

ARTICLE

Concentration functions and entropy bounds for discrete log-concave distributions

Sergey G. Bobkov¹, Arnaud Marsiglietti^{2,*} and James Melbourne³

¹School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA, ²University of Florida, Department of Mathematics, Gainesville, FL 32611, USA and ³Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA

*Corresponding author. Email: a.marsiglietti@ufl.edu

(Received 23 July 2020; revised 2 March 2021; accepted 26 April 2021; first published online 27 May 2021)

Abstract

Two-sided bounds are explored for concentration functions and Rényi entropies in the class of discrete log-concave probability distributions. They are used to derive certain variants of the entropy power inequalities.

2020 MSC Codes: Primary 60E15; Secondary 94A17, 39B62

1. Introduction

Given a random variable X , its concentration function is defined by

$$Q(X; \lambda) = \sup_x \mathbb{P}\{x \leq X \leq x + \lambda\}, \quad \lambda \geq 0. \quad (1.1)$$

Of a large interest is also the particular value

$$M(X) = Q(X; 0) = \sup_x \mathbb{P}\{X = x\}. \quad (1.2)$$

These quantities are related to other important characteristics of probability distributions such as the Shannon and Rényi entropies. Our original goal was to find effective two-sided bounds on $Q(X, \lambda)$ in the class of discrete log-concave distributions, and thus to explore a number of similarities with the well studied continuous setting.

Let us recall that an integer-valued random variable X (also called discrete) is said to have a (discrete) log-concave distribution, if its probability function $f(k) = \mathbb{P}\{X = k\}$ has an integer supporting interval $\text{supp}(f) = \{k \in \mathbb{Z} : f(k) > 0\}$, and

$$f(k)^2 \geq f(k-1)f(k+1) \quad \text{for all } k \in \mathbb{Z}. \quad (1.3)$$

Many classical discrete distributions belong to this class: discrete uniform, Bernoulli, binomial and convolutions of Bernoulli distributions with arbitrary parameters, Poisson, geometric, negative binomial, etc., (cf. [7] and references therein). It is, therefore, interesting to know how basic characteristics of such distributions are connected with each other: variance, moments, concentration

function, entropies, in analogy with usual log-concave distributions in the continuous setting (for which the densities with respect to the linear Lebesgue measure are log-concave). While a number of challenging questions about this class are still open, in this paper, we develop several techniques (reduction to the continuous setting, rearrangement) and describe some of the results in this direction. In particular, we prove:

Theorem 1.1. *If the random variable X has a discrete log-concave distribution, then*

$$\frac{1}{\sqrt{1 + 12 \operatorname{Var}(X)}} \leq M(X) \leq \frac{2}{\sqrt{1 + 4 \operatorname{Var}(X)}}. \tag{1.4}$$

Moreover, if the distribution of X is symmetric about a point, then the above upper bound may be sharpened to

$$M(X) \leq \frac{1}{\sqrt{1 + 2 \operatorname{Var}(X)}}. \tag{1.5}$$

The inequality (1.5) becomes an equality for two-sided geometric distributions, that is, for densities $f(k) = c_p p^{|k|}$, $k \in \mathbb{Z}$, with an arbitrary value $p \in (0, 1)$, where $c_p = \frac{1-p}{1+p}$ is a normalising constant. As for the lower bound in (1.4), it does not need any log-concavity assumption; here, the factor 12 in front of the variance of X is optimal, as can be seen on the example of discrete uniform distributions.

Relations similar to (1.4)–(1.5) with involved parameter λ may also be stated for the corresponding concentration functions (see Section 8). Let us, however, turn to information-theoretic applications of Theorem 1.1 such as entropy power inequalities (EPIs).

If the random variable X has an absolutely continuous distribution with density $f(x)$ with respect to the Lebesgue measure (the continuous setting), the differential Rényi entropy power of a given order $\alpha > 0$, $\alpha \neq 1$, is defined by

$$N_\alpha(X) = \left(\int_{-\infty}^{\infty} f(x)^\alpha dx \right)^{-\frac{2}{\alpha-1}}, \tag{1.6}$$

while the limit case $N(X) = \lim_{\alpha \rightarrow 1} N_\alpha(X)$ represents the Shannon differential entropy power. As is well known, this functional is subadditive on convolutions, that is, it satisfies a fundamental EPI

$$N(S_n) \geq N(X_1) + \dots + N(X_n), \tag{1.7}$$

where $S_n = X_1 + \dots + X_n$ is the sum of independent continuous random variables. A more general relation such as

$$N_\alpha(S_n) \geq c_\alpha (N_\alpha(X_1) + \dots + N_\alpha(X_n)), \tag{1.8}$$

with $c_\alpha = \alpha^{\frac{1}{\alpha-1}}$, $\alpha > 1$, is also true for Rényi entropy powers. We refer an interested reader to [3], cf. also [15], [25], [5], [13], [18] and [14] for other variants and extensions of (1.7).

If X takes only integer values with density $f(k)$ with respect to the counting measure, the Rényi entropy power is defined similarly to (1.6) as

$$N_\alpha(X) = e^{2H_\alpha(X)} = \left(\sum_{k \in \mathbb{Z}} f(k)^\alpha \right)^{-\frac{2}{\alpha-1}}. \tag{1.9}$$

Here

$$H_\alpha(X) = -\frac{1}{\alpha - 1} \log \sum_{k \in \mathbb{Z}} f(k)^\alpha,$$

is the classical Rényi entropy, with the limit case

$$H_1(X) = H(X) = - \sum_{k \in \mathbb{Z}} f(k) \log f(k).$$

Both in the discrete and continuous settings, these entropies are monotone in the sense that they may only increase when adding an independent summand to a given random variable. However, while the EPI (1.7) quantifies this property in the continuous setting, not much is known so far about the discrete random variables. The inequality (1.7) has been verified to be true in [10] for the family of the (symmetric) binomial distributions (cf. also [8], [17]). In the general discrete case, it was shown in [9] that

$$H(X_1 + X_2) \geq \frac{1}{2} H(X_1) + \frac{1}{2} H(X_2) + g(H(X_1), H(X_2)),$$

for some positive doubly increasing function g on \mathbb{R}_+^2 such that $g(x_1, x_2) \rightarrow \frac{1}{8}$ as $x_1, x_2 \rightarrow \infty$.

Anyhow, as far as we know, the inequality (1.7) is no longer true in the general discrete case. It seems the nature of the discrete entropy power is different, since for example necessarily $N_\alpha(X) \geq 1$. The functionals

$$\Delta_\alpha(X) = N_\alpha(X) - 1,$$

reflecting the size of the variance more accurately (in analogy with the ‘continuous’ entropy power), seem to be more appropriate and therefore more suitable from the point of view of EPIs. This can be seen from the following assertion.

Theorem 1.2. *If the independent random variables X_k ($1 \leq k \leq n$) have symmetric discrete log-concave distributions, then*

$$\frac{1}{c_\alpha} \sum_{k=1}^n \Delta_\alpha(X_k) \leq \Delta_\alpha(S_n) \leq c_\alpha \sum_{k=1}^n \Delta_\alpha(X_k), \tag{1.10}$$

for any $\alpha > 1$ with $c_\alpha = \frac{2(3\alpha-1)}{\alpha-1}$. Moreover, in the case $1 < \alpha \leq 2$, the symmetry assumption may be dropped, while the constant may be improved to the value $c_\alpha = \frac{3\alpha-1}{\alpha-1}$.

In this connection, let us mention a recent paper [21] dealing with similar problems. In particular, it is shown there that for (not necessarily symmetric) Bernoulli summands X_k and all $\alpha \geq 2$, the left inequality in (1.10) remains to hold with $c_\alpha = \frac{6(\alpha-1)}{\alpha} \leq 6$. It was also observed that such a lower bound is no longer true for $\alpha = 1$ with any fixed positive constant in place of c_α . It would be interesting to explore whether or not it is possible to remove the symmetry assumption in Theorem 1.2 for the whole range of α 's.

