A More General Maximal Bernstein-type Inequality

Péter Kevei *
MTA-SZTE Analysis and Stochastics Research Group
Bolyai Institute, Aradi vértanúk tere 1, 6720 Szeged, Hungary
e-mail: kevei@math.u-szeged.hu

David M. Mason†
University of Delaware
213 Townsend Hall, Newark, DE 19716, USA
e-mail: davidm@udel.edu

May 24, 2012

Abstract

We extend a general Bernstein-type maximal inequality of Kevei and Mason (2011) for sums of random variables.

Keywords: Bernstein inequality, dependent sums, maximal inequality, mixing, partial sums.

AMS Subject Classificiation: MSC 60E15; MSC 60F05; MSC 60G10.

1 Introduction

Let $X_1, X_2, \ldots$ be a sequence of random variables, and for any choice of $1 \leq k \leq l < \infty$ we denote the partial sum $S(k,l) = \sum_{i=k}^{l} X_i$, and define $M(k,l) = \max\{|S(k,k)|, \ldots, |S(k,l)|\}$. It turns out that under a variety of assumptions the partial sums $S(k,l)$ will satisfy a generalized Bernstein-type inequality of the following form: for suitable constants $A > 0$, $a > 0$, $b \geq 0$ and $0 < \gamma < 2$ for all $m \geq 0$, $n \geq 1$ and $t \geq 0$,

$$\mathbf{P}\{|S(m+1, m+n)| > t\} \leq A \exp\left\{-\frac{at^2}{n + bt^\gamma}\right\}. \tag{1.1}$$

Kevei and Mason [2] provide numerous examples of sequences of random variables $X_1, X_2, \ldots$, that satisfy a Bernstein-type inequality of the form (1.1). They show, somewhat unexpectedly, without any additional assumptions, a modified version of it also holds for $M(1+m, n+m)$ for all $m \geq 0$ and $n \geq 1$. Here is their main result.

Theorem 1.1. *Assume that for constants $A > 0$, $a > 0$, $b \geq 0$ and $\gamma \in (0,2)$, inequality (1.1) holds for all $m \geq 0, n \geq 1$ and $t \geq 0$. Then for every $0 < c < a$ there exists a $C > 0$ depending only on A, a, b and γ such that for all $n \geq 1$, $m \geq 0$ and $t \geq 0$,*

$$\mathbf{P}\{M(m+1, m+n) > t\} \leq C \exp\left\{-\frac{ct^2}{n + bt^\gamma}\right\}. \tag{1.2}$$

*Supported by the TAMOP–4.2.1/B–09/1/KONV–2010–0005 project.
†Research partially supported by NSF Grant DMS–0503908.

There exists an interesting class of Bernstein-type inequalities that are not of the form (1.1). Here are two motivating examples.

Example 1. Assume that $X_1, X_2, \dots,$ is a stationary Markov chain satisfying the conditions of Theorem 6 of Adamczak [1] and let f be any bounded measurable function such that $Ef(X_1) = 0$. His theorem implies that for some constants $D > 0$, $d_1 > 0$ and $d_2 > 0$ for all $t \geq 0$ and $n \geq 1$,

$$P\{|S_n(f)| \geq t\} \leq D^{-1} \exp\left(-\frac{Dt^2}{nd_1 + td_2 \log n}\right), \tag{1.3}$$

where $S_n(f) = \sum_{i=1}^{n} f(X_i)$, and D/d_1 is related to the limiting variance in the central limit theorem.

Example 2. Assume that $X_1, X_2, \dots,$ is a strong mixing sequence with mixing coefficients $\alpha(n)$, $n \geq 1$, satisfying for some $d > 0$, $\alpha(n) \leq \exp(-2dn)$. Also assume that $EX_i = 0$ and for some $M > 0$, $|X_i| \leq M$, for all $i \geq 1$. Theorem 2 of Merlevède, Peligrad and Rio [4] implies that for some constant $D > 0$ for all $t \geq 0$ and $n \geq 1$,

$$P\{|S_n| \geq t\} \leq D \exp\left(-\frac{Dt^2}{nv^2 + M^2 + tM(\log n)^2}\right), \tag{1.4}$$

where $S_n = \sum_{i=1}^{n} X_i$ and $v^2 = \sup_{i>0}\left(Var(X_i) + 2\sum_{j>i} |cov(X_i, X_j)|\right)$.

