A New Probabilistic Anti-Concentration Inequality For Independent Non-Negative Random Variables

Metin Duerr (Bozkurt)[*] [†]

May 29, 2025

Abstract

Let $X_1, ..., X_n$ be arbitrary non-negative independent random variables. We prove the anti-concentration inequality

$$\mathbb{P}\left(f\left(X\right) < \mu + \delta\right) \geq \min\left\{\frac{\delta}{\mu+\delta}\,,\, \exp\left(\frac{-m}{c}\right)\right\},$$

where $f(X)$ is a combination of a sum and a product of the random variables X_i (possibly only the sum or only the product), and μ is the expected value of $f(X)$. The expected value μ may possibly be bounded by a finite constant m and δ may possibly depend on n, which determines c. In particular, for $f(X) = \sum_{i=1}^{n} X_i$ and $\mu = 1$ we prove a generalized version of an intriguing conjecture by Feige [3]. We show by a simple example how this inequality finds use in mathematical finance.

1 Introduction

Most probabilistic inequalities, e.g., Markov's inequality, bound the probability mass above a given constant of a given random variable X. Moreover, additional conditions might be imposed, e.g., the existence of the variance of X. In contrast, our inequality gives a lower bound on the probability mass of X below a certain threshold and does not assume that the higher moments of X exist.

This research was motivated by a conjecture of Feige [3]. For arbitrary non-negative independent random variables $X_1, ..., X_n$ with the respective expected values μ_i at most one and any $\delta > 0$ the conjecture states

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i < \mu + \delta\right) \geq \min\left\{\frac{\delta}{1+\delta}\,,\, \lambda\right\}, \tag{1}$$

where $\lambda = \exp(-1)$. In [3] Feige proved a bound of approximately $\lambda = 0.0769$ by applying an algorithm, which reduces the support of a random variable to a discrete set with two elements, thereby enabling a case analysis. We use a similiar idea for which we construct a one-step algorithm described in section 4.1.

Feige's bound was further improved by Garnett [4] to $\lambda = 0.14$ by considering the first four

[*]My legal name is Bozkurt, however, due to the political associations with this name I prefer to be referenced by the surname of my future wife, i.e., Duerr.

[†]metin.duerr.research@proton.me

moments. The best known bound the author was able to find is $\lambda = 0.1798$ due to Guo et al. [5], who use an optimization approach paired with the Berry-Essen Theorem.
Moreover, we want to note that Feige's conjecture was proven true for discrete log-concave distributions by Alqasem et al. [1] and for identically distributed random variables by Egozcue et al. [2].
We prove our inequality by an analytical approach. To this end, we study minimizers of random variables and arrive at a first bound of

$$\frac{\delta}{\mu+\delta}. \tag{2}$$

Next, we consider the highest possible lower bound for a combination of a sum and a product of arbitrary non-negative independent random variables $X_1, ..., X_n$ and obtain the bound

$$\exp\left(\frac{-m}{c}\right) \tag{3}$$

under some conditions on μ and δ.

2 Notation

Throughout this paper we will denote by $\widehat{X}, X, X_1, ..., X_n$ arbitrary non-negative random variables and write r.v. as a shorthand. Moreover, we write $\widehat{\mu}, \mu, \mu_1, ..., \mu_n$ for their respective expected values. We always assume that the random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. When we write $f(X)$, we mean a combination of a sum and a product of the random variables X_i, i.e., for an indexing $\mathbf{I}$ and $\mathbf{J}$ of random variables we define

$$f(X) := \sum_{i\in\mathbf{I}} X_i + \prod_{j\in\mathbf{J}} X_j. \tag{4}$$

Lastly, by

$$\lim_{\mathbf{I}\cup\mathbf{J}\to\infty} \tag{5}$$

we mean that the set $\mathbf{I}\cup\mathbf{J}$ grows arbitrarily large.

