Statistics and Probability Letters 129 (2017) 12–16

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On maximal tail probability of sums of nonnegative, independent and identically distributed random variables

Tomasz Łuczak [a], Katarzyna Mieczkowska [a], Matas Šileikis [b,*]

[a] Adam Mickiewicz University, Collegium Mathematicum, ul. Umultowska 87, 61-614 Poznań, Poland

[b] Charles University, KAM MFF, Malostranské nám. 25, 118 00 Praha 1, Czech Republic

ARTICLE INFO

Article history:
Received 19 January 2017
Accepted 26 April 2017
Available online 10 May 2017

Keywords:
I.i.d. random variables
Tail probability
Markov's inequality

ABSTRACT

We consider the problem of finding the optimal upper bound for the tail probability of a sum of k nonnegative, independent and identically distributed random variables with given mean x. For $k = 1$ the answer is given by Markov's inequality and for $k = 2$ the solution was found by Hoeffding and Shrikhande in 1955. We show that the solution for $k = 3$ as well as for general k, provided $x \le 1/(2k-1)$, follows from recent results of extremal combinatorics.

1. Introduction

In this note we consider, for $k = 1, 2, \ldots$ and $x \ge 0$, the quantity

$$m_k(x) = \sup_{\mathbf{X}} \mathbb{P}(X_1 + \cdots + X_k \ge 1), \tag{1}$$

where the supremum is taken over all random vectors $\mathbf{X} = (X_1, \ldots, X_k)$ of nonnegative, independent and identically distributed (further *i.i.d.*) random variables X_i such that $\mathbb{E}(X_i) \le x$ for $i = 1, \ldots, k$.

For $k = 1$ and $x < 1$ the solution $m_1(x) = \max\{x, 1\}$ is given by Markov's inequality. In the case of two variables the problem was solved by Hoeffding and Shrikhande (1955) who showed that

$$m_2(x) = \begin{cases} 2x - x^2 & \text{for } x < 2/5; \\ 4x^2 & \text{for } 2/5 \le x < 1/2; \\ 1, & \text{for } x \ge 1/2. \end{cases}$$

We conjecture that the following generalization holds.

Conjecture 1. *For every positive integer $k \ge 2$ and $x \ge 0$ we have*

$$m_k(x) = \begin{cases} 1 - (1-x)^k & \text{if } x < x_0(k), \\ (kx)^k & \text{if } x_0(k) \le x < 1/k, \\ 1, & \text{if } x \ge 1/k, \end{cases} \tag{2}$$

where $x_0(k)$ is the root of the polynomial $p_k(x) = (kx)^k + (1-x)^k - 1$ in $(0, 1/k)$.

* Corresponding author.
E-mail addresses: tomasz@amu.edu.pl (T. Łuczak), kaska@amu.edu.pl (K. Mieczkowska), matas.sileikis@gmail.com (M. Šileikis).

http://dx.doi.org/10.1016/j.spl.2017.04.024
0167-7152/

In Łuczak and Mieczkowska (2014) it was noted that $1/k - 1/(2k^2) < x_0(k) < 1/k - 2/(5k^2)$.

If Conjecture 1 is correct, $m_k(x)$ is attained in the first case by a two-point distribution $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = x$, in the second case by a two-point distribution $\mathbb{P}(X_i = 1/k) = 1 - \mathbb{P}(X_i = 0) = kx$, and in the third case by the degenerate distribution $\mathbb{P}(X_i = x) = 1$.

The case of not necessarily identically distributed X_i's has also been studied. Let

$$s_k(x) = \sup_{\mathbf{X}} \mathbb{P}\left(X_1 + \cdots + X_k \geq 1\right),$$

where the supremum is taken over all vectors of independent, nonnegative random variables with common mean x. Clearly, $m_k(x) \leq s_k(x)$. In 1966 Samuels (1966) conjectured the maximum value of the tail probability in terms of arbitrary means $\mathbb{E}X_i$, $i = 1, \ldots, k$. For simplicity, we consider the case of equal means.

