A DISCRETE COMPLEMENT OF LYAPUNOV’S INEQUALITY AND ITS INFORMATION THEORETIC CONSEQUENCES

BY JAMES MELBOURNE\textsuperscript{1}, GERARDO PALAFOX-CASTILLO\textsuperscript{2}

\textsuperscript{1}Probabilidad y Estadística, Centro de Investigación en Matemáticas, james.melbourne@cimat.mx
\textsuperscript{2}Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León, gerardo.palafoxcslt@uanl.edu.mx

We establish a reversal of Lyapunov’s inequality for monotone log-concave sequences, settling a conjecture of Havrilla-Tkocz and Melbourne-Tkocz. A strengthened version of the same conjecture is disproved through counter example. We also derive several information theoretic inequalities as consequences. In particular sharp bounds are derived for the varentropy, Rényi entropies, and the concentration of information of monotone log-concave random variables. Moreover, the majorization approach utilized in the proof of the main theorem, is applied to derive analogous information theoretic results in the symmetric setting, where the Lyapunov reversal is known to fail.

1. Introduction. In this paper we prove the following reversal of Lyapunov’s inequality\textsuperscript{1}, conjectured in [22] and [38],

\textbf{Theorem 1.1.} For $x$, a monotone, log-concave sequence in $\ell_1$, the function

$$ t \mapsto \log \left( t \sum_i x_i^t \right) $$

is strictly concave for $t \in (0, \infty)$.

This is anticipated by affirmative results in the continuous setting dating back to Cohn [12] on $\mathbb{R}$, and Borell [10] in $\mathbb{R}^d$.

\textbf{Theorem 1.2 (Cohn [12], Borell [10]).} For $f$ in $L_1(\mathbb{R})$ such that $\log f$ is concave, then

$$ t \mapsto \log \left( t \int_{\mathbb{R}} f^t(x) dx \right), $$

is concave. Moreover, if $g$ is a non-negative concave function with compact support and $\gamma > 0$,

$$ t \mapsto \log \left( (t + \gamma) \int_{\mathbb{R}} g^{\frac{t}{\gamma}}(x) dx \right), $$

is concave.

\textit{MSC2020 subject classifications:} Primary 94A15; secondary 62B10.

\textit{Keywords and phrases:} Log-concave sequences, Reverse Lyapunov, Concentration of information, Rényi Entropy.

\textsuperscript{1}By Lyapunov’s inequality we refer to the fact that $p \mapsto \log \|f\|_p^p$ is convex in $p$ for a general measurable function $f$ and measure.
However in contrast to the continuous setting, the requirement that \( x \) is monotone cannot be dropped\(^2\), see [38] for examples. Moreover, we will also provide a counter example to a strengthening of Theorem 1.1 conjectured in [22, 38], further differentiating the continuous and discrete settings.

The main novelty in the proof is to establish a majorization between the distribution function of a monotone log-concave sequence and its geometric counterpart. Though we will not expound upon this outside of its application to this proof, it can be understood as a second order analog of the distributional majorization lemma utilized in [37, 39]. This alongside some further reductions, leaves one needing only the special case of a geometric sequence, which can be approached with direct computation, to complete the proof of Theorem 1.1.

Log-concave sequences arise naturally in many different mathematical disciplines. For example, a fact that goes back to Newton (see [48]), a positive sequence that forms the coefficients of a real rooted polynomial is log-concave. In combinatorics, when \( I(n) \) is the number of independent sets of size \( n \) in a matroid, \( I(n) \) is log-concave [1], see [11, 23] for more on recent developments on log-concave poset inequalities. In convex geometry, by the Alexandrov-Fenchel inequalities, (see [13, 46, 52] for recent proofs), the sequence of intrinsic volumes associated to a convex body is log-concave [45]. This effort also fits within a general pursuit, developing discrete analogs for the continuous convexity theory, which in recent investigation has connected information theory and convex geometry (see [30] for background). One instantiation is the effort to understand the behavior of the entropy of discrete variables under independent summation, see [9, 24, 29, 32, 38]. Another is the pursuit of discrete Brunn-Minkowski type inequalities [18, 19, 20, 25, 35, 40, 47]. In fact, in information theoretic language, the Brunn-Minkowski inequality can be understood as a “Rényi entropy power” inequality, see [5, 6, 16, 26, 27, 31, 34, 42, 43, 49, 50].

We will see that Theorem 1.1 yields several information theoretic consequences for discrete monotone variables. We obtain sharp bounds on the varentropy to be compared to its continuous analog [17], and utilize this to derive concentration of information, analogous to [3, 17]. We give sharp reversals of the monotonicity of Rényi entropy general parameters, augmenting the recently obtained comparisons for the \( \infty \)-Rényi entropy given in [38], and as a consequence we obtain a sharp reverse entropy power type inequality for iid variables, tightening a result from [38]. We mention that this reverse entropy power is the discrete analog of an entropic Rogers-Shephard inequality pursued by Madiman and Kontoyannis in [28]. We also obtain a sharp comparison between the value of a log-concave sequence at its mean, and the value at its mode which we compare with the classical result of Darroch [15] for Bernoulli sums. We will also obtain as a corollary of our arguments that for the monotone log-concave variables of a fixed \( p \)-Rényi entropy, the geometric distribution has maximal \( q \)-Rényi entropy for \( q \geq p \) and minimal \( q \)-Rényi entropy for \( q \leq p \).

As mentioned, Theorem 1.1 can fail without the assumption of monotonicity. In particular, symmetric variables do not necessarily satisfy the conclusion of Theorem 1.1. However we will demonstrate that the majorization techniques used are robust enough to be applied in the symmetric case, and we use them to deliver sharp Rényi entropy comparisons, varentropy bounds, and concentration of information results in the symmetric setting. We also establish the “symmetric geometric” distribution as the maximal (resp. minimal) \( q \)-Rényi entropy distribution for fixed \( p \)-Rényi entropy among discrete symmetric log-concave variables for \( q \geq p \) (resp. \( q \leq p \)).

Let us outline the paper. In Section 2, we will define notation and derive applications of the main theorem. In Section 3 we give the proof of Theorem 1.1. In Section 4 we give a counter

\(^2\)Note that for continuous variables on \( \mathbb{R} \), proving the result for monotone variables is equivalent to the general result since log-concavity is preserved under rearrangement, see for example [36].
example to the strengthening of Theorem 1.1 conjectured in [22, 38], while in Section 5, we derive analogs of the consequences in Section 2 for symmetric log-concave variables. In the Appendix we recall some elementary results from the theory of majorization for the convenience of the reader.

2. Applications.

2.1. Definitions. For a real valued random variable $Y$, we let $\mathbb{E}Y$ denote its expectation, and denote its variance $\text{Var}(Y) := \mathbb{E}Y^2 - (\mathbb{E}Y)^2$.