The purpose of this note to establish the following extended version of Theorem 1.1 that will show that a maximal version of inequalities (1.3) and (1.4) also holds.

Theorem 1.2. *Assume that there exist constants $A > 0$ and $a > 0$ and a sequence of non-decreasing non-negative functions $\{g_n\}_{n\geq 1}$ on $(0, \infty)$, such that for all $t > 0$ and $n \geq 1$, $g_n(t) \leq g_{n+1}(t)$ and for all $0 < \rho < 1$*

$$\lim_{n\to\infty} \inf\left\{\frac{t^2}{g_n(t)\log t} : g_n(t) > \rho n\right\} = \infty, \tag{1.5}$$

where the infimum of the empty set is defined to be infinity, such that for all $m \geq 0$, $n \geq 1$ and $t \geq 0$,

$$P\{|S(m+1, m+n)| > t\} \leq A \exp\left\{-\frac{at^2}{n + g_n(t)}\right\}. \tag{1.6}$$

Then for every $0 < c < a$ there exists a $C > 0$ depending only on A, a and $\{g_n\}_{n\geq 1}$ such that for all $n \geq 1$, $m \geq 0$ and $t \geq 0$,

$$P\{M(m+1, m+n) > t\} \leq C \exp\left\{-\frac{ct^2}{n + g_n(t)}\right\}. \tag{1.7}$$

Note that condition (1.5) trivially holds when the functions g_n are bounded, since the corresponding sets are empty sets. However, in the interesting cases g_n's are not bounded, and in this case the condition basically says that $g_n(t)$ increases slower than t^2.

Essentially the same proof shows that the statement of Theorem 1.2 remains true if in the numerator of (1.6) and (1.7) the function t^2 is replaced by a regularly varying function at infinity $f(t)$ with a positive index. In this case the t^2 in condition (1.5) must be replaced by $f(t)$. Since we do not know any application of a result of this type, we only mention this generalization.

Proof. Choose any $0 < c < a$. We prove our theorem by induction on n. Notice that by the assumption, for any integer $n_0 \geq 1$ we may choose $C > An_0$ to make the statement true for all $1 \leq n \leq n_0$. This remark will be important, because at some steps of the proof we assume that n is large enough. Also since the constants A and a in (1.6) are independent of m, we can without loss of generality assume $m = 0$.
Assume the statement holds up to some $n \geq 2$. (The constant C will be determined in the course of the proof.)
Case 1. Fix a $t > 0$ and assume that

$$g_{n+1}(t) \leq \alpha\, n, \tag{1.8}$$

for some $0 < \alpha < 1$ be specified later. (In any case, we assume that $\alpha n \geq 1$.) Using an idea of [5], we may write for arbitrary $1 \leq k < n$, $0 < q < 1$ and $p + q = 1$ the inequality

$$\begin{aligned} P\{M(1, n+1) > t\} \leq & P\{M(1, k) > t\} + P\{|S(1, k+1)| > pt\} \\ & + P\{M(k+2, n+1) > qt\}. \end{aligned}$$

Let

$$u = \frac{n + g_{n+1}(qt) - q^2 g_{n+1}(t)}{1 + q^2}.$$

Note that $u \leq n - 1$ if $0 < \alpha < 1$ is chosen small enough depending on q, for n large enough. Notice that

$$\frac{t^2}{u + g_{n+1}(t)} = \frac{q^2 t^2}{n - u + g_{n+1}(qt)}. \tag{1.9}$$

Set

$$k = \lceil u \rceil. \tag{1.10}$$

Using the induction hypothesis and (1.6), keeping in mind that $1 \leq k \leq n - 1$, we obtain

$$\begin{aligned} P\{M(1, n+1) > t\} \leq & C \exp\left\{-\frac{ct^2}{k + g_k(t)}\right\} + A \exp\left\{-\frac{ap^2t^2}{k + 1 + g_{k+1}(pt)}\right\} \\ & + C \exp\left\{-\frac{cq^2t^2}{n - k + g_{n-k}(qt)}\right\} \\ \leq & C \exp\left\{-\frac{ct^2}{k + g_{n+1}(t)}\right\} + A \exp\left\{-\frac{ap^2t^2}{k + 1 + g_{n+1}(pt)}\right\} \\ & + C \exp\left\{-\frac{cq^2t^2}{n - k + g_{n+1}(qt)}\right\}. \end{aligned} \tag{1.11}$$