3 Main Results

Consider a r.v. X with expected value $\mu \geq 0$ and a given $\delta \geq 0$. Since the object of our study is the probability

$$\mathbb{P}\left(X < \mu+\delta\right) \tag{6}$$

we may consider a transformed, discrete version $\widehat{X}$ of the random variable X, such that,

$$\mathbb{P}\left(X < \mu+\delta\right) = \mathbb{P}\left(\widehat{X} < \mu+\delta\right), \tag{7}$$

and the expected values are the same, i.e.,

$$\mathbb{E}(X) = \mathbb{E}(\widehat{X}) = \mu. \tag{8}$$

If we furthermore impose that $\widehat{X}$ has exactly two realisations $0 \leq \alpha < \mu+\delta$ and $\beta \geq \mu+\delta$, then we have the advantage that the r.v. $\widehat{X}$ is easier to study and all statements about its probability mass concerning the bound $\mu+\delta$ hold true for the probability mass of X as well. We obtain such a transformation via a straightforward algorithm described in section 4.1, which leads to

the key Corollary 4.4. By this Corollary we find our first lower bound for the probability 6, which is Proposition 4.7.
Next, we study under which conditions on μ and δ this lower bound converges to a value between zero and one, excluding those two numbers. With the help of Lemma 4.10 we thus find our next lower bound, which is Proposition 4.11. Moreover, we prove Feige's Conjecture in Corollary 4.12. Proposition 4.7 and Proposition 4.11 prove our main Theorem.

Theorem 3.1. *Let $(X_k)_{k\in\mathbb{N}}$ be a sequence of arbitrary non-negative random variables and let $\delta > 0$. Then we have*

$$\mathbb{P}\left(\lim_{\mathbf{I}\cup\mathbf{J}\to\infty}\sum_{i\in\mathbf{I}}X_i+\prod_{j\in\mathbf{J}}X_j<\mu+\delta\right)\geq\frac{\delta}{\mu+\delta}. \tag{9}$$

Moreover, if there exists a constant m such that $\mu \leq m$ for all n and $\delta = cn + s > 0$ for some values c and s, then we have

$$\mathbb{P}\left(\lim_{\mathbf{I}\cup\mathbf{J}\to\infty}\sum_{i\in\mathbf{I}}X_i+\prod_{j\in\mathbf{J}}X_j<\mu+\delta\right)\geq\min\left\{\frac{\delta}{\mu+\delta}\,,\,\exp\left(\frac{-m}{c}\right)\right\}. \tag{10}$$

4 Bounds And Proofs

4.1 The First Bound

Let X be a r.v. with expected value $\mu > 0$, and let $\delta \geq 0$. The assumption $\mu > 0$ is not a restriction, since for $\mu = 0$ it is the case that all the probability mass is at 0, i.e.,

$$\mathbb{P}(X<\mu+\delta)=1. \tag{11}$$

We know that the variable X has at least one realisation below μ, which may be zero, and at least one realisation above μ. Otherwise, there would again be nothing to prove. Thus, for every $\delta \geq 0$ there exists a probability $p \in (0, 1]$ such that

$$\mathbb{P}\left(X<\mu+\delta\right)=p. \tag{12}$$

The following one-step algorithm describes how to obtain a discrete r.v. $\widehat{X}$ from X, such that the equation 12 is true for $\widehat{X}$ and $\mathbb{E}(\widehat{X}) = \mu$.

Step 1:
If X already is a discrete r.v. with two realisations, return X.
Else, to find two realisations α, β of $\widehat{X}$, such that

$$\alpha\hat{p}+\beta(1-\hat{p})=\mu, \tag{13}$$

we regard β as a function of α and $\hat{p}$ and re-write the equation 13 to read

$$\beta(\alpha,\hat{p})=\frac{\mu-\alpha\hat{p}}{1-\hat{p}}. \tag{14}$$

For the equation 12 to hold true, we choose $\hat{p} = p$ and $0 \leq \alpha < \mu + \delta$. Note that we have

$$\beta(\alpha,\hat{p})\geq\mu+\delta\quad\text{if and only if}\quad\alpha\leq\mu+\delta-\frac{\delta}{p}\leq\mu+\delta. \tag{15}$$

Return

$$\widehat{X} = \begin{cases} \alpha, & p, \\ \frac{\mu-\alpha p}{1-p}, & 1-p. \end{cases} \tag{16}$$

End of Step 1.