**Definition 1.** Let $(E, \mu)$ be a measure space and $X$ an $E$ valued random variable with density function $f$ such that $\mathbb{P}(X \in A) = \int_A f\,d\mu$. We define the information content $I_X : E \to \mathbb{R}$, as $I_X(x) = -\log f(x)$.

To avoid confusion, in sections where we discuss the information content random variable $I_X(X)$, we will avoid the usual abuse of notation and write $H(f)$ for the entropy of a variable $X$. Conversely, when there is no risk of confusion, and we consider a single variable $X$, we will omit the subscript and write $I$ for the information content. We write $H(\mu)(X) = H(\mu)(f) := \mathbb{E}I(X)$ in the general case. For example, when $E$ is discrete, and $\mu$ is the counting measure,

$$\mathbb{E}I(X) = H(X)$$

is just the Shannon entropy of $X$. Observe that when $\mu$ is a probability measure given by a random variable $Y$, then the expectation of the information content is given by the relative entropy (or Kullback-Leibler divergence), $H(\mu)(f) = -D(X||Y)$, and the varentropy measures the deviation of $-I(f)$ from $D(X||Y)$.

In physical applications, it may be more natural to write the density of $X$ in terms of a potential $E$, $f(x) = e^{-E(x)}$, in which case, $\mathbb{E}I(X)$ reflects the average energy of a system, and $V(X)$ the average fluctuation.

**Definition 2.** For a random variable $X$ taking values on a measure space $(E, \mu)$ with density function $f$, define the varentropy functional,

$$V(X) = \mathbb{E}(\log f(X) - \mathbb{E}\log f(X))^2.$$  

Unless specified, we will consider $E = \mathbb{Z}$ and $\mu$ the standard counting measure, so that the density function $x_n := \mathbb{P}(X = n)$ of a variable $X$, can be expressed as a non-negative sequence.

**Definition 3.** A non-negative sequence $x_i$, indexed over $\mathbb{Z}$, is log-concave when

$$x_i^2 \geq x_{i-1}x_{i+1}$$

and $i \leq j \leq k$ with $x_ix_k > 0$ implies $x_j > 0$.

We consider a sequence to be increasing when $x_ix_{i+1} > 0$ implies $x_{i+1} \geq x_i$, and decreasing when $x_ix_{i+1} > 0$ implies $x_{i+1} \leq x_i$. A sequence is monotone when it is either increasing or decreasing.

**Definition 4.** A discrete random variable $X$ in $\mathbb{Z}$ with probability mass function $f : \mathbb{Z} \to [0, 1]$ is log-concave when the sequence $x_i := f(i) := \mathbb{P}(X = i)$ is log-concave. The variable $X$ is monotone when the sequence $x_i$ is monotone.

We say a non-negative sequence $x_i$ belongs to $\ell_p$ when $\sum_{i \in \mathbb{Z}} x_i^p < \infty$. Note that when $x_i$ is log-concave, $x_i$ belonging to $\ell_1$ implies that $x_i$ belongs to $\ell_p$ for all $p \in (0, \infty)$. 

2.2. Varentropy bounds.

**Theorem 2.1.** For a monotone, log-concave discrete random variable $X$ in $\mathbb{Z}$, 
\[ V(X) < 1. \]

**Proof.** Let $X$ have a probability mass function $f$, and define $\Psi(t) = \log \left( t \sum_n f^t(n) \right)$, then
\[ \Psi'(t) = \frac{\sum_n \log f(n) f^t(n)}{\sum_n f^t(n)} + \frac{1}{t} \]
and
\[ \Psi''(t) = \frac{\left( \sum_n f^t(n) \right) \left( \sum_n \log^2 f(n) f^t(n) \right) - \left( \sum_n \log f(n) f^t(n) \right)^2}{\left( \sum_n f^t(n) \right)^2} - \frac{1}{t^2} \]

By Theorem 1.1, $\Psi$ is concave, thus $\Psi''(1) = V(X) - 1 < 0$, and our result follows. \qed

The bound is sharp, the varentropy of a geometric distribution $Z_p$ with parameter $p$, meaning $\mathbb{P}(Z_p = k) = p^k (1 - p)$ for $k = 0, 1, \ldots$, can be explicitly computed as $V(Z_p) = \frac{p}{(1-p)^2} \log^2(p)$ which tends to 1 with $p \to 1$.

2.3. Rényi entropy comparisons.

**Definition 5.** For $X$ a random variable on $\mathbb{Z}$, and $p \in (0, 1) \cup (1, \infty)$ define
\[ H_p(X) := \frac{\log \left( \sum_i x_i^p \right)}{1 - p}, \]
where $x_i := \mathbb{P}(X = i)$. Let $H_1(X) := H(X) = -\sum_i x_i \log x_i$ and $H_\infty(X) = -\log \|x\|_\infty$ where $\|x\|_\infty := \max_i x_i$ and $H_0(X) = \#\{i : x_i > 0\}$.

**Theorem 2.2.** When $X$ is a monotone and log-concave variable taking values in $\mathbb{Z}$ then $p > q > 0$ implies,
\[ H_p(X) > H_q(X) + \log \left( \frac{c(p)}{c(q)} \right), \]
where $c(\alpha) := \alpha^{-\frac{1}{\alpha - 1}}$.

**Proof.** Let $x_i = \mathbb{P}(X = i)$. We prove the case $p > q > 1$, the other cases can be treated similarly. Letting $\lambda = \frac{q - 1}{p - 1}$, $q = \lambda p + (1 - \lambda) 1$, so that by the strict concavity given by Theorem 1.1,
\[ \log \left( q \sum_i x_i^q \right) > \lambda \log \left( p \sum_i x_i^p \right) + (1 - \lambda) \log \left( 1 \sum_i x_i^1 \right), \]
and our result follows for $p, q \notin \{1, \infty\}$ from this inequality. Owing to log-concavity there is no difficulty obtaining the limiting cases through continuity. \qed

Note that when $X_\lambda$ has geometric distribution $\mathbb{P}(X = n) = (1 - \lambda) \lambda^n$ for parameter $\lambda \in (0, 1)$, its Rényi entropy can be computed directly,
\[ H_p(X_\lambda) = \log \left( \frac{(1 - \lambda)^p}{1 - \lambda^p} \right)^{\frac{1}{p-1}}. \]
Hence,

\[ H_p(X_\lambda) - H_q(X_\lambda) = \log \left( \frac{(1 - \lambda)^p}{1 - \lambda^p} \right)^{1/p} \quad - \log \left( \frac{(1 - \lambda)^q}{1 - \lambda^q} \right)^{1/q} \]

\[ = \log \left( \frac{1 - \lambda}{1 - \lambda^p} \right)^{1/p} \quad - \log \left( \frac{1 - \lambda}{1 - \lambda^q} \right)^{1/q}, \]

which tends to \( \log \left( \frac{1}{p^{1/p}} \right) \) with \( \lambda \to 1 \). Thus we see that Theorem 2.2 is sharp.

