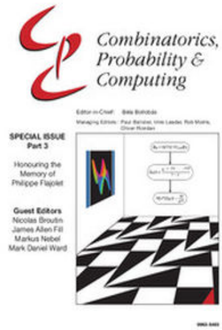

A Sharp Uniform Bound for the Distribution of Sums of Bernoulli Trials

JEAN-BERNARD BAILLON, ROBERTO COMINETTI and JOSÉ VAISMAN

Combinatorics, Probability and Computing / Volume 25 / Issue 03 / May 2016, pp 352 - 361
DOI: 10.1017/S0963548315000127, Published online: 07 July 2015

Combinatorics, Probability and Computing (2016) 25, 352–361.
doi:10.1017/S0963548315000127

A Sharp Uniform Bound for the Distribution of Sums of Bernoulli Trials

JEAN-BERNARD BAILLON[1], ROBERTO COMINETTI[2] and JOSÉ VAISMAN[3]

[1]SAMM–EA 4543, Université de Paris 1, 75013 Paris, France
(e-mail: baillon@univ.paris1.fr)
[2]Departamento de Ingeniería Industrial, Universidad de Chile, Avenida República 701, Santiago, Chile
(e-mail: cominetti.roberto@gmail.com)
[3]Departamento de Ingeniería Matemática, Universidad de Chile, Avenida Blanco Encalada 2120, Santiago, Chile
(e-mail: hellovaisman@gmail.com)

Received 3 July 2012; revised 10 May 2015; first published online 7 July 2015

In this note we establish a uniform bound for the distribution of a sum $S_n = X_1 + \cdots + X_n$ of independent non-homogeneous Bernoulli trials. Specifically, we prove that $\sigma_n \mathbb{P}(S_n = j) \leqslant \eta$, where σ_n denotes the standard deviation of S_n, and η is a universal constant. We compute the best possible constant $\eta \sim 0.4688$ and we show that the bound also holds for limits of sums and differences of Bernoullis, including the Poisson laws which constitute the worst case and attain the bound. We also investigate the optimal bounds for n and j fixed. An application to estimate the rate of convergence of Mann's fixed-point iterations is presented.

2010 *Mathematics subject classification*: Primary 60C05
Secondary 60E05, 60E15

1. Introduction

Let $S_n = X_1 + \cdots + X_n$ be a sum of independent non-homogeneous Bernoulli trials with success probabilities p_i. The distribution of S_n is known to be unimodal and bell-shaped with mean $\mu_n = \sum_{i=1}^n p_i$ and variance $\sigma_n^2 = \sum_{i=1}^n p_i(1 - p_i)$. Its mode is either $\lfloor \mu_n \rfloor$ or $\lceil \mu_n \rceil$ or both [9, 26], and the same holds for the median [18]. In this paper we investigate how large the modal probability can be. More precisely, we establish a uniform upper bound

$$\sigma_n \mathbb{P}(S_n = j) \leqslant \eta \tag{1.1}$$

for all n, j and p_i, and we prove that the best possible constant is

$$\eta = \max_{\lambda \geqslant 0} \sqrt{2\lambda}\, e^{-2\lambda} \sum_{k=0}^{\infty} \left(\frac{\lambda^k}{k!} \right)^2 \sim 0.4688. \tag{1.2}$$

The existence of a universal bound (1.1) can be established using tools related to the local limit theorem [10, 13, 23]. It also follows as a special case of the Kolmogorov–Rogozin concentration inequality [25], which states a more general bound valid for discrete random variables X_i with σ_n replaced by $\sqrt{\sum_{i=1}^{n}(1-\psi_i)}$, where $\psi_i = \max_x \mathbb{P}(X_i = x)$. For sums of Bernoullis this is equivalent to (1.1), so our contribution is mainly the computation of the optimal constant η, as well as the identification of the role of the Poisson law in the worst case. Namely, the local limit theorem shows that for a wide range of random variables the limit as $n \to \infty$ in (1.1) exists and equals $1/\sqrt{2\pi}$. Since η exceeds this value, it follows that the worst-case situation is not associated with random variables obeying the central limit theorem. It is then natural to expect that the worst case may have to do with the Poisson law, and that is what actually happens. In fact the expression (1.2) is just

$$\eta = \max_{\lambda \geqslant 0} \sqrt{2\lambda}\, \mathbb{P}(N_\lambda = N'_\lambda), \tag{1.3}$$

where N_λ and N'_λ are independent Poisson variables with parameter λ. Since the Poisson law also happens to be extremal in other bounds such as Rosenthal's inequality (see [12, 17, 27, 28]), a natural question is whether (1.1) might hold for more general sums of random variables.