One may complement these relations with similar ones for the entropy powers in analogy with (1.8). To this aim, we involve an additional condition on the variances of the summands.

Theorem 1.3. *If the random variables X_k have discrete log-concave distributions, and $\text{Var}(X_k) \geq \sigma^2 > 0$ for all $k \leq n$, then for any $\alpha \geq 1$,*

$$N_\alpha(S_n) \geq \frac{1}{c_\sigma} \sum_{k=1}^n N_\alpha(X_k), \tag{1.11}$$

where $c_\sigma = 2\pi e (1 + \frac{1}{12\sigma^2})$. In fact, without any condition on the variances, this lower bound may be reversed to the form

$$N_\alpha(S_n) \leq -\frac{\pi e}{6} (3n - 1) + 2\pi e \sum_{k=1}^n N_\alpha(X_k). \tag{1.12}$$

The paper is organised as follows. We start with general upper bounds for the Rényi entropy powers in terms of variance and recall known results in the continuous setting (including an important theorem due to Moriguti). Such results are used in Section 3 to derive similar upper bounds for discrete random variables. In Section 4, we turn to the notion of discrete log-concavity and first recall its relationship with purely algebraic problems. Sections 5 and 6 connect discrete log-concave distributions with usual (continuous) log-concave measures on the real line. In this way, one may derive a number of interesting relations in the discrete setting, although not always with sharp constants. As a sharpening approach, in Section 7, we develop rearrangement arguments, which allow to complete the proof of Theorem 1.1. A more general form of this theorem is considered in Section 8 in terms of concentration functions. Proof of Theorems 1.2–1.3 is postponed to Section 9, and in the last Section 10, we conclude the exposition with remarks on Bernoulli sums.

2. Maximum to Rényi entropy subject to variance constraint

In the discrete case, the M -functional (1.2) may be viewed as a member in the hierarchy of Rényi entropies. More precisely, letting $\alpha \rightarrow \infty$, the definition (1.9) leads to the identity

$$N_\infty(X) = M(X)^{-2}. \tag{2.13}$$

Convention. In the continuous case, if a random variable X has density $f(x)$, put

$$M(X) = \text{ess sup}_x f(x).$$

Thus, we use the same notation for two formally different objects, in analogy with N_α . As a consequence, the formula (2.1) remains true both in the continuous and discrete setting, as follows from the definition (1.6).

First, let us state one elementary general relation connecting the M -functional to variance.

Proposition 2.1. *Given a continuous random variable X with a fixed variance, the M -functional is minimised for the uniform distribution on a finite interval. Equivalently,*

$$M^2(X) \text{Var}(X) \geq \frac{1}{12}. \tag{2.14}$$

The equality here is attained if and only if X has a uniform distribution on a finite interval.

This relation is well known, and here we recall a simple argument. One may assume that X has a finite second moment with $M(X) = 1$ (by homogeneity). Then, the non-negative function $u(x) = \mathbb{P}\{|X - \mathbb{E}X| \geq x\}$ is Lipschitz (therefore absolutely continuous) and satisfies $u(0) = 1$, $u'(x) \geq -2$ a.e., so that $u(x) \geq 1 - 2x$ for all $x \geq 0$. This gives

$$\text{Var}(X) = 2 \int_0^\infty xu(x) dx \geq 2 \int_0^{1/2} xu(x) dx \geq 2 \int_0^{1/2} x(1 - 2x) dx = \frac{1}{12}.$$

Now, let us restate (2.14) as a homogeneous inequality (with respect to X)

$$N_\infty(X) \leq 12 \text{Var}(X). \tag{2.15}$$

Importantly, this relation may be extended to Rényi entropies of all orders $\alpha \geq 1$. The problem of maximisation of $N_\alpha(X)$ when the variance is fixed was considered and solved by Moriguti in 1952 (cf. also [12] and [6] for a multidimensional extension).

Proposition 2.2. ([22]). *Let $1 < \alpha < \infty$. Given a continuous random variable X with a fixed variance, the Rényi entropy power $N_\alpha(X)$ is maximised for the distribution whose density f_α is supported on the interval $(-1, 1)$ and is proportional there to $(1 - x^2)^{\frac{1}{\alpha-1}}$. Equivalently,*

$$N_\alpha(X) \leq A_\alpha \text{Var}(X), \tag{2.16}$$

where the constant A_α corresponds to f_α .

The inequality (2.16) is affine invariant, so it is equivalent to the formally weaker relation $N_\alpha(X) \leq A_\alpha \mathbb{E}X^2$, which was actually considered in [22].

Thus, the extremal density has the form $f_\alpha(x) = c_\alpha(1 - x^2)^{\frac{1}{\alpha-1}}$, $|x| < 1$, in which the normalising constant is given by

$$c_\alpha = \frac{1}{B(\frac{\alpha}{\alpha-1}, \frac{1}{2})} = \frac{\Gamma(\frac{3\alpha-1}{2(\alpha-1)})}{\Gamma(\frac{\alpha}{\alpha-1})\Gamma(\frac{1}{2})}. \tag{2.17}$$

In this case, as was already noted in [22],

$$\text{Var}(X) = \frac{\alpha - 1}{3\alpha - 1}. \tag{2.18}$$

Let us compute the constant A_α . Putting $\beta = \frac{2\alpha-1}{\alpha}$, we have

$$\int_{-\infty}^{\infty} f_\alpha(x)^\alpha dx = c_\alpha^\alpha \int_{-1}^1 (1 - x^2)^{\frac{\alpha}{\alpha-1}} dx = c_\alpha^\alpha \int_{-1}^1 (1 - x^2)^{\frac{1}{\beta-1}} dx = \frac{c_\alpha^\alpha}{c_\beta}.$$

Hence, according to the definition (1.6) of the Rényi entropy power,

$$A_\alpha = \frac{N_\alpha(X)}{\text{Var}(X)} = \frac{1}{\text{Var}(X)} \left(\int_{-\infty}^{\infty} f_\alpha(x)^\alpha dx \right)^{-\frac{2}{\alpha-1}} = \frac{3\alpha - 1}{\alpha - 1} \left(\frac{c_\beta}{c_\alpha^\alpha} \right)^{\frac{2}{\alpha-1}}. \tag{2.19}$$

For example, $A_2 = \frac{125}{9} \sim 13.888\dots$ Although the expression in (2.19) is rather complicated, one can show that $A_\alpha \rightarrow 2\pi e \sim 17.079$ as $\alpha \rightarrow 1$ (by Stirling’s formula), while $A_\alpha \rightarrow 12$ as $\alpha \rightarrow \infty$, so that Proposition 2.2 includes (2.14). This can also be seen by noting that the extremal distribution approaches the uniform distribution on the interval $(-1, 1)$ for large values of α , while after a linear transformation it approaches the standard normal distribution for α approaching 1. In the latter case, we arrive at another well-known relation

$$N(X) \leq 2\pi e \text{Var}(X), \tag{2.20}$$

where an equality is attained for all non-degenerate normal laws. Since the function $\alpha \rightarrow N_\alpha(X)$ is non-increasing, we also conclude that A_α is a decreasing function in α . In particular, (2.8) yields a slightly weaker variant of (2.16) with a universal constant, namely

$$N_\alpha(X) \leq 2\pi e \text{Var}(X). \tag{2.21}$$

It holds for all $\alpha \geq 1$, and an equality is attained for $\alpha = 1$ and all normal laws.