The following result is actually a consequence of the Rényi entropy comparison derived in [38]. It does not need the assumption of monotonicity. The result should be compared to the classical result of Darroch [15], that states that for independent sums of Bernoulli random variables the distance between the mean and mode is no greater than 1, see also [29, 41, 51] for background and recent developments on such variables. For the larger class of log-concave variables such a result is impossible. For example, a geometric distribution has mode at 0, but can have arbitrarily large expectation. However, the result below demonstrates that the value of any log-concave distribution at its mean approximates up to an absolute constant \( e \), the value of the distribution at its mode.

**Corollary 2.3.** For \( X \) with log-concave density function \( f \) with support \( A \subseteq \mathbb{Z} \),

\[ \max \{ f([\|\mathbb{E}X\|]), f([\mathbb{E}X]) \} \geq e^{-1} \|f\|_{\infty} \]

where \([\cdot]\) and \([\cdot]\) denote the usual floor and ceiling.

Note the inequality is sharp in the sense that the constant \( e^{-1} \) cannot be improved, as can be seen by choosing a geometric distribution with large, integer valued mean.

In the proof, a slightly stronger result is obtained. It in fact holds that if \( g \) is a log-concave function in the continuous sense, such that \( g(n) = f(n) \) when \( n \) is an integer, then

\[ g(\mathbb{E}X) \geq e^{-1}\|f\|_{\infty}. \]

The smallest such log-concave function, is the log-affine interpolation which we denote by \( \tilde{f} \). It is characterized by the fact that \( \tilde{f}(n) = f(n) \) when \( n \) is an integer, while \( \log \tilde{f} \) is affine on \([n, n+1] \), and can be written explicitly

\[ \tilde{f}(x) = \begin{cases} 
1 - (x-[x])([x]) & \text{for } x \in \text{co}(A), \\
0 & \text{otherwise.} 
\end{cases} \]

Thus the corollary allows a small strengthening,

\[ f^{1-t}([\|\mathbb{E}X\|])f^t([\mathbb{E}X]) \geq e^{-1}\|f\|_{\infty} \]

with \( t = \mathbb{E}X - \|\mathbb{E}X\| \).

**Proof.** If \( \tilde{f} \) denotes the log-affine interpolation of \( f \) as in (2) then \( \log \tilde{f}(x) \) is a concave function for \( x \) in the convex hull of \( A \). In particular \( \tilde{f}(\mathbb{E}X) \) is well defined. Thus, by Jensen’s inequality we have

\[ H(f) = -\mathbb{E} \log f(X) = -\mathbb{E} \log \tilde{f}(X) \geq -\log \tilde{f}(\mathbb{E}X). \]

By Theorem 1.3 of [38], \( H(f) \leq H(\infty) + 1 \), and inserting the inequality into exponentials we have

\[ \exp(-\log \|f\|_{\infty} + 1) \geq \exp(-\log \tilde{f}(\mathbb{E}X)), \]
and

\[ \frac{e}{\|f\|_\infty} \geq \frac{1}{\tilde{f}(\mathbb{E}X)}, \]

which yields the result since

\[ \max\{f([\mathbb{E}X]), f([\mathbb{E}X])\} \geq \tilde{f}(\mathbb{E}X). \]

\[ \blacksquare \]

2.4. Concentration of information content.

**Theorem 2.4.** For \( X \sim f \) monotone log-concave variable on \( \mathbb{Z} \), for \( t > 0 \)

\[ \mathbb{P}(I(X) \geq H(f) + t) \leq (1 + t)e^{-t}, \]

and when \( t \leq 1 \),

\[ \mathbb{P}(I(X) \leq H(f) - t) \leq (1 - t)e^t. \]

Note that when \( t = 1 \), we obtain \( \mathbb{P}(I(X) \leq H(X) - 1) = 0 \), implying that \(- \log \|f\|_\infty = H_\infty(f) > H(f) - 1\) recovers the sharp comparison of min-entropy and Shannon entropy above. The inequality \( H_\infty(X) \geq H(X) - 1 \) holds without the monotonicity assumption, see [38].

The following is a general and elementary technique for deriving concentration of the information content based on uniform bounds on the varentropy of the “canonical ensemble”. In [17], it is assumed that \( X \) takes values in \( \mathbb{R}^d \), and has a density with respect to the Lebesgue measure. We include the proof adapted from [17] below, for the convenience of the reader.

**Lemma 2.5 (Fradelizi-Madiman-Wang [17]).** For a random variable \( X \) on \( E \) with density \( f \in L^\alpha(\mu) \) for all \( \alpha > 0 \), and \( X_\alpha \sim \int f^{\alpha} d\mu \) satisfying \( V(X_\alpha) \leq K \), then for \( t > 0 \)

\[ \mathbb{P}(I(X) - H(\mu)(f) \geq t) \leq e^{-Kr(t/K)}, \]

and

\[ \mathbb{P}(I(X) - H(\mu)(f) \leq -t) \leq e^{-Kr(-t/K)}, \]

where \( r(t) = t - \log(1 + t) \) for \( t \geq -1 \) and is infinite otherwise.

The proof is a combination of results from [17], Theorem 3.1 and Corollary 3.4 in particular.

**Proof.** Observe that the function \( F(\alpha) = \log \int f^{\alpha}(x) d\mu(x) \) is infinitely differentiable\(^3\)

\[ K = \sup_{\alpha > 0} V(X_\alpha) = \sup_{\alpha > 0} 2\alpha^2 F''(\alpha). \]

\(^3\)Indeed, the \( n \)-th derivative of \( \alpha \mapsto f^{\alpha}(x) \), \( f^{\alpha}(\log f)^n \) is measurable as the composition of a measurable function \( f \), with a continuous function \( x^{\alpha}(\log x)^n \), and that further, and \( |f^{\alpha}(\log f)^n| \leq \frac{1}{2} \{ f > 1 \} f^{\alpha + \epsilon} C(n, \epsilon) + \frac{1}{2} \{ f < 1 \} f^{\alpha - \epsilon} c(n, \epsilon) \) for \( \alpha' \in (\alpha - \epsilon/2, \alpha + \epsilon/2) \), where \( C \) and \( c \) are uniform bounds on \( (\log x)^n/x^{\epsilon/2} \) for \( x \geq 1 \) and \( x^{\epsilon/2}|\log x|^n \) for \( x \leq 1 \) respectively, so that the requisite domination exists for Lebesgue dominated convergence to pass the derivative and integrals.
By applying $F''(t) \leq K/t^2$ to the Taylor expansion,

$$F(\alpha) = F(1) + (\alpha - 1)F'(1) + \int_1^\alpha (\alpha - t)F''(t)\,dt$$

yields

(5) $$F(\alpha) = F(1) + (\alpha - 1)F'(1) + K(\alpha - 1 - \log \alpha).$$

With the substitution $\alpha = 1 - \beta$, and the insertion of $F(1) = 0$, and $F'(1) = -H(\mu)(X)$ we can rewrite (5) as

(6) $$\mathbb{E} \left( e^{\beta(I(X)-H(\mu)(f))} \right) \leq e^{K\beta(1-\beta)}.$$  

For $\beta, t > 0$, taking exponentials and applying Markov’s inequality,

$$\mathbb{P}(I(X) - H(\mu)(f) \leq -t) \leq e^{-\beta(I(X)-H(\mu)(f))} e^{-\beta t} \leq e^{K(r(\beta) - \frac{2t}{K})}.$$  

Standard calculus allows minimization over $\beta$ and yields, $\inf_\beta r(\beta) - \frac{2t}{K} = -r(-t/K)$ which gives (4). Applying the same ideas yields (3) as well.