The inequality (1.1) complements the large deviation bounds that provide estimates of the form $\mathbb{P}(|S_n - \mathbb{E}(S_n)| \geqslant t) \leqslant f(nt^2)$ with $f(x) \to 0$ as $x \to \infty$, usually at an exponential rate (see [1, 2, 5, 6, 11, 15, 16, 22, 24]). In contrast, (1.1) does not give such fast asymptotic rates but it can be used to bound $\mathbb{P}(S_n = j)$ for all values of j including values close to the mean $\mathbb{E}(S_n)$. This has already proved useful in establishing an optimality guarantee for an approximation algorithm in discrete stochastic optimization (see [7]). In this paper we present another application to the rate of convergence of fixed-point iterations for non-expansive maps. In both settings, a sharp constant η is relevant as it yields better bounds.

The paper is organized as follows. In Section 2 we show the sharp uniform bound (1.1) to be valid for all n, j and p_i, and we briefly discuss some extensions to more general distributions including sums and differences of Bernoullis, as well as their limits, which cover all Poisson distributions and more. In Section 3 we investigate more closely the optimal bounds for fixed n and j. Finally, in Section 4 we show how (1.1) allows us to establish the rate of convergence for fixed-point iterations.

2. A sharp uniform bound

Theorem 2.1. *Let $S_n = X_1 + \cdots + X_n$ denote a sum of independent Bernoulli trials where $\mathbb{P}(X_i = 1) = p_i$, and let $\sigma_n^2 = \sum_{i=1}^{n} p_i(1-p_i)$ be its variance. Then*

$$\sigma_n\, \mathbb{P}(S_n = j) \leqslant \eta, \tag{2.1}$$

where

$$\eta = \max_{\lambda > 0} \sqrt{2\lambda}\, \mathbb{P}(N_\lambda = N'_\lambda),$$

with N_λ and N'_λ independent Poisson variables of parameter λ. This bound is sharp, and we have more explicitly $\eta = \max_{x \geqslant 0} \sqrt{x}\, e^{-x} I_0(x) \sim 0.4688$, where $I_0(x)$ is the modified Bessel *function*

$$I_0(x) = \sum_{k=0}^{\infty} \left(\frac{x^k}{2^k k!} \right)^2 = \frac{1}{\pi} \int_0^{\pi} \exp(x \cos \theta)\, d\theta.$$

Proof. Consider the generating function

$$\phi(z) = \mathbb{E}(z^{S_n}) = \sum_{j=0}^{n} \mathbb{P}(S_n = j) z^j.$$

Integrating $\phi(z)/z^{j+1}$ along the unit circle $\mathcal{C}$ in the complex plane, we get

$$\mathbb{P}(S_n = j) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\phi(z)}{z^{j+1}}\, dz = \frac{1}{2\pi} \int_0^{2\pi} \phi(e^{i\theta}) e^{-ij\theta}\, d\theta, \tag{2.2}$$

so that taking absolute value it follows that

$$\mathbb{P}(S_n = j) \leqslant \frac{1}{2\pi} \int_0^{2\pi} |\phi(e^{i\theta})|\, d\theta. \tag{2.3}$$

The independence of the X_i yields

$$\phi(z) = \mathbb{E}\left[\prod_{i=1}^{n} z^{X_i} \right] = \prod_{i=1}^{n} [(1 - p_i) + p_i z],$$

from which we obtain

$$|\phi(e^{i\theta})| = \prod_{i=1}^{n} \sqrt{1 + 2p_i(1 - p_i)(\cos \theta - 1)}.$$

