Estimating Sums of Independent Random Variables

Paweł Burzyński

under the direction of
Mr. Chiheon Kim
Department of Mathematics
Massachusetts Institute of Technology

Research Science Institute
August 2, 2016

Abstract

The paper deals with a problem proposed by Uriel Feige in 2005: if $X_1, \ldots, X_n$ is a set of independent nonnegative random variables with expectations equal to 1, is it true that $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) \geqslant \frac{1}{e}$? He proved that $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) \geqslant \frac{1}{13}$. In this paper we prove that infimum of the $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right)$ can be achieved when all random variables have only two possible values, and one of them is 0. We also give a partial solution to the case when all random variables are equally distributed. We prove that the inequality holds when n goes to infinity and provide numerical evidence that the probability decreases when n increases.

Summary

This paper deals with a problem in probability theory that is applicable to analysis of the stock exchange. To illustrate the main idea with a simple example, imagine the following situation. You buy a random stock for one dollar, and then sell it the following day. You can repeat the whole procedure as many times as you wish, but the only information you have is that the expected price of your stock is one dollar. Your goal is to earn at least one dollar, and you want to find the minimum chance that you fail. In fact, an Israeli computer scientist Uriel Feige, asked this question in 2005. Numerical evidence shows that the chance of failure is at least about 37 % . We show that the minimum possible chance of failure, can be achieved when there are at most two possible scenarios every day. We also give the partial solution in the case, where we additionally assume that the stock exchange behaves identically every day.

1 Introduction

1.1 Historical outline

The theory of probability emerged as a result of a conversation between famous French mathematician Blaise Pascal and a friend of his, who was interested in gaming and gambling. He asked the following question involving a popular dice game.

Example 1. *Imagine that you throw a pair of dice 24 times. Should you bet even money on the occurrence of at least one double six ?*

This problem led to a correspondence between Blaise Pascal and Pierre de Fermat, which resulted in the formulation of first principles of probability theory. Because of people's interest in gambling games, probability theory soon became popular. In 1812, Pierre de Laplace introduced new probabilistic ideas, and applied them to many scientific and practical problems in his book *Théorie Analytique des Probabilités* [1]. Wide applications of probability theory to statistics, genetics, psychology, economics, and engineering led to its rapid development. Some of the most important mathematicians who contributed to this development were Chebyshev, Markov, and von Mises. The most challenging problem in probability theory was to develop a definition of probability precise enough for use in mathematics and comprehensive enough to be applicable to a wide range of phenomena. This problem was solved by the Russian mathematician Andrey Nikolaevich Kolmogorov, who formulated axioms of the modern probability theory [2]. This paper deals with one of the fundamental issues in probability theory, which is a random variable.

1.2 Motivation and background

A *random variable* is associated with a set of its possible values, and the probabilities related to them. If we denote by X the random variable taking value x_1 with probability p_1, value

x_2 with probability p_2, and so on, up to value x_k with probability p_k, then the *mathematical expectation* of X, denoted by $\mathbb{E}[X]$, is defined as

$$\mathbb{E}[X] = \sum_{i=1}^{k} x_i p_i.$$

Let us illustrate this definition in the case of a die.

Example 2. *A roll of a die is a random variable, whose possible values are integers from* 1 *to* 6*, each with probability* $\frac{1}{6}$*. Thus, the expectation of* X *is*

$$\mathbb{E}[X] = \sum_{i=1}^{6} \frac{1}{6} i = \frac{21}{6} = 3.5.$$

This paper deals with sums of independent random variables. Our purpose is to bound the probability that the sum of values of n independent random variables is less than the sum of their expectation increased by 1. Let us return to rolling a die.

