

On a Conjecture of Feige for Discrete Log-Concave Distributions

Abdulmajeed Alqasem, Heshan Aravinda, Arnaud Marsiglietti, James Melbourne

Abstract

A remarkable conjecture of Feige (2006) asserts that for any collection of n independent non-negative random variables X_1, X_2, \dots, X_n , each with expectation at most 1,

$$\mathbb{P}(X < \mathbb{E}[X] + 1) \geq \frac{1}{e},$$

where $X = \sum_{i=1}^n X_i$. In this paper, we investigate this conjecture for the class of discrete log-concave probability distributions and we prove a strengthened version. More specifically, we show that the conjectured bound $1/e$ holds when X_i 's are independent discrete log-concave with arbitrary expectation.

Keywords: Feige's conjecture, Sum of random variables, Small deviations, Log-concave.

1 Introduction

Motivated by the problem of estimating the average degree of a graph, Feige [8] investigated the probabilistic quantity, $\mathbb{P}(X < \mathbb{E}[X] + 1)$, where X is the sum of n independent non-negative random variables X_1, X_2, \dots, X_n , with $\mathbb{E}[X_i] \leq 1$ for each i . Classical inequalities such as Markov's and Chebyshev's inequalities yield no useful information about this probability. Chebyshev's inequality does not play a role in this case since there is no assumption on the variance of X , while Markov's inequality implies $\mathbb{P}(X < \mathbb{E}[X] + 1) \geq \frac{1}{\mathbb{E}[X] + 1}$, which is essentially useless when n is large. Using an approach based on a case analysis, Feige managed to prove that

$$\mathbb{P}(X < \mathbb{E}[X] + 1) \geq \frac{1}{13}. \quad (1)$$

However, as Feige noted, one may take a collection of n number of i.i.d. random variables such that for each i , $X_i = n + 1$ with probability $\frac{1}{n+1}$ and $X_i = 0$, otherwise. Then,

$$\mathbb{P}(X < \mathbb{E}[X] + 1) = \left(1 - \frac{1}{n+1}\right)^n.$$

Based on this, it was conjectured that one could replace $1/13$ with $1/e$. The improvement on inequality (1) was first made by He, Zhang and Zhang (see [17]) by establishing the lower bound $1/8$. Later, Garnett (see [12]) improved the lower bound to $7/50$. The current best bound in this direction is 0.1798 by Guo, He, Ling and Liu (see [16]).

Feige's inequality has many applications in computer science and combinatorics including the weighted max-cut problem (see [17]), approximating the average degree of a graph in sublinear-time (see [8], [14]), and the connection of a conjecture of Manickam, Miklós, and Singhi with matchings and fractional covers of hypergraphs (see [3]). We refer the reader to [2], [11] and [35] for more applications.

Our goal in this article is to prove that the conjectured lower bound holds when the collection $\{X_i\}_{i=1}^n$ is independent discrete log-concave. Recall that an integer-valued random variable X is said to be log-concave if its probability mass function p satisfies

$$p(k)^2 \geq p(k-1)p(k+1),$$

for all $k \in \mathbb{Z}$ and X has contiguous support. For example, Bernoulli, discrete uniform, binomial, negative binomial, geometric, hypergeometric and Poisson distributions are all log-concave. Many sequences in combinatorics are log-concave (or conjectured to be log-concave), see, e.g., [24], [31], [32], [6], [1]. The surveys [33] and [5] provide more details about the notion of log-concavity in the context of combinatorics. Recent developments on discrete log-concavity in probabilistic setting include log-Sobolev type inequalities [18], discrete Rényi entropy inequalities [26], [20], [25], concentration bounds and moments inequalities [4], [22].

Our main result is as follows.

Theorem 1.1. *Let X be a discrete log-concave random variable. Then,*

$$\mathbb{P}(X < \mathbb{E}[X] + 1) \geq e^{-1}. \tag{2}$$

The lower bound e^{-1} is sharp among the class of discrete log-concave random variables as can be seen by taking a random variable with probability mass function $p(k) = Cn^{k/n}$, $k \in \{1, \dots, n\}$, with n sufficiently large, where C is the normalizing constant; the details are carried out after the proof of Theorem 1.1 in Remark 2.2.