Note that α is a realisation of $\widehat{X}$, which we can choose at will as long as $0 \leq \alpha < \mu + \delta$. Thus, for any r.v. X its transformed variable $\widehat{X}$ is not unique by this algorithm. This leads us to the following Definition and Lemmas.

Definition 4.1. Let X denote an arbitrary non-negative random variable with expected value $\mu > 0$ and let $\delta \geq 0$. Let these three statements be conditions

$$\mathbb{P}(X < \mu + \delta) = p, \tag{CI}$$

$$\mathbb{E}(X) = \mu, \tag{CII}$$

$$X \text{ is a r.v. with two realisations } \; 0 \leq \alpha < \mu + \delta \; \text{ and } \; \beta \geq \mu + \delta. \tag{CIII}$$

We define the set ${}_{1,2}\Gamma_p^{\mu,\delta}$ to be

$${}_{1,2}\Gamma_p^{\mu,\delta} := \{X \mid X \text{ fulfills conditions (CI) and (CII)}\}. \tag{17}$$

We define the set ${}_{1,2,3}\Gamma_p^{\mu,\delta}$ to be

$${}_{1,2,3}\Gamma_p^{\mu,\delta} := \{X \mid X \text{ fulfills conditions (CI) and (CII) and (CIII)}\}. \tag{18}$$

Lemma 4.2. *Let X denote an arbitrary non-negative random variable with expected value $\mu > 0$ and let $\delta \geq 0$. Then there exists a mapping*

$$\widetilde{T} : {}_{1,2}\Gamma_p^{\mu,\delta} \to {}_{1,2,3}\Gamma_p^{\mu,\delta}. \tag{19}$$

Proof. The algorithm described in the beginning of this section guarantees that such a mapping $\widetilde{T}$ exists. Moreover, neither of the two sets ${}_{1,2,3}\Gamma_p^{\mu,\delta}$ and ${}_{1,2}\Gamma_p^{\mu,\delta}$ is empty. □

Lemma 4.3. *Let*

$$\widetilde{T} : {}_{1,2}\Gamma_p^{\mu,\delta} \to {}_{1,2,3}\Gamma_p^{\mu,\delta} \tag{20}$$

be a mapping and let $\widehat{X} \in {}_{1,2,3}\Gamma_p^{\mu,\delta}$. Then there exists a mapping

$$T : {}_{1,2}\Gamma_p^{\mu,\delta} \to \widehat{X}. \tag{21}$$

Proof. Let $\widetilde{T}$ be defined by the algorithm described in the beginning of this section. Restricting the image of $\widetilde{T}$ to be the element $\widehat{X}$ gives us the mapping T, which can be seen as fixing α in the algorithm before defining the mapping $\widetilde{T}$. □

Corollary 4.4. *Let X be a r.v. with $X \in {}_{1,2}\Gamma_p^{\mu,\delta}$. Then there exists a mapping*

$$\mathcal{T}^{-1} : \widehat{X} \to X, \tag{22}$$

where $\widehat{X} \in {}_{1,2,3}\Gamma_p^{\mu,\delta}$.

Proof. This in the inverse of the mapping T in Lemma 4.3 restricted to X, i.e.,

$$\mathcal{T}^{-1} = (T|_X)^{-1} : \widehat{X} \to X. \tag{23}$$

□

By Corollary 4.4 it suffices to find a r.v. $\widehat{X} \in {}_{1,2,3}\Gamma_p^{\mu,\delta}$ with the lowest possible value p to bound every r.v. $X \in {}_{1,2,3}\Gamma_p^{\mu,\delta}$, i.e., to find a *minimizer.*

Definition 4.5. Let X be a r.v. with $X \in {}_{1,2}\Gamma_p^{\mu,\delta}$ and let $\mathcal{T}$ be a mapping

$$\mathcal{T} : X \to \widehat{X}, \tag{24}$$

If for every other mapping

$$\mathcal{H} : X \to \widehat{X} \tag{25}$$

we have

$$\mathbb{P}(\mathcal{T}(X) < \mu + \delta) \leq \mathbb{P}(\mathcal{H}(X) < \mu + \delta), \tag{26}$$

we say that $\mathcal{T}$ is a *minimizer* of X.