PROOF OF THEOREM 2.4. If $X \sim f$, is log-concave and monotone, then $X_\alpha \sim f_\alpha := f^\alpha / \sum_n f^\alpha(n)$ is as well. Hence by Theorem 2.1, $V(X_\alpha) \leq 1$. Applying Lemma 2.5 with $K = 1$ yields the result.

2.5. Renyi Entropy Power Reversals. The entropy power inequality, is a fundamental inequality in information theory that gives a sharp lower bound on the amount of entropy increase in summation of continuous independent variables, explicitly taking $\mu$ to be the Lebesgue measure on $\mathbb{R}^d$, and denoting for $X$ with density $f$ with respect to $\mu$,

$$N(X) = e^{\frac{d}{2}H(\mu)(f)},$$

Shannon’s entropy power inequality states that

$$N(X + Y) \geq N(X) + N(Y)$$

for independent random vectors $X$ and $Y$. More generally, super-additivity properties of the Rényi entropy have been studied, extending the Shannon’s EPI, see [7, 26, 27, 30, 34, 42, 43]. We consider a Rényi Entropy Power reversal to be any non-trivial upper bound on the entropy of a sum of random variables, see [2, 4, 8, 9, 14, 53, 54].

THEOREM 2.6. For $X, Y$ iid, log-concave, and monotone on $\mathbb{Z}$, and $\alpha \in [2, \infty]$

$$H_\alpha(X - Y) \leq H_\alpha(X) + \log 2.$$  

The inequality is a sharp improvement for monotone log-concave variables of Theorem 6.2 of [38], where it is proven that $H_\alpha(X - Y) \leq H_\alpha(X) + \alpha^{\frac{1}{\alpha}} \log 2$ for $X$ and $Y$ iid and log-concave. To see that the constant 2 cannot be improved, take $X$ to have density $f(n) = (1 - p)p^n$ so that for $n \geq 0$, $f_{X-Y}(n) = \frac{1-p}{1+p}p^n$. Taking the limit with $p \to 1$ shows the inequality to be sharp. An alternative motivation for the inequality is its relationship to an entropic generalization conjectured by Madiman and Kontoyannis [28] of the Rogers-Shephard inequality from convex geometry [44], see also [38] for further discussion.
The proof relies on an elementary trick, known to specialists, that $H_2(X) = H_\infty(X - Y)$ holds for iid variables $X$ and $Y$. We include a proof for completeness and emphasize that this equality is independent of any property of the distribution$^4$.

**Lemma 2.7.** For $X$ and $Y$ iid on $\mathbb{Z}$,

\begin{equation}
H_2(X) = H_\infty(X - Y).
\end{equation}

**Proof.** Let $f$ denote the shared distribution of $X$ and $Y$ and $f_{X-Y}$ the distribution of $X - Y$. We compute directly,

\[
\sum_k f^2(X = k) = \sum_k \mathbb{P}(X = k) \mathbb{P}(Y = k) = \mathbb{P}(X - Y = 0).
\]

After taking logarithms, this shows that

\begin{equation}
H_2(X - Y) = -\log f_{X-Y}(0),
\end{equation}

thus the result follows from demonstrating that $f_{X-Y}(0) = \|f_{X-Y}\|_\infty$. To this end, we recall the elementary rearrangement inequality (see for instance [33]) that for non-negative sequences $x, y \in \ell_2$,

\[
\sum_i x_iy_i \leq \sum_i x_i^\downarrow y_i^\downarrow
\]

where $x^\downarrow$ is the sequence $x$ rearranged in decreasing order. If we denote $\tau_nf(k) = f(n + k)$ then

\[
f_{X-Y}(n) = \sum_k \tau_nf(k) f(k) \leq \sum_k (\tau_nf)^\downarrow(k) f^\downarrow(k).
\]

However since $\tau_nf$ is just a translation of $f$, $(\tau_nf)^\downarrow = f^\downarrow$ and since $\sum_k (f^\downarrow)^2(k) = \sum_k f^2(k) = f_{X-Y}(0)$ our result follows. \qed

When $\alpha \leq 2$ a constant depending on $\alpha$ can be found using only monotonicity of Rényi, see Theorem 6.2 in [38].

**Proof of Theorem 2.6.** Recalling the notation of Theorem 2.2, $c(\alpha) = \alpha^{\frac{1}{\alpha - 1}}$, with $c(\infty) := 1$, and then applying this same theorem twice, the result is as follows.

\[
H_\alpha(X - Y) \leq H_\infty(X - Y) + \log \frac{c(\alpha)}{c(\infty)}
\]

\[
= H_2(X) + \log \frac{c(\alpha)}{c(\infty)}
\]

\[
\leq H_\alpha(X) + \log \frac{c(2)}{c(\alpha)} + \log \frac{c(\alpha)}{c(\infty)}
\]

\[
= H_\alpha(X) + \log 2.
\]

$^4$The proof is given for iid log-concave $X$ and $Y$ in [38].
3. Proof of Theorem 1.1. For \( x : \mathbb{Z} \rightarrow [0, \infty) \), a monotone, log-concave, \( \ell_1 \) sequence, we denote \( S = \{ i \in \mathbb{Z} : x_i > 0 \} \), and \( \Phi_x : (0, \infty) \rightarrow \mathbb{R} \),

\[
\Phi_x(t) := \log \left( t \sum_{i \in S} x_i^t \right).
\]

To prove that \( \Phi_x \) is always strictly concave, we will first start with some reductions. For \( x \) a log-concave sequence and \( p > q \) we wish to prove,

\[
\Phi_x((1 - s)p + sq) - (1 - s)\Phi_x(p) - s\Phi_x(q) \geq 0, \quad s \in [0, 1].
\]

If we denote by \( x^q \), the monotone log-concave sequence \( (x^q)_i = (x^q)_i \) and \( \bar{p} = p/q \), then by algebraic manipulation the left hand side of (9) is exactly

\[
\Phi_{x^q}((1 - s)\bar{p} + 1) - (1 - s)\Phi_{x^q}(\bar{p}) - s\Phi_{x^q}(1) \geq 0.
\]

Additionally observe that for a constant \( c > 0 \), with \( cx \) denoting the sequence \( (cx)_i = cx_i \) that \( \Phi_{cx}(t) = \Phi_x(t) + t \log c \). Thus we can and will without loss of generality assume that \( \sum_i x_i = 1 \) and need only prove that for \( p > 1 \), and \( s \in (0, 1) \)

\[
\Phi_x((1 - s)p + s) \geq (1 - s)\Phi_x(p) + s.
\]

For the proof of this result we will derive the following lemma.