Example 3. *Imagine that you roll two dice. We denote them as random variables* X_1 *and* X_2*. Notice that the sum of their expectations is equal to 7. The probability that the sum of numbers on two dice is less then 8 is*

$$\frac{1+2+3+4+5+6}{6 \cdot 6} = \frac{21}{36} = \frac{7}{12}.$$

Of course, the random variables can be different from each other (for example rolling a die and throwing a coin). Moreover, for a given random variable X, its values may be very different from its expectation, and the probability that value of X exceeds its expectation may be arbitrarily close to 1. The distribution of probability of the random variables may also depend on their number. In this paper, we only assume that their expected values are equal to 1 (notice that a die does not follow this assumption).

Let us illustrate the main idea with an example of the stock exchange.

Example 4. *Imagine you buy a random stock for one dollar, and then sell it the following day. You can repeat the whole procedure as many times as you wish, but the only information you have is that the expected price of your stock is one dollar. Your goal is to earn at least one dollar, and you would like to know the least possible chance of failure.*

In this example, we treat a price of your stock as a random variable. Suppose, you repeat the procedure n times. One possible scenario is that every day your stock is worth $n+1$ dollars, with probability $\frac{1}{n+1}$, and nothing otherwise. You fail if and only if you lose every single day. Thus, the probability that you fail is equal to $(1-\frac{1}{n+1})^n$, which converges to $\frac{1}{e}$, while n goes to infinity. We presume, that this is the boundary case. Let us formulate this in the following conjecture, stated by Uriel Feige [3].

Conjecture. *Let $X_1, \dots, X_n$ be a collection of nonnegative random variables with expectations equal to 1. Then the following inequality holds:*

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) \geqslant \frac{1}{e}.$$

1.3 Previous results and applications

In his paper, Uriel Feige proves that the answer to our question is more than $\frac{1}{13}$, and states the conjecture that it is equal to $\frac{1}{e}$. He also shows how to apply the inequality to find those edges in a network that belong to many shortest paths [3]. In his paper, John H. Elton investigates the case when all the random variables are identically distributed, and can have only integer values [4]. Samuels proves the Conjecture when the number of random variables is at most four [5].

1.4 Statement of results

In section 2, we formally define notations introduced in the introduction . In section 3, we present how to prove that the minimum $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right)$ can be achieved when when all random variables have at most two possible values, one of which is 0. In section 4, we generalize John H. Elton's result to the real valued case, when n goes to the infinity.

2 Preliminaries

Formally, a discrete random variable X is associated with a set V_X that denotes its possible values, and a probability density function $p_X : V_X \to \mathbb{R}$ such that $p_X(v) = \mathbb{P}(X = v)$ denotes the probability that X has value v. The expectation of X is the average of all its possible values weighted by the possibility of obtaining them. More formally, we write $\mathbb{E}[X] = \sum_{v \in V_X} v p_X(v)$. In our paper, we consider independent random variables $X_1, \ldots, X_n$, such that $\mathbb{E}[X_i] = 1$ and $X_i \geqslant 0$ for any $1 \leqslant i \leqslant n$. Let a be a real number and $a \geqslant 1$. We define $\mathcal{X}_a$ to be the following distribution:

$$\mathcal{X}_a = \begin{cases} a & \text{with probability } \frac{1}{a} \\ 0 & \text{with probability } \frac{a-1}{a}. \end{cases}$$

3 Reduction of the random variables

In this section, we show that $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right)$ can be minimized with the additional assumption that each X_i's distribution is in the form of $\mathcal{X}_{a_i}$ for some $a_i \geqslant 1$. To achieve this, we introduce three Lemmas.