Using the inequality $1 + y_i \leqslant \exp(y_i)$ with $y_i = 2p_i(1 - p_i)(\cos \theta - 1)$, and setting

$$x = \sum_{i=1}^{n} p_i(1 - p_i),$$

we deduce

$$|\phi(e^{i\theta})| \leqslant \prod_{i=1}^{n} \exp(y_i/2) = \exp(x(\cos \theta - 1)),$$

and since $\sigma_n = \sqrt{x}$ we conclude

$$\sigma_n\, \mathbb{P}(S_n = j) \leqslant \sqrt{x} \frac{1}{2\pi} \int_0^{2\pi} \exp(x(\cos \theta - 1))\, d\theta = \sqrt{x}\, e^{-x} I_0(x) \leqslant \eta.$$

In order to show that the bound is sharp, consider a sum of $n = 2a$ Bernoullis, half of them with $p_i = \lambda/a$ and the other half with $p_i = 1 - \lambda/a$, so that $S_n = U + V$ with $U \sim B(a, \lambda/a)$ and $V \sim B(a, 1 - \lambda/a)$ independent binomials. Note that $V \stackrel{d}{=} a - U'$ with U' an independent copy of U, and for $j = a$ we get

$$\sigma_n\, \mathbb{P}(S_n = a) = \sqrt{2\lambda \left(1 - \frac{\lambda}{a}\right)}\, \mathbb{P}(U = U').$$

Since U and U' converge as $a \to \infty$ to independent Poisson variables N_λ and N'_λ with parameter λ, this expression tends to

$$\sqrt{2\lambda}\,\mathbb{P}(N_\lambda = N'_\lambda) = \sqrt{2\lambda}\,e^{-2\lambda}\sum_{k=0}^{\infty}\left(\frac{\lambda^k}{k!}\right)^2 = \sqrt{2\lambda}\,e^{-2\lambda}I_0(2\lambda)$$

which proves that the bound is sharp. □

Remarks. **(1)** The optimal bound

$$\eta = \max_{\lambda>0}\sqrt{2\lambda}\,\mathbb{P}(N_\lambda = N'_\lambda)$$

is approximately $\eta \sim 0.468822355499$, which is attained for $\lambda \sim 0.39498893$.

(2) The proof above shows that the bound η is asymptotically attained for a sum of two binomials with different success probabilities $p = \lambda/a$ and $p' = 1 - \lambda/a$. As a matter of fact, allowing for two different binomials is essential since for a single binomial $S_n \sim B(n, x)$ we have the sharper bound

$$\sigma_n\,\mathbb{P}(S_n = j) = \sqrt{n\,x(1-x)}\binom{n}{j}x^j(1-x)^{n-j} \leqslant \frac{1}{\sqrt{2e}} \tag{2.4}$$

with $1/\sqrt{2e} \sim 0.4289 < \eta$. To prove (2.4) we note that for n and j given, the maximum over $x \in [0,1]$ is attained at $x = (j+1/2)/(n+1)$, so that replacing this value all we must show is that $C^n_j \leqslant 1/\sqrt{2e}$, where

$$C^n_j = \binom{n}{j}\frac{\sqrt{n}\big(j+1/2\big)^{j+1/2}(n-j+1/2)^{n-j+1/2}}{(n+1)^{n+1}}.$$

Now $C^n_{j+1}/C^n_j = H(n-j)/H(j+1)$, where

$$H(x) = x\left(x - \frac{1}{2}\right)^{x-1/2}\Big/\left(x + \frac{1}{2}\right)^{x+1/2}$$

is decreasing, so that C^n_j decreases for $j \leqslant (n-1)/2$ and increases afterwards. Hence C^n_j is maximal at $j = 0$ or $j = n$, and then the conclusion follows since

$$C^n_0 = C^n_n = \frac{1}{\sqrt{2}}\sqrt{\frac{n}{n+1/2}}\left(1 - \frac{1}{2(n+1)}\right)^{n+1} \leqslant \frac{1}{\sqrt{2}}\exp\left(-\frac{1}{2}\right) = \frac{1}{\sqrt{2e}}.$$

2.1. Extension to more general distributions

As a consequence of Theorem 1 we see that (1.1) still holds for any random variable $S_n = \sum_{i=1}^n \pm X_i$ that can be expressed as *sums and differences* of independent Bernoullis. Moreover, the bound remains true for limits of such variables, which includes all Poisson distributions as well as infinite series $S^\infty = \sum_{i=1}^\infty X_i$ of independent Bernoullis with $\sum_{i=1}^\infty p_i < \infty$.

