Probab. Theory Relat. Fields 126, 51–60 (2003)
Digital Object Identifier (DOI) 10.1007/s00440-002-0252-0

Michael J. Klass · Krzysztof Nowicki

An optimal bound on the tail distribution of the number of recurrences of an event in product spaces

Received: 8 June 2001 / Revised version: 30 October 2002
Published online: 15 April 2003 – © Springer-Verlag 2003

Abstract. Let $X_1, X_2, \dots$ be independent random variables and a a positive real number. For the sake of illustration, suppose A is the event that $|X_{i+1} + \dots + X_j| \geq a$ for some integers $0 \leq i < j < \infty$. For each $k \geq 2$ we upper-bound the probability that A occurs k or more times, i.e. that A occurs on k or more disjoint intervals, in terms of $P(A)$, the probability that A occurs at least once.

More generally, let $\mathbf{X} = (X_1, X_2, \dots) \in \Omega = \prod_{j\geq 1} \Omega_j$ be a random element in a product probability space $(\Omega, \mathcal{B}, P = \otimes_{j\geq 1} P_j)$. We are interested in events $A \in B$ that are (at most contable) unions of finite-dimensional cylinders. We term such sets sequentially searchable. Let $L(A)$ denote the (random) number of disjoint intervals $(i, j]$ such that the value of $X_{(i,j]} = (X_{i+1}, \dots, X_j)$ ensures that $\mathbf{X} \in A$. By definition, for sequentially searchable A, $P(A) \equiv P(L(A) \geq 1) = P(\mathcal{N}_{-\ln(P(A^c))} \geq 1)$, where $\mathcal{N}_\gamma$ denotes a Poisson random variable with some parameter $\gamma > 0$. Without further assumptions we prove that, if $0 < P(A) < 1$, then $P(L(A) \geq k) < P(\mathcal{N}_{-\ln(P(A^c))} \geq k)$ for all integers $k \geq 2$.

An application to sums of independent Banach space random elements in l^∞ is given showing how to extend our theorem to situations having dependent components.

1. Introduction

Let $X_1, X_2, \dots, X_n$ be independent symmetric random elements taking values in a Banach space $(B, \|.\|)$. Suppose that $\|X_j\|_\infty \leq 1$ for all $j \geq 1$. Let $S_n = \sum_{j=1}^n X_j$ and $S_n^* = \max_{1\leq k\leq n} \|\sum_{j=1}^k X_j\|$ and introduce stopping times $T_i, i = 0, 1, \dots$ where $T_0 = 0$ and

$$T_{j+1} = \inf\{k \in (T_j, n] : \|S_k - S_{T_j}\| \geq a\}$$

M.J. Klass*: University of California, Departments of Mathematics and Statistics, Berkeley, CA 94720-3840, USA

K. Nowicki†: Lund University, Department of Statistics, Box 743, S-220 07 Lund, Sweden. e-mail: Krzysztof.Nowicki@stat.lu.se

*Supported by NSF Grant DMS-99-72417.

†Supported by the Swedish Research Council.

Mathematics Subject Classification (2000): Primary 60E15, 60G50

Key words or phrases: Tail probability inequalities – Hoffmann-Jørgensen inequality – Poisson bounds – Number of event recurrences – Number of entrance times – Product spaces

for all $j \geq 0$, where here and throughout the paper we will use the convention $\inf \emptyset = \infty$. Using T_1, T_2 Hoffmann-Jørgensen (1974) derived the famous inequality

$$P(S_n^* \geq 2a + 1) \leq 2P^2(S_n^* \geq a).$$

Setting

$$L = \sup\{m \geq 0 : T_m < \infty\}$$

extends the Hoffmann-Jørgensen inequality to upper-bounds of $P(S_n^* \geq ka+k-1)$ in terms of $P(L \geq k)$, i.e.,

$$P(S_n^* \geq ka + k - 1) \leq P(L \geq k) \leq 2^{k-1} P^k(S_n^* \geq a).$$

In Klass and Nowicki (2000) a much improved upper-bound of $P(L \geq k)$ was obtained, to wit

$$P(L \geq k) \leq \frac{\gamma^{k-1}}{k!}(-\ln(1 - \lambda^*))^k,$$

where $\lambda^* = P(S_n^* \geq a)$ and $\gamma = 1$ if the X_i are non-negative or $\gamma = 2$ if the X_i are symmetric.