Since the sum of independent discrete log-concave random variables is log-concave (see [9]), Theorem 1.1 can be extended to $X = \sum_{i=1}^n X_i$, where X_i 's are independent discrete log-concave. This proves a stronger version of Feige's conjecture as the optimal lower bound holds without any assumption on expectation and on the sign of the random variables. Note that this cannot be true for distributions that are not log-concave in general, this can be seen by taking, for example, a random variable X such that $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = m) = p$, with m sufficiently large, and p sufficiently close to 1. Thus, in general, the constraint on the expectation is necessary.

Inequality (2) has been established for binomial and the sum of independent Bernoulli random variables by Garnett [13]. Theorem 1.1 extends Garnett's result to the whole class of log-concave probability sequences. The following stronger inequality has been proved for Poisson distribution by Teicher [34],

$$\mathbb{P}(X \leq \mathbb{E}[X]) > e^{-1}.$$

However, Teicher's inequality does not hold for all log-concave random variables. This can be seen by taking a truncated Poisson distribution with the parameter equals to 5, supported on $\{0, 1, 2\}$ so that $\mathbb{P}(X \leq \mathbb{E}[X]) < e^{-1}$.

In the special case where $\mathbb{E}[X] \in \mathbb{Z}$, Theorem 1.1 yields the following result.

Corollary 1.2. *Let X be a discrete log-concave random variable. If $\mathbb{E}[X] \in \mathbb{Z}$, then*

$$\mathbb{P}(X \leq \mathbb{E}[X]) \geq e^{-1}. \tag{3}$$

For specific distributions, the lower bound in corollary 1.2 can be improved to $1/2$ (when $\mathbb{E}[X]$ is an integer), as shown for the sum of independent Bernoulli (see [30]), Poisson (see [34]), and can easily be verified for the discrete uniform distribution. However, this is not

true for the whole class of log-concave distributions, as can be seen by taking a random variable supported on $\{1, \dots, 8\}$ with probability mass function $p(k) = Cp^k$, $k \in \{1, \dots, 8\}$, where p is chosen so that $\mathbb{E}[X] = 6$ and C is the normalizing constant. For this distribution, $\mathbb{P}(X \leq \mathbb{E}[X]) < \frac{1}{2}$.

In the continuous setting, a result of Grünbaum [15] (see also [23, Lemma 2.4]) states that if X is a continuous log-concave random variable then the following stronger inequality holds,

$$\mathbb{P}(X \leq \mathbb{E}[X]) \geq \frac{1}{e}. \quad (4)$$

Recall that a real-valued random variable is log-concave if it has a probability density function f (with respect to Lebesgue measure) satisfying

$$f((1 - \lambda)x + \lambda y) \geq f(x)^{1-\lambda}f(y)^\lambda,$$

for all $x, y \in \mathbb{R}$ and $\lambda \in [0, 1]$. Examples include Gaussian, exponential, and uniform over an interval. See, e.g., [19], [29], [7], for properties and applications of continuous log-concave distributions. Since the sum of independent continuous log-concave random variables is log-concave (see [28]), inequality (4) applies to $X = \sum_{i=1}^n X_i$, where X_i 's are independent continuous log-concave, in particular Feige's conjecture holds for this class of random variables.

Let us note that Theorem 1.1 implies inequality (4) for continuous log-concave random variables. Indeed, suppose that X is log concave with support $[0, 1]$ and density function f . For integer $n \geq 1$, define X_n to be the discrete log-concave random variable on $\{1, \dots, n\}$ with probability mass function

$$\mathbb{P}(X_n = k) = \frac{f(k/n)}{\sum_{j=1}^n f(j/n)}.$$

One can check that X_n/n converges to X in distribution, and together with Theorem 1.1 it follows that

$$\mathbb{P}(X \leq \mathbb{E}[X]) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(\frac{X_n}{n} \leq \frac{\mathbb{E}[X_n]}{n} + \frac{1}{n}\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(X_n \leq \mathbb{E}[X_n] + 1) \geq \frac{1}{e}.$$

The proof of the result for general X follows by approximation on compact sets and rescaling.