Corollary 2.2. *Let $S = (X + S^\infty_+) - (Y + S^\infty_-)$ with X, Y independent Poisson and S^∞_+, S^∞_- convergent series of independent Bernoullis. Then for all $j \in \mathbb{Z}$ we have $\sigma_S\,\mathbb{P}(S = j) \leqslant \eta$.* □

A natural question is whether such uniform bounds hold for more general distributions. In particular, it would be interesting to characterize the distributions that can be obtained as limits of sums and differences of Bernoullis, beyond those in Corollary 2.2. In this respect we recall the fundamental result of Kintchine [19] (see also Gnedenko and Kolmogorov [14, Theorem 2, p. 115]) which characterizes the limit distributions for sums of independent variables. The latter may or may not be Bernoullis, so this general result provides only necessary conditions for our more specific question.

3. Optimal bounds for fixed n and j

The bound η in (1.1) holds uniformly for all possible values of n and j, and is attained asymptotically with $n \to \infty$. This observation motivates the following two questions: (i) Can we get sharper bounds when n and j are fixed? and (ii) At which rate do these sharper bounds tend towards η?

In order to address these questions we consider the function $R_j^n(p) = \sigma_n P_j^n$ with

$$\sigma_n = \sqrt{\sum_{i=1}^{n} p_i(1-p_i)} \quad \text{and} \quad P_j^n = \mathbb{P}(S_n = j) = \sum_{|A|=j} \prod_{i\in A} p_i \cdot \prod_{i\notin A}(1-p_i),$$

so that the corresponding optimal bound for fixed n and j is given by

$$\eta_j^n = \max_{p\in[0,1]^n} R_j^n(p).$$

Clearly this maximum is attained and, since $R_j^n(p)$ is a symmetric function, any permutation of an optimal solution remains optimal. It can be shown that all optimal solutions lie in the interior with $0 < p_i < 1$, and that there is an optimal p which takes at most two different values $p_i \in \{\alpha, \beta\}$ for all $i = 1, \ldots, n$. The proof of these facts, which is somewhat technical and is omitted here, can be found in [4].

Figure 1 shows the typical profile of the bound η_j^n. Note the symmetry property $\eta_j^n = \eta_{n-j}^n$ which follows directly by replacing each p_i by $(1-p_i)$. The following result shows that the optimal bounds η_j^n can be bracketed between simpler 'centred' bounds of the form η_a^{2a}.

Proposition 3.1. *Let $0 \leqslant j \leqslant n$ and set $k = \min(j, n-j)$. Then $\eta_k^{2k} \leqslant \eta_j^n \leqslant \eta_n^{2n}$.*

Proof. In view of the symmetry $\eta_j^n = \eta_{n-j}^n$ it suffices to consider the case $j \leqslant n-j$ so that $k = j$. Taking $(p_1, \ldots, p_{2j})$ an optimal solution for η_j^{2j} and completing this vector with $n-j$ zeros, we have

$$\eta_j^{2j} = R_j^{2j}(p_1, \ldots, p_{2j}) = R_j^n(p_1, \ldots, p_{2j}, 0, \ldots, 0) \leqslant \eta_j^n.$$

Similarly, if $(p_1, \ldots, p_n)$ is optimal for η_j^n, then completing this vector with $n-j$ ones and n additional zeros we get

$$\eta_j^n = R_j^n(p_1, \ldots, p_n) = R_n^{2n}(p_1, \ldots, p_n, 1, \ldots, 1, 0, \ldots, 0) \leqslant \eta_n^{2n}. \qquad \square$$