Lemma 3.1. *The random variable X_1 can be replaced by another random variable X_1', which* *takes only two possible values, such that*

$$\mathbb{P}\left(X_1' + \sum_{i=2}^{n} X_i < n+1\right) \leqslant \mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right).$$

Proof. Let us fix the random variables $X_2, \ldots, X_n$. Let $V_{X_1} = \{v_1, \ldots, v_k\}$ be a set of all possible values of X_1 with a probability density function $p_{X_1} : V_{X_1} \to \mathbb{R}$. Let $a_1, \ldots, a_k$ be $a_i = \mathbb{P}\left(\sum_{i=2}^{n} X_i < n+1-v_i\right)$ and $y_1, \ldots, y_k$ be $y_i = p_{X_1}(v_i)$. We will treat a_i as constants and y_i as variables. We want to choose the values of y_i $(i = 1, \ldots, k)$ so that $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right)$ is minimized. This problem can be formulated in terms of linear programing as follows:

$$\begin{aligned}
\text{Minimize} \quad & \sum_{i=1}^{k} a_i y_i \\
\text{subject to} \quad & \mathbb{E}[X_1] = \sum_{i=1}^{k} v_i y_i = 1, \\
& \sum_{i=1}^{k} y_i = 1, \\
& y_i \geqslant 0.
\end{aligned}$$

In Lemma 3.2, we prove that there is an optimal solution in the case where at most two y_i are not equal to 0. To complete the proof, we choose $V_{X_1} = V_{X_1'}$ and $p_{x_1'}(v_i) = y_i$. $\square$

Lemma 3.2. *Let* $y = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix}$ *and I be a k by k identity matrix. Then the following linear* *programing problem has an optimal solution y, where at most two of y_i are not equal to 0.*

$$\begin{aligned}
&\textit{Minimize} && \begin{bmatrix} a_1 & \dots & a_k \end{bmatrix} y \\
&\textit{subject to} && \begin{bmatrix} v_1 & \dots & v_k \\ 1 & \dots & 1 \end{bmatrix} y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\
& && I\ y \geqslant \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.
\end{aligned}$$

Proof. Any row vector is considered to be a constraint. A *feasible* solution to a linear program is such y that satisfies all the constraints. A constraint of a linear program is *active* for a certain solution y if it meets with equality. A solution to a linear program is *basic* if it satisfies the following conditions:

- All the equality constraints are active.
- There are k linearly independent active constraints.

Note that the basic solutions are the vertices of a polyhedron. It is known that if there exists an optimal feasible solution and a basic feasible solution, then there exists a solution that is both basic and optimal [6]. Since the objective function is equal to $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right)$, which is between 0 and 1, there exists an optimal solution. If $v_i = 1$ for some i, we can choose y, such that $y_i = 1$ and $y_j = 0$ for any $j \neq i$, which is a basic feasible solution. If there exist no such v_i then there must exist such v_i and v_j that $v_i < 1 < v_j$. We choose such y that

$$\begin{cases} y_i + y_j = 1 \\ v_i y_i + v_j y_j = 1 \end{cases},$$

and $y_l = 0$ if $l \neq i$ and $l \neq j$. Since the vectors $\begin{bmatrix} v_i \\ v_j \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are linearly independent, y is a basic feasible solution. Hence, there is an optimal and basic feasible solution y. Note that at least k of the constraints are active for y, but there are only two equality constraints. Thus, at most two of the inequalities do not meet the equality. □

Lemma 3.3. *If X_1 is a random variable with only two possible values, we can replace one of them by 0, and increase the other one so that $\mathbb{P}(\sum_{i=1}^n X_i < n+1)$ does not increase, and $\mathbb{E}[X_1]$ will remain 1.*

Proof. Let a, b, p be real numbers such that $a \geqslant b \geqslant 0$ and

$$X_1 = \begin{cases} a & \text{with probability } p \\ b & \text{with probability } 1-p. \end{cases}$$

Now we replace X_1 with a variable $X_1' \sim \mathcal{X}_{1/p}$. Since $\mathbb{E}[X_1] = ap + b(1-p) = 1$ and $0 \leqslant p \leqslant 1$ then $b \leqslant a \leqslant \frac{1}{p}$.

