in (0, 1)$. We notice that this behavior (which is simply $\sim n$ if we adopt the original Rényi choice $\log_2$) is the same as for the two other cases considered in this paper, Eqs. (3.18) and (4.28), and also for the binomial and Laplace de Finetti distributions considered in [15]. We will come back to this important point in the conclusion.

6   Conclusions

In this paper our main interest is the extensivity property of different entropies constructed from generalized binomial distributions. We analyse the behavior of three entropies, mainly the Boltzmann-Gibbs, Tsallis, and Rényi ones for the three examples of generalized binomial distributions presented in [4], recalling that our point of view is strictly microscopic. For that sake we examined the asymptotic behavior of the deformed probability distributions in question, which are those whose generating functions are the q-exponential, the exponential of the Lambert function and the exponential of a second-degree polynomial:

the probabilities obtained are respectively the Polya distribution, a product of modified Abel polynomials and a product of Hermite polynomials.

As could be expected, the Tsallis entropy is not extensive for the three probability distributions considered. The results found for the other two entropies are interesting: the Rényi entropy is extensive for the three probability distributions and, which is surprising, the Boltzmann-Gibbs one is extensive for two cases, those related to the q-exponential and to the Hermite polynomials, but not when the probability distribution is given by modified Abel polynomials. This example of non-extensivity of Boltzmann-Gibbs is a result that deserves further investigation, as it has so far been considered as the universally extensive entropy. As to the Rényi entropy an important aspect of the result found here is that for all the three studied distributions its asymptotic value at large n is the same, $n \log 2$, and therefore does not depend on its parameter q.

Actually, this extensivity is probably due to the nature of the three distributions examined here, which are smooth deformations of the binomial one. We have shown in [15] that both Boltzmann-Gibbs and Rényi are extensive for the binomial case. Deformations of the binomial distribution introduce correlations, and these correlations may or not be strong enough to substantially modify the asymptotic behaviors. The fact that extensivity holds for Rényi and for its BG limit at $q = 1$ when the deformed probability is either the Polya distribution or a product of Hermite polynomials indicates that in these cases the related correlations are weak. Otherwise, the behavior of the deformed probability given as products of modified Abel polynomials is different as the Boltzmann-Gibbs limit of the Rényi entropy is asymptotically not extensive. This distribution deserves a further investigation on the correlations it introduces and we might expect them to be stronger than the two former mentioned cases; this issue will be the subject of future work. Due to this exceptionality of the modified Abel polynomials case we illustrated it here with a concrete and non trivial probabilistic model.

A   Axiomatic(s) for entropies

As a complement to the introduction and since the content of the paper is strongly concerned with entropy, we remind in this appendix, through different sets of postulates, the senses which can be given to this mathematical entity. Entropy is at the same time an information theory concept and a physical quantity as well - physical in the sense that it should be accessible to measurement, and *which acts*, according to Boltzmann, *as a link between the microscopic and the macroscopic worlds*.

It is worthy to start with the way Shannon introduced it in [5]:

> We have represented a discrete information source as a Markoff process. Can we define a quantity which will measure, in some sense, how much information is "produced" by such a process, or better, at what rate information is produced? Suppose we have a set of n possible events whose probabilities of occurrence are $p_1, p_2, \ldots, p_n$. These

probabilities are known but that is all we know concerning which event will occur. Can we find a measure of how much "choice" is involved in the selection of the event or of how uncertain we are of the outcome? If there is such a measure, say $H(p_1, p_2, \ldots, p_n)$, it is reasonable to require of it the following properties:

S1 H should be continuous in the p_k.

S2 If all the p_k are equal, $p_k = 1/n$, then H should be a monotonic increasing function of n. With equally n likely events there is more choice, or uncertainty, when there are more possible events.

S3 If a choice be broken down into two successive choices, the original H should be the weighted sum of the individual values of H.

Then (Theorem) the only H satisfying the three above assumptions is of the form:

$$H \equiv H(\mathcal{P}) = -K \sum_{k=1}^{n} p_k \log p_k \equiv \langle -K \log p_k \rangle. \qquad (A.1)$$

where K is a positive constant and $\mathcal{P} = (p_1, p_2, \ldots, p_n)$.

The above identity means that if we look at $k \mapsto -K \log p_k = y_k$ as a random variable Y, then the entropy H is its expected value with respect to the distribution $k \mapsto p_k$, $H = \langle Y \rangle$. Information theory uses $\log_2$ instead of log and (A.1) with $K = 1$ is the average number of bits needed to describe any random variable with the same probability distribution.