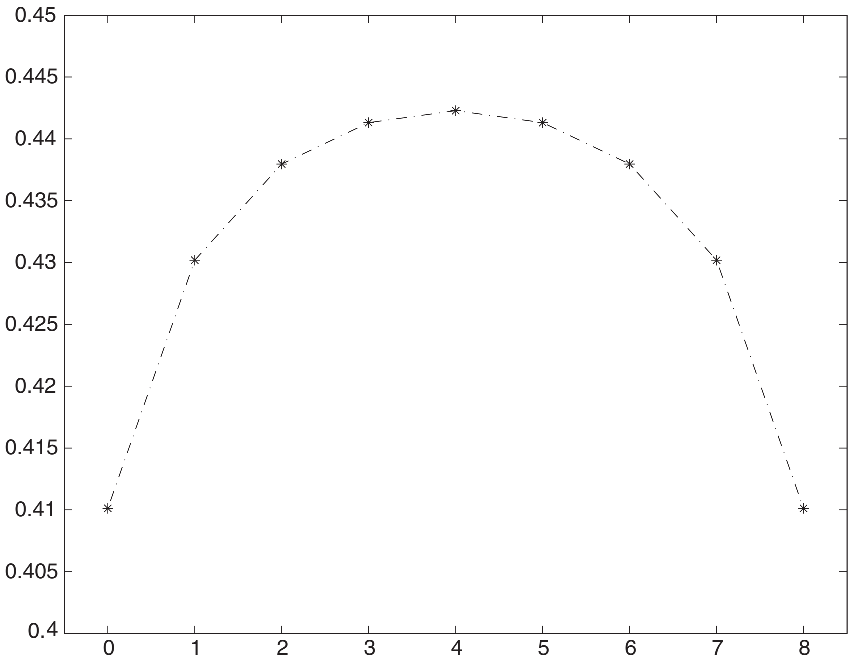

Figure 1. Profile of η_j^n for $j = 0, \ldots, n$ (here $n = 8$).

According to this result, in order to estimate η_j^n it suffices to study η_a^{2a}. The next result gives an alternative analytic expression for η_a^{2a} which provides an estimate in terms of the uniform bound η.

Theorem 3.2. *For all $a \geqslant 1$ we have*

$$\eta_a^{2a} = \max_{0 \leqslant z \leqslant 1/4} \sqrt{2az} \int_0^\pi \frac{1}{\pi} (1 - 4z \sin^2 \xi)^a \, d\xi. \tag{3.1}$$

Moreover,

$$\eta_a^{2a} > \left(1 - \frac{1}{2a}\right) \eta.$$

Proof. Proceeding as in the proof of Theorem 2.1, setting $z_i = p_i(1 - p_i) \in [0, 1/4]$, we have

$$\mathbb{P}(S_n = j) \leqslant \frac{1}{2\pi} \int_0^{2\pi} \prod_{i=1}^n \sqrt{1 + 2z_i(\cos\theta - 1)} \, d\theta,$$

so using the change of variables $\theta = 2\xi$ we get

$$R_j^n(p) \leqslant \Phi_n(z) := \sqrt{\sum_{i=1}^n z_i} \int_0^\pi \frac{1}{\pi} \prod_{i=1}^n \sqrt{1 - 4z_i \sin^2 \xi} \, d\xi. \tag{3.2}$$

Hence, denoting by ϕ_n the maximum of $\Phi_n(z)$ for $z \in [0,1/4]^n$, we have $\eta_j^n \leqslant \phi_n$. We will prove that for $n = 2a$ and $j = a$ we have $\eta_a^{2a} = \phi_{2a}$. To this end we observe that for $n = 2a$ the map $\Phi_n(z)$ has a unique maximizer $z \in [0,1/4]^n$, which is of the form $z_i = \bar{z}$ for all $i = 1, \ldots, n$ with $\bar{z} \in (0, 1/4)$. The proof of this fact is somewhat technical and we refer once again to [4]. From this it follows that

$$\phi_{2a} = \max_{0 \leqslant z \leqslant 1/4} \sqrt{2az} \int_0^\pi \frac{1}{\pi}(1 - 4z \sin^2 \xi)^a \, d\xi. \tag{3.3}$$

Now, from (3.2) we have $\eta_a^{2a} \leqslant \phi_{2a}$. To prove the reverse inequality take $\alpha \in (0, 1/2)$ the unique solution of $\alpha(1-\alpha) = \bar{z}$ and consider the vector p with half of the p_i equal to α and the other half $1 - \alpha$, so that equation (2.2) gives

$$\eta_a^{2a} \geqslant R_a^{2a}(p) = \Phi_{2a}(\bar{z}, \ldots, \bar{z}) = \phi_{2a}.$$

This proves our claim $\eta_a^{2a} = \phi_{2a}$ and establishes (3.1).