The bound of $\frac{1}{k!}(-\ln(1 - \lambda^*))^k$ suggested to us the possible presence of a tail probability of a hidden Poisson variable. Nevertheless, the factor 2^{k-1} in the symmetric case seemed strange.

To find the concealed Poisson variable we modify the definition of λ^*. We consider $\lambda = P(\bigcup_{0 \leq i < j \leq n}\{\|S_j - S_i\| \geq a\})$. Notice that for both non-negative and symmetric random variables λ is at most $\gamma\lambda^*$ and these two quantities can be arbitrarily close. Defining the T_j's and L as before, we can now investigate $P(L \geq k)$ in terms of this λ. We write L as $L(X_1, X_2, ..., X_n)$ since it depends on these components.

Letting $n \to \infty$ produces a natural regenerative setting primed for an inductive argument: Once $n \to \infty$ and T_1 has been obtained, the event now denoted $\{L(X_1, X_2, ...) \geq k\}$ depends only on the occurrence of the event $\{L(X_{1+T_1}, X_{2+T_1}, ...) \geq k - 1 \mid T_1\}$. Then, assuming the probability of this event to be upper-bounded by the relevant Poisson random variable, the probability $P(L \geq k)$ also turns out to be Poisson bounded.

Writing the event in this manner transports us into a much more general context. What counts is only that we have constructed events depending on disjoint intervals of the X_j's, that the X_j's are independent, that we are given the numerical value of the chance that one or more intervals produces an occurrence of an event (or so-called "signal"), and want to know how to upperbound the chance that such a signal occurs at least k times.

2. Results in the independent setting

Let $(\Omega_j, \mathcal{B}_j, P_j)_{j\geq 1}$ be a sequence of probability spaces and let $(\Omega, \mathcal{B}, P)$ denote the usual product measure space. Elements of Ω will be denoted by $\mathbf{X} = (X_1, X_2, ...)$ (the X_j are thus independent, taking values in Ω_j according to probability measure P_j). For $0 \leq i < j < \infty$, put $X_{(i,j]} = (X_{i+1}, ..., X_j)$ (which is measurable with respect to $\mathcal{B}_{(i,j]} = \mathcal{B}_{i+1} \otimes \cdots \otimes \mathcal{B}_j$).

We say that $A \in \mathcal{B}$ is sequentially searchable if it is an (at most countable) union of finite-dimensional cylinders. These are the sets that can be determined by inspection of $X_{(i,j]}$ for some $1 \leq i < j < \infty$. (Recall that cylinder sets are obtained by taking finite intersections of $\mathcal{B}$-measurable sets generated by the σ-field of the individual component random variables X_i. Thus, they are generated by finite intersections of sets of the form $\{\mathbf{x} \in R^\infty : x_i \in V\}$ for real valued Borel measurable sets $V \subset R$.)

Note that not all $A \in \mathcal{B}$ are sequentially searchable. For example if $P_j(X_j = 0) = p_j$, then the singleton $A = \{(0, 0, \dots)\}$ is not sequentially searchable (and has probability $\prod_{j\geq 1} p_j$ that can be arbitrary). On the other hand, for any measurable set $\tilde{A} \in \mathcal{B}$ and $\epsilon > 0$ there is a sequentially searchable set A such that the probability of the symmetric difference of these two sets is less than ϵ.