Remark 4.6. We have stated and proved Corollary 4.4 only for $\mathcal{T}$, however, we may go through this section again and convince ourselves that we can define mappings of the form

$$\widetilde{T} : {}_{1,2}\Gamma_p^{\mu,\delta} \to {}_{1,2,3}\Gamma_q^{\mu,\delta}. \tag{27}$$

with $q < p$ and all the Lemmas and Corollary 4.4 hold true analogously.

We now have all the tools gathered to prove our first lower bound.

Proposition 4.7. *Let X denote any arbitrary non-negative random variable with expected value $\mu > 0$ and let $\delta > 0$. We have the bound*

$$\mathbb{P}(X < \mu + \delta) \geq \frac{\delta}{\mu + \delta}. \tag{28}$$

Proof. Let $\mathcal{T}$ be a transformation of X, i.e., $\mathcal{T}(X) = \widehat{X}$. For $\mathcal{T}$ to be minimizer of X we need the probability p to be the smallest possible value, i.e., by condition CII and condition CIII we need to minimize

$$p(\alpha, \beta) = \frac{\beta - \mu}{\beta - \alpha}. \tag{29}$$

Note that $p(\alpha, \beta)$ is a strictly increasing function in α and β, since

$$\begin{aligned} \frac{\partial}{\partial \alpha} \frac{\beta - \mu}{\beta - \alpha} &= \frac{\beta - \mu}{(\beta - \alpha)^2} > 0, \\ \frac{\partial}{\partial \beta} \frac{\beta - \mu}{\beta - \alpha} &= \frac{1}{(\beta - \alpha)^3} > 0. \end{aligned} \tag{30}$$

Thus, p is minimized if $\alpha = 0$ and $\beta = \mu + \delta$, leading us to

$$\mathbb{P}(X < \mu + \delta) \geq \mathbb{P}(\mathcal{T}(X) < \mu + \delta) = \frac{\delta}{\mu + \delta}. \tag{31}$$

This concludes our proof. ∎

Remark 4.8. Notice that we may introduce a condition on δ, i.e., for

$$\mathbb{P}(\alpha) = \frac{\delta}{\mu + \delta} \leq 1 - \frac{\delta}{\mu + \delta} = \frac{\mu}{\mu + \delta} = \mathbb{P}(\beta) \tag{32}$$

to be true, we must have $\delta \leq \mu$. Thus, for $\delta > \mu$ we introduce a random variable

$$\widetilde{X} = \begin{cases} 0, & \frac{\mu}{\mu+\delta}, \\ \mu + \delta, & \frac{\delta}{\mu+\delta}. \end{cases} \tag{33}$$

Then we see that

$$\mathbb{P}(\mathcal{T}(X) < \mu + \delta) \geq \mathbb{P}(\widetilde{X} < \mu + \delta). \tag{34}$$

Proposition 4.7 might then be written as

$$\mathbb{P}(X < \mu + \delta) \geq \min \left\{ \frac{\delta}{\mu + \delta}, \frac{\mu}{\mu + \delta} \right\}. \tag{35}$$

4.2 The Second Bound

So far we were only concerned with one r.v. X. In this section we will consider sums, products and the mixture of sums and products of random variables. To begin, let $(X_k)_{k \in \mathbb{N}}$ be any sequence of random variables. We seek to bound the sum

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i < \mu + \delta \right), \tag{36}$$

where $\delta > 0$ and $\mu = \sum_{i=0}^{n} \mu_i$ by the independence of the random variables. Note that by equation 12

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i < \mu + \delta \right) = p > 0. \tag{37}$$

Since the sum of random variables is again a random variable we can bound p by Proposition 4.7

$$p \geq \left[\frac{\delta}{\mu + \delta} \right]^n. \tag{38}$$

To expound some more, the lowest probability a sum of random variables can take is one realisation of the vector $X = (X_1, ..., X_n)$, where each element X_i takes on the lowest possible probability p_i given by Proposition 4.7. This reasoning holds true for products and mixtures of sums and products of random variables as well, giving us the next Lemma.