Conjecture 2 (*Samuels, 1966*). *For every positive integer k and $x \geq 0$ we have*

$$s_k(x) = \begin{cases} 1 - \min\limits_{t=0}^{k-1} \left(1 - \dfrac{x}{1-tx}\right)^{k-t} & \text{if } x < 1/k, \\ 1, & \text{if } x \geq 1/k. \end{cases} \tag{3}$$

If Conjecture 2 holds, $s_k(x)$ is attained by one of the random vectors

$$\mathbf{X}^t = (X_1^t, \ldots, X_k^t), \quad t = 0, \ldots, k-1,$$

where for $i \leq t$ we have $\mathbb{P}(X_i^t = x) = 1$ and for $i = t+1, \ldots, k$ we have $\mathbb{P}(X_i^t = 1 - tx) = 1 - \mathbb{P}(X_i^t = 0) = x/(1-tx)$. Thus $\mathbb{P}(X_1^t + \cdots + X_k^t \geq 1) = 1 - \left(1 - \frac{x}{1-tx}\right)^{k-t}$.

Therefore, if true, (3) implies Conjecture 1 only when the minimum in (3) is attained by $t = 0$. Computer-generated graphs suggest that the minimum is attained by $t = 0$ when $x < x_1(k)$ and by $t = k-1$ when $x \geq x_1(k)$, where $x_1(k) \in (0, 1/k)$ is the solution of $(1-x)^k = 1 - x/(1-(k-1)x)$. In Alon et al. (2012, Proposition 2.1) it was shown that

$$\min_{t=0}^{k-1} \left(1 - \frac{x}{1-tx}\right)^{k-t} = (1-x)^k, \quad 0 \leq x \leq 1/(k+1), \tag{4}$$

so $x_1(k) > 1/(k+1)$. Samuels (1966, 1968) confirmed (3) for $k = 3, 4$. Analysing numerically the graphs of functions $s_3(x)$ and $s_4(x)$ one can see that for $k = 3, 4$ we have

$$m_k(x) = s_k(x) = 1 - (1-x)^k, \quad x \in [0, x_1(k)], \tag{5}$$

and $x_1(3) = 0.27729\ldots, x_1(4) = 0.21737\ldots$.

Moreover, Samuels (1969) proved that for $k \geq 5$

$$m_k(x) = s_k(x) = 1 - (1-x)^k, \quad x \leq 1/(k^2 - k).$$

Our main result follows.

Theorem 3. *Conjecture 1 holds for $x \geq 0$ when $k = 3$, and for $x < \frac{1}{2k-1}$, when $k \geq 4$.*

Of course, for $k = 4$, Samuels' result (5) gives a wider range, but for $k = 3$ and $k \geq 5$ Theorem 3 gives an improvement.

The proof of Theorem 3 is based on reduction of Conjecture 1 to the fractional version of so-called Erdős' Conjecture on matchings in hypergraphs. The conjecture is introduced in the next section. Given this reduction, Theorem 3 follows from results by Łuczak and Mieczkowska (2014) as well as Frankl (2012, 2013). The proof is given in Section 3.

In fact the two conjectures are equivalent, but here we use only one implication. For the proof of the other one see Łuczak et al. (2016).

2. The hypergraph matching problem

A *k-uniform hypergraph* $H = (V, E)$ is a set of *vertices* V together with a family E of k-element subsets of V, called *edges*. A *matching* is a family of disjoint edges of H, and the *matching number* of H, denoted by $\nu(H)$, is the size of the largest matching in H.

In Erdős (1965) Erdős stated the following.

Conjecture 4 (*Erdős, 1965*). *Let $H = (V, E)$ be a k-uniform hypergraph, $|V| = n$, $\nu(H) = s$. If $n \geq ks + k - 1$, then*

$$|E| \leq \max\left\{\binom{n}{k} - \binom{n-s}{k}, \binom{ks+k-1}{k}\right\}. \tag{6}$$

Note that the equality in Conjecture 4 holds either when H consists of all edges intersecting a given subset $S \subset V$, $|S| = s$, or when H consists of edges contained in a given subset $T \subset V$, $|T| = ks+k-1$. We denote these two families of hypergraphs by $Cov_{n,k}(s)$ and $Cl_{n,k}(ks+k-1)$, respectively.