3. Discrete case: Bounds on Rényi entropy via variance

In order to derive similar relations in the discrete case, one may apply (2.16) to the random variable $\tilde{X} = X + U$, where U is independent of X and has a uniform distribution on the interval $(-\frac{1}{2}, \frac{1}{2})$. In terms of the probability function $f(k) = \mathbb{P}\{X = k\}$, \tilde{X} has density

$$\tilde{f}(x) = \sum_{k \in \mathbb{Z}} f(k) 1_{(k-\frac{1}{2}, k+\frac{1}{2})}(x), \quad x \in \mathbb{R}.$$

It follows that

$$\begin{aligned} h_\alpha(\tilde{X}) &\equiv -\frac{1}{\alpha - 1} \log \int_{-\infty}^{\infty} \tilde{f}(x)^\alpha dx \\ &= -\frac{1}{\alpha - 1} \log \sum_{k \in \mathbb{Z}} f(k)^\alpha = H_\alpha(X), \end{aligned}$$

and therefore $N_\alpha(\tilde{X}) = N_\alpha(X)$. Since $\text{Var}(\tilde{X}) = \frac{1}{12} + \text{Var}(X)$, from (2.16) we therefore obtain:

Proposition 3.1. *Let $1 \leq \alpha \leq \infty$. For any integer-valued random variable X having finite variance,*

$$N_\alpha(X) \leq A_\alpha \left(\frac{1}{12} + \text{Var}(X) \right), \tag{3.22}$$

where the constant A_α is described in (2.5) and (2.7). In particular,

$$N_\alpha(X) \leq 2\pi e \left(\frac{1}{12} + \text{Var}(X) \right).$$

For $\alpha = \infty$, we have $A_\infty = 12$, and (3.22) yields:

Corollary 3.2. *For any integer-valued random variable X having finite variance,*

$$1 \leq N_\infty(X) \leq 1 + 12 \text{Var}(X). \tag{3.23}$$

In view of (2.13), the above upper bound is exactly the lower bound (1.4) in Theorem 1.1. Note that it may also be obtained with a similar argument on the basis of Proposition 2.1, thus without referring to Moriguti’s theorem (cf. Proposition 8.2 below).

Since in general $N_\alpha(X) \geq 1$, while the variance may take any prescribed value, the constant $\frac{1}{12}$ may not be removed from (3.22). Nevertheless, one may ask the following question: Is it possible to replace $\frac{1}{12}$ in (3.22) with $1/A_\alpha$ at the expense of an additional factor in front of the variance in analogy with (2.16)? The answer is affirmative in some sense for $\alpha > 1$, if we allow the factor depend on α .

Indeed, generalising the previous argument, let us apply (2.16) to random variables of the form $\tilde{X} = X + U$, assuming that U is independent of X and has density $g(x)$ supported on the unit interval. Then, \tilde{X} has density

$$\tilde{f}(x) = f(k) g \left(x - k + \frac{1}{2} \right), \quad k - \frac{1}{2} < x < k + \frac{1}{2}, \quad k \in \mathbb{Z}.$$

If $\alpha > 1$, it follows that

$$\int_{-\infty}^{\infty} \tilde{f}(x)^\alpha dx = \sum_{k \in \mathbb{Z}} f(k)^\alpha \int_{-\infty}^{\infty} g(x)^\alpha dx,$$

and therefore

$$N_\alpha(\tilde{X}) = N_\alpha(X) N_\alpha(U).$$

Since $\text{Var}(\tilde{X}) = \text{Var}(X) + \text{Var}(U)$, an application of (2.16) yields

$$N_\alpha(X) N_\alpha(U) \leq A_\alpha (\text{Var}(X) + \text{Var}(U)),$$

that is,

$$N_\alpha(X) \leq \frac{A_\alpha \text{Var}(U)}{N_\alpha(U)} + \frac{A_\alpha}{N_\alpha(U)} \text{Var}(X).$$

Here, according to Proposition 2.2, the first term on the right-hand side is minimised and is equal to 1 when $U = \frac{1}{2} Z$ where Z has density f_α . In that case, for the second term we have

$$\frac{A_\alpha}{N_\alpha(U)} = \frac{1}{\text{Var}(U)} = \frac{4}{\text{Var}(Z)} = \frac{4(3\alpha - 1)}{\alpha - 1},$$

where we recalled the identity (2.18). Thus, we arrive at the following relation, which contains (3.23) for $\alpha = \infty$.

Proposition 3.3. *Let $1 < \alpha \leq \infty$. For any integer-valued random variable X having finite variance, we have*

$$1 \leq N_\alpha(X) \leq 1 + \frac{4(3\alpha - 1)}{\alpha - 1} \text{Var}(X). \tag{3.24}$$

4. Log-concave sequences

We now turn to lower bounds for the Rényi entropies. This cannot be performed in terms of variances in the entire class of discrete probability distributions, so some extra hypotheses are needed. As it turns out, the class of discrete log-concave distributions perfectly fits our aims. First, let us recall that a sequence $\{a_k\}_{k \in \mathbb{Z}}$ of non-negative numbers is called log-concave, if

$$a_k^2 \geq a_{k-1} a_{k+1}, \tag{4.25}$$

for all $k \in \mathbb{Z}$. Similarly, a finite sequence $\{a_k\}_{k=m}^n$ of non-negative numbers is log-concave, if this inequality is fulfilled whenever $m + 1 \leq k \leq n - 1$. Defining $a_k = 0$ for $k < m$ and $k > n$, we then obtain an infinite log-concave sequence.

Log-concave sequences appear in a purely algebraic framework. The following classical result goes back to Newton.

Proposition 4.1. *Suppose that a polynomial $P(z) = \sum_{k=0}^n \binom{n}{k} a_k z^k$ has only real zeros over the field \mathbb{C} of complex numbers. Then (4.25) holds true for all $1 \leq k \leq n - 1$.*

Following [26], one may give a proof based on the following observation of independent interest due to Gauss–Lucas: For any polynomial P over \mathbb{C} of degree $\deg(P) \geq 1$, the zeros of its derivative P' belong to the convex hull of the zeros of P .

The sequences that produce polynomials with real roots give an interesting subset of the class of sequences satisfying a log-concavity inequality. One should note that such sequences need not have contiguous support. For example, $z^2 - 1$ has roots ± 1 and corresponds to the sequence $\{1, 0, -1\}$.

Let us now mention the following important theorem due to Hoggar [11].

Proposition 4.2. *If the coefficients of two polynomials are positive and form log-concave sequences, then so does their product.*

Note that for $P(z) = \sum_{k=0}^n a_k z^k$ and $Q(z) = \sum_{k=0}^m b_k z^k$ of degrees n and m respectively,

$$R(z) = P(z)Q(z) = \sum_{k=0}^{n+m} c_k z^k, \quad c_k = \sum_{k_1+k_2=k} a_{k_1} b_{k_2}.$$

Here, the coefficients $\{c_k\}_{k=0}^{n+m}$ appear as the convolution of the sequences $\{a_k\}_{k=0}^n$ and $\{b_k\}_{k=0}^m$. Thus, Proposition 4.2 tells us that the class of finite positive log-concave sequences is closed under the convolution operation.

The assumption about the positivity may not be removed in this conclusion. For a counterexample, one may take the polynomials

$$P(z) = 1 + z, \quad Q(z) = 1 + z^3, \quad R(z) = 1 + z + z^3 + z^4,$$

which correspond to the sequences $\{1, 1\}$, $\{1, 0, 0, 1\}$, $\{1, 1, 0, 1\}$. Here, the first two sequences are log-concave, while the third one is not. With this in mind, the notion of a log-concave discrete distribution on \mathbb{Z} should be introduced as in Introduction.

Recalling the definition (1.3), as an immediate consequence from Proposition 4.2, we obtain:

Corollary 4.3. *The class of (discrete) log-concave probability distributions on \mathbb{Z} is closed under the convolution operation.*

5. From discrete to continuous log-concave measures

Let us turn to basic properties of discrete log-concave distributions. Some of them can be obtained by employing known results from the theory of ‘continuous’ log-concave functions. To this aim, we describe a simple construction, which allows one to associate with a discrete log-concave distribution defined by a probability function a certain log-concave function on the real line.