Lemma 4.9. *Let $(X_k)_{k \in \mathbb{N}}$ be a sequence of random variables. Then we have*

$$\mathbb{P}\left(\sum_{i \in \mathbf{I}} X_i + \prod_{j \in \mathbf{J}} X_j \right) \geq \left[\frac{\delta}{\mu + \delta} \right]^n, \tag{39}$$

where $\mathbf{I}$ and $\mathbf{J}$ is any indexing of the random variables with $|\mathbf{I}| + |\mathbf{J}| = n$.

Suppose that in Lemma 4.9 the expected value μ grows arbitrarily large or δ does not depend in some fashion on n. Then the Lemma 4.9 gives us a trivial bound of zero. The next Lemma gives us a condition on δ with which we find a bound away from zero.

Lemma 4.10. *Let $f(n) > cn + s$ and $g(n) < cn + s$ be a functions dependent on n, where $g(n)$ is non-linear. Moreover, let $\mu \leq m$ for some constant m. Then we have*

(a) $\left(\frac{f(n)}{\mu + f(n)} \right)^n \geq \exp\left(\frac{-m}{c} \right)$ *for large n.*

(b) $\left(\frac{g(n)}{\mu+g(n)}\right)^n$ *tends to zero for large* n.

Proof. (a) Since $f(n) > cn + s$, we have

$$\left(\frac{f(n)}{\mu + f(n)}\right)^n \geq \left(\frac{cn+s}{\mu + cn + s}\right)^n \geq \left(\frac{cn+s}{m + cn + s}\right)^n. \tag{40}$$

The right hand side of equation 40 tends to $\exp\left(\frac{-m}{c}\right)$ for large n.
(b) We may write $g(n) = cn^\kappa + s$ for some $\kappa < 1$. Thus

$$\left(\frac{g(n)}{\mu + g(n)}\right)^n = \left(\frac{cn^\kappa + s}{cn^\kappa + s + \mu}\right)^n. \tag{41}$$

The right hand side of equation 41 tends to zero for large n.

□

Let us now achieve our second bound by considering growing n.

Proposition 4.11. *Let* $(X_k)_{k\in\mathbb{N}}$ *be a sequence of random variables and let* $\delta = cn + s > 0$ *for some values* c *and* s. *Moreover, let* $\mu \leq m$ *for all* n. *Then we have*

$$\mathbb{P}\left(\lim_{\mathbf{I}\cup\mathbf{J}\to\infty} \sum_{i\in\mathbf{I}} X_i + \prod_{j\in\mathbf{J}} X_j < \mu + \delta\right) \geq \exp\left(\frac{-m}{c}\right). \tag{42}$$

Proof. By Lemma 4.9 it suffices to study the growth of

$$\left[\frac{\delta}{\mu+\delta}\right]^n \tag{43}$$

for large n. We have

$$\lim_{n\to\infty}\left[\frac{cn+s}{\mu+cn+s}\right]^n \geq \lim_{n\to\infty}\left[\frac{cn+s}{m+cn+s}\right]^n = \exp\left(\frac{-m}{c}\right). \tag{44}$$

□

Corollary 4.12. *Let* $X_1, ..., X_n$ *be arbitrary non-negative independent random variables, with expectations* $\mu_1, ..., \mu_n$ *respectively, where* $\mu_i \leq 1$ *for every* i. *Let* $X = \sum_{i=1}^n X_i$, *and let* μ *denote the expectation of* X. *Then for every* $\delta > 0$

$$\mathbb{P}(X < \mu + \delta) \geq \min\left\{\frac{\delta}{1+\delta}, \exp\left(\frac{-1}{c}\right)\right\}. \tag{45}$$