In order to prove the inequality

$$\eta_a^{2a} > \left(1 - \frac{1}{2a}\right)\eta,$$

let $x_0 \sim 0.78997786$ be the point where the maximum $\eta = \sqrt{x_0}\, e^{-x_0} I_0(x_0)$ is attained. Taking $z = x_0/2a$ in (3.3) and using the inequality $(1 - y/a)^{a-1} > e^{-y}$, which holds for all $a \geqslant 1$ and $y \in (0, 2x_0)$, we obtain

$$\phi_{2a} \geqslant \frac{\sqrt{x_0}}{\pi} \int_0^\pi \left(1 - \frac{2x_0 \sin^2 \xi}{a}\right)^a d\xi > \frac{\sqrt{x_0}}{\pi} \int_0^\pi \left(1 - \frac{2x_0 \sin^2 \xi}{a}\right) e^{-2x_0 \sin^2 \xi} \, d\xi.$$

Using the change of variables $\theta = 2\xi$, the latter can be expressed in terms of the modified Bessel functions $I_0(\cdot)$ and $I_1(\cdot)$ as

$$\begin{aligned} \phi_{2a} &> \frac{\sqrt{x_0}}{2\pi} \int_0^{2\pi} \left(1 + \frac{x_0}{a}(\cos\theta - 1)\right) e^{x_0(\cos\theta - 1)} \, d\theta \\ &= \sqrt{x_0}\, e^{-x_0} \left[\left(1 - \frac{x_0}{a}\right) I_0(x_0) + \frac{x_0}{a} I_1(x_0)\right]. \end{aligned}$$

Now, $I_1(\cdot) = I_0'(\cdot)$, while by optimality the derivative of $\sqrt{x}\, e^{-x} I_0(x)$ vanishes at x_0, which yields $x_0 I_1(x_0) = x_0 I_0'(x_0) = (x_0 - 1/2) I_0(x_0)$, and therefore we get

$$\phi_{2a} > \left(1 - \frac{1}{2a}\right) \sqrt{x_0}\, e^{-x_0} I_0(x_0) = \left(1 - \frac{1}{2a}\right)\eta. \qquad \square$$

Remark. Using the change of variables $t = \sin^2 \xi$, the integral in (3.3) can be expressed equivalently in terms of the hypergeometric function ${}_2F_1$, namely

$$\eta_a^{2a} = \phi_{2a} = \max_{0 \leqslant z \leqslant 1/4} \sqrt{2az}\, {}_2F_1\left(-a, \frac{1}{2}; 1; 4z\right).$$

As a direct consequence of the previous results, we obtain the following.

Corollary 3.3. *For each $n \geqslant 1$ and $0 < j < n$, let $k = \min(j, n-j)$. Then*

$$\left(1 - \frac{1}{2k}\right)\eta < \eta_j^n < \eta. \tag{3.4}$$

In particular, η_a^{2a} converges to η at rate $O(1/a)$. □

This result shows that the optimal bounds η_j^n are close to η, except when j is close to the extreme values 0 and n. In this respect, we note that for these extreme values a direct calculation gives explicitly

$$\eta_0^n = \eta_n^n = \sqrt{\frac{n}{2n+1}}\left(1 - \frac{1}{2(n+1)}\right)^{n+1},$$

whose limit for $n \to \infty$ is $1/\sqrt{2e}$, which is strictly smaller than η.