A (partially) different set of axioms, which involve conditional probabilities, was established by Khinchin [20] in view of characterizing the Shannon entropy (A.1). Here we also use the notation $H[\xi] \equiv H(\mathcal{P})$ where $\xi = (x_1, x_2, \ldots, x_n)$ is a random variable with probability distribution $\mathcal{P}$, i.e. p_k is the probability that ξ assumes the value x_k.

K1 H is symmetrical in its arguments.

K2 The uniform distribution $p_k = 1/n$ has maximal $H = K \log n$.

K3 If $\mathcal{Q} = (q_1, q_2, \ldots, q_m)$ is a probability distribution with $m > n$, $q_k = p_k$ for $1 \leq k \leq n$ and $q_k = 0$ for $n + 1 \leq k \leq m$, then $H(\mathcal{P}) = H(\mathcal{Q})$.

K4 For any random variables ξ and η, $H[\xi, \eta] = H[\xi] + \sum_{k=1}^{n} p_k H[\eta \,|\, \xi = x_k]$, which means that the joint entropy is the sum of the entropy of one variable, plus the average value of the entropy of the other variable, once the first is given.

According to Rényi in [9] the Shannon entropy is characterized by another (partially different) set of postulates (Fadeev):

F1=K1 H is symmetrical in its arguments.

F2 $H(p, 1 - p)$ is a continuous function of p for $0 \leq p \leq 1$.

F3 $H(1/2, 1/2) = 1$.

F4 $H(tp_1, (1-t)p_1, p_2, \ldots, p_n) = H(p_1, p_2, \ldots, p_n) + p_1 H(t, 1-t)$ for any distribution $\mathcal{P} = (p_1, p_2, \ldots, p_n)$ and for $0 \leq t \leq 1$.

The Shannon entropy is also the only one which satisfies these four postulates. On the other hand, there are many quantities other than (A.1) that satisfy F1, F2 and F3, plus the property of additivity

$$H(\mathcal{P} * \mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q}), \tag{A.2}$$

where $\mathcal{P} * \mathcal{Q}$ is the direct product of the distributions $\mathcal{P}$ and $\mathcal{Q}$. The fundamental property (A.2) is weaker than the Shannon S3. Rényi in [9] gave the following example which now bears the name of *Rényi entropy*:

$$H_q(p_1, p_2, \ldots, p_n) = \frac{1}{1-q} \log_2 \left(\sum_{k=1}^{n} p_k^q \right) \equiv \log_2 \left(\left\langle p_k^{q-1} \right\rangle \right)^{1/(1-q)}, \tag{A.3}$$

where $q > 0$ and $q \neq 1$, which is one of the entropies examined in this paper. Usually $\log_2$ is replaced by log. This family of entropies goes to the Shannon entropy as $q \to 1$.

To dispel any remnant ambiguity regarding the definition of both the above entropies if one wants to impose additivity, Rényi defined a set of 5 postulates that characterize completely these quantities. First, he extended his considerations to *incomplete* distributions, i.e. sequences $\mathcal{P} = (p_1, p_2, \ldots, p_n)$ of non-negative numbers such that their *weights*

$$\mathtt{w}(\mathcal{P}) := \sum_{k=1}^{n} p_k \tag{A.4}$$

are positive and ≤ 1, but not necessarily equal to 1. The Rényi postulates for the entropy function $H(\mathcal{P})$ are

R1 $H(\mathcal{P})$ is a symmetric functions of the elements of $\mathcal{P}$.

R2 If $\{p\}$ denotes the generalized probability distribution consisting of the single probability p then $H(\{p\})$ is a continuous function of p for $0 < p \leq 1$ (not necessarily in 0).

R3 $H(\{1/2\}) = 1$.

R4 Additivity holds for any pair of incomplete distributions, $H(\mathcal{P} * \mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q})$.

R5 There exists a strictly monotonic and continuous function $y = g(x)$ such that for two incomplete distributions $\mathcal{P} = (p_1, p_2, \ldots, p_m)$ and $\mathcal{Q} = (q_1, q_2, \ldots, q_n)$ with $\mathtt{w}(\mathcal{P}) + \mathtt{w}(\mathcal{Q}) \leq 1$, we have the g-mean value formula

$$H(\mathcal{P} \cup \mathcal{Q}) = g^{-1} \left[\frac{\mathtt{w}(\mathcal{P}) \, g\left(H(\mathcal{P})\right) + \mathtt{w}(\mathcal{Q}) \, g\left(H(\mathcal{Q})\right)}{\mathtt{w}(\mathcal{P}) + \mathtt{w}(\mathcal{Q})} \right]. \tag{A.5}$$

By adding some considerations involving conditional probability, Rényi proved that there are only two possible solutions for the function g.