To each searchable set and $i < j$, we associate the maximal $\sigma(X_{i+1}, \dots, X_j)$-measurable cylinder $A^\infty_{(i,j]}$ contained in A. In other words, $A^\infty_{(i,j]}$ is the largest event depending on the values of $(X_{i+1}, \dots, X_j)$ such that observing the values of these $j - i$ variables ensures that $X \in A$. When $(i, j] \subset (i', j']$, then $A^\infty_{(i,j]} \subset A^\infty_{(i',j']}$. We will also use the finite-dimensional projection of $A^\infty_{(i,j]}$

$$A_{(i,j]} = \{x_{(i,j]} \ : \ \mathbf{x} \in A^\infty_{(i,j]}\}.$$

Note that $A_{(i,j]} \in \mathcal{B}_{(i,j]}$ and $A^\infty_{(i,j]} = \{\mathbf{x} \in \mathcal{B} \ : \ x_{(i,j]} \in A_{(i,j]}\}$. Clearly

$$A = \bigcup_{0\leq i<j<\infty} A^\infty_{(i,j]}.$$

Given the single quantity $\lambda(A) := P(A)$, we wish to upper-bound (for each $k \geq 2$) the probability that there are k or more disjoint intervals $(i, j]$ such that $A^\infty_{(i,j]}$ occurs.

Set $\tau_0 = 0$ and define inductively (for all $m \geq 0$)

$$\tau_{m+1} = \min\{k > \tau_m \ : \ A^\infty_{(\tau_m,k]} \text{ occurs}\}$$

and put

$$L \equiv L(A) = \sup\{m \ : \ \tau_m < \infty\}.$$

In terms of random variables

$$\tau_{m+1} = \min\{k > \tau_m \ : \ X_{(i,k]} \in A_{(i,k]} \text{ for some } \tau_m \leq i < k\}.$$

We intend to prove the following theorem:

Theorem 1. *With the definitions and assumptions given above, suppose* $0 < \lambda(A) < 1$ *and put* $\gamma = -\ln(1-\lambda)$. *Then for every integer* $k \geq 0$,

$$P(L \geq k) \leq P(\mathcal{N}_\gamma \geq k)$$

with strict inequality for every $k \geq 2$, *where* $\mathcal{N}_\gamma$ *denotes a Poisson random variable with parameter* $\gamma > 0$.

Equivalently, if the event $\{L \geq 1\}$ has probability upper-bounded by $P(\mathcal{N}_u \geq 1)$, the probability of the event $\{L \geq k\}$ is then strictly upper-bounded by $P(\mathcal{N}_u \geq k)$ for all integers $k \geq 2$.

Let us note that this theorem is sharp: For each $0 < \lambda < 1$, the stated inequality cannot be improved by replacing $P(\mathcal{N}_\gamma \geq k)$ by $a_k P(\mathcal{N}_\gamma \geq k) + b_k$ for any sequence of constants a_k and b_k independent of the distribution of $\mathbf{X}$ and the subset A of $\mathcal{B}$: Suppose for instance that $(\mathcal{N}(t), t \geq 0)$ is a standard Poisson process so that in particular $P(\mathcal{N}(\gamma) \neq 0) = \lambda$. For all $n \geq 1$, define the random vector $\mathbf{X}_n = (X_{n,1}, \ldots, X_{n,n})$ by

$$X_{n,i} = 1_{\{\mathcal{N}(i\gamma/n) \neq \mathcal{N}((i-1)\gamma/n)\}}.$$

Clearly, for each fixed n, the variables $X_{n,1}, \ldots, X_{n,n}$ are independent. Let $A = \{\mathcal{N}(\gamma) \neq 0\}$ so that $A = A_n = \{\exists i \in \{1, \ldots, n\} \ : \ X_{n,i} \geq 1\}$ and $L(A_n)$ is the number of intervals $(i\gamma/n, (i+1)\gamma/n]$ where $\mathcal{N}$ increases. Clearly, when $n \to \infty$, the events $\{L(A_n) \geq k\}$ converge (up to negligible sets) to $\{\mathcal{N}(\gamma) \geq k\}$. Hence,

$$\lim_{n \to \infty} P(L(A_n) \geq k) = P(\mathcal{N}_\gamma \geq k).$$

Proof of the theorem. The exact distribution of L is determined by the quantities

$$q_{(i,j]} = P(A^\infty_{(i,j]} \setminus A^\infty_{(i,j-1]}) = P(A^\infty_{(i,j]}) - P(A^\infty_{(i,j-1]}).$$