Notice that we chose k to make the first and third terms in (1.11) almost equal, and since by (1.10)

$$\frac{t^2}{k + g_{n+1}(t)} \leq \frac{q^2 t^2}{n - k + g_{n+1}(qt)}$$

the first term is greater than or equal to the third.
First we handle the second term in formula (1.11), showing that whenever $g_{n+1}(t) \leq \alpha n$,

$$\exp\left\{-\frac{ap^2t^2}{k + 1 + g_{n+1}(pt)}\right\} \leq \exp\left\{-\frac{ct^2}{n + 1 + g_{n+1}(t)}\right\}.$$

For this we need to verify that for $g_{n+1}(t) \leq \alpha n$,

$$\frac{ap^2}{k+1+g_{n+1}(pt)} > \frac{c}{n+1+g_{n+1}(t)}, \tag{1.12}$$

which is equivalent to

$$ap^2(n+1+g_{n+1}(t)) > c(k+1+g_{n+1}(pt)).$$

Using that

$$k = \lceil u \rceil \leq u+1 = 1 + \frac{1}{1+q^2}\left[n + g_{n+1}(qt) - q^2 g_{n+1}(t)\right],$$

it is enough to show

$$\begin{aligned} &n\left(ap^2 - \frac{c}{1+q^2}\right) + ap^2 - 2c \\ &+ \left[g_{n+1}(t)ap^2 - g_{n+1}(pt)c - \frac{c}{1+q^2}\left(g_{n+1}(qt) - q^2 g_{n+1}(t)\right)\right] > 0. \end{aligned}$$

Note that if the coefficient of n is positive, then we can choose α in (1.8) small enough to make the above inequality hold. So in order to guarantee (1.12) (at least for large n) we only have to choose the parameter p so that $ap^2 - c > 0$, which implies that

$$ap^2 - \frac{c}{1+q^2} > 0 \tag{1.13}$$

holds, and then select α small enough, keeping mind that we assume $\alpha n \geq 1$ and $k \leq n-1$. Next we treat the first and third terms in (1.11). Because of the remark above, it is enough to handle the first term. Let us examine the ratio of $C\exp\{-ct^2/(k+g_{n+1}(t))\}$ and $C\exp\{-ct^2/(n+1+g_{n+1}(t))\}$. Notice again that since $u+1 \geq k$, the monotonicity of $g_{n+1}(t)$ and $g_{n+1}(t) \leq \alpha n$ implies

$$\begin{aligned} n+1-k \geq n-u &= n - \frac{n+g_{n+1}(qt) - q^2 g_{n+1}(t)}{1+q^2} \\ &\geq \frac{q^2 n - (1-q^2)g_{n+1}(t)}{1+q^2} \\ &\geq n\frac{q^2 - \alpha(1-q^2)}{1+q^2} \\ &=: c_1 n. \end{aligned}$$

At this point we need that $0 < c_1 < 1$. Thus we choose α small enough so that

$$q^2 - \alpha(1-q^2) > 0. \tag{1.14}$$

Also we get using $g_{n+1}(t) \leq \alpha n$ the bound

$$(n+1+g_{n+1}(t))(k+g_{n+1}(t)) \leq 2n^2(1+\alpha)^2 =: c_2 n^2,$$

which holds if n large enough. Therefore, we obtain for the ratio

$$\exp\left\{-ct^2\left(\frac{1}{k+g_{n+1}(t)} - \frac{1}{n+1+g_{n+1}(t)}\right)\right\} \leq \exp\left\{-\frac{cc_1t^2}{c_2 n}\right\} \leq \mathrm{e}^{-1},$$

whenever $cc_1t^2/(c_2n) \geq 1$, that is $t \geq \sqrt{c_2n/(cc_1)}$. Substituting back into (1.11), for $t \geq \sqrt{c_2n/(cc_1)}$ and $g_{n+1}(t) \leq \alpha n$ we obtain