The proof of our main theorem is postponed to the next section.

2 Proof of Theorem 1.1

The key idea of the proof of Theorem 1.1 is to reduce the problem to truncated geometric distributions. This is due to the identification of the extreme points of the convex hull of a subset of discrete log-concave probability distributions satisfying a linear constraint and the standard use of Krein-Milman's theorem, developed by the third and fourth named authors in [21] (see also [27]). For completeness, we recall the main argument.

Let $M, N \in \mathbb{Z}$. Denote by $\mathcal{P}(\{M, \dots, N\})$ the set of all discrete probability measures supported on $\{M, \dots, N\}$. Let $h: \{M, \dots, N\} \rightarrow \mathbb{R}$ be an arbitrary function. Consider $\mathcal{P}_h(\{M, \dots, N\})$ the set of all log-concave distributions \mathbb{P}_X in $\mathcal{P}(\{M, \dots, N\})$ satisfying the constraint $\mathbb{E}[h(X)] \geq 0$, that is,

$$\mathcal{P}_h(\{M, \dots, N\}) = \{\mathbb{P}_X \in \mathcal{P}(\{M, \dots, N\}) : X \text{ log-concave, } \mathbb{E}[h(X)] \geq 0\}.$$

The following theorem describes the shape of the extreme points of $\text{Conv}(\mathcal{P}_h(\{M, \dots, N\}))$.

Theorem 2.1 ([21]). *If $\mathbb{P}_X \in \text{Conv}(\mathcal{P}_h(\{M, \dots, N\}))$ is an extreme point, then its probability mass function p satisfies*

$$p(k) = Cp^k 1_{\{m, \dots, n\}}(k), \quad (5)$$

for some $C, p > 0$, $m, n \in \{M, \dots, N\}$.

One can therefore deduce by the (finite-dimensional version of the) Krein-Milman theorem that the supremum of any convex functional over the set $\mathcal{P}_h(\{M, \dots, N\})$ is attained at probability distributions of the form (5) (see [21]).

The convex (linear) functional that will be considered in the proof of Theorem 1.1 is of the form $\Phi: \mathbb{P}_X \mapsto \mathbb{P}_X(A)$ for a fixed Borel set $A \subset \mathbb{R}$.

Proof of Theorem 1.1. First, by an approximation argument, it is enough to prove inequality (2) for compactly supported log-concave random variables, say on $\{M, \dots, N\}$, for arbitrary $M \leq N$. Next, according to Theorem 2.1, it is enough to consider log-affine distributions supported in $\{M, \dots, N\}$, that is, distributions of the form

$$p(k) = Cp^k, \quad k \in \{m, \dots, n\},$$

for all $p > 0$ and $M \leq m \leq n \leq N$. Finally, if X is log-affine on $\{m, \dots, n\}$, then $\tilde{X} = X - m + 1$ is log-affine supported on $\{1, \dots, \tilde{n}\}$, where $\tilde{n} = n - m + 1$, and

$$\mathbb{P}(X \geq \mathbb{E}[X] + 1) = \mathbb{P}(\tilde{X} \geq \mathbb{E}[\tilde{X}] + 1).$$

Therefore, it is enough to prove the desired inequality for log-affine distributions supported on $\{1, \dots, n\}$, for arbitrary $n \geq 1$. In the following, we consider X with probability mass function of the form

$$p(k) = Cp^k, \quad k \in \{1, \dots, n\},$$

for $n \geq 1$ and $p > 0$, and we will establish inequality (2) for such random variables. If $p = 1$, then

$$\mathbb{P}(X < \mathbb{E}[X] + 1) \geq \frac{1}{2} \geq \frac{1}{e}.$$

When $p \neq 1$, we have

$$C = \frac{1-p}{p(1-p^n)}.$$

Therefore,

$$\mathbb{E}[X] = Cp \left[\sum_{k=0}^n p^k \right]' = \frac{np^{n+1} - (n+1)p^n + 1}{(1-p)(1-p^n)} = \frac{1}{1-p} - \frac{np^n}{1-p^n}.$$