Now, we have

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) &= p\,\mathbb{P}\left(\sum_{i=2}^{n} X_i < n+1-a\right) + (1-p)\,\mathbb{P}\left(\sum_{i=2}^{n} X_i < n+1-b\right) \\ &\geqslant p\,\mathbb{P}\left(\sum_{i=2}^{n} X_i < n+1-\frac{1}{p}\right) + (1-p)\,\mathbb{P}\left(\sum_{i=2}^{n} X_i < n+1-\frac{1}{p}\right) \\ &\geqslant p\,\mathbb{P}\left(\sum_{i=2}^{n} X_i < n+1-\frac{1}{p}\right) \\ &= \mathbb{P}\left(X_1' + \sum_{i=2}^{n} X_i < n+1\right). \end{aligned}$$

□

Theorem 1. *Let $X_1, \dots, X_n$ be random variables with $\mathbb{E}[X_i] = 1$ and $X_i \geq 0$. Then*

$$\mathbb{P}\left(\sum_{i=1}^{n} X'_i < n+1\right) \leqslant \mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right),$$

where $X'_i \sim \mathcal{X}_{a_i}$ for some choice of $a_1 \dots a_n$.

Proof. If there exists any X_i such that $X_i \neq \mathcal{X}_a$ for any a, we can swap it with X_1 and then apply Lemma 3.3 and Lemma 3.1. We repeat this step until there is no such X_i. $\square$

From now on we can assume that every random variable has only two possible values, and one of them is 0.

4 Identically distributed case

In this section, we investigate the case when all variables are identically distributed. More formally we assume that $X_1, \dots, X_n \sim \mathcal{X}_a$ where a is some real number. Then, $\sum_{i=1}^{n} X_i < n+1$ if and only if less than $\lceil \frac{n+1}{a} \rceil$ of X_i's are equal to a. Thus, if we decrease the value of a while $\lceil \frac{n+1}{a} \rceil$ remain unchanged, then $\mathbb{P}(\sum_{i=1}^{n} X_i < n+1)$ decreases because $\mathbb{P}(X_i \neq a)$ decreases. Therefore, we can assume that $a = \frac{n+1}{k}$, where k is an integer between 1 and n. This allows us to give an explicit formula,

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) = \sum_{i=0}^{k-1} \binom{n}{i} \left(\frac{k}{n+1}\right)^{i} \left(\frac{n+1-k}{n+1}\right)^{n-i}.$$

In this section, we prove that when k is fixed, the limit of the above sum is greater than $\frac{1}{e}$, when n goes to infinity. Before we prove this, we introduce two lemmas.

Lemma 4.1. *For any positive integers k, i where $i \leqslant k$, we have*

$$\lim_{n\to\infty} \frac{n!}{(n-i)!} \cdot \frac{(n+1-k)^{n-i}}{(n+1)^n} = \frac{1}{e^k}.$$

Proof. Note that

$$\begin{aligned}
\lim_{n\to\infty} \frac{n!}{(n-i)!} \cdot \frac{(n+1-k)^{n-i}}{(n+1)^n} &= \lim_{n\to\infty} \frac{n!}{(n-i)!(n+1)^i} \cdot \lim_{n\to\infty} \frac{(n+1-k)^{n-i}}{(n+1)^{n-i}} && (1)\\
&= \lim_{n\to\infty} \frac{(n+1-k)^{n-i}}{(n+1)^{n-i}} && (2)\\
&= \lim_{n\to\infty} \left(\frac{n+1-k}{n+1}\right)^n \cdot \lim_{n\to\infty} \left(\frac{n+1-k}{n+1}\right)^{-i} && (3)\\
&= \lim_{n\to\infty} \left(\frac{n+1-k}{n+1}\right)^n && (4)\\
&= \frac{1}{e^k}. && (5)
\end{aligned}$$

Equality (2) holds because $\frac{n!}{(n-i)!}$ and $(n+1)^i$ are polynomials with equal leading coefficients. Equality (4) holds because

$$\lim_{n\to\infty} \left(\frac{n+1-k}{n+1}\right)^{-i} = \lim_{n\to\infty} \left(1 - \frac{k}{n+1}\right)^{-i} = 1.$$

Equality (5) holds in virtue of the following induction on k.