In terms of random variables, $q_{(i,j]}$ is the probability that the smallest j' in $(i, j]$ for which $X_{(i',j']} \in A_{(i',j']}$ for some $i \leq i' < j'$ is $j' = j$. Thus

$$\begin{aligned}
&P(L \geq k) \\
&\quad = \sum_{0=i_0<i_1<\ldots<i_k<\infty} P(\tau_1 = i_1, \ldots, \tau_k = i_k) \\
&\quad = \sum_{0=i_0<i_1<\ldots<i_k<\infty} P(\tau_1 = i_1) P(\tau_2 = i_2 \mid \tau_1 = i_1) \ldots P(\tau_k = i_k \mid \tau_{k-1} = i_{k-1}) \\
&\quad (\text{since } P(\tau_j = i_j \mid \tau_1 = i_1, \ldots, \tau_{j-1} = i_{j-1}) = P(\tau_j = i_j \mid \tau_{j-1} = i_{j-1})) \\
&\quad = \sum_{0=i_0<i_1<\ldots<i_k<\infty} q_{(i_0,i_1]} q_{(i_1,i_2]} \cdots q_{(i_{k-1},i_k]} \\
&\quad (\text{since } P(\tau_j = i_j \mid \tau_{j-1} = i_{j-1}) = q_{(i_{j-1},i_j]}).
\end{aligned}$$

The right hand side above expresses the tail probability of L in terms of infinitely many unknowns $q_{(i,j]}$. With regard to these infinitely many unknowns, we have assumed that only one piece of information is given, namely λ, which is expressible as $\lambda = \sum_{j=1}^{\infty} q_{(0,j]}$.

We prove our theorem by induction. The case $k = 0$ is trivial. The case $k = 1$ holds since our choice of λ clearly shows that

$$P(\mathcal{N}_{-\ln(1-\lambda)} \geq 1) = \lambda = P(L \geq 1).$$

Suppose now that the result holds for $k - 1$.

Introduce $L_{(i)}$ where $L = L_{(0)}$ and $L_{(i)}$ denotes the generic L variable defined on the random vector $(X_{i+1}, X_{i+2}, ...)$ using the same fixed set A. Thus

$$P(L \geq k) = \sum_{i=1}^{\infty} q_{(0,i]} P(L_{(i)} \geq k - 1). \tag{1}$$

Inserting the inductive hypothesis

$$P(L_{(i)} \geq k - 1) \leq P(\mathcal{N}_{-\ln(1-\lambda_{>i})} \geq k - 1),$$

where

$$\lambda_{>i} \equiv P(L_{(i)} \geq 1) = \sum_{j=i+1}^{\infty} q_{(i,j]},$$

into (1) we obtain

$$P(L \geq k) \leq \sum_{i=1}^{\infty} q_{(0,i]} P(\mathcal{N}_{-\ln(1-\lambda_{>i})} \geq k - 1).$$

Upperbounding $\lambda_{>i}$ in terms of the partial sums of $q_{(0,j]}$ provides further simplification. To obtain it let $\lambda_0 = 0$ and $\lambda_i = \sum_{j=1}^{i} q_{(0,j]}$ and observe that

$$\lambda - \lambda_i = P(A \setminus A^{\infty}_{(0,i]}) \geq P(\{A^{\infty}_{(0,i]}\}^c \cap \{L_{(i)} \geq 1\}) = (1 - \lambda_i)\lambda_{>i},$$

the latter equality holding by independence. Since $q_{(0,i]} = \lambda_i - \lambda_{i-1}$, and since tail probabilities of the form $P(\mathcal{N}_\gamma \geq j)$ are increasing in γ, it follows for $k \geq 2$ that

$$\begin{aligned}
P(L \geq k) &\leq \sum_{i=1}^{\infty} (\lambda_i - \lambda_{i-1}) P(\mathcal{N}_{-\ln(1-\frac{\lambda-\lambda_i}{1-\lambda_i})} \geq k - 1) \\
&< \sum_{i=1}^{\infty} \int_{\lambda_{i-1}}^{\lambda_i} P(\mathcal{N}_{\ln((1-y)/(1-\lambda))} \geq k - 1) dy \\
&= \int_0^{\lambda} P(\mathcal{N}_{\ln((1-y)/(1-\lambda))} \geq k - 1) dy.
\end{aligned}$$