On the other hand,

$$\mathbb{P}(X < \mathbb{E}[X] + 1) = \sum_{k=1}^{\lceil \mathbb{E}[X] \rceil} Cp^k = \frac{1-p^{\lceil \mathbb{E}[X] \rceil}}{1-p^n}.$$

Thus we are left to prove that

$$\frac{1-p^{\lceil \mathbb{E}[X] \rceil}}{1-p^n} \geq \frac{1}{e}.$$

Since

$$\frac{1-p^{\lceil \mathbb{E}[X] \rceil}}{1-p^n} \geq \frac{1-p^{\mathbb{E}[X]}}{1-p^n},$$

it suffices to show that for all positive integer n and every positive real number $p \neq 1$,

$$\frac{1 - p^{\frac{1}{1-p} - \frac{np^n}{1-p^n}}}{1 - p^n} \geq \frac{1}{e}.$$

Substituting $x = p^n$, the above inequality is equivalent to

$$\frac{1 - x^{\frac{1}{n(1-x^{1/n})} - \frac{x}{1-x}}}{1 - x} \geq \frac{1}{e}.$$

Note that the left hand side is non-increasing in n since

$$\frac{d}{dt} \left[t(1 - x^{1/t}) \right] = 1 - x^{1/t} + x^{1/t} \log(x^{1/t}) \geq 0,$$

where the substitution $y = x^{1/t}$ makes the inequality obvious. Taking the limit in n , it suffices to prove that for $x \neq 1$,

$$g(x) := \frac{1 - x^{-\frac{1}{\log(x)} - \frac{x}{1-x}}}{1 - x} \geq \frac{1}{e}.$$

Note that $\lim_{x \rightarrow +\infty} g(x) = \frac{1}{e}$, so the result would follow from

$$0 \geq g'(x) = \frac{e(x-1) - x^{\frac{x}{x-1}} \log(x)}{e(x-1)^3}.$$

To prove that $g'(x) \leq 0$, it suffices to prove that

$$h(x) := \frac{x^{\frac{x}{x-1}} \log(x)}{e(x-1)} \geq 1.$$

Computing,

$$h'(x) = \frac{x^{\frac{1}{x-1}} \left((x-1)^2 - x \log(x)^2 \right)}{e(x-1)^3},$$

the result will follow from $k(x) := (x-1)^2 - x \log(x)^2 \geq 0$ as this will give $h'(x) \leq 0$ for $x < 1$ and $h'(x) \geq 0$ for $x > 1$, so that $h(x) \geq h(1) = 1$. Since k is convex, as $k''(x) = \frac{2}{x}(x-1 - \log(x)) \geq 0$, and $k'(1) = 0$, we deduce that k takes its minimum value 0 when $x = 1$, completing the proof. \square

The next remark comments on the sharpness of Theorem 1.1.

Remark 2.2. For a discrete log-concave random variable X , one trivially has the extension to $t \geq 1$,

$$\mathbb{P}(X \geq \mathbb{E}[X] + t) \leq \mathbb{P}(X \geq \mathbb{E}[X] + 1) \leq 1 - \frac{1}{e}.$$

Let us see that these inequalities are sharp for any $t \geq 1$. Fix $n \geq 2$ and let X be log-affine on $\{1, \dots, n\}$ with parameter $p = n^{\frac{1}{n}}$. Following the computation of the proof of Theorem 1.1, one has

$$\frac{n - n^{\frac{\mathbb{E}[X]+t-1}{n}}}{n-1} \leq \mathbb{P}(X \geq \mathbb{E}[X] + t) \leq \frac{n - n^{\frac{\mathbb{E}[X]+t}{n}}}{n-1}.$$