**Lemma 3.1.** For \( x \) a non-Dirac, monotone log-concave probability sequence, \( p > 1 \), and \( q \in (1, p) \), there exists a \( \lambda \in (0, 1) \) such that the sequence \( z \) given by \( z_k = (1 - \lambda)\lambda^k \) satisfies

\[
\sum_i x_i^p = \sum_i z_i^p,
\]

and

\[
\sum_i x_i^q \geq \sum_i z_i^q.
\]

As we will see Lemma 3.1 reduces our problem to proving (11) for the geometric distribution. To prove the lemma, we establish a majorization between the distribution function of a monotone log-concave variable and its geometric counterpart.

**Proposition 3.2.** For a sequence \( x \), define \( F_x(t) := \#\{ i : x_i > t \} \). Let \( x \) be a log-concave, non-increasing sequence, and \( z_k = Cp^k \) for \( C > 0 \) and \( p \in (0, 1) \). Then there exist a finite interval \( I \) such that \( F_x(t) \leq F_z(t) \) if \( t \in I \) and \( F_x(t) \geq F_z(t) \) if \( t \notin I \).

**Proof.** Define \( a := \min\{ k : x_k \geq z_k \} \), and \( b := \sup_k \{ k : x_k \geq z_k \} \). It follows from the log-concavity of \( x \) and the log-affinity of \( z \) that \( \{ k : x_k \geq z_k \} \) is a discrete interval. Thus, the interval\(^5\)

\[
[a, b] = \{ k : x_k \geq z_k \}.
\]

Let \( I = [z_b, x_a] \), with \( z_b = 0 \) in the case \( b = +\infty \). Let \( t \in I \). Two cases will be considered: \( t < z_a \) and \( z_a \leq t \). First assume \( z_b \leq t < z_a \). Let \( m = \min\{ i : z_i \leq t \} = F_z(t) \). See that \( a < m \leq b \): since \( z_b \leq t \), then \( m \leq b \) because \( m \) is the minimum index such that \( z \) satisfies such inequality. Also, if \( m \leq a \), then we have \( z_m \geq z_a \) because \( z \) is decreasing, which gives

\(^5\)With the interpretation that \([a, b] = [a, \infty) \cap \mathbb{Z}\) when \( b = \infty \).
us both \( z_m \leq t \) by definition of \( m \) and \( z_m > t \) because \( z_a > t \). This is a contradiction, thus \( a < m \). Finally, since \( x_i \) is non-increasing and \( a < m \leq b \), we must have

\[
(14) \quad z_m \leq t < z_{m-1} \leq x_{m-1} \leq x_{m-2} \leq \cdots \leq x_0.
\]

From (14) we see that \( F_x(t) \geq m = F_x(t) \). Now, suppose \( z_a \leq t < x_a \). Since \( z_a \leq t \) then \( F_x(t) \leq a \). Now, since \( t < x_a \) and \( x_i \) is non-increasing, so \( x_0 \geq x_1 \geq \cdots \geq x_a > t \) and thus \( F_x(t) \geq a + 1 \). Therefore \( F_x(t) \leq a < a + 1 \leq F_x(t) \). \hfill \Box

Note that without the assumption of monotonicity, the existence of an interval \( I \) such that \( F_x(t) \leq F_x(t) \) can fail. For instance if one takes \( z_k = p^k \) for \( k \geq 0 \) and \( p \in (0, 1) \), and we define for \( \delta > 0 \), \( x \) a sequence taking the following 5 non-zero values,

\[
\{p^3 - \delta, p - \delta, 1 + \delta, p - \delta, p^3 - \delta\}.
\]

For \( \delta \) small enough that \((1 + \delta)(p^3 - \delta) \leq (p - \delta)^2\), \( x \) is log-concave, but

\[
\{F_x \geq F_x\} = [p^4, p^3 - \delta] \cup [p^3, p - \delta] \cup [p, \infty).
\]

The following is a standard fact that holds for general measure spaces. It follows from the layer-cake representation of a non-negative function, a change of variables, and an application of Fubini-Tonelli.

**Proposition 3.3.** Let \( X \) be a random variable on the non-negative integers and the sequence \( x_i := \mathbb{P}(X = i) \), then for \( t > 0 \), \( F_x(\lambda) \) as defined in Proposition 3.2 satisfies

\[
(15) \quad \sum_i x_i^t = t \int_0^\infty \lambda^{t-1} F_x(\lambda) d\lambda.
\]

In particular, \( F_x \) is a probability distribution function on \((0, \infty)\) when \( x \) is a log-concave probability sequence.

**Lemma 3.4.** If \( U, V \) are non-negative random variables with densities \( f, g \) respectively, such that \( \mathbb{E}(U) = \mathbb{E}(V) \), and \( f \leq g \) on an interval \( I \), and \( f \geq g \) outside \( I \), then

\[
\mathbb{E}(w(U)) \geq \mathbb{E}(w(V))
\]

for any convex function \( w \). The inequality reverses if \( w \) is concave.

The proof of Lemma 3.4 is classical, and given as Theorem A.9 in the Appendix for completeness.

**Theorem 3.5.** If \( U, V \) are non-negative random variables with densities \( f \) and \( g \) respectively, that satisfy \( \mathbb{E}(U^p) = \mathbb{E}(V^p) < \infty \) for \( p > -1 \) and \( f \leq g \) on an interval \( I \), and \( f \geq g \) outside of \( I \), then

\[
\mathbb{E}(w(U^p)) \geq \mathbb{E}(w(V^p))
\]

for any convex function \( w \). The inequality reverses if \( w \) is concave.

**Proof.** The \( p = 0 \) case is trivial. We will see that when \( p \neq 0 \) the result follows directly from Lemma 3.4. Indeed, \( U^p \) has density \( \tilde{f}(x) = f(x^\frac{1}{p})x^{-\frac{1-p}{p}}|p|^{-1} \) while \( V^p \) has density \( \tilde{g}(x) = g(x^\frac{1}{p})x^{-\frac{1-p}{p}}|p|^{-1} \) so that \( U^p \) and \( V^p \) satisfy the hypothesis of Lemma 3.4 for the interval \( I^p := \{w : w = x^p, x \in I\} \). \hfill \Box
PROOF OF LEMMA 3.1. For $p > 1$, and $x$ not a point mass, $0 < \sum x_i^p < \sum x_i = 1$. Then, observe that $\Psi(\lambda) := \sum_{k=0}^{\infty} ((1 - \lambda)\lambda^k)^p = \left(\frac{1 - \lambda}{\lambda}\right)^p$. By the intermediate value theorem, since $\Psi(0) = 1$ and $\lim_{\lambda \to 1} \Psi(\lambda) \to 0$ as $\lambda \to 1$ (L’Hospital), there exists $\lambda$ such that (12) holds.

