Chapter 2
Maximal Inequalities for Multi-Indexed Sums of Independent Random Variables

Bounds for distributions and moments of sums of independent random variables play a crucial role in the proofs of limit theorems (especially for almost sure convergence). One of the first results of this kind is *Chebyshev's inequality*.

Theorem 2.1 (P. L. Chebyshev) *Let* $\{X_k\}$ *be independent random variables*, $n \ge 1$, $S_n = X_1 + \cdots + X_n$. *Assume that, for all* $k \le n$,

$$\mathsf{E}[X_k] = 0, \quad \mathsf{E}\left[X_k^2\right] = \sigma_k^2 < \infty.$$

Then for all $x > 0$

$$\mathsf{P}(|S_n| \ge x) \le \frac{\sigma_1^2 + \cdots + \sigma_n^2}{x^2}.$$

This result holds even if the maximum of cumulative sums substitutes S_n. The corresponding result is called *Kolmogorov's inequality*.

Theorem 2.2 (A. N. Kolmogorov) *Let* $\{X_k\}$ *be independent random variables*, $n \ge 1$, *and* $S_n = X_1 + \cdots + X_n$. *Assume that, for all* $k \le n$,

$$\mathsf{E}[X]_k = 0, \quad \mathsf{E}[X]_k^2 = \sigma_k^2 < \infty.$$

Then for all $x > 0$

$$\mathsf{P}\left(\max_{k \le n} |S_k| \ge x\right) \le \frac{\sigma_1^2 + \cdots + \sigma_n^2}{x^2}. \tag{2.1}$$

Kolmogorov's inequality can also be proved for weighted sums S_k/b_k (in this form, it is called *Hájek–Rényi's inequality*).

Theorem 2.3 (J. Hájek, A. Rényi) *Let* $\{X_k\}$ *be independent random variables*, $n \ge 1$, *and* $S_n = X_1 + \cdots + X_n$. *Assume that, for all* $k \le n$,

O. Klesov, *Limit Theorems for Multi-Indexed Sums of Random Variables*,
Probability Theory and Stochastic Modelling 71, DOI 10.1007/978-3-662-44388-0_2

$$\mathsf{E}[X_k] = 0, \quad \mathsf{E}\left[X_k^2\right] = \sigma_k^2 < \infty.$$

Also let $\{b_k\}$ be an increasing sequence of positive numbers. Then for all $x > 0$

$$\mathsf{P}\left(\max_{k\le n} \frac{|S_k|}{b_k} \ge x\right) \le \frac{1}{x^2}\sum_{k=1}^{n} \frac{\sigma_k^2}{b_k^2}. \tag{2.2}$$

Another generalization of Theorem 2.2, called *Doob's inequality*, provides a bound for the second moment of the maximum of cumulative sums.

Theorem 2.4 (J. Doob) *Let $\{X_k\}$ be independent random variables, $n \ge 1$, and $S_n = X_1 + \cdots + X_n$. Assume that, for all $k \le n$,*

$$\mathsf{E}[X_k] = 0, \quad \mathsf{E}\left[X_k^2\right] = \sigma_k^2 < \infty.$$

Then

$$\mathsf{E}\left[\max_{k\le n} S_k^2\right] \le 4(\sigma_1^2 + \cdots + \sigma_n^2). \tag{2.3}$$

Yet another classical result is *Lévy's inequality* for symmetric random variables.

Theorem 2.5 (P. Lévy) *Let $\{X_k\}$ be independent symmetric random variables. Then*

$$\mathsf{P}\left(\max_{k\le n} S_k \ge x\right) \le 2\mathsf{P}(S_n \ge x) \tag{2.4}$$

for all $x > 0$.

The methods used to prove Theorems 2.2–2.5 for $d = 1$ are similar and based on the *first time* when the sequence $S_1, \ldots, S_n$ crosses a level x from below (in other words, enters an interval $[x, \infty]$). Since the "first time" when a field $\{S(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ crosses a level x is not well defined for $d > 1$, the methods of proof of the above inequalities also change for $d > 1$.

The analogs of the above inequalities for multi-indexed sums are collected and proved in this chapter. These results will be used in the later chapters of the book.

2.1 Estimates for Moments of Multi-Indexed Sums

In this section, we consider a sequence $\{X_k, 1 \le k \le n\}$ of independent random variables and a field $\{X(\mathbf{n}), \mathbf{n} \in \mathbf{N}^d\}$ of independent random variables. Some of the results presented below also hold for certain classes of dependent random variables.