Thus to prove that $1 - \frac{1}{e}$ is optimal among log-concave distributions it is enough to prove that for $\lambda \geq 0$,

$$\lim_{n \rightarrow +\infty} \frac{n - n^{\frac{\mathbb{E}[X] + \lambda}{n}}}{n - 1} = 1 - \frac{1}{e}.$$

The result follows since

$$\frac{n}{n-1} - \frac{1}{n^{\frac{n+1}{n}} - n} = 1 - \frac{1}{\log(n)} + O\left(\frac{1}{n}\right),$$

therefore, using $\mathbb{E}[X] = \frac{n^2}{n-1} - \frac{1}{n^{1/n-1}}$, we have

$$\frac{n - n^{\frac{\mathbb{E}[X] + \lambda}{n}}}{n - 1} = \frac{n}{n-1} \left(1 - n^{\frac{\lambda}{n}} n^{-\frac{1}{\log(n)} + O\left(\frac{1}{n}\right)}\right) \xrightarrow{n \rightarrow +\infty} 1 - \frac{1}{e}.$$

References

- [1] Anari, N., Liu K., Gharan, S.O., Vintant, C. *Log-Concave Polynomials III: Mason's Ultra-Log-Concavity Conjecture for Independent Sets of Matroids*. Preprint. arXiv:1811.01600.
- [2] Arieli, I., Babichenko, Y., Peretz, R., Young, H. P. *The speed of innovation diffusion in social networks*. *Econometrica*, 88(2), 569-594, 2020.
- [3] Alon, N., Huang, H., Sudakov, B. *Nonnegative k-sums, fractional covers, and probability of small deviations*. *Journal of Combinatorial Theory, Series B*, 102(3), 784-796, 2012.
- [4] Aravinda, H., Marsiglietti, A., Melbourne, J. *Concentration inequalities for ultra log-concave distributions*. *Studia Mathematica*, 2022. doi: 10.4064/sm210605-2-10.
- [5] Brenti, F. *Log-concave and unimodal sequences in algebra, combinatorics, and geometry: an update*. In *Jerusalem combinatorics '93*, volume 178 of *Contemp. Math.* pages 71–89. Amer. Math. Soc., Providence, RI, 1994.
- [6] Brändén, P. *Unimodality, log-concavity, real-rootedness and beyond*. *Handbook of enumerative combinatorics, Discrete Math. Appl.*, CRC Press:437–483, 2015.
- [7] Colesanti, A. *Log-concave functions, Convexity and concentration*. 487-524, *IMA Vol. Math. Appl.*, 161, Springer, New York, 2017.
- [8] Feige, U. *On sums of independent random variables with unbounded variance and estimating the average degree in a graph*. *SIAM Journal on Computing*, 35(4), 964-984, 2006.
- [9] Fekete, M., Pólya, G. *Über ein problem von Laguerre*. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 34(1):89–120, 1912.
- [10] Fradelizi, M., Guédon, O. *The extreme points of subsets of s-concave probabilities and a geometric localization theorem*. *Discrete and Computational Geometry* 31 (2004), 327-335.
- [11] Frankl, P., Kupavskii, A. *The Erdős Matching Conjecture and concentration inequalities*. arXiv preprint arXiv:1806.08855.