- The function g is linear, $g(x) = ax + b$, and then the corresponding entropy is Shannon for incomplete distributions

$$H(\mathcal{P}) = -\frac{\sum_{k=1}^{n} p_k \log_2 p_k}{\sum_{k=1}^{n} p_k} . \tag{A.6}$$

- It is exponential, $g_q(x) = 2^{(q-1)x}$, $q > 0$, $q \neq 1$, and then the entropy is Rényi for incomplete distributions

$$H_q(\mathcal{P}) = \frac{1}{1-q} \log_2 \left[\frac{\sum_{k=1}^{n} p_k^q}{\sum_{k=1}^{n} p_k} \right] . \tag{A.7}$$

The first case is the limit as $q \to 1$ of the second one.

Finally, we have as well considered the Tsallis entropy which is also a deformation of (A.1):

$$S_q(p_1, p_2, \ldots, p_n) = \frac{1}{q-1} \left(1 - \sum_{k=1}^{n} p_k^q \right) \equiv \left\langle \frac{1 - p_k^{q-1}}{q-1} \right\rangle . \tag{A.8}$$

This entropy also goes to the Shannon entropy as $q \to 1$. While it satisfies F1 and F2, the Tsallis entropy does not satisfies F3 and has the deformed additivity property

$$S_q(\mathcal{P} * \mathcal{Q}) = S_q(\mathcal{P}) + S_q(\mathcal{Q}) + (1-q)S_q(\mathcal{P}) S_q(\mathcal{Q}) . \tag{A.9}$$

More precisely, Abe [21] has proved that this entropy is characterized by three postulates adapted from the Shannon-Khinchin axioms.

A1 $S_q(\mathcal{P})$ is continuous with respect to all its arguments and takes its maximum for the equiprobability distribution $p_k = 1/n$.

A2 If $\mathcal{Q} = (q_1, q_2, \ldots, q_m)$ is a probability distribution with $m > n$, $q_k = p_k$ for $1 \leq k \leq n$ and $q_k = 0$ for $n + 1 \leq k \leq m$, then $S_q(\mathcal{P}) = S_q(\mathcal{Q})$.

A3 For any random variables ξ and η,

$$S_q[\xi, \eta] = S_q[\xi] + S_q[\eta \,|\, \xi] + (1-q)S_q[\xi] \, S_q[\eta \,|\, \xi] .$$

B Asymptotic formulas

From the Stirling formula,

$$n! \sim \sqrt{2\pi} \, e^{-n} \, n^{n+1/2} \quad \text{at large } n , \tag{B.1}$$

we derive the asymptotic behavior of binomial coefficient at large n,

$$\binom{n}{k = nx} \sim \frac{1}{\sqrt{2\pi n x(1-x)}} \, e^{n\mathcal{C}(x)} \tag{B.2}$$

where the function $\mathcal{C}(x) := -x \log x - (1-x) \log(1-x)$. From this expression, we can check that the summation formula

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n \tag{B.3}$$

keeps its validity at large n. Indeed, with $k = nx$, $0 < x < 1$ and replacing the above sum $\sum_{k=0}^{n}$ by the integral $\int_0^1 n \mathrm{d}x$, leads to

$$\sum_{k=0}^{n} \binom{n}{k} \sim \sqrt{\frac{n}{2\pi}} \int_0^1 [x(1-x)]^{-1/2} \, e^{n\mathcal{C}(x)} \, \mathrm{d}x \,. \tag{B.4}$$

Then we apply the Laplace's method for evaluating the above integral. Laplace's approximation formula (with suitable conditions on the functions involved) reads

$$\int_a^b h(x) \, e^{nA(x)} \, \mathrm{d}x \approx \sqrt{\frac{2\pi}{n|A''(x_0)|}} \, h(x_0) \, e^{nA(x_0)} \ \text{ as } n \to \infty \,, \tag{B.5}$$

where $A'(x_0) = 0$ for $x_0 \in [a,b]$, $A''(x_0) < 0$ and h is positive. Here, we have

$$\mathcal{C}'(x) = -\log(x) + \log(1-x) \quad \mathcal{C}''(x) = -\frac{1}{x(1-x)} \,.$$