We will indicate all such cases where the random variables $\{X_k\}$ or $\{X(\mathbf{n})\}$ are not necessarily independent but still admit the corresponding inequality.

Let $n > 1$ and put $S_1 = X_1$, $S_n = X_1 + \cdots + X_n$. Assume that $r > 0$ and

$$\mathsf{E}\left[|X_k|^r\right] < \infty, \quad 1 \le k \le n.$$

Similar assumptions are made concerning the random fields $\{X(\mathbf{n}), \mathbf{n} \in \mathbf{N}^d\}$, that is we assume that

$$\mathsf{E}\left[|X(\mathbf{k})|^r\right] < \infty, \quad \mathbf{k} \le \mathbf{n}.$$

2.1.1 The Bahr–Esseen and Dharmadhikari–Jogdeo Inequalities

The triangle inequality implies that

$$|S_n|^r \le |X_1|^r + \cdots + |X_n|^r$$

for $0 < r \le 1$. Passing to the expectations on both sides of the latter inequality we get

$$\mathsf{E}\left[|S_n|^r\right] \le \sum_{k=1}^{n} \mathsf{E}\left[|X_k|^r\right]. \tag{2.5}$$

Inequality (2.5) holds for $r = 2$, as well, if $\mathsf{E}[X_k] = 0$, $1 \le k \le n$. Moreover, it becomes an equality in this case. The case $1 < r < 2$ is more delicate.

Theorem 2.6 (B. Bahr, C.-G. Esseen) *Let X_k, $k \le n$, be independent random variables with $\mathsf{E}[X_k] = 0$, $k \le n$. If $1 \le r \le 2$, then*

$$\mathsf{E}\left[|S_n|^r\right] \le \left(2 - \frac{1}{n}\right) \sum_{k=1}^{n} \mathsf{E}\left[|X_k|^r\right]. \tag{2.6}$$

For convenience, we refer to the constant on the right-hand side of inequalities (2.5) and (2.6) as the *Bahr–Esseen constant*. Thus the Bahr–Esseen constant equals 1 if $0 < r \le 1$ or $r = 2$ and $2 - \frac{1}{n}$ if $1 < r < 2$.

The case $r > 2$ is the most complicated one.

Theorem 2.7 (S.W. Dharmadhikari, K. Jogdeo) *Let X_k, $k \le n$, be independent random variables with $\mathsf{E}[X_k] = 0$, $k \le n$. If $r \ge 2$, then*

$$\mathsf{E}\left[|S_n|^r\right] \le J_r n^{\frac{r}{2}-1} \sum_{k=1}^{n} \mathsf{E}\left[|X_k|^r\right], \tag{2.7}$$

where

$$J_r = \frac{r(r-1)}{2} \max\left(1, 2^{r-3}\right)\left(1 + \frac{2}{r} K_{2m}^{(r-2)/2m}\right), \quad K_{2m} = \sum_{i=1}^{m} \frac{i^{2m-1}}{(i-1)!},$$

and where the number m *is defined by the condition* $2m \le r < 2m + 2$.

For convenience, we refer to J_r as the *Dharmadhikari–Jogdeo constant*.

Remark 2.1 Any attempt to prove inequality (2.7) by "elementary" reasoning (with the help of Hölder's inequality, for example) is doomed to failure. Indeed, let $r > 2$ and p be such that $\frac{1}{p} + \frac{1}{r} = 1$. Then Hölder's inequality implies that

$$\left|\sum_{k=1}^{n} X_k\right| \le n^{1/p}\left(\sum_{k=1}^{n} |X_k|^r\right)^{1/r},$$

whence

$$\mathsf{E}\left[|S_n|^r\right] \le n^{r/p} \sum_{k=1}^{n} \mathsf{E}\left[|X_k|^r\right].$$

This bound is worse than (2.7) as far as the order of growth of n is concerned, since

$$\frac{r}{p} = r - 1 > \frac{r}{2} - 1.$$

The preceding inequalities can be rewritten for multi-indexed sums as follows.