To evaluate the latter-most integral we notice that, for a Poisson process $(\mathcal{N}_t, t \geq 0)$,

$$P(\mathcal{N}_t \geq k) = \int_0^t e^{-s} P(\mathcal{N}_{t-s} \geq k - 1) ds$$

for all $k \geq 1$. Letting $t = -\ln(1 - \lambda)$ and $y = 1 - e^{-s}$, a change of variables yields

$$P(\mathcal{N}_{-\ln(1-\lambda)} \geq k) = \int_0^{\lambda} P(\mathcal{N}_{\ln((1-y)/(1-\lambda))} \geq k - 1) dy.$$

Hence

$$P(L \geq k) < P(\mathcal{N}_{-\ln(1-\lambda)} \geq k).$$

□

Let $X_1, X_2, \ldots$ be independent real valued random variables such that $S_n = X_1 + \ldots + X_n$ converges to a finite valued random variable S almost surely. For $0 \leq i < j < \infty$ let $S_{(i,j]} \equiv S_j - S_i = X_{i+1} + \ldots + X_j$ and for any real $a > 0$ let $\lambda \equiv \lambda(a) = P(\bigcup_{0 \leq i < j < \infty} \{S_{(i,j]} \geq a\})$.

The following is then an immediate corollary of Theorem 1:

Corollary 2. *In addition to the preceding conditions and notations, suppose that the X_j's are real taking values not exceeding* 1*. Then, for every integer $k \geq 1$ and every positive real a,*

$$P(\max_{1 \leq j < \infty} S_j \geq ka + (k-1)) \leq P(L \geq k) \leq P(\mathcal{N}_{-\ln(1-\lambda)} \geq k) \tag{2}$$

where $\mathcal{N}_\lambda$ denotes a Poisson variable with parameter $\lambda > 0$.

The distribution inequalities of Theorem 1 also imply an infinitude of expectation inequalities such as:

Corollary 3. $1 - \lambda(A)$ *is smaller than the following quantities (for all $a > 0$):*

$$\exp(-EL(A)) \text{ and } (E\exp(aL(A)))^{-1/(\exp(a)-1)}.$$

Proof. One just has to compare $EL(A)$ and $E\exp(aL(A))$ with the corresponding quantities for the Poisson variable $\mathcal{N}_\gamma$. □

Corollary 4. *Theorem 1 extends to an independent increment process in continuous time as long as its corresponding filtrations are right continuous.*

3. Generalization to dependent sequences

Though formally Theorem 1 holds only for product spaces with independent coordinates, its application can be substantially extended. For instance, well-known results on tangent and conditionally independent sequences can allow us to approximate a probability in question in terms of a mixture of independent variables. We illustrate this with a Banach space result from l^∞ wherein the random variables, although not necessarily their components, are independent.

Let $(B, \|.\|)$ denote l^∞. Let $\{\overrightarrow{X_j} = (X_{j1}, X_{j2}, \ldots)\}$ be a sequence of independent random variables on B. Then, let $\mathcal{F}_0$ be the trivial σ-field and, for $k \geq 1$, $\mathcal{F}_k$ denote the σ-field generated by $\{X_{jk} : j \geq 1\}$ and $\mathcal{F}_{k-1}$, and let $\mathcal{F}_\infty = \bigvee_{k=1}^\infty \mathcal{F}_k$.