Let $x$ be log-concave, non-increasing with $\sum x_i = 1$, let $z$ be geometric and let $p$ be such that $\sum x_i^p = \sum z_i^p$. Let $U$ be a random variable with density $F_x$, and $V$ be a random variable with density $F_z$. Since $\sum x_i^p = \sum z_i^p$ then $\frac{1}{p} \sum x_i^p = \frac{1}{p} \sum z_i^p$, which implies $\mathbb{E}(V^{p-1}) = \mathbb{E}(U^{p-1})$ by Proposition 3.3. With $p > 1$ and $q \in (1, p)$, we have that $g(x) = x^{\frac{q-1}{p}}$ is concave, thus $\mathbb{E}(g(V^{p-1})) \geq \mathbb{E}(g(U^{p-1}))$ by Proposition 3.2 and Theorem 3.5. Thus $\mathbb{E}(V^{q-1}) \geq \mathbb{E}(U^{q-1})$ and, multiplying both sides by $q$ and using Proposition 3.3, we get $\sum x_i^q \geq \sum z_i^q$. \hfill \Box

The last ingredient of the proof of Theorem 1.1 is to prove it in the special case that the sequence is geometric.

PROPOSITION 3.6. Let $z = (z_k)$ be a geometric distribution, i.e., $z_k = (1 - \lambda)\lambda^k$ for $\lambda \in (0, 1)$ and $k \in \{0, 1, \ldots\}$. Then

$$
\Phi_z(t) = \log \left[ t \sum z_i^t \right]
$$

is a concave function in $(0, +\infty)$.

PROOF. See that

$$
\Phi_z(t) = \log [t(1 - \lambda)^t] + \log \left[ \sum (\lambda^t)^i \right]
$$

$$
= \log t + t \log(1 - \lambda) - \log(1 - \lambda^t),
$$

thus

$$
\Phi_z''(t) = -\frac{1}{t^2} + \frac{\lambda^t}{(1 - \lambda^t)^2} \log^2 \lambda
$$

$$
= \frac{\lambda^t \log^2 \lambda - (1 - \lambda^t)^2}{((1 - \lambda^t)t)^2},
$$

so $\Phi_z''(t) \leq 0$ if and only if $\lambda^t \log^2 \lambda - (1 - \lambda^t)^2 \leq 0$, which is equivalent to $-\log(\lambda^t) \leq \frac{1}{\lambda^t} - \sqrt{\lambda^t}$. Making $x = 1/\sqrt{\lambda^t}$ gives us $2\log(x) \leq x - \frac{1}{x}$, which is readily seen true by observing the first derivative has a fixed sign. \hfill \Box

PROOF OF THEOREM 1.1. By the aforementioned reductions, let $x_i$ be a non-increasing log-concave probability sequence, and $s \in (0, 1)$. Then there exists, by Lemma 3.1, a geometric distribution $z_i$ such that $\sum x_i^p = \sum z_i^p$ and moreover for all $q \in (1, p)$,

$$
\sum x_i^q \geq \sum z_i^q.
$$

Taking $q = s + (1 - s)p$, we have by Lemma 3.1

$$
\Phi_z(s + (1 - s)p) \geq \Phi_z(s + (1 - s)p).
$$
By Proposition 3.6 Φ_z is concave, and hence
\[ Φ_z((1-s)p + s) \geq (1-s)Φ_z(p) + s. \]
Then by hypothesis, Φ_x(p) = Φ_z(p), and Φ_x(1) = Φ_z(1) = 0. Compiling these results gives the following sequence of equalities and inequalities,
\[ Φ_x(s + (1-s)p) \geq Φ_z(s + (1-s)p) \geq (1-s)Φ_z(p) + s = (1-s)Φ_x(p) + s. \]
Hence (11) holds, and we have concavity for any Φ_x.

\[ \square \]

**Corollary 3.7.** For X a monotone log-concave random variable, and Z a geometric random variable such that
\[ H_p(X) = H_p(Z), \]
then for \( q \geq p, \)
\[ H_q(X) \geq H_q(Z), \]
while
\[ H_q(X) \leq H_q(Z) \]
for \( q \leq p. \)

The proof is omitted as it is the same as the symmetric case which is given in detail in Section 5.

**4. Extensions.** A natural generalization\(^6\) was first conjectured in an early version of [22], and reiterated in [38].

**Question 1.** For \( γ > 0 \) and a positive monotone concave sequence \((y_n)_{n=1}^N\) then the function
\[ Φ_y(t) := \log \left( \left(t + \frac{γ}{2}\right) \sum_{n=1}^{N} \frac{1}{y_n} \right) \]
is concave for \( t > -γ. \)

While the continuous analogue of Question 1 is true, the following counterexample precludes an affirmative answer in the discrete case. Let \( N = 2, y = \{λ, 1+λ\} \) and consider the points, \( \{0, γ, 2γ\} \subseteq (-γ, ∞). \) Concavity of \( Φ_y \) would imply,
\[ \exp Φ_y(γ) ≥ \exp (Φ_y(0)) \exp (Φ(2γ)), \]
which is,
\[ 4γ^2(2λ + 1)^2 ≥ 6γ^2(λ^2 + (1+λ)^2). \]
Taking the limit with \( λ \to 0 \) would imply \( 4 ≥ 6. \)

---

\(^6\)If one writes \( M_γ(u,v) = \left(\frac{u^γ + v^γ}{2}\right)^{\frac{1}{γ}} \) for \( γ > 0, \) and considers a sequence \( x_n \) to be \( γ \)-concave when its support is a contiguous interval and \( x_n \geq M_γ(x_{n+1}, x_{n-1}), \) then \( x_n = y_n^{1/γ} \) is \( γ \)-concave iff \( y_n \) is concave. For \( u,v > 0, \) we define \( M_0(u,v) \lim_{γ→0} M_γ(u,v) = √{uv} \) and \( x_n \) is a log-concave sequence iff it is 0-concave. With this terminology, Theorem 1.1 affirms that if \( x \) is a monotone \( γ \)-concave sequence \( t→\log \left( \left(t + \frac{γ}{2}\right) \sum_{n=1}^{N} \frac{1}{x_n^γ} \right) \) is log-concave when \( γ = 0. \) Question 1 asks if this result can be extended to \( γ > 0. \)
5. Symmetric Variables. A random variable on \(\mathbb{Z}\) can be symmetric about a point \(m \in \mathbb{Z}\) (\(f(m+n) = f(m-n)\)) or it could be symmetric about \(n + \frac{1}{2}\) for \(n \in \mathbb{Z}\). For example \(\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2}\) is symmetric about \(0 + \frac{1}{2}\). In this case, when a log-concave sequence \((x_i)_{i \in \mathbb{Z}}\) is symmetric about a point \(n + \frac{1}{2}\),

\[
\log \left( t \sum_i x_i^t \right) = \log \left( t \sum_{i>n} x_i^t \right) + \log 2
\]

is concave by Theorem 1.1 as \((x_i)_{i>n}\) is monotone and log-concave. Thus, we have the following corollary.