Theorem 2.8 *Let* $\{X(\mathbf{n})\}$ *be a field of independent random variables. Let* $\mathfrak{L} \subset \mathbf{N}^d$ *be an arbitrary finite subset. Let either* $0 < r < 1$ *or* $r \ge 1$ *and* $\mathsf{E}[X_k] = 0$ *for all* $\mathbf{k} \in \mathfrak{L}$. *Then*

$$\mathsf{E}\left[\left|\sum_{\mathbf{k}\in\mathfrak{L}} X(\mathbf{k})\right|^r\right] \le C_r \sum_{\mathbf{k}\in\mathfrak{L}} \mathsf{E}\left[|X(\mathbf{k})|^r\right], \quad 0 < r \le 2, \tag{2.8}$$

$$\mathsf{E}\left[\left|\sum_{\mathbf{k}\in\mathfrak{L}} X(\mathbf{k})\right|^r\right] \le C_r \left(\operatorname{card}(\mathfrak{L})\right)^{\frac{r}{2}-1} \sum_{\mathbf{k}\in\mathfrak{L}} \mathsf{E}\left[|X(\mathbf{k})|^r\right], \quad r > 2, \tag{2.9}$$

where C_r *equals the Bahr–Esseen constant for* $0 < r \le 2$ *and the Dharmadhikari–Jogdeo constant for* $r > 2$, *that is*

$$C_r = \begin{cases} 1, & 0 < r \le 1 \text{ or } r = 2, \\ 2, & 1 < r < 2, \\ J_r, & r > 2. \end{cases} \tag{2.10}$$

2.1.2 Rosenthal's Inequality

Let $r \geq 2$. Another bound for the expectation $\mathsf{E}\left[|S_n|^r\right]$ is written in terms of the moments of the random variables $\{X_k\}$ only and without any multiplicative terms such as $n^{\frac{r}{2}-1}$.

Theorem 2.9 (H. Rosenthal) *Let X_k, $k \leq n$, be independent random variables. Assume that $r \geq 2$ and*

$$\mathsf{E}\left[X_k\right] = 0, \quad 1 \leq k \leq n.$$

Then

$$\mathsf{E}\left[|S_n|^r\right] \leq R_r \left[\sum_{k=1}^{n} \mathsf{E}\left[|X_k|^r\right] + B_n^{r/2} \right], \tag{2.11}$$

where $B_n = \sum_{k=1}^{n} \mathsf{E}\left[X_k^2\right]$, $R_r = 2^{r^2}$.

For convenience, we refer to R_r as the *Rosenthal constant*.

Rosenthal's inequality for multi-indexed sums is written as follows.

Theorem 2.10 *Let $\{X(\mathbf{n})\}$ be a field of independent random variables. Assume that $r \geq 2$. Let $\mathfrak{L} \subset \mathbf{N}^d$ be an arbitrary finite set and let*

$$\mathsf{E}\left[X(\mathbf{k})\right] = 0, \quad \mathbf{k} \in \mathfrak{L}.$$

Then

$$\mathsf{E}\left[\left|\sum_{\mathbf{k}\in\mathfrak{L}} X(\mathbf{k})\right|^r\right] \leq R_r \left[\sum_{\mathbf{k}\in\mathfrak{L}} \mathsf{E}\left[|X(\mathbf{k})|^r\right] + B^{r/2}(\mathfrak{L}) \right], \tag{2.12}$$

where

$$B(\mathfrak{L}) = \sum_{\mathbf{k}\in\mathfrak{L}} \mathsf{E}\left[X^2(\mathbf{k})\right].$$

2.1.3 The Marcinkiewicz–Zygmund Inequality

Yet another bound for moments of sums of independent random variables is due to Marcinkiewicz and Zygmund.

Theorem 2.11 (J. Marcinkiewicz, A. Zygmund) *Let X_i, $1 \leq i \leq n$, be independent random variables. Let $r \geq 2$. Assume that $\mathsf{E}\left[X_i\right] = 0$ and $\mathsf{E}\left[|X_i|^r\right] < \infty$, $1 \leq i \leq n$. Then*

$$Z_r' \mathsf{E}\left[\left(\sum_{i=1}^{n} X_i^2\right)^{r/2}\right] \leq \mathsf{E}\left[|S_n|^r\right] \leq Z_r'' \mathsf{E}\left[\left(\sum_{i=1}^{n} X_i^2\right)^{r/2}\right], \tag{2.13}$$

where Z_r' *and* Z_r'' *are universal constants that depend only on* r.

We refer to the constants Z_r' and Z_r'' as the *upper and lower Marcinkiewicz–Zygmund constants*.

An analog of inequality (2.13) for multi-indexed sums is given below.