4. An application to fixed-point iterations

Let us illustrate how Theorem 2.1 can be used to study the rate of convergence of fixed-point iterations [20, 21]. Namely, let $(E, \|\cdot\|)$ be a normed vector space and $T : E \to E$ a non-expansive map, that is, $\|T(x) - T(y)\| \leqslant \|x - y\|$ for all $x, y \in E$, with a non-empty set of fixed points $\mathrm{Fix}(T)$. Consider the Krasnosel'skiĭ–Mann iteration

$$x_n = (1 - \alpha_n)x_{n-1} + \alpha_n T x_{n-1}$$

with $x_0 \in E$ given and $0 < \alpha_n < 1$. In [3], Baillon and Bruck conjectured the existence of a universal constant C such that

$$\|x_n - Tx_n\| \leqslant C\frac{\mathrm{dist}(x_0, \mathrm{Fix}(T))}{\sqrt{\sum_{i=1}^n \alpha_i(1-\alpha_i)}}, \tag{4.1}$$

proving this bound with $C = 2/\sqrt{\pi} \sim 1.1284$ for $\alpha_i \equiv \alpha$ constant. The general case with non-constant α_i was recently settled in [8] with this same C, while Vaisman [29] proved that it holds with $C = 1$ when E is a Hilbert space. Here we use Theorem 2.1 to find a slightly improved bound for affine maps in general normed spaces.

Proposition 4.1. *Let $T(x) = a + Lx$ with $L : E \to E$ linear and non-expansive. Then* (4.1) *holds with $C = 2\eta \sim 0.9376$, where η is given by* (1.2). □

Proof. A simple inductive argument shows that $x_n = \sum_{j=0}^n \pi_j^n T^j x_0$, where the coefficients π_j^n satisfy the recursion $\pi_j^n = \alpha_n \pi_{j-1}^{n-1} + (1-\alpha_n)\pi_j^{n-1}$. Notice that $\pi_j^n = \mathbb{P}(S_n = j)$, where $S_n = X_1 + \cdots + X_n$ is a sum of independent Bernoullis with $\mathbb{P}(X_i = 1) = \alpha_i$. In particular $\sum_{j=0}^n \pi_j^n = 1$, so that for each $y \in \mathrm{Fix}(T)$ we have $x_n - y = \sum_{j=0}^n \pi_j^n T^j(x_0 - y)$, and since $x_n - Tx_n = (x_n - y) - T(x_n - y)$ the triangle inequality implies

$$\|x_n - Tx_n\| \leqslant \sum_{j=0}^{n+1} |\pi_j^n - \pi_{j-1}^n| \|T^j(x_0 - y)\| \leqslant \|x_0 - y\| \sum_{j=0}^{n+1} |\pi_j^n - \pi_{j-1}^n|.$$

Since the distribution of S_n is unimodal, the latter sum is $2\max_j \pi_j^n$. The conclusion follows by using (1.1) and taking infimum over $y \in \mathrm{Fix}(T)$. □

Acknowledgements

We are indebted to Professor David McDonald for helpful discussions on the connection of our main result with the local limit theorem, as well as to Professor Michel Weber for pointing out the Kolmogorov–Rogozin inequality. We also thank an anonymous referee for very useful suggestions that contributed to improvement of the presentation. Roberto Cominetti was supported by FONDECYT grant 1100046 (CONICYT–Chile), as well as Nucleo Milenio Información y Coordinación en Redes ICM/FIC P10-024F. This work was completed during a visit by Jean-Bernard Baillon to Universidad de Chile, which was supported by FONDECYT grant 1130564.

References

[1] Alon, N. and Spencer, J. (1992) *The Probabilistic Method*, Wiley.

[2] Azuma, K. (1967) Weighted sums of certain dependent random variables. *Tohoku Math. J.* **19** 357–367.

[3] Baillon, J. B. and Bruck, R. (1996) The rate of asymptotic regularity is $O(1/\sqrt{n})$. In *Theory and Applications of Nonlinear Operators of Accretive and Monotone Type* (A. G. Karsatos, ed.), Vol. 178 of *Lecture Notes in Pure and Applied Mathematics*, CRC Press, pp. 51–81.

[4] Baillon, J-B., Cominetti, R. and Vaisman, J. (2013) A sharp uniform bound for the distribution of sums of Bernoulli trials. `arXiv:0806.2350v4`

[5] Bernstein, S. N. (1937) On certain modifications of Chebyshev's inequality. *Doklady Akad. Nauk SSSR* **17** 275–277.