- [12] Garnett, B. *Small deviations of sums of independent random variables*. Journal of Combinatorial Theory, Series A, 169, 105119, 2020.
- [13] Garnett, B. *Small deviations of sums of random variables (Doctoral dissertation, Rutgers University-Graduate School-New Brunswick)*, 2016.
- [14] Gonen, M., Ron, D., Shavitt, Y. *Counting stars and other small subgraphs in sublinear-time*. SIAM Journal on Discrete Mathematics, 25(3), 1365-1411, 2011.
- [15] Grünbaum, B. *Partitions of mass-distributions and of convex bodies by hyperplanes*. Pacific J. Math. 10 (1960), 1257-1261.
- [16] Guo, J., He, S., Ling, Z., Liu, Y. *Bounding probability of small deviation on sum of independent random variables: Combination of moment approach and Berry-Esseen theorem*. arXiv preprint arXiv:2003.03197, 2020.
- [17] He, S., Zhang, J., Zhang, S. *Bounding probability of small deviation: A fourth moment approach*. Mathematics of Operations Research, 35(1), 208-232, 2010.
- [18] Johnson, O., T. *A discrete log-Sobolev inequality under a Bakry-Émery type condition*. Annales de l'Institut Henri Poincaré B (Probability and Statistics), vol 53/4, 2017, 1952-1970.
- [19] Klartag, B., Milman, V. D. *Geometry of log-concave functions and measures*. Geom. Dedicata 112 (2005) 169-182.
- [20] Madiman, M., Melbourne, J., Roberto, C. *Bernoulli sums and Rényi entropy inequalities*. Preprint, arXiv:2103.00896.
- [21] Marsiglietti, A., Melbourne, J. *Geometric and functional inequalities for log-concave probability sequences*. Preprint, arXiv:2004.12005.
- [22] Marsiglietti, A., Melbourne, J. *Moments, Concentration, and Entropy of Log-Concave Distributions*. Preprint, arXiv:2205.08293.
- [23] Marsiglietti, A., Pandey, P. *On the equivalence of statistical distances for isotropic convex measures*. Preprint, arXiv:2112.09009.
- [24] Mason, J. H. *Matroids: unimodal conjectures and Motzkin's theorem*. In Combinatorics (Proc. Conf. Combinatorial Math., Math. Inst., Oxford, 1972), pages 207-220. Inst. Math. Appl., Southend, 1972.
- [25] Melbourne, J., Palafox-Castillo, A. *Discrete complement of Lyapunov's inequality and its information theoretic consequences*. Preprint, arXiv:2111.06997.
- [26] Melbourne, J., Tkocz, T. *Reversals of Rényi Entropy Inequalities under Log-Concavity*. IEEE Trans. Inform. Theory 67 (2021), no. 1, 45-51.
- [27] Nayar, P., Słobodianiuk, S. *Ultra-log-concavity and discrete degrees of freedom*. Preprint. arXiv:2205.04069.
- [28] Prékopa, A. *On logarithmic concave measures and functions*. Acta Sci. Math. (Szeged) 34 (1973), 335-343.

- [29] Saumard, A, Wellner, J. A. *Log-concavity and strong log-concavity: A review*. Statist. Surv. 8: 45-114 (2014).
- [30] Siegel, A. *Median bounds and their application*. J. Algorithms, 38, pp. 184–236. 2001.
- [31] Stanley, R. P. *Two combinatorial applications of the Aleksandrov-Fenchel inequalities*. Journal of Combinatorial Theory, Series A, 31(1):56–65, 1981.
- [32] Stanley, R. P. *Positivity problems and conjectures in algebraic combinatorics*. Mathematics: frontiers and perspectives, 295:319, 2000.
- [33] Stanley, R. P. *Log-concave and unimodal sequences in algebra, combinatorics, and geometry*. In *Graph theory and its applications: East and West* (Jinan, 1986), volume 576 of Ann. New York Acad. Sci., pages 500–535. New York Acad. Sci., New York, 1989.
- [34] Teicher, H. *An inequality on Poisson probabilities*. The Annals of Mathematical Statistics, 26(1), 147-149, 1955.
- [35] Vondrák, J. *Probabilistic methods in combinatorial and stochastic optimization (Doctoral dissertation)*, Massachusetts Institute of Technology), 2005.

Abdulmajeed Alqasem
 Department of Mathematics
 University of Florida
 Gainesville, FL 32611, USA
 aalqasem@ufl.edu

Heshan Aravinda
 Department of Mathematics
 University of Florida
 Gainesville, FL 32611, USA
 heshanaravinda.p@ufl.edu

Arnaud Marsiglietti
 Department of Mathematics
 University of Florida
 Gainesville, FL 32611, USA
 a.marsiglietti@ufl.edu

James Melbourne
Probabilidad y Estadística
Centro de Investigación en Matemáticas (CIMAT)
Guanajuato, Gto 36023, México
james.melbourne@cimat.mx