Given a random variable X with a discrete log-concave distributions defined by a probability function $f(k)$, denote by Δ the smallest closed interval containing $\text{supp}(f)$. Let us extend the function

$$V(k) = -\log f(k),$$

linearly on every segment $[k, k + 1] \subset \Delta$, and put $V = \infty$ outside Δ (if Δ is not the whole real line). Then V is finite and convex on Δ . To see this, assume that $|\Delta| \geq 2$, and for $[k - 1, k + 1] \subset \Delta$, rewrite (4.25) as

$$V(k) - V(k - 1) \leq V(k + 1) - V(k).$$

But $V(k) - V(k - 1) = V'(x)$ on $(k - 1, k)$, which means that V has a non-decreasing (Radon-Nikodym) derivative. As a result, we obtain a log-concave function

$$f(x) = e^{-V(x)}, \quad x \in \mathbb{R},$$

which coincides with $f(k)$ on all integers k . Let us call it a log-piecewise linear extension of the sequence $f(k)$.

In general, $f(x)$ does not need be a probability density on the line. Nevertheless, since

$$\int_k^{k+1} f(x) dx \leq \max\{f(k), f(k + 1)\} \leq f(k) + f(k + 1),$$

after summation over all $k \in \mathbb{R}$, we conclude that f is integrable. Let us emphasise this fact once more.

Proposition 5.1. *The restrictions of densities of finite (continuous) log-concave measures from \mathbb{R} to \mathbb{Z} describe the whole class of probability functions of discrete log-concave measures.*

Returning to the log-piecewise linear extension of the sequence $f(k)$ and using an additional property $f(k) \leq 1$, an immediate consequence of the integrability of f is that, with some constant $c > 0$ depending on the distribution of X ,

$$f(k) \leq 2 e^{-c|k|} \quad \text{for all } k \in \mathbb{Z}.$$

In particular, there is a point of maximum for this sequence, which is called a mode (using a probabilistic language). It is also a mode for the function $f(x)$. So, it is a mode of X .

Let $m \in \mathbb{Z}$ be a mode of X . Then necessarily $f(k)$ is non-increasing in $k \geq m$ and is non-decreasing in $k \leq m$. Hence

$$\int_k^{k+1} f(x) dx \geq f(k + 1) \quad \text{for } k \geq m, \quad \int_{k-1}^k f(x) dx \geq f(k - 1) \quad \text{for } k \leq m.$$

Performing summation over all k , we get

$$\int_{-\infty}^{\infty} f(x) dx \geq 1 - f(m).$$

Also, since V is linear on each interval $[k, k + 1] \subset \Delta$, $f(x)$ is convex, and therefore its graph lies below the segment connecting the points $(k, f(k))$ and $(k + 1, f(k + 1))$. Hence

$$\int_k^{k+1} f(x) dx \leq \frac{f(k) + f(k + 1)}{2}.$$

Performing summation over all k , we arrive at:

Proposition 5.2. *If m is mode of X and $|\Delta| \geq 2$, then*

$$1 - f(m) \leq \int_{-\infty}^{\infty} f(x) dx \leq 1.$$

Similar bounds hold true for the second moment. In particular, one may easily derive:

Proposition 5.3. *Let X be a discrete random variable with a log-concave density f with mode at m . If $f(m) \leq \frac{1}{2}$, then the log-piecewise extension of the density satisfies*

$$\frac{1}{3} \mathbb{E} (X - m)^2 \leq \int_{-\infty}^{\infty} (x - m)^2 f(x) dx \leq 3 \mathbb{E} (X - m)^2.$$

Moreover, the left inequality is also true when $f(m) \geq \frac{1}{2}$.

6. Second moment and maximum of density

Propositions 5.2 and 5.3 may be used to extend a number of results about continuous log-concave distributions to the discrete setting. Let us remind basic relations for the class of log-concave densities on the line in the usual continuous setting. One important feature in this case is that the general lower bound as in Proposition 2.2 may be reversed.

Keeping X to denote a discrete random variable, let Z be a (continuous) random variable with a log-concave density g . It is supported on some closed interval $\Delta \subset \mathbb{R}$, finite or not, which contains a point of maximum m of g , that is the mode of Z (which may be one of the endpoints of Δ when this interval is bounded from the left or from the right). Hence $M(Z) = g(m)$. Together with (2.14), it is known that

$$\frac{1}{12} \leq M^2(Z) \text{Var}(Z) \leq 1, \tag{6.26}$$

where an inequality on the right-hand side is achieved for the one-sided exponential distribution with density $g(x) = e^{-x} 1_{(0,\infty)}(x)$ (cf. Proposition 2.2 in [4]). In fact, if g is symmetric about m , the upper bound can be improved to

$$M^2(Z) \text{Var}(Z) \leq \frac{1}{2}, \tag{6.27}$$

which is attained for the two-sided exponential distribution with density $g(x) = \frac{1}{2} e^{-|x|}$.

There is a similar assertion about the second moment when Z is centred at the mode.

Proposition 6.1. *If a (continuous) random variable Z has a log-concave density g with mode at the point m , then*

$$M^2(Z) \mathbb{E} (Z - m)^2 \leq 2.$$

Combining this relation with the lower bound in (6.26), we also get

$$\mathbb{E} (Z - m)^2 \leq 24 \text{Var}(Z).$$

Let us now describe an immediate application of Proposition 6.1 to the discrete setting. Suppose that we are given an integer-valued random variable X with discrete log-concave distribution, and denote by $f(x)$ the log-piecewise linear extension of the probability function $f(k) = \mathbb{P}\{X = k\}$, $k \in \mathbb{Z}$. Let Z be a random variable whose density g is proportional to f , that is, $g(x) = \frac{1}{B} f(x)$, $B = \int_{-\infty}^{\infty} f(x) dx$. Recall that, by Proposition 5.2, $B \leq 1$. Hence, by Propositions 6.1 applied to g , we get that

$$f(m)^2 \int_{-\infty}^{\infty} (x - m)^2 f(x) dx \leq 2B^3.$$

On the other hand, according to Proposition 5.3, the above integral dominates $\frac{1}{3} \mathbb{E} (X - m)^2$. Combining the two inequalities, we arrive at:

Proposition 6.2. *If X is a discrete random variable with a log-concave density f with mode at m , then*

$$M^2(X) \mathbb{E} (X - m)^2 \leq 6.$$

In particular,

$$M^2(X) \text{Var}(X) \leq 6. \tag{6.28}$$

Proof of Proposition 6.1. We may assume that $m = 0$ and $g(0) = 1$. Moreover, let $a = \mathbb{P}\{Z < 0\}$ and $b = \mathbb{P}\{Z > 0\}$ be positive.

By the log-concavity, $g(x) = \exp\{-V(x)\}$ for some convex function $V: \mathbb{R} \rightarrow [0, \infty]$ with $V(0) = 0$. Putting $\tilde{g}(x) = \exp\{-\tilde{V}(x)\}$ with $\tilde{V}(x) = x/b$ for $x \geq 0$ and $\tilde{V}(x) = -x/a$ for $x \leq 0$, we obtain another log-concave probability density such that

$$\int_{-\infty}^0 g(x) dx = \int_{-\infty}^0 \tilde{g}(x) dx = a, \quad \int_0^{\infty} g(x) dx = \int_0^{\infty} \tilde{g}(x) dx = b.$$

Since \tilde{V} is linear on $[0, \infty)$ and $\tilde{V}(0) = 0$, necessarily $V(x) \leq \tilde{V}(x)$ on the interval $0 \leq x < x_0$ and $V(x) \geq \tilde{V}(x)$ on the half-axis $x > x_0$ for some $x_0 > 0$. Equivalently, $g(x) \geq \tilde{g}(x)$ for $0 \leq x < x_0$ and $g(x) \leq \tilde{g}(x)$ for $x > x_0$. Introduce the distribution functions

$$G(x) = \mathbb{P}\{Z \leq x\} = \int_{-\infty}^x g(y) dy \quad \text{and} \quad \tilde{G}(x) = \int_{-\infty}^x \tilde{g}(y) dy.$$

It follows that the function $\psi(x) = G(x) - \tilde{G}(x)$ is vanishing at the origin and at infinity, and its derivative is non-negative on $[0, x_0)$ and non-positive on (x_0, ∞) . Hence $\psi(x) \geq 0$ for all $x \geq 0$. This implies that

$$\int_0^{\infty} x^2 g(x) dx - \int_0^{\infty} x^2 \tilde{g}(x) dx = -2 \int_0^{\infty} x \psi(x) dx \leq 0.$$