$$P\{M(1, n+1) > t\}$$

$$\leq \left(\frac{2}{\mathrm{e}}C + A\right)\exp\{-ct^2/(n+1+g_{n+1}(t))\} \leq C\exp\{-ct^2/(n+1+g_{n+1}(t))\},$$

where the last inequality holds for $C > A\mathrm{e}/(\mathrm{e}-2)$.
Next assume that $t < \sqrt{c_2n/(cc_1)}$. In this case choosing C large enough we can make the bound > 1, namely

$$C\exp\left\{-\frac{ct^2}{n+1+g_{n+1}(t)}\right\} \geq C\exp\left\{-\frac{cc_2n}{cc_1n}\right\} = C\mathrm{e}^{-c_2/c_1} \geq 1,$$

if $C > \mathrm{e}^{c_2/c_1}$.

Case 2. Now we must handle the case $g_{n+1}(t) > \alpha n$. Here we apply the inequality

$$P\{M(1, n+1) > t\} \leq P\{M(1, n) > t\} + P\{|S(1, n+1)| > t\}.$$

Using assumption (1.6) and the induction hypothesis, we have

$$\begin{aligned} P\{M(1, n+1) > t\} &\leq C\exp\left\{-\frac{ct^2}{n+g_n(t)}\right\} + A\exp\left\{-\frac{at^2}{n+1+g_{n+1}(t)}\right\} \\ &\leq C\exp\left\{-\frac{ct^2}{n+g_{n+1}(t)}\right\} + A\exp\left\{-\frac{at^2}{n+1+g_{n+1}(t)}\right\}. \end{aligned}$$

We will show that the right side $\leq C\exp\{-ct^2/(n+1+g_{n+1}(t))\}$. For this it is enough to prove

$$\begin{aligned} &\exp\left\{-ct^2\left(\frac{1}{n+g_{n+1}(t)} - \frac{1}{n+1+g_{n+1}(t)}\right)\right\} \\ &+ \frac{A}{C}\exp\left\{-\frac{t^2(a-c)}{n+1+g_{n+1}(t)}\right\} \leq 1. \end{aligned} \tag{1.15}$$

Using the bound following from $g_{n+1}(t) > \alpha n$ and recalling that $\alpha n \geq 1$ and $0 < \alpha < 1$, we get

$$\frac{t^2}{(n+g_{n+1}(t))(n+1+g_{n+1}(t))} \geq \frac{\alpha^2 t^2}{(1+\alpha)(1+2\alpha)g_{n+1}(t)^2} =: c_3\frac{t^2}{g_{n+1}(t)^2},$$

and

$$\frac{t^2(a-c)}{n+1+g_{n+1}(t)} \geq \frac{t^2}{g_{n+1}(t)}\frac{\alpha(a-c)}{1+2\alpha} =: \frac{t^2}{g_{n+1}(t)}c_4.$$

Choose $\delta > 0$ so small such that $0 < x \leq \delta$ implies $\mathrm{e}^{-cc_3x^2} \leq 1 - \frac{cc_3}{2}x^2$.
For $t/g_{n+1}(t) \geq \delta$ the left-hand side of (1.15) is less then

$$\mathrm{e}^{-cc_3\delta^2} + \frac{A}{C},$$

which is less than 1, for C large enough.

For $t/g_{n+1}(t) \leq \delta$ by the choice of δ the left-hand side of (1.15) is less then

$$1 - \frac{cc_3}{2} \frac{t^2}{g_{n+1}(t)^2} + \frac{A}{C} \exp\left\{ -\frac{t^2}{g_{n+1}(t)} c_4 \right\},$$

which is less than 1 if

$$\frac{cc_3}{2} \frac{t^2}{g_{n+1}(t)^2} > \frac{A}{C} \exp\left\{ -\frac{t^2}{g_{n+1}(t)} c_4 \right\}.$$

By (1.5), for any $0 < \eta < 1$ and all large enough n, $g_{n+1}(t) 1\left\{ g_{n+1}(t) > \alpha n \right\} \leq \eta t^2$, so that for all large n, whenever $g_{n+1}(t) > \alpha n$, we have

$$\frac{t^2}{g_{n+1}(t)^2} \geq t^{-2},$$

and again by (1.5) for all large n, whenever $g_{n+1}(t) > \alpha n$, $t^2/g_{n+1}(t) \geq (3/c_4) \log t$. Therefore for all large n, whenever $g_{n+1}(t)\, \alpha n$,

$$\exp\left\{ -\frac{t^2}{g_{n+1}(t)} c_4 \right\} \leq t^{-3},$$

which is smaller than $t^{-2} \frac{Ccc_3}{2A}$, for t large enough, i.e. for n large enough. The proof is complete. □