Now, $x = 1/2$ is the only root of $\mathcal{C}'(x) = 0$ in the interval $0 < x < 1$, and this corresponds to the maximum of $\mathcal{C}(x)$ in that interval: $\mathcal{C}(1/2) = \log 2$. Moreover, $\mathcal{C}''(1/2) = 4$. Thus,

$$\sqrt{\frac{n}{2\pi}} \int_0^1 [x(1-x)]^{-1/2} \, e^{n\mathcal{C}(x)} \, \mathrm{d}x \sim 2^n \,.$$

Acknowledgments

References

[1] Curado EMF, Gazeau JP and Rodrigues LMCS, *On a Generalization of the Binomial Distribution and Its Poisson-like Limit*, 2012 *J. Stat. Phys.* **146** 264-280 1

[2] Bergeron H, Curado EMF, Gazeau JP and Rodrigues LMCS, *Generating functions for generalized binomial distributions*, 2012 *J. Math. Phys.* **53** 103304-1-22 1, 5

[3] Bergeron H, Curado EMF, Gazeau JP and Rodrigues LMCS, *Generalized binomial distributions*, 2013 in *Group 29: Physical and Mathematical Aspects of Symmetries* eds. J.P. Gazeau, M. Ge and C. Bai, Proceedings of the XXIXth International Colloquium on Group Theoretical Methods in Physics, 19–25 August 2012, Tianjin, China, World Scientific, 265-270 1

[4] Bergeron H, Curado EMF, Gazeau JP and Rodrigues LMCS, *Symmetric generalized binomial distribution*, 2013 *J. Math. Phys.* **54** 123301-1-18. arXiv: 1308.4863v1 [math-ph] 1, 1, 2, 2, 6

[5] Shannon CE, *A Mathematical Theory of Communication*, 1948 *The Bell System Technical Journal* **27** 379–423 & 623–656 1, A

[6] Shannon CE and Weaver W, *The Mathematical Theory of Communication*, 1949 Urbana, University of Illinois Press 1

[7] Tsallis C, *Possible generalization of Boltzmann-Gibbs statistics*, 1988 *J. Stat. Phys.* **52** 479-487 1

[8] Tsallis C, *Nonadditive entropy and nonextensive statistical mechanics-an overview after 20 years*, 2009 *Braz. J. Phys.* **39** 337-356 3.1

[9] Rényi A, *On measures of entropy and information*, 1960 in *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability* **1** 547-561 1, A, A

[10] Johnson N, Kotz S and Kempf AW, *Univariate Discrete Distributions*, 1992 Second Edition, John Wiley and Sons, p.239 3.1

[11] Pólya distribution. *Encyclopedia of Mathematics* http://www.encyclopediaofmath.org/ 3.1

[12] Hanel R, Thurner S, and Tsallis C, *Limit distributions os scale-invariant probabilistic models of correlated random variables with the q-Gaussian as an explicit example*, 2009 *Eur. Phys. J. B*, **72** 263-268 3.1

[13] Prato D and Tsallis C, *Nonextensive foundation of Lévy distributions*, 1999 *Phys. Rev. E* **60** 2398-2401 3.2

[14] Ruiz G and Tsallis C, *private communication* 2014. 3.2

[15] Bergeron H, Curado EMF, Gazeau JP and Rodrigues LMCS, *Entropy extensivity is necessary but not sufficient for thermodynamic predictions, submitted for publication* 3.4, 5.3, 6

[16] Corless RM, Gonnet GH, Hare DEG, Jeffrey DJ and Knuth DE, *On the Lambert W function*, 1996 *Adv. Computational Maths.* **5** 329-359 4.1

[17] Roman S, *The Abel Polynomials*, 1984 § 4.1.5 in *The Umbral Calculus*, New York, Academic Press, p. 29-30 and 72-75 4.1

[18] Gradshteyn IS and Ryzhik IM, *Table of Integrals, Series, and Products*, 2007 Edited by A. Jeffrey and D. Zwillinger, Academic Press, New York, 7th edition 4.3

[19] Magnus W, Oberhettinger F and Soni RP, *Formulas and Theorems for the Special Functions of Mathematical Physics*, 1966 Springer-Verlag, Berlin, 3rd Edition 4.2, 5.1

[20] Khinchin AI, *Mathematical Foundations of Information Theory*, 1956 New York, Dover. Translated by R. A. Silverman and M. D. Friedman from 1953 *Uspekhi Matematicheskikh Nauk* **7** 3-20 and 1956 **9** 17–75 A

[21] Abe S, *Axioms and uniqueness theorem for Tsallis entropy*, 2000 *Phys. Lett. A* **271** 74-79 A