By the same argument, a similar inequality holds when integrating over the negative half-axis. Hence

$$\mathbb{E} Z^2 = \int_{-\infty}^{\infty} x^2 g(x) dx \leq \int_{-\infty}^{\infty} x^2 \tilde{g}(x) dx = 2(a^3 + b^3) \leq 2(a + b)^3 = 2. \quad \square$$

Remark. In terms of the Rényi entropy powers, the inequalities (6.26) and (6.27) take the form

$$\frac{1}{12} N_{\infty}(Z) \leq \text{Var}(Z) \leq N_{\infty}(Z), \quad \text{Var}(Z) \leq \frac{1}{2} N_{\infty}(Z). \tag{6.29}$$

Here, the first inequality is general and corresponds to (2.16) in the limit case $\alpha = \infty$, while the upper bounds on the variance hold for log-concave probability measures (symmetric in the second case). Similarly to (2.16), these upper bounds may be extended to the whole range of α at the expense of certain α -dependent factors. Although the extremal log-concave distributions are not known so far, one particular case may easily be settled. Given a random variable Z with a log-concave distribution, let us apply the second upper bound in (6.29) to $\tilde{Z} = Z - Z'$, where Z' is an independent copy of Z . Since $N_{\infty}(\tilde{Z}) = N_2(Z)$, we then obtain that

$$\text{Var}(Z) \leq \frac{1}{4} N_2(Z).$$

Here, we have an equality for the exponential distribution, and thus the conclusion that amongst continuous log-concave variables on the real line, the exponential distribution has minimal 2-Rényi entropy, often referred to as collision entropy.

7. Sharpenings via rearrangement arguments

The inequality (6.28) may be sharpened and generalised by involving more delicate arguments based on the so-called rearrangement of densities. Similarly to the continuous setting, these arguments allow one to explore an extremal role of a discrete counterpart of the exponential distributions. More precisely, when a probability mass function $f: \mathbb{Z} \rightarrow [0, \infty)$ can be written as $f(k) = Cp^{|k|}$ for some $0 \leq p < 1$, we will call it a symmetric two-sided geometric distribution. Here, the symmetry property refers to the identity $f(-k) = f(k)$.

Let us start with basic definitions.

Definition (Decreasing rearrangement). For a probability function $f: \mathbb{Z} \rightarrow [0, \infty)$, denote by f^\downarrow its decreasing rearrangement. Explicitly, f^\downarrow is defined on \mathbb{N} and satisfies $f^\downarrow(k) \geq f^\downarrow(k + 1)$ for all $k \geq 0$ and there exists a bijection $\tau: \mathbb{N} \rightarrow \mathbb{Z}$ such that $f^\downarrow(k) = f(\tau(k))$.

Definition (Majorisation). Given probability functions f and g on \mathbb{Z} , we say that f majorises g and write $f \succ g$, when $\sum_{k=0}^n f^\downarrow(k) \geq \sum_{k=0}^n g^\downarrow(k)$ for all $n \geq 0$.

We extend this notion to random variables by writing $X \succ Y$ when $f \succ g$ holds for their respective probability functions.

Definition (Schur-Concavity). A functional Φ defined on a given family of probability functions on \mathbb{N} is called Schur-convex, when $f \succ g$ implies $\Phi(f) \geq \Phi(g)$. A functional Φ is Schur-concave when $-\Phi$ is Schur-convex.

We extend Schur-convexity/concavity to random variables by asking that Schur-convexity/concavity hold for the respective probability mass functions.

Lemma 7.1 ([20]). Any symmetric log-concave probability function f majorises the symmetric two-sided geometric distribution g with the same maximum as f .

A more general result was actually first proven in [19]. We include a proof of Lemma 7.1 for completeness.

Proof. Let $q \in [0, 1]$. Define, for $k \in \mathbb{Z}$, $g_q(k) = q^{|k|}f(0)$ with the convention that $0^0 = 1$. Note that

$$\sum_{k \in \mathbb{Z}} g_0(k) = f(0) \leq \sum_{k \in \mathbb{Z}} f(k) = 1,$$

and $\sum_{k \in \mathbb{Z}} g_1(n) = \infty$. Since the function $q \mapsto \sum_{k \in \mathbb{Z}} g_q(k)$ is continuous and strictly increasing on $[0, 1]$, we deduce by the intermediate value theorem that there exists a unique $q^* \in [0, 1)$ such that

$$\sum_{k \in \mathbb{Z}} g_{q^*}(k) = \sum_{k \in \mathbb{Z}} f(k) = 1.$$

By construction, g_{q^*} is the probability function of a two-sided geometric distribution. Denote by the same letters f and g_q the corresponding log-piecewise linear extensions. In particular,

$$\max_{x \in \mathbb{R}} g_{q^*}(x) = g_{q^*}(0) = f(0) = \max_{x \in \mathbb{R}} f(x).$$

Thus, the function $L = -\log g_{q^*}$ is affine on $[0, \infty)$, while $V = -\log f$ is convex, with $L(0) = V(0)$. Hence, there must exist $l \geq 0$ such that $V \leq L$ on $[0, l]$ and $V \geq L$ on $[l, \infty)$, that is, $f \geq g_{q^*}$ on $[0, l]$ and $f \leq g_{q^*}$ on $[l, \infty)$. Hence, for all $n \geq 0$, $\sum_{k=0}^n f^\downarrow(k) \geq \sum_{k=0}^n g_{q^*}^\downarrow(k)$. \square

Lemma 7.2 (Schur-concavity of moments on symmetric log-concaves). *If the function $u: [0, \infty) \rightarrow [0, \infty)$ is non-decreasing, then the functional $\Phi(f) = \mathbb{E} u(|X|)$ is Schur-concave on the set of symmetric discrete log-concave probability functions.*

Proof. By linearity of Φ with respect to u , one may assume that $u(x) = 1_{(\lambda, \infty)}(x)$ for some integer $\lambda \geq 0$. Suppose that X and Y are symmetric log-concave variables such that $f_X \succ f_Y$. By symmetry, unimodality and the assumed majorisation, we have

$$\begin{aligned} \mathbb{E} u(|X|) &= \mathbb{P}\{|X| > \lambda\} = 1 - \mathbb{P}\{|X| \leq \lambda\} = 1 - \sum_{k=-\lambda}^{\lambda} f_X(k) \\ &= 1 - \sum_{k=0}^{2\lambda} f_X^\downarrow(k) \leq 1 - \sum_{k=0}^{2\lambda} f_Y^\downarrow(k) = \mathbb{P}\{|Y| > \lambda\} = \mathbb{E} u(|Y|), \end{aligned}$$

and the result follows. \square

We are prepared to settle the upper bound of Theorem 1.1 in the symmetric case (and in somewhat more general form).

Proposition 7.3. *Let $\Psi: [0, \infty) \rightarrow [0, \infty)$ be a non-decreasing function. For any symmetric discrete log-concave random variable X ,*

$$M(X) \Psi(\text{Var}(X)) \leq \sup_{0 \leq p < 1} \left[\frac{1-p}{1+p} \Psi\left(\frac{2p}{(1-p)^2}\right) \right].$$

In particular,

$$M^2(X) (1 + 2 \text{Var}(X)) \leq 1. \tag{7.30}$$

The latter inequality becomes an equality for all two-sided geometric distributions.