Base case $k = 1$: If $k = 1$ then

$$\lim_{n\to\infty} \left(\frac{n}{n+1}\right)^n = \lim_{n\to\infty} \frac{1}{\left(1+\frac{1}{n}\right)^n} = \frac{1}{e}$$

Inductive hypothesis: Suppose the theorem holds for all values of k up to some m, $m \geqslant 1$.

Inductive step: Let $k = m + 1$. We have

$$\begin{aligned}
\lim_{n\to\infty}\left(\frac{n+1-k}{n+1}\right)^n &= \lim_{n\to\infty}\left(\frac{n+1-(k-1)}{n+1}\right)^n \cdot \lim_{n\to\infty}\left(\frac{n+1-k}{n+2-k}\right)^n \\
&= \frac{1}{e^{k-1}} \cdot \lim_{n\to\infty}\left(1-\frac{1}{n+2-k}\right)^n \\
&= \frac{1}{e^{k-1}} \cdot \lim_{n\to\infty}\left(1-\frac{1}{n+2-k}\right)^{n+2-k} \cdot \lim_{n\to\infty}\left(1-\frac{1}{n+2-k}\right)^{k-2} \\
&= \frac{1}{e^k}.
\end{aligned}$$

The theorem holds for $k = m + 1$. By the principle of mathematical induction, the theorem holds for all $k \in \mathbb{N}$. □

Lemma 4.2. *For any integer k with $k \geqslant 1$, we have*

$$\sum_{i=0}^{k-1} \frac{k^i}{i!} \geqslant \frac{1}{e^{k-1}}.$$

Proof. This proof is a simplification of the result from K. P. Choi's paper [7]. Using the Taylor expansion of function e^x and the integral formula for the residue, we get

$$e^k = \sum_{i=0}^{\infty} \frac{k^i}{i!} = \sum_{i=0}^{k-1} \frac{k^i}{i!} + \int_0^k e^t \cdot \frac{(k-t)^{k-1}}{(k-1)!} dt.$$

Function $\Gamma(k) = \int_0^\infty e^{-t} t^{k-1} dt$ is an extension of factorial function and $\Gamma(k) = (k-1)!$ when k is a positive integer. Thus,

$$\int_0^\infty \frac{e^{-t} t^{k-1} dt}{(k-1)!} = 1.$$

By replacing t with $k-t$, we obtain

$$\begin{aligned}
\sum_{i=0}^{k-1}\frac{k^i}{i!} &= e^k - \int_0^k e^{k-t}\cdot\frac{t^{k-1}}{(k-1)!}dt \\
&= e^k\left(1-\int_0^k \frac{e^{-t}t^{k-1}dt}{(k-1)!}\right) \\
&= e^k\left(\int_0^\infty \frac{e^{-t}t^{k-1}dt}{(k-1)!} - \int_0^k \frac{e^{-t}t^{k-1}dt}{(k-1)!}\right) \\
&= \frac{e^k}{(k-1)!}\int_k^\infty e^{-t}t^{k-1}dt.
\end{aligned}$$

Let

$$a_k = e^{-k}\sum_{i=0}^{k-1}\frac{k^i}{i!} = \frac{1}{(k-1)!}\int_k^\infty e^{-t}t^{k-1}dt.$$

Notice $a_1 = \frac{1}{e}$. We prove that a_k is increasing sequence by showing that $a_k - a_{k-1} \geqslant 0$.

$$\begin{aligned}
(k-1)!(a_k - a_{k-1}) &= \int_k^\infty e^{-t}t^{k-1}dt - (k-1)\int_{k-1}^\infty e^{-t}t^{k-2}dt && (6)\\
&= \int_k^\infty e^{-t}t^{k-1}dt - \int_{k-1}^\infty e^{-t}(t^{k-1})'dt && (7)\\
&= \int_k^\infty e^{-t}t^{k-1}dt - \left(e^{-t}t^{k-1}\Big|_{k-1}^\infty + \int_{k-1}^\infty e^{-t}t^{k-1}dt\right) && (8)\\
&= e^{-(k-1)}(k-1)^{k-1} - \int_{k-1}^k e^{-t}t^{k-1}dt && (9)\\
&= \int_{k-1}^k e^{-(k-1)}(k-1)^{k-1} - e^{-t}t^{k-1}dt && (10)\\
&\geqslant 0. && (11)
\end{aligned}$$