Since the k^{th} component random variables $X_{1k}, X_{2k}, \ldots$ are independent, then according to Kwapień and Woyczyński (1989), conditional on $\mathcal{F}_\infty$, there exist (possibly by enlarging the probability space) mutually independent $\tilde{X}_{jk}$ for all $j, k \geq 1$ such that for each $k \geq 1$

$$\mathcal{L}(\tilde{X}_{1k}, \tilde{X}_{2k}, \ldots \mid \mathcal{F}_\infty) = \mathcal{L}(X_{1k}, X_{2k}, \ldots \mid \mathcal{F}_{k-1}).$$

Take any sequence of reals a_{jk} which is $\mathcal{F}_{k-1}$-measurable and let $S_{(i,j],k} = X_{i+1,k} + \ldots + X_{j,k}$, $\tilde{S}_{(i,j],k} = \tilde{X}_{i+1,k} + \ldots + \tilde{X}_{j,k}$, $T_{(i,j],k} = S_{(i,j],k} - a_{(i,j],k}$ and $\tilde{T}_{(i,j],k} = \tilde{S}_{(i,j]} - a_{(i,j],k}$, where $a_{(i,j],k} = a_{i+1,k} + \ldots + a_{j,k}$. For example, we could set $a_{jk} = 0$ or $a_{jk} = E(X_{j,k} \mid \mathcal{F}_{k-1})$. Let a_n be any constant such that

$$\tilde{y}_n(a_n) \equiv P(\bigcup_k \bigcup_{0 \le i < j \le n} \{\tilde{T}_{(i,j],k} \ge a_n\} \mid \mathcal{F}_\infty) < 1 \text{ a.s.} \tag{3}$$

Let $\tau_{0,k} = 0$. Having defined $\tau_{0,k}, \tau_{1,k}, \ldots, \tau_{q,k}$, let

$$\tau_{q+1,k} = \begin{cases} \min\{j \in (\tau_{q,k}, n] : T_{(i,j],k} \ge a_n \text{ for some } \tau_{q,k} \le i < j\} & \text{if such } j \text{ exists} \\ \tau_{q+1,k} = \infty & \text{otherwise.} \end{cases}$$

Let $L_k = \max\{q \ge 0 : \tau_{q,k} < \infty\}$ and define $\tilde{L}_k$ as above. Observe that $L_k \le n$ and $P(L_k \ge j \mid \mathcal{F}_{k-1}) = P(\tilde{L}_k \ge j \mid \mathcal{F}_{k-1}) = P(\tilde{L}_k \ge j \mid \mathcal{F}_\infty)$ for all $k, j \ge 1$. Hence, setting $\tilde{\lambda}_{n,k}(a_n) = P(\tilde{L}_k \ge 1 \mid \mathcal{F}_\infty)$

$$\begin{aligned}
P(\bigcup_{k=1}^\infty \{L_k \ge q\}) &\le 2P(\bigcup_{k=1}^\infty \{\tilde{L}_k \ge q\}) \\
&\quad \text{(by a result of Hitczenko (1994), see also Kwapień and Woyczyński (1989))} \\
&\le 2E \sum_{k=1}^\infty P(\tilde{L}_k \ge q \mid \mathcal{F}_\infty) \\
&\le 2E \sum_{k=1}^\infty P(\mathcal{N}_{-\ln(1-\tilde{\lambda}_{n,k}(a_n))} \ge q) \text{ (by Theorem 1)} \\
&= 2E \sum_{k=1}^\infty \tilde{\lambda}_{n,k}(a_n) \frac{P(\mathcal{N}_{-\ln(1-\tilde{\lambda}_{n,k}(a_n))} \ge q)}{\tilde{\lambda}_{n,k}(a_n)} \\
&\le 2E \max_{j \ge 1} \frac{P(\mathcal{N}_{-\ln(1-\tilde{\lambda}_{n,j}(a_n))} \ge q)}{\tilde{\lambda}_{n,j}(a_n)} \sum_{k=1}^\infty \tilde{\lambda}_{n,k}(a_n).
\end{aligned} \tag{4}$$