Proof. Given X a symmetric log-concave random variable, let Y be the two-sided geometric distribution with same maximum (necessarily attained at 0). Then by Lemma 7.1, we have $X \succ Y$, so that by Lemma 7.2, $\text{Var}(X) \leq \text{Var}(Y)$. Thus

$$M(X) \Psi(\text{Var}(X)) \leq M(Y) \Psi(\text{Var}(Y)).$$

Note that for a two-sided geometric distribution with probability mass function $g(k) = Cp^{|k|}$, $0 \leq p < 1$, one has

$$\begin{aligned} C^{-1} &= \sum_{k \in \mathbb{Z}} p^{|k|} = \frac{1+p}{1-p}, & M(Y) &= C = \frac{1-p}{1+p}, \\ \text{Var}(Y) &= \mathbb{E}Y^2 = \frac{2p}{(1-p)^2}. \end{aligned}$$

Hence

$$M(Y) \Psi(\text{Var}(Y)) \leq \sup_{p \in (0,1)} \left[\frac{1-p}{1+p} \Psi\left(\frac{2p}{(1-p)^2}\right) \right].$$

In particular, for $\Psi(x) = \sqrt{1 + 2x}$, the function under the above supremum is equal to 1 for any value of the parameter p , and we obtain (7.30) together with the assertion about the extremal role of two-sided geometric distributions. \square

While (7.30) is a full analog of the inequality (6.27) from the continuous setting under the symmetry hypothesis, the more general inequality (6.26) also extends to the discrete setting. The next assertion provides the upper bound of Theorem 1.1 in the non-symmetric case.

Proposition 7.4. *For any discrete log-concave random variable X ,*

$$M^2(X) (1 + 4 \operatorname{Var}(X)) \leq 4. \tag{7.31}$$

This bound is asymptotically attained for one-sided geometric distributions. In particular,

$$M^2(X) \operatorname{Var}(X) \leq 1. \tag{7.32}$$

Proof. We employ another relation,

$$H_\alpha(X) \leq H_\infty(X) + \log\left(\alpha^{\frac{1}{\alpha-1}}\right), \quad 0 < \alpha < \infty, \tag{7.33}$$

which was recently obtained in [20] using the rearrangement arguments. This inequality holds for any discrete log-concave random variable X and is asymptotically attained for the geometric distributions with probability functions $f(k) = (1 - p) p^k, k \geq 0$, as $p \rightarrow 1$. In particular, for $\alpha = 2$, it takes the form $H_2(X) \leq H_\infty(X) + \log 2$, which is the same as

$$M(X) \leq 2 \sum_{k \in \mathbb{Z}} f(k)^2, \tag{7.34}$$

in terms of the probability function $f(k)$ of X .

Note that, if Y is an independent copy of X , the random variable $X - Y$ will have a symmetric discrete log-concave distribution with

$$M(X - Y) = \mathbb{P}\{X - Y = 0\} = \sum_{k \in \mathbb{Z}} f(k)^2, \tag{7.35}$$

so that, by (7.34),

$$M^2(X) \leq 4 M^2(X - Y). \tag{7.36}$$

On the other hand, by Proposition 7.3, $M^2(X - Y) (1 + 2 \operatorname{Var}(X - Y)) \leq 1$, that is,

$$M^2(X - Y) (1 + 4 \operatorname{Var}(X)) \leq 1.$$

It remains to apply this inequality in (7.36). \square

Note that for the geometric distribution with probability function $f(k) = (1 - p) p^k, k \geq 0$, we have $M^2(X) (1 + 4 \operatorname{Var}(X)) = (1 + p)^2$. Hence, (7.2) is asymptotically attained for $p \rightarrow 1$.

8. Concentration functions

Turning to applications, first let us relate the concentration function to the M -functional.

Lemma 8.1. *For any random variable X ,*

$$Q(X; \lambda) = \lambda M(X + U_\lambda), \quad \lambda > 0, \tag{8.37}$$

where the random variable U_λ is independent of X and has a uniform distribution on the interval $(0, \lambda)$. Analogously, in the discrete setting, we have

$$Q(X; \lambda) = (\lambda + 1) M(X + U_\lambda), \quad \lambda = 0, 1, 2, \dots, \tag{8.38}$$

assuming that the random variable U_λ is independent of X and has a discrete uniform distribution on the integer interval $\{0, 1, \dots, \lambda\}$.

Proof. The first claim follows from the fact that the random variable $X_\lambda = X + U_\lambda$ has an absolutely continuous distribution with density

$$f_\lambda(x) = \frac{1}{\lambda} \mathbb{P}\{x - \lambda \leq X \leq x\} \quad \text{a.e.}$$

According to the definition (1.1), this formula yields

$$M(X_\lambda) = \text{ess esup}_x f_\lambda(x) = \frac{1}{\lambda} Q(X; \lambda).$$

For the second claim, we similarly have that X_λ takes integer values with probabilities

$$f_\lambda(k) = \frac{1}{\lambda + 1} \mathbb{P}\{k - \lambda \leq X \leq k\}, \quad k \in \mathbb{Z}.$$

Since the supremum in (1.1) is attained for some integer value $x = k$, Lemma 8.1 follows. □

One can now apply the lower bound of Proposition 2.2, cf. (2.14), to the random variables X_λ in (8.37). This leads to a corresponding lower bound for the concentration function (which is actually known, cf. [4]).

Proposition 8.2. *For any random variable X ,*

$$Q(X; \lambda) \geq \frac{\lambda}{\sqrt{\lambda^2 + 12 \text{Var}(X)}}, \quad \lambda > 0. \tag{8.39}$$

As a consequence, applying (8.39) with $\lambda \uparrow 1$ in the discrete setting, we arrive at the lower bound (1.4) in Theorem 1.1. Note that since in general $M(X) \leq 1$, while the variance may take any prescribed value (already within specific families such as Poisson distributions), the constant 1 may not be removed from the square root in (1.4).

More generally, if the random variable X is discrete and $\lambda \geq 0$ is integer, we apply (8.39) with $\lambda' \uparrow \lambda + 1$, so that to get in the limit

$$Q(X; \lambda) \geq \frac{\lambda + 1}{\sqrt{(\lambda + 1)^2 + 12 \text{Var}(X)}}. \tag{8.40}$$

For the upper bound, one may start from the identity (8.38) and apply the upper bounds of Theorem 1.1 to the random variables $X_\lambda = X + U_\lambda$ in place of X , assuming that the random variable U_λ is independent of X and has a discrete uniform distribution on the integer interval $\{0, 1, \dots, \lambda\}$. Note that X_λ is symmetric around the point $\lambda/2$, if X is symmetric about the origin. Since

$$\text{Var}(U_\lambda) = \frac{\lambda(\lambda + 2)}{12},$$

together with (8.40) we are led to the following statement.

Proposition 8.3. *If the random variable X has a log-concave discrete distribution, then, for any integer $\lambda \geq 0$,*

$$\frac{\lambda + 1}{\sqrt{1 + \lambda(\lambda + 2) + 12 \text{Var}(X)}} \leq Q(X; \lambda) \leq \frac{2(\lambda + 1)}{\sqrt{1 + \frac{\lambda(\lambda + 2)}{3} + 4 \text{Var}(X)}}.$$

Moreover, if the distribution of X is symmetric about a point, then the upper bound may be sharpened to

$$Q(X; \lambda) \leq \frac{\lambda + 1}{\sqrt{1 + \frac{\lambda(\lambda+2)}{6} + 2 \operatorname{Var}(X)}}.$$

Here, the value $\lambda = 0$ returns us to the statement of Theorem 1.1.

9. Proof of Theorems 1.2 and 1.3

Let X_1, \dots, X_n be independent discrete random variables, and $S_n = X_1 + \dots + X_n$.