Proof. We set $m = 1$ in Proposition 4.11 and also use Proposition 4.7. □

4.3 Proof Of Our Main Theorem

Proof. Proposition 4.7 and Proposition 4.11 put together prove our Theorem 3.1. □

5 An Example In Financial Mathematics

Suppose we have invested into n stocks of type A at time $t_0 = 0$. At some future time $t_1 > t_0$ the value of each individual stock is higher than at time t_0. From the time t_1 onward let $X_k(t)$ describe the profit we would gain if we sold the stock k at time $t \geq t_1$, where we immediately sell the stock once the profit drops to zero, i.e., the random variables $X_k(t)$ will never be negative, excluding transaction costs. We know from historical data, that a portfolio of type A stocks has an expected value of 1.05 units of currency per invested currency for the time $T > t_1$. Thus, we may ask ourselves, what the probability is that our portfolio will outperform this expected value by an amount of $\delta > 0$ at time T. This probability is given by

$$1 - \min\left\{\frac{\delta}{1.05 + \delta}, \exp\left(\frac{-1.05}{c}\right)\right\} \geq \mathbb{P}\left(\sum_{i=1}^{n} X_i(T) > 1.05 + \delta\right). \tag{46}$$

If we choose a δ of 1.1 independent of c (or set c to be infinite), our excess profit probability, so to speak, is bounded by

$$0.51163 \geq \mathbb{P}\left(\sum_{i=1}^{n} X_i(T) > 2.15\right). \tag{47}$$

6 Conclusion

In this paper we have proved an anti-concentration inequality, which is similiar in nature to a conjecture by Samuels [6].

Conjecture 6.1. *Let $X_1, ..., X_n$ be arbitrary non-negative independent random variables, with expectations $\mu_1, ..., \mu_n$ respectively, where $\mu_1 \geq ... \geq \mu_n$. Then for every $\lambda > \sum_{j=1}^{n} \mu_j$ there is some $1 \leq i \leq n$ such that $\mathbb{P}\left(\sum_{j=1}^{n} X_j < \lambda\right)$ is minimized when the random variables X_j are distributed as follows:*

- *For $j > i$, $X_j = \mu_j$ with probability 1.*
- *For $j \leq i$, $X_j = \lambda - \sum_{k=i+1}^{n} \mu_k$ with probability $\frac{\mu_j}{\lambda - \sum_{k=i+1}^{n} \mu_k}$, and $X_j = 0$ otherwise.*

Feige in his paper [3] discusses this conjecture and how it relates to his own conjecture. The author of this paper has not thought about the relationships of Conjecture 6.1 and Theorem 3.1. Moreover, the author is interested to see whether a theorem of similiar nature to the one in this paper can be developed for dependent random variables and whether there can be conditions put on random variables which may have negative values and still allow lower bounds. The author has not spent any time pondering these questions.

References

[1] A. Alqasem, H. Aravinda, A. Marsiglietti, and J. Melbourne. On a Conjecture of Feige for Discrete Log-Concave Distributions. *SIAM Journal on Discrete Mathematics*, 38(1):93–102, 2024.

[2] M. Egozcue and L. F. García. A simple proof of Feige's conjecture for identically distributed random variables. *https://ssrn.com/abstract=5188125*, 2025.

[3] U. Feige. 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.

[4] B. Garnett. Small Deviations of Sums of Independent Random Variables. *Journal of Combinatorial Theory, Series A*, 169(3), 2018.

[5] J. Guo, S. He, Z. Ling, and Y. Liu. Bounding probability of small deviation on sum of independent random variables: Combination of moment approach and Berry-Essen theorem, 2020.

[6] S. M. Samuels. On a Chebyshev-type inequality for sums of independent random variables. *Ann. Math. Statist.*, 37:248–259, 1966.