[6] Chernoff, H. (1952) A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[7] Cominetti, R., Correa, J., Rothvoß, T. and San Martín, J. (2010) Optimal selection of customers for a last-minute offer. *Oper. Res.* **58** 878–888.

[8] Cominetti, R., Soto, J. and Vaisman, J. (2014) On the rate of convergence of Krasnosel'skiĭ–Mann iterations and their connection with sums of Bernoullis. *Israel J. Math.* **199** 757–772.

[9] Darroch, J. N. (1964) On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* **35** 1317–1321.

[10] Davis, B. and McDonald, D. (1995) An elementary proof of the local central limit theorem. *J. Theoret. Probab.* **8** 693–701.

[11] Feller, W. (1943) Generalization of a probability limit theorem of Cramér. *Trans. Amer. Math. Soc.* **54** 361–372.

[12] Figiel, T., Hitczenko, P., Johnson, W. B., Schechtman, G. and Zinn, J. (1997) Extremal properties of Rademacher functions with applications to the Kintchine and Rosenthal inequalities. *Trans. Amer. Math. Soc.* **349** 997–1027.

[13] Gamkrelidze, N. G. (1988) Application of a smoothness function in the proof of a local limit theorem (in Russian). *Teor. Veroyatnost. i Primenen.* **33** 373–376. Translation in *Theory Probab. Appl.* **33** (1989) 352–355.

[14] Gnedenko, B. V. and Kolmogorov, A. N. (1954) *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley.

[15] Godbole, A. P. and Hitczenko, P. (1998) Beyond the method of bounded differences. In *Microsurveys in Discrete Probability* (D. Aldous and J. Propp, eds), Vol. 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, AMS, pp. 43–58.

[16] Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

[17] Ibragimov, R. and Sharakhmetov, S. (1997) On an exact constant for the Rosenthal inequality. *Teor. Veroyatnost. i Primenen.* **42** 341–350.

[18] Jogdeo, K. and Samuels, S. M. (1968) Monotone convergence of binomial probabilities and a generalization of Ramanujan's equation. *Ann. Math. Statist.* **39** 1191–1195.
[19] Kintchine, A. Y. (1937) Zur Theorie der unbeschränkt teilbaren Verteilungsgesetze. *Rec. Mat* (*Mat. Sbornik*) N.S. **2** 79–120.
[20] Krasnosel'ski, M. A. (1955) Two remarks on the method of successive approximations. *Uspekhi Mat. Nauk* **10** 123–127.
[21] Mann, W. R. (1953) Mean value methods in iteration. *Proc. Amer. Math. Soc.* **4** 506–510.
[22] McDiarmid, C. (1989) On the method of bounded differences. In *Surveys in Combinatorics, 1989* (J. Siemons, ed.), Vol. 141 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, pp. 148–188.
[23] McDonald, D. (1980) On local limit theorem for integer-valued random variables. *Theory Probab. Appl.* **24** 613–619.
[24] Petrov, V. (1995) *Limit Theorems of Probability Theory*, Vol. 4 of *Oxford Studies in Probability*, Clarendon Press.
[25] Rogozin, B. A. (1961) An estimate for concentration functions. *Theory Probab. Appl.* **6** 94–97.
[26] Samuels, S. M. (1965) On the number of successes in independent trials. *Ann. Math. Statist.* **36** 1272–1278.
[27] Schechtman, G. (2007) Extremal configurations for moments of sums of independent positive random variables. In *Banach Spaces and their Applications in Analysis* (B. Randrianantoanina *et al.*, eds), de Gruyter, pp. 183–191.
[28] Utev, S. A. (1985) Extremal problems in moment inequalities. In *Limit Theorems of Probability Theory*, Trudy Inst. Mat., Nauka Sibirsk. Otdel. Novosibirsk, pp. 56–75.
[29] Vaisman, J. (2005) *Convergencia Fuerte del Método de Medias Sucesivas para Operadores Lineales No-Expansivos*, Memoria de Ingeniería Civil Matemática, Universidad de Chile.