**Corollary 5.1.** For \((x_i)_{i \in \mathbb{Z}}\) an \(\ell_1\) log-concave sequence, symmetric about a point \(n + \frac{1}{2}\),

\[
t \mapsto \log \left( t \sum_i x_i^t \right)
\]

is concave in \(t\). Moreover, if \(X\) is a random variable satisfying \(\mathbb{P}(X = i) = x_i\) then

\[V(X) < 1,\]

and

\[H_p(X) > H_q(X) + \log \left( \frac{p^{p-1}}{q^{q-1}} \right).\]

If \(f\) denotes the density of \(X\) then,

\[\mathbb{P}(I(X) \geq H(f) + t) \leq (1 + t)e^{-t},\]

and when \(t \leq 1\),

\[\mathbb{P}(I(X) \leq H(f) - t) \leq (1 - t)e^t.\]

**Remark 1.** See that this implies Theorem 2.4 is valid for sequences symmetric about a point \(n + \frac{1}{2}\).

However, in the case that \((x_i)\) is symmetric about a point \(n \in \mathbb{Z}\), the concavity of (19) is known to fail. In spite of this, we show in the sequel that arguments from the proof of Theorem 1.1 are able to recover sharp bounds on the varentropy and the Rényi entropy in this setting.

**Definition 6.** A sequence \(z\) is symmetric geometric when there exists \(\lambda \in (0, 1)\) and \(C > 0\) such that

\[z_n = C\lambda^{\lfloor n \rfloor}\]

for \(n \in \mathbb{Z}\). When \(C = \frac{1-\lambda}{1+\lambda}\) the sequence defines a probability distribution. A random variable \(Z\) is symmetric geometric when

\[\mathbb{P}(Z = n) = \frac{1-\lambda}{1+\lambda}\lambda^{\lfloor n \rfloor}.\]

Given \(p \in (0, \infty)\), and \(X\) symmetric and log-concave, there exists \(Z\) symmetric geometric, such that \(\sum_n f_X^p(n) = \sum_n f_Z^p(n)\).
**PROPOSITION 5.2.** Let \( x_i \) be a non-trivial probability distribution over \( \mathbb{Z} \). When \( \sum_i x_i^p < \infty \), there exists a symmetric geometric sequence \( z_i = \frac{1 - \lambda}{1 + \lambda} \lambda^{|i|} \) such that \( \sum x_i^p = \sum z_i^p \).

**PROOF.** Note for fixed \( p > 1 \), \( 0 \leq x_i^p \leq x_i \), so \( 0 < \sum_i x_i^p < \sum_i x_i = 1 \) so that the assumption of finiteness is unnecessary. A symmetric geometric sequence \( z_i \) with parameter \( \lambda \), satisfies

\[
S(\lambda) := \sum_i z_i^p = \sum_i \left( \frac{1 - \lambda}{1 + \lambda} \lambda^{|i|} \right)^p = \frac{(1 - \lambda)^p}{(1 + \lambda)^p} \left( 1 + \lambda^p \right).
\]

Clearly \( S(0) = 1 \), while \( \lim_{\lambda \to 1} \frac{(1 - \lambda)^p}{(1 + \lambda)^p} = 0 \). Also \( \lim_{\lambda \to 1} \frac{1 + \lambda^p}{(1 + \lambda)^p} = \frac{1}{2p-1} \). Therefore \( \lim_{\lambda \to 1} S(\lambda) = 0 \). By the intermediate value theorem, since \( S \) is continuous for \( \lambda \in (0, 1) \), there must be a \( \lambda \in (0, 1) \) such that \( S(\lambda) = \sum x_i^p = \sum z_i^p \).

A similar approach will handle the case that \( p \in (0, 1) \). \( \square \)

**PROPOSITION 5.3.** If \( x_i \) is a symmetric log-concave sequence for \( i \geq \mathbb{Z} \) and \( z \) is symmetric geometric, then there exists a finite interval \( I \) such that

\[
F_x(t) \geq F_z(t) \quad t \in I
\]

\[
F_x(t) \leq F_z(t) \quad t \notin I
\]

**PROOF.** We know the result to be true for \( x^* = (x_i)_{i \geq 0} \) and \( z^* = (z_i)_{i \geq 0} \) by Proposition 3.2. Now, see that

\[
2F_x - 1 = F_x
\]

and

\[
2F_z - 1 = F_z.
\]

Furthermore, \( F_x \geq F_z \) if and only if \( 2F_x - 1 \geq 2F_z - 1 \) if and only if \( F_{x^*} \geq F_{z^*} \). Similarly for \( F_x \leq F_z \). Therefore the same interval \( I \) given by Proposition 3.2 satisfies our desired inequalities. \( \square \)

**LEMMA 5.4.** Let \( X \) be log-concave, symmetric about a point \( n \in \mathbb{Z} \) with mass function \( x_i \). Then there exists a symmetric geometric distribution \( Z \), such that \( H_p(X) = H_p(Z) \) and

\[
H_q(X) \geq H_q(Z)
\]

for \( \infty \geq q \geq p > 0 \), and

\[
H_q(X) \leq H_q(Z)
\]

for \( 0 < q \leq p \leq \infty \).

**PROOF.** Fix \( p, q \in (0, 1) \cup (1, \infty) \) and choose \( Z \) to be the symmetric geometric distribution from Proposition 5.2 with mass function \( z \) satisfying \( H_q(X) = H_p(Z) \). We derive our conclusion from fact that by Proposition 5.3, the random variable \( U^{p-1} \), where \( U \) has density \( F_z \), is majorized in the convex order by \( V^{p-1} \), where \( V \) has density \( F_z \).

To this end let us write our conclusion in a more the compact form. It suffices to prove

\[
(q-p)H_q(X) \geq (q-p)H_q(Z).
\]
DISCRETE COMPLEMENT OF LYAPUNOV’S INEQUALITY

Since \( H_q(X) \geq H_q(Z) \) if and only if \( (q-1) \sum_i x_i^q \leq (q-1) \sum_i z_i^q \), our conclusion is equivalent to

\[
(q-p)(q-1) \sum_i x_i^q \leq (q-p)(q-1) \sum_i z_i^q.
\]

Thus, after algebra and recalling Proposition 3.3 this is equivalent to

\[
\frac{(q-p)(q-1)}{(p-1)^2} E U^{q-1} \leq \frac{(q-p)(q-1)}{(p-1)^2} E V^{q-1}.
\]  

(22)

Note that since \( x \mapsto x^s \) is convex for \( s \) \((s-1) > 0 \) and is concave otherwise,

\[
\phi_s(x) = s(s-1)x^s
\]

is convex for \( s \neq 0, 1 \). Thus writing (22) as

\[
E \phi_{s-1}(U^{p-1}) \leq E \phi_{s-1}(V^{p-1}),
\]

the result follows.