Equality (8) follows from integration by parts. Inequality (11) holds because $f(t) = e^{-t}t^{k-1}$ is a decreasing function on $[k-1, k]$. $\square$

Now we prove the following theorem.

Theorem 2. *Fix $k > 1$. Let $X_1, \ldots, X_n$ be identically distributed random variables, with* *distribution* $\mathcal{X}_{\frac{n+1}{k}}$. *Then,*

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) \geqslant \frac{1}{e}.$$

Proof. This theorem follows from Lemma 4.1 and Lemma 4.2. Note that

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) &= \lim_{n\to\infty} \sum_{i=0}^{k-1} \binom{n}{i} \left(\frac{k}{n+1}\right)^{i} \left(\frac{n+1-k}{n+1}\right)^{n-i} && (12)\\
&= \lim_{n\to\infty} \sum_{i=0}^{k-1} \frac{k^i}{i!} \cdot \frac{n!}{(n-i)!} \cdot \frac{(n+1-k)^{n-i}}{(n+1)^n} && (13)\\
&= \sum_{i=0}^{k-1} \frac{k^i}{i!} \cdot \frac{1}{e^k} && (14)\\
&\geqslant \frac{1}{e}. && (15)
\end{aligned}$$

Equality (14) follows from Lemma 4.1. Inequality (15) follows form Lemma 4.2. □

Let $f_k(n) = \mathbb{P}(\sum_{i=1}^{n} X_i < n+1)$. We conjecture that $f_k(n)$ is a decreasing function in n, which implies that $\mathbb{P}(\sum_{i=1}^{n} X_i < n+1) \geqslant \frac{1}{e}$. We provide some numerical evidences supporting it (see figure 1-3).

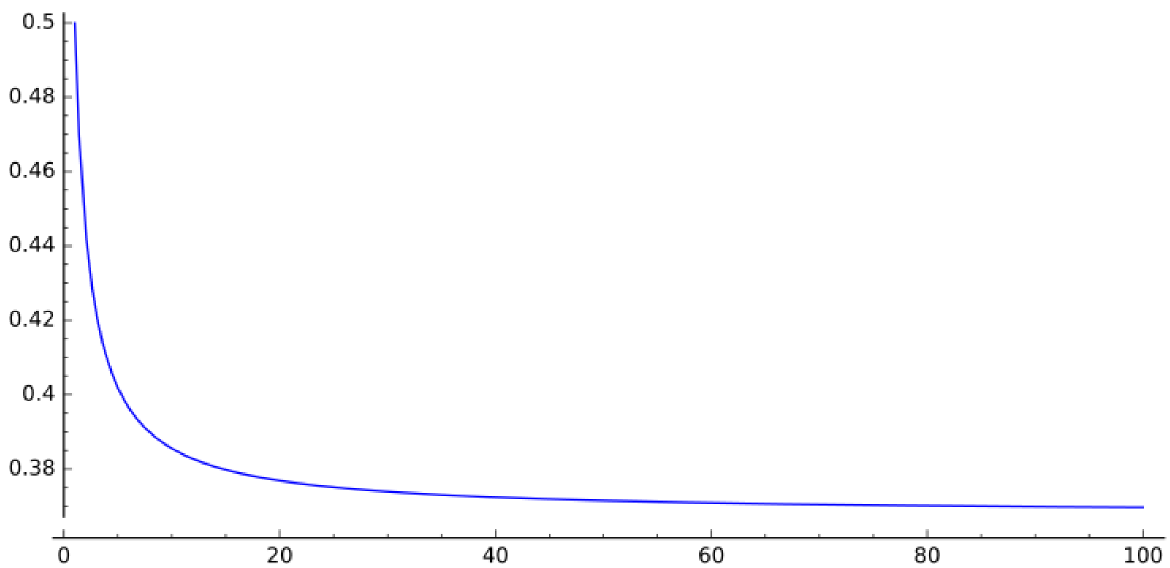

Figure 1: Plot of the $f_1(n)$ function .