By choosing $g_n(t) = bt^\gamma$ for all $n \geq 1$ we see that Theorem 1.2 gives Theorem 1.1 as a special case. Also note that Theorem 1.2 remains valid for sums of Banach space valued random variables with absolute value $|\cdot|$ replaced by norm $||\cdot||$. Theorem 1.2 permits us to derive the following maximal versions of inequalities (1.3) and (1.4).

Application 1. In Example 1 one readily checks that the assumptions of Theorem 1.2 are satisfied with $A = D^{-1}$ and $a = D/d_1$

$$g_n(t) = \left(\frac{td_2}{d_1} \right) \log n.$$

We get the maximal version of inequality (1.3) holding for any $0 < c < 1$ and all $n \geq 1$ and $t > 0$

$$P\left\{ \left| \max_{1 \leq m \leq n} S_n(f) \right| \geq t \right\} \leq C \exp\left(-\frac{cDt^2}{nd_1 + td_2 \log n} \right), \tag{1.16}$$

for some constant $C \geq D^{-1}$ depending on c, D^{-1}, D/d_1 and $\{g_n\}_{n \geq 1}$.

Application 2. In Example 2 one can verify that the assumptions of the Theorem 1.2 hold with $A = D$ and $a = D/v^2$ and

$$g_n(t) = \frac{M^2}{v^2} + \left(\frac{tM}{v^2} \right) (\log n)^2,$$

which leads to the maximal version of inequality (1.4) valid for any $0 < c < 1$ and all $n \geq 1$ and $t > 0$

$$P\left\{ \max_{1 \leq m \leq n} |S_m| \geq t \right\} \leq C \exp\left(-\frac{cDt^2}{nv^2 + M^2 + tM (\log n)^2} \right) \tag{1.17}$$

for some constant $C \geq D$ depending on c, D/v^2 and $\{g_n\}_{n\geq 1}$. See Corollary 24 of Merlevède and Peligrad [3] for a closely related inequality that holds for all $n \geq 2$ and $t > K \log n$ for some $K > 0$.

Remark There is a small oversight in the published version of the Kevei and Mason paper. Here are the corrections that fix it.

1. Page 1057, line -9: Replace "$1 \leq k \leq n$" by "$1 \leq k < n$".
2. Page 1057, line -7: Replace this line with
$\leq \mathbf{P}\{M(1,k) > t\} + \mathbf{P}\{S(1,k+1) > pt\} + \mathbf{P}\{M(k+2,n+1) > qt\}$.
3. Page 1058: Replace "$k + bp^\gamma t^\gamma$" by "$k + 1 + bp^\gamma t^\gamma$" in equations (2.4) and (2.5), as well as in line -13.
4. Page 1058: Replace "$ap^2 - c$" by "$ap^2 - 2c$" in line -9.

Acknowledgment

We thank a referee for a careful reading of the manuscript and a number of useful comments.

References

[1] R. Adamczak, *A tail inequality for suprema of unbounded empirical processes with applications to Markov chains.* Electron. J. Probab. **13** (2008), 1000–1034.

[2] P. Kevei and D.M. Mason, *A note on a maximal Bernstein inequality.* Bernoulli **17** (2011), 1054–1062.

[3] F. Merlevède and M. Peligrad, *Rosenthal-type inequalities for the maximum of partial sums of stationary processes and examples.* Ann. Probab. To appear.

[4] F. Merlevède, M. Peligrad, M. and E. Rio, *Bernstein inequality and moderate deviations under strong mixing conditions.* In: High Dimensional Probability V: The Luminy Volume, C. Houdré, V. Koltchinskii, D. M. Mason and M. Peligrad, eds., (Beachwood, Ohio, USA: IMS, 2009), 273–292.

[5] F.A. Móricz, R.J. Serfling and W.F. Stout, *Moment and probability bounds with quasisuperadditive structure for the maximum partial sum.* Ann. Probab. **10** (1982), 1032–1040.