Also note that

$$\sum_{k=1}^\infty \tilde{\lambda}_{n,k}(a_n) \le -\ln \prod_{k=1}^\infty (1 - \tilde{\lambda}_{n,k}(a_n)) = -\ln(1 - \tilde{y}_n(a_n)).$$

Therefore,

$$P\left(\bigcup_{k=1}^\infty L_k \ge q\}\right) \le -2E\left(\left(\max_{j \ge 1} \frac{P(\mathcal{N}_{-\ln(1-\tilde{\lambda}_{n,j}(a_n))} \ge q)}{\tilde{\lambda}_{n,j}(a_n)}\right) \ln(1 - \tilde{y}_n(a_n))\right).$$

Thus we have obtained the upper bound stipulated in Theorem 5 (below) for the probability of q or more norm exceedances of a given level in terms of the probability of at least a single such norm exceedance and the largest of the componentwise ratios of the single-component conditional exceedance probabilities for q or more instances relative to one or more.

Moreover (3) could be extended by letting a_n depend upon k as well as n. Furthermore (4) could be alternatively upper-bounded in a variety of ways, multiplying and dividing by $(\tilde{\lambda}_{n,k}(a_n))^{1/p}$ followed by an application of Hölder's inequality.

Theorem 5. *Let $\overrightarrow{X}_j = (X_{j1}, X_{j2}, ...)$ be independent random elements taking values in l^∞. Let a_{jk} be reals which are $\mathcal{F}_{k-1}$-measurable as previously described. Let $V_1 \geq V_2 \geq ... \geq V_{q-1}$ denote the $q-1$ largest order statistics from $\sup_{k\geq 1} \overline{X}_{1k}$, $\sup_{k\geq 1} \overline{X}_{2k}$,..., $\sup_{k\geq 1} \overline{X}_{nk}$ where $\overline{X}_{jk} = X_{jk} - a_{jk}$. Suppose a_n satisfies (3). Then*

$$P\left(\max_{1\leq r\leq n} \sup_{k\geq 1} \sum_{j=1}^{r} \overline{X}_{jk} \geq qa_n + \sum_{j=1}^{q-1} V_j\right) \leq \min\{P\left(\mathcal{N}_{-\ln(1-\lambda_n(a_n))} \geq q\right),$$

$$-2E\left(\left(\sup_{j\geq 1} \frac{P(\mathcal{N}_{-\ln(1-\tilde{\lambda}_{n,j}(a_n))} \geq q)}{\tilde{\lambda}_{n,j}(a_n)}\right) \ln(1-\tilde{y}_n(a_n))\right)\} \tag{5}$$

where

$$\lambda_n(a_n) = P\left(\bigcup_{0\leq i<j\leq n} \bigcup_{k\geq 1} \{X_{i+1,k} + ... + X_{j,k} - a_{(i,j],k} \geq a_n\}\right),$$

$$\tilde{\lambda}_{n,k}(a_n) = P\left(\bigcup_{0\leq i<j\leq n} \{\tilde{X}_{i+1,k} + ... + \tilde{X}_{j,k} - a_{(i,j],k} \geq a_n\} \mid \mathcal{F}_{k-1}\right)$$

$$= P\left(\bigcup_{0\leq i<j\leq n} \{\tilde{X}_{i+1,k} + ... + \tilde{X}_{j,k} - a_{(i,j],k} \geq a_n\} \mid \mathcal{F}_{\infty}\right)$$

and $\tilde{y}_n(a_n)$ is defined as in (3). Moreover, if the components are independent, the factor of 2 *can be dropped.*

Remark 6. When there is no "overshoot" beyond a_n, the order statistics V_j, whose use was first suggested by Talagrand (1989), may be deleted. Not only is the right hand side in (5) often much smaller than the Hoffman-Jørgensen bound, it can be much smaller than our improvement $P(\mathcal{N}_{-\ln(1-\lambda_n(a_n))} \geq q)$, which is obtained by direct application of Theorem 1 along the lines of Corollary 2. □

The degree to which Theorem 5 can improve upon the Hoffman-Jørgensen inequality beyond the first part of the right hand side of (5) is perhaps best illustrated by the following family of examples, where (for simplicity) we assume independence of components. For this family of examples the upper-bound obtained from Theorem 5 is essentially sharp: It produces the exact asymptotic value of the tail probability in question.