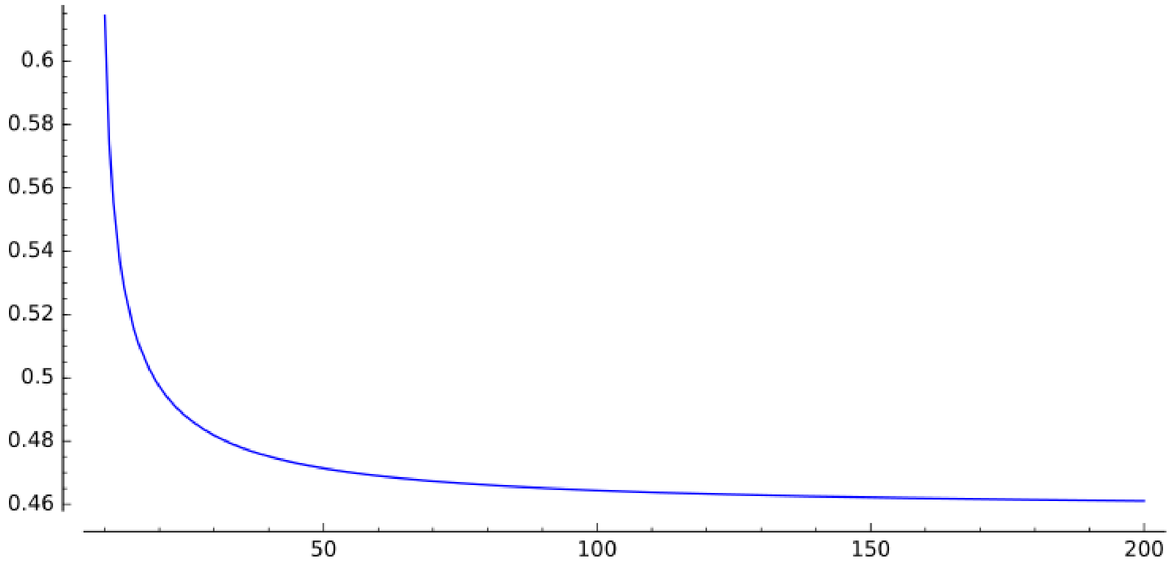

Figure 2: Plot of the $f_{10}(n)$ function.