The limiting cases with \( q \) or \( p \in \{1, \infty\} \) can be easily handled using continuity and monotonicity of the Rényi entropy as a function of \( \alpha \mapsto H_\alpha(X) \) and as a function of the parameter \( \lambda \) of a symmetric geometric distribution \( Z_\lambda, \lambda \to H_\alpha(Z_\lambda) \).

\[\Box\]

We exclude the 0-Rényi entropy in the above for cleaner exposition, as when \( H_0(X) < \infty \) there is no symmetric geometric distribution \( Z \) with matching 0-Rényi entropy. In this case one can still conclude (trivially) that for \( p \neq q \) a symmetric log-concave variable \( X \) there exists a symmetric geometric distribution \( Z \) such that

\[
|H_p(X) - H_q(X)| \leq |H_p(Z) - H_q(Z)|
\]

as the right hand side of the above is always infinite.

**Theorem 5.5.** For a discrete log-concave random variable \( X \) symmetric about a point \( n \in \mathbb{Z} \) and \( p \geq q > 0 \), we have

\[
H_q(X) - H_p(X) \leq C(q, p) := \sup_Z H_q(Z) - H_p(Z),
\]

where the supremum is taken over all \( Z \) symmetric-geometric.

**Proof.** For any \( p \neq 1 \) and \( q \leq p \) we have, by Lemma 5.4, a symmetric geometric \( Z \) with \( H_p(X) = H_p(Z) \) and

\[
H_q(X) \leq H_q(Z),
\]

which implies

\[
H_q(X) - H_p(X) \leq H_q(Z) - H_p(Z) \leq \sup_Z H_q(Z) - H_p(Z).
\]

\[\Box\]

**Theorem 5.6.** For \( X \) log-concave and symmetric, we have for the varentropy,

\[
V(X) \leq V_S := \sup_Z V(Z),
\]

where the supremum is taken over all \( Z \) symmetric-geometric.
PROOF. Let $\Psi_X(t) = \log \sum_i x_i^{t+1}$ where $x_i = P(X = i)$. Observe that

- $\Psi_X(0) = 0$,
- $\Psi_X'(0) = -H(X)$,
- $\Psi_X''(0) = V(X)$.

Choose $Z$ to be a symmetric geometric distribution satisfying $H(Z) = H(X)$. By Lemma 5.4, $H_{1+t}(Z) \leq H_{1+t}(X)$ for $t > 0$, which corresponds to $\Psi_Z(t) \geq \Psi_X(t)$ for $t > 0$. By Taylor expansion,

$$\Psi''_X(0) = \lim_{t \downarrow 0} \frac{\Psi_X(t) - \Psi_X'(0)t}{t^2/2} \leq \lim_{t \downarrow 0} \frac{\Psi_Z(t) - \Psi_Z'(0)t}{t^2/2} = \Psi''_Z(0),$$

since the Taylor expansions are identical up to linear terms. It follows that $\Psi''_Z(0) = V(Z) \geq V(X) = \Psi''_X(0)$. $\square$

Note that if one expresses the distribution of a symmetric-geometric variable $Z_\lambda$ as $\frac{1-\lambda}{1+\lambda} \lambda^{|k|}$, its varentropy has the closed form expression,

$$V(Z_\lambda) = \frac{2\lambda(\lambda^2 + 1) \log^2 \lambda}{(\lambda^2 - 1)^2}.$$  

By the Theorem $V_S = \max_{\lambda \in [0,1]} V(Z_\lambda)$, and thus we have numerically $V_S \approx 1.16923$. This is used for the following corollary.

**Corollary 5.7.** For a discrete log-concave random variable $X$ symmetric on $\mathbb{Z}$ with the entropy $H = \mathbb{E}I(X)$ and $t \geq 0$, we have

$$\mathbb{P}(I(X) - H \geq t) \leq \left(1 + \frac{t}{V}\right)^V e^{-t},$$

and

$$\mathbb{P}(I(X) - H \leq -t) \leq \left(1 - \frac{t}{V}\right)^V e^t,$$

where $V := V_S \approx 1.16923$ is defined in Theorem 5.6.

**Proof.** The result follows from combining Lemma 2.5 and Theorem 5.6. $\square$

**APPENDIX: MAJORIZATION**

The following theorem is a well known characterization of the convex order, see [33] for proof and further background.

**Theorem A.8.** If $X$ and $Y$ are random variables on $[0, \infty)$ such that $\mathbb{E}X = \mathbb{E}Y < \infty$, then

$$\mathbb{E}\varphi(Y) \geq \mathbb{E}\varphi(X)$$

holds for all convex functions $\varphi$, if it holds for all $\varphi$ of the form $\varphi(x) = [x-t]_+$ for $t \in (0, \infty)$.
When $X$ and $Y$ satisfy (23) we say that $Y$ majorizes $X$ in the convex order, or that $Y$ majorizes $X$ for short, and write $Y \succ X$.

**Theorem A.9.** For non-negative random variables $X \sim f$ and $Y \sim g$ with densities taking values on $[0, \infty)$ such that $\mathbb{E}X = \mathbb{E}Y < \infty$, if there exists an interval $I \subseteq [0, \infty)$ such that $g \leq f$ on $I$, and $g \geq f$ on $[0, \infty) - I$, then $Y \succ X$.

**Proof.** For $t \in [0, \infty)$ define $\Psi(t) = \mathbb{E}[Y - t]_+ - \mathbb{E}[X - t]_+$. By assumption $\mathbb{E}X = \mathbb{E}Y$, and hence $\Psi(0) = 0$. By monotone convergence, $\lim_{t \to \infty} \Psi(t) = 0$. Computing the derivatives of $\Psi$, one obtains $\Psi'(t) = \mathbb{P}(X > t) - \mathbb{P}(Y > t)$, and $\Psi''(t) = g(t) - f(t)$. Observe that $\Psi'(0) = 0$, $\lim_{t \to \infty} \Psi'(t) = 0$, and $0 = \mathbb{E}Y - \mathbb{E}X = \int_0^\infty \Psi'(t)dt$. Thus, $\Phi'$ must be both positive and negative or it is exactly 0 and the problem is trivial. As such $\Phi''$ is positive, negative, and then positive. It follows that $\Phi'$ is positive and then negative, and hence $\Phi \geq 0$ and our result follows. 

**Funding.** The second author was supported fully by CONACyT, Mexico scholarship number 779613.

**REFERENCES**