Example 7. Take positive integers n and g_n and a real $\gamma_n > 0$. For $1 \le j \le n$ let $\overline{X}_j = (X_{j1}, X_{j2}, ...)$ where $X_{jk} \sim \mathcal{N}_{\gamma_n}$ for $1 \le k \le g_n$ and $X_{jk} \equiv 0$ for $k > g_n$. Further, assume that $\lim_{n\to\infty} g_n(n\gamma_n)^q = 0$ whenever $q \ge 1$ is used below. Direct calculations yield

$$P(\|\overline{S}_n\| \ge q) \sim \begin{cases} 1 - \exp(-n\gamma_n g_n) & \text{if } q = 1 \\ g_n \frac{(n\gamma_n)^q}{q!} & \text{if } q \ge 2. \end{cases}$$

This follows since

$$P(\|\overline{S}_n\| \ge 1) = 1 - (P(S_{n1} = 0))^{g_n} = 1 - \exp(-n\gamma_n g_n)$$

and, for $q \ge 2$,

$$P(\|\overline{S}_n\| \ge q) = 1 - (1 - P(S_{n1} \ge q))^{g_n} \sim g_n \frac{(n\gamma_n)^q}{q!}.$$

If we approximate $P(\|\overline{S}_n\| \ge q)$ by Theorem 1 "directly" as in Corollary 2, we learn only that

$$P(\|\overline{S}_n\| \ge q) \le P(\mathcal{N}_{-\ln(1-P(\|\overline{S}_n\|=0))} \ge q),$$

which is asymptotic to

$$(\exp(-n\gamma_n g_n)) \sum_{j=q}^{\infty} \frac{(n\gamma_n g_n)^j}{j!}.$$

If $\limsup_{n\to\infty} n\gamma_n g_n$ is positive this upper-bound does not tend to zero for any $q \ge 2$ despite the fact that the true probability is asymptotic to $\frac{g_n(n\gamma_n)^q}{q!}$, which by assumption tends to 0 as $n \to \infty$.

Using the second upper-bound in Theorem 5 and noting that we have restricted attention to the independent case:

$$-\ln(1 - \tilde{y}_n(a_n)) = -\ln P(\|\overline{S}_n\| = 0) = n\gamma_n g_n$$

and

$$\sup_{j\ge1} \frac{P(\mathcal{N}_{-\ln(1-\tilde{\lambda}_{n,j}(a_n))} \ge q)}{\tilde{\lambda}_{n,j}(a_n)} \sim \frac{(n\gamma_n)^{q-1}}{q!}.$$

Multiplying these two quantities gives the RHS of (5) in the independent case. Indeed, this result is asymptotically identical to the true probability.

Acknowledgements. The authors would like to thank the associate editor and the referee. Their insightful comments and high standards led to considerable improvement in the presentation of this paper.

References

1. Hitczenko, P.: On domination of sums of random variables by sums of conditionally independent ones. The Annals of Probability **22**(1), 453–468 (1994)
2. Hoffmann-Jørgensen, J.: Sums of independent Banach space valued random variables. Studia Mathematica **52**, 159–186 (1974)
3. Klass, M.J., Nowicki, K.: An improvement of Hoffmann-Jørgensen's inequality. The Annals of Probability **28**(2), 851–862 (2000)
4. Kwapień, S., Woyczyński, W.A.: Tangent sequences of random variables: Basic inequalities and their appliations. In Proceedings of the International Conference on Almost Everywhere Convergence in Probability and Ergodic Theory (G. A. Edgar and L. Sucheston, eds.) 1989, pp. 237–265, Academic Press, Boston
5. Talagrand, M.: Isoperimetry and integrability of the sum of independent Banach-space valued random variables. The Annals of Probability **17**(4), 1546–1570 (1989)