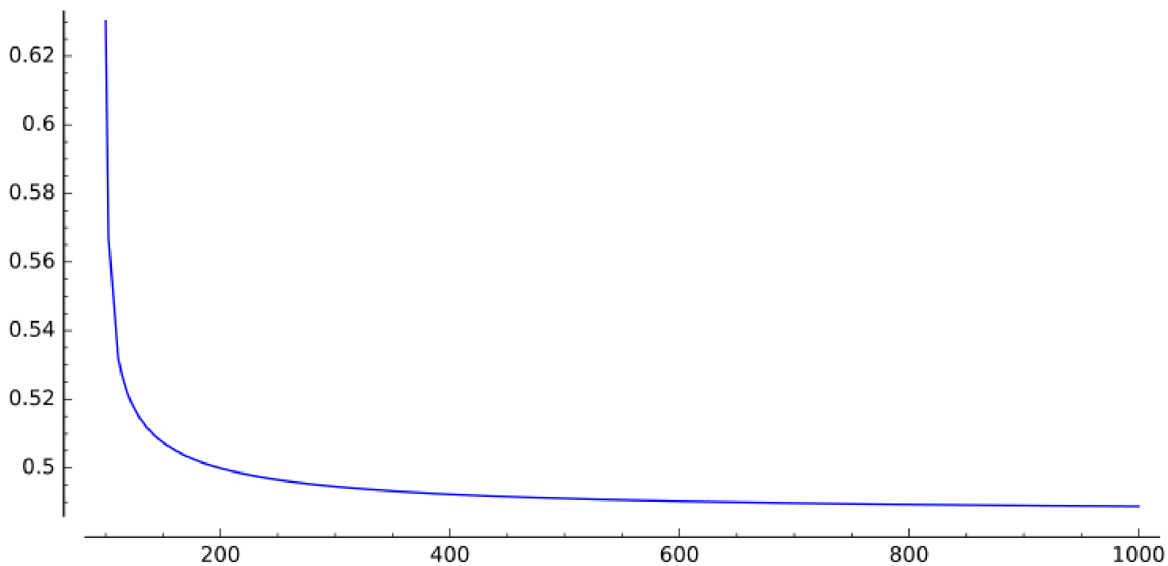

Figure 3: Plot of the $f_{100}(n)$function.

5 Conclusion

The results of our research could hopefully contribute to the study of the behavior of sums of independent random variables. We have proved that $\mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right)$ can be minimized when all random variables have at most two possible values, one of which is 0. However, our reduction does not take into account the assumption that random variables are identically distributed. If the latter assumption holds, we have proved that

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_i < n+1\right) \geqslant \frac{1}{e}.$$

As a natural extension of this project, we could generalize the result of the identically distributed case by scaling all the random variables by integer k tending to infinity. More formally, let us formulate the following conjecture.

Conjecture. *Let $a_1, \ldots, a_n$ be real numbers greater than one. Then,*

$$\lim_{k\to\infty} \mathbb{P}\left(\sum_{i=1}^{n} \sum_{j=1}^{k} X_{ij} < kn+1\right) \geqslant \frac{1}{e},$$

where X_{ij} is a random variable with distribution $\mathcal{X}_{ka_i}$.

If $\mathbb{P}(\sum_{i=1}^{n} \sum_{j=1}^{k} X_{ij} < kn+1)$ is a decreasing function of n, then together with conjecture it implies the Feige's conjecture.

6 Acknowledgments

I would like to thank my mentor Chiheon Kim for giving me many inspirational ideas and for choosing the project for me, which was close to my interests, enjoyable, and educational. I also would like to thank Dr. Jenny Sendova for making this text more approachable for a larger scientific audience. I would like to thank Dr. Tanya Khovanova for helping with important advice on math writing, Prof. David Jerison, Dr. Slava Gerovitch, and Prof. Ankur Moitra for general math support. I would like to thank Fragkiskos Koufogiannis, Dennis Tseng, and Michelle Xu for reviewing my paper, Brad Sturt for advice on linear programing. I would like to thank the Center for Excellence in Education and RSI for providing the opportunity to do research at MIT and Polish Children's Fund for sponsoring me.

References

[1] P. de Laplace. Théorie analytique des probabilités. *Paris, Ve. Courcier*, Paris 1814.

[2] A. N. Kolmogorov. Foundations of probability theory. *Chelsea*, New York, 1950.

[3] U. Feige. On sums of independent random variables with unbounded variance, and estimating the average degree in a graph. *Department of Computer Science and Applied Mathematics The Weizmann Institute.*

[4] J. H. Elton. Notes on feige's gumball machines problem. 25 Aug 2009.

[5] S. M. Samuels. On a chebyshev-type inequality for sums of independent random variables. *Ann. Math. Statist. Vol 37*, page 248–259, 1966.

[6] J. N. T. Dimitris Bertsimas. Introduction to linear optimization, theorem 2.3. 1997.

[7] K. P. Choi. On the medians of gamma distributions and an equation of ramanujan. *Proc. Amer. Math. Soc. 121*, 1994.