Proof of Theorem 1.2. First, consider the symmetric case. By Proposition 3.3, for any integer-valued random variable X having a finite variance,

$$\Delta_\alpha(X) \leq \frac{4(3\alpha - 1)}{\alpha - 1} \operatorname{Var}(X).$$

On the other hand, if the distribution of X is symmetric about a point and log-concave, the inequality (1.5) of Theorem 1.1 yields a lower bound

$$\Delta_\infty(X) \geq 2 \operatorname{Var}(X). \tag{9.41}$$

Since the function $\alpha \rightarrow \Delta_\alpha(X)$ is non-increasing, the two inequalities give

$$2 \operatorname{Var}(X) \leq \Delta_\alpha(X) \leq \frac{4(3\alpha - 1)}{\alpha - 1} \operatorname{Var}(X). \tag{9.42}$$

Applying this to X_k 's, as well as to $X = S_n$, we then get

$$\Delta_\alpha(S_n) \geq 2 \operatorname{Var}(S_n) = 2 \sum_{k=1}^n \operatorname{Var}(X_k) \geq \frac{\alpha - 1}{2(3\alpha - 1)} \sum_{k=1}^n \Delta_\alpha(X_k).$$

In fact, up to an α -dependent constant, this upper bound may be reversed, since, by (9.2), we also have

$$\begin{aligned} \Delta_\alpha(S_n) &\leq \frac{4(3\alpha - 1)}{\alpha - 1} \operatorname{Var}(S_n) \\ &\leq \frac{4(3\alpha - 1)}{\alpha - 1} \sum_{k=1}^n \operatorname{Var}(X_k) \leq \frac{2(3\alpha - 1)}{\alpha - 1} \sum_{k=1}^n \Delta_\alpha(X_k). \end{aligned}$$

This proves the desired relations in (1.10).

Now, suppose that $1 < \alpha \leq 2$ and drop the symmetry assumption. If Y is an independent copy of X , the random variable $\tilde{X} = X - Y$ will have a symmetric discrete log-concave distribution. Hence, (9.41) is applicable to \tilde{X} . On the other hand, the identity (7.35) reads as $N_\infty(\tilde{X}) = N_2(X)$, that is, $\Delta_\infty(\tilde{X}) = \Delta_2(X)$. Hence, from (9.1) we get that

$$\Delta_\alpha(X) \geq \Delta_2(X) \geq 4 \operatorname{Var}(X).$$

Combining this with the upper bound in (9.42), we arrive at its sharpened variant

$$4 \operatorname{Var}(X) \leq \Delta_\alpha(X) \leq \frac{4(3\alpha - 1)}{\alpha - 1} \operatorname{Var}(X). \tag{9.43}$$

It remains to apply this to X_k 's, as well as to $X = S_n$ similarly as above, and then we obtain the lower bound as in (1.10) with an improved constant. The proof of the upper bound is based on (9.3) and is also similar. □

Remark. In the case $\alpha = \infty$, (1.10) takes the form

$$\frac{1}{6} \sum_{k=1}^n \Delta_\infty(X_k) \leq \Delta_\infty(S_n) \leq 6 \sum_{k=1}^n \Delta_\infty(X_k).$$

Here, the lower bound reminds the relation

$$N_\infty(S_n) \geq \frac{1}{2} \sum_{k=1}^n N_\infty(X_k),$$

for the continuous setting (without any assumptions on the shape of distributions, cf. [2, 16]). Since the functions $\alpha \rightarrow \Delta_\alpha(X)$ are non-increasing, it implies that

$$\Delta_1(S_n) \geq \frac{1}{6} \sum_{k=1}^n \Delta_\infty(X_k).$$

However, one cannot estimate $\Delta_\infty(X_k)$ in terms of $\Delta_1(X_k)$. This can be seen on the example of the two-sided geometric distribution with probability function $f(k) = Cp^{|k|}$, $k \in \mathbb{Z}$, for small $0 < p < 1$. In this case,

$$\Delta_\infty(X) = 2 \operatorname{Var}(X) = \frac{4p}{(1-p)^2},$$

which corresponds to the equality in (1.5), while

$$H(X) = -\log(1-p) + \log(1+p) - \frac{2p \log p}{(1-p)(1+p)}.$$

As $p \rightarrow 0$, we have $H(p) \sim 2p \log(1/p)$, hence

$$\begin{aligned} \Delta(X) &= N(X) - 1 = e^{2H(X)} - 1 \\ &\sim 2H(X) \sim 4p \log(1/p) \sim 2 \operatorname{Var}(X) \log(1/p). \end{aligned}$$

Proof of Theorem 1.3. Suppose that $\operatorname{Var}(X_k) \geq \sigma^2$ for all $k \leq n$ with some $\sigma > 0$.

By Proposition 3.1, if X is a discrete random variable with $\operatorname{Var}(X) \geq \sigma^2$, we have

$$N_\alpha(X) \leq 2\pi e \left(1 + \frac{1}{12\sigma^2}\right) \operatorname{Var}(X),$$

for any $\alpha \geq 1$. In addition, according to the upper bound (1.4) of Theorem 1.1,

$$N_\infty(X) \geq \frac{1}{4} + \operatorname{Var}(X) \geq \operatorname{Var}(X). \tag{9.44}$$

Since $\alpha \rightarrow N_\alpha(X)$ is a non-increasing function, these bounds imply that

$$\operatorname{Var}(X) \leq N_\infty(X) \leq N_\alpha(X) \leq 2\pi e \left(1 + \frac{1}{12\sigma^2}\right) \operatorname{Var}(X). \tag{9.45}$$

Being applied first to $X = S_n$ and then to each $X = X_k$, (9.5) yields

$$N_\alpha(S_n) \geq \operatorname{Var}(S_n) = \sum_{k=1}^n \operatorname{Var}(X_k) \geq \frac{1}{2\pi e \left(1 + \frac{1}{12\sigma^2}\right)} \sum_{k=1}^n N_\alpha(X_k).$$

This proves the first assertion (1.11) of the theorem.

By a similar argument, the resulting bound may be reversed, even without any condition on variances, and actually further strengthened. Indeed, applying (9.45) to $X = S_n$ and using the first inequality in (9.44) in the weaker form $N_\alpha(X_k) \geq \frac{1}{4} + \operatorname{Var}(X_k)$, we have

$$\begin{aligned}
 N_\alpha(S_n) &\leq 2\pi e \left(\frac{1}{12} + \text{Var}(S_n) \right) \\
 &= 2\pi e \left(\frac{1}{12} + \sum_{k=1}^n \text{Var}(X_k) \right) \leq 2\pi e \left(\frac{1}{12} + \sum_{k=1}^n \left(N_\alpha(X_k) - \frac{1}{4} \right) \right),
 \end{aligned}$$

and (1.12) immediately follows. □

10. Remarks on Bernoulli sums

Finally, let us illustrate Theorem 1.1 on the example of the Bernoulli sums

$$S_n = X_1 + \dots + X_n,$$

where the independent summands X_k take the values 1 and 0 with probabilities p_k and q_k respectively. By Hoggar’s theorem on preservation of log-concavity under convolutions, the distribution of S_n is discrete log-concave (4.3). This property is not so obvious on the basis of the explicit expression for the probability function

$$f_n(k) = \mathbb{P}\{S_n = k\} = \sum p_1^{\varepsilon_1} q_1^{1-\varepsilon_1} \dots p_n^{\varepsilon_n} q_n^{1-\varepsilon_n}, \quad k = 0, 1, \dots, n,$$

where the summation is running over all 0 – 1 sequences $\varepsilon_1, \dots, \varepsilon_n$ such that $\varepsilon_1 + \dots + \varepsilon_n = k$.

One of the challenging problems about S_n has been to how effectively estimate from above the maximum $M(S_n) = \max_k f_n(k)$ in terms of p_k ’s. In this particular model, there exist several different approaches to the problem. First, let us describe the standard approach in Probability Theory that is irrelevant to log-concavity. In general, the concentration function of a random variable X may be bounded in terms of the characteristic function $v(t) = \mathbb{E} e^{itX}$ by virtue of Esseen’s bound

$$Q(X; \lambda) \leq \left(\frac{96}{95} \right)^2 \lambda \int_{-1/\lambda}^{1/\lambda} |v(t)| dt, \quad \lambda > 0, \tag{10.46}$$

cf., for example, [23], [24]. For the characteristic function of S_n we have $v(t) = \prod_{k=1}^n (q_k + p_k e^{it})$. Since

$$|q + p e^{it}|^2 = p^2 + q^2 + 2pq \cos t = 1 - 4pq \sin^2(t/2) \leq \exp\{-4pq \sin^2(t/2)\},$$

it follows that

$$|v(t)| \leq \exp \left\{ -2 \sum_{k=1}^n p_k q_k \sin^2(t/2) \right\} = \exp \left\{ -2\sigma^2 \sin^2(t/2) \right\},$$

where $\sigma^2 = \text{Var}(S_n)$ ($\sigma > 0$). Using $|\sin x| \geq \frac{2}{\pi} |x|$ for $|x| \leq \pi/2$ and choosing $\lambda = 1/\pi$ in (10.46), this bound yields

$$\begin{aligned}
 Q(S_n; 0) &\leq \left(\frac{96}{95} \right)^2 \frac{1}{\pi} \int_{-\pi}^{\pi} \exp \left\{ -2\sigma^2 \sin^2(t/2) \right\} dt \\
 &\leq \left(\frac{96}{95} \right)^2 \frac{1}{\pi} \int_{-\infty}^{\infty} \exp \left\{ -\frac{2\sigma^2}{\pi^2} t^2 \right\} dt = \left(\frac{96}{95} \right)^2 \frac{1}{\sigma \sqrt{2}} \sqrt{\pi}.
 \end{aligned}$$

That is, simplifying the numerical constant, we arrive at

$$M(S_n) \leq \frac{1.28}{\sqrt{\text{Var}(S_n)}}, \quad \text{Var}(S_n) = \sum_{k=1}^n p_k q_k.$$

As we now see, the constant 1.28 in this inequality can be improved by virtue of the upper bound (1.4), which yields

$$M(S_n) \leq \frac{2}{\sqrt{1 + 4\text{Var}(S_n)}} \leq \frac{1}{\sqrt{\text{Var}(S_n)}}. \tag{10.47}$$

Moreover, the best universal constant $c > 0$ in

$$M(S_n) \leq \frac{c}{\sqrt{\text{Var}(S_n)}} \tag{10.48}$$

is actually better than 1. As was shown in [1], the optimal constant is given by

$$c = \max_{\lambda \geq 0} \left[\sqrt{2\lambda} e^{-2\lambda} \sum_{k=0}^{\infty} \left(\frac{\lambda^k}{k!} \right)^2 \right] \sim 0.4688.$$

One should however mention that the first inequality in (10.47) is better than (10.48) for small values of $\text{Var}(S_n)$, namely when

$$\text{Var}(S_n) < \frac{c^2}{4(1 - c^2)} \sim 0.0704.$$

So, it makes sense to consider improvements in the form $M(S_n) \leq \Psi(\text{Var}(S_n))$.

Note that the upper bound in (1.4) also provides a similar lower bound which may equivalently be rewritten as

$$N_{\infty}(S_n) \geq \frac{1}{4} + \text{Var}(S_n).$$

In [21] this inequality is sharpened and is extended to all α -entropy powers with $\alpha \geq 2$ as

$$N_{\alpha}(S_n) \geq 1 + 2\beta \text{Var}(S_n), \quad \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

Acknowledgements

The authors would like to thank the referee for careful reading and valuable comments. The research of S.B. was supported by the NSF grant DMS-1855575.

References

- [1] Baillon, J.-B., Cominetti, R. and Vaisman, J. (2016) A sharp uniform bound for the distribution of sums of Bernoulli trials. *Combin. Probab. Comput.* **25** 352–361.
- [2] Bobkov, S. G. and Chistyakov, G. P. (2012) Bounds for the maximum of the density of the sum of independent random variables. (*Russian*) *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **408**, *Veroyatnost i Statistika* **18** 62–73, 324; translation in *J. Math. Sci. (N.Y.)* **199** (2014) 100–106.
- [3] Bobkov, S. G. and Chistyakov, G. P. (2015) Entropy power inequality for the Rényi entropy. *IEEE Trans. Inform. Theory* **61** 708–714.
- [4] Bobkov, S. G. and Chistyakov, G. P. (2015) On concentration functions of random variables. *J. Theor. Probab.* **28** 976–988.
- [5] Bobkov, S. G. and Marsiglietti, A. Variants of the entropy power inequality. *IEEE Trans. Inform. Theory* **63** 7747–7752.
- [6] Costa, J., Hero, A. and Vignat, C. (2003) On solutions to multivariate maximum α -entropy problems. In *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, EMMCVPR 2003*, pp. 211–226.
- [7] Johnson, O. and Goldschmidt, C. (2006) Preservation of log-concavity on summation. *ESAIM Probab. Stat.* **10** 206–215 (electronic).
- [8] Johnson, O. and Yu, Y. (2010) Monotonicity, thinning, and discrete versions of the entropy power inequality. *IEEE Trans. Inform. Theory* **56** 5387–5395.

- [9] Haghghatshoar, S., Abbe, E. and Telatar, I. E. (2014) A new entropy power inequality for integer-valued random variables. *IEEE Trans. Inform. Theory* **60** 3787–3796.
- [10] Harremoës, P. and Vignat, V. (2003) An entropy power inequality for the binomial family. *J. Inequal. Pure Appl. Math.* **4** 1–15.
- [11] Hoggar, S. G. (1974) Chromatic polynomials and logarithmic concavity. *J. Combin. Theory Ser. B* **16** 248–254.
- [12] Kapur, J. N. (1988) Generalised Cauchy and Student's distributions as maximum-entropy distributions. *Proc. Nat. Acad. Sci. India Sect. A* **58** 235–246.
- [13] Li, J. (2018) Rényi entropy power inequality and a reverse. *Studia Math.* **242** 303–319.
- [14] Li, J., Marsiglietti, A. and Melbourne, J. (2020) Further investigations of Rényi entropy power inequalities and an entropic characterization of s -concave densities. In *Geometric Aspects of Functional Analysis: GAFA Israel Seminar (2017–2019)*, vol. 2256 of Lecture Notes in Mathematics (B. Klartag and E. Milman, eds), Springer.
- [15] Madiman, M. and Barron, A. R. (2007) Generalized entropy power inequalities and monotonicity properties of information. *IEEE Trans. Inform. Theory* **53** 2317–2329.
- [16] Madiman, M., Melbourne, J. and Xu, P. (2017) Rogozin's convolution inequality for locally compact groups. [arXiv:1705.00642](https://arxiv.org/abs/1705.00642)
- [17] Madiman, M., Wang, L. and Woo, J. O. (2019) Majorization and Rényi entropy inequalities via sperner theory. *Disc. Math. (AEGT 2017 Special issue edited by S. Cioaba, R. Coulter, E. Fiorini, Q. Xiang, F. Pfender)* **342** 2911–2923.
- [18] Marsiglietti, A. and Melbourne, J. (2019) On the entropy power inequality for the Rényi entropy of order $[0, 1]$. *IEEE Trans. Inform. Theory* **65** 1387–1396.
- [19] Melbourne, J. and Tkocz, T. (2020) On the Rényi entropy of log-concave sequences. In *IEEE International Symposium on Information Theory (ISIT)*, pp. 2292–2296.
- [20] Melbourne, J. and Tkocz, T. (2020) Reversal of Rényi entropy inequalities under log-concavity. *IEEE Trans. Inform. Theory* **67** 45–51.
- [21] Madiman, M., Melbourne, J. and Roberto, C. (2020) Bernoulli sums and Rényi entropy inequalities. Preprint.
- [22] Moriguti, S. (1952) A lower bound for a probability moment of any absolutely continuous distribution with finite variance. *Ann. Math. Statistics* **23** 286–289.
- [23] Petrov, V. V. (1975) Sums of independent random variables. In *Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82* (Translated from the Russian by A. A. Brown) Springer-Verlag, New York-Heidelberg, x+346 pp. Russian ed.: Moscow, Nauka, 1972, 414 pp.
- [24] Petrov, V. V. (1987) *Limit Theorems for Sums of Independent Random Variables*. (Russian) Moscow, Nauka, 320 pp.
- [25] Ram, E. and Sason, I. (2016) On Rényi Entropy Power Inequalities. *IEEE Trans. Inform. Theory* **62** 6800–6815.
- [26] Stanley, R. (1989) Log-concave and unimodal sequences in algebra, combinatorics, and geometry. *Ann. New York Acad. Sci.* **576** 500–535.