Theorem 2.12 *Let* $\{X(\mathbf{n})\}$ *be a field of independent random variables. Let* $\mathfrak{L} \subset \mathbf{N}^d$ *be a finite set and* $r \geq 2$. *Assume that* $\mathsf{E}\left[|X(\mathbf{k})|^r\right] < \infty$ *and* $\mathsf{E}\left[X(\mathbf{k})\right] = 0$ *for all* $\mathbf{k} \in \mathfrak{L}$. *Then*

$$Z_r' \mathsf{E}\left[\left(\sum_{\mathbf{k}\in\mathfrak{L}} X^2(\mathbf{k})\right)^{r/2}\right] \leq \mathsf{E}\left[\left|\sum_{\mathbf{k}\in\mathfrak{L}} X(\mathbf{k})\right|^r\right] \leq Z_r'' \mathsf{E}\left[\left(\sum_{\mathbf{k}\in\mathfrak{L}} X^2(\mathbf{k})\right)^{r/2}\right], \tag{2.14}$$

where Z_r' *and* Z_r'' *are the upper and lower Marcinkiewicz–Zygmund constants.*

2.1.4 Skorokhod's Inequality

An interesting generalization of inequality (2.11) can be obtained from *Skorokhod's inequality*.

Theorem 2.13 (A. V. Skorokhod) *Let* $\{X_i\}$ *be independent bounded random variables*, $|X_i| \leq 1$. *Assume that* $\mathsf{P}(|S_n| \geq a) \leq 1/8e$ *for some* a. *Then*

$$\mathsf{E}\left[|S_n|^r\right] \leq \mathscr{S}_r (a+1)^r \tag{2.15}$$

for any positive integer r, *where* $\mathscr{S}_r$ *is a universal constant that depends only on* r.

We refer to $\mathscr{S}_r$ as the *Skorokhod constant*.

Remark 2.2 One can estimate the number a in Theorem 2.13 with the help of the Markov–Chebyshev inequality. Indeed, if $t \geq 1$, then $a = \left(8e\mathsf{E}\left[|S_n|^t\right]\right)^{1/t}$ fits the restriction in Theorem 2.13. For this a, we get from inequality (2.15)

$$\mathsf{E}\left[|S_n|^r\right] \leq \mathscr{S}_r \left(\left(8e\mathsf{E}\left[|S_n|^t\right]\right)^{1/t} + 1\right)^r.$$

Thus if n is such that $\mathsf{E}\left[|S_n|^t\right] \geq 1/8e$, then

$$\mathsf{E}\left[|S_n|^r\right] \leq 2^r \mathscr{S}_r (8e)^{r/t} \left(\mathsf{E}\left[|S_n|^t\right]\right)^{r/t}. \tag{2.16}$$

The restriction $|X_i| \leq 1$, $i \leq n$, in Theorem 2.13 can easily be replaced by a more general one $|X_i| \leq b$, $i \leq n$, for some constant $b > 0$. This is proved by passing to the random variables $X_k' \stackrel{\text{def}}{=} X_k/b$ and applying inequality (2.16) for X_k'. It is worth mentioning that inequality (2.16) does not change at all in this case.

If $r \le t$, then inequality (2.16) is too rough, since

$$\left(\mathsf{E}\left[|S_n|^r\right]\right)^{1/r} \le \left(\mathsf{E}\left[|S_n|^t\right]\right)^{1/t}$$

by *Lyapunov's inequality*. Otherwise, namely if $r > t$, then (2.16) is a new result.

Put, for example, $t = 2$ in inequality (2.16). Then, for $r > 2$,

$$\mathsf{E}\left[|S_n|^r\right] \le \mathscr{S}'_r \left(\sum_{i=1}^n \mathsf{E}\left[X_i^2\right]\right)^{r/2}$$

if $\mathsf{E}\left[S_n^2\right] \ge 1/8e$, where $\mathscr{S}'_r = 2^r \mathscr{S}_r (8e)^{r/2}$. The latter inequality, in contrast to Rosenthal's inequality (2.11), does not contain higher moments of the random variables $\{X_i\}$ on the right-hand side.

Inequality (2.16) can be used for values other than $t = 2$. For example, let $1 \le t \le 2$. Assuming that $\mathsf{E}[X_i] = 0$, $i \le n$, we derive from the Bahr–Esseen inequality (2.6) that

$$\mathsf{E}\left[|S_n|^r\right] \le \mathscr{S}_{r,t} \left(\sum_{i=1}^n \mathsf{E}\left[|X_i|^t\right]\right)^{r/t},$$

where $\mathscr{S}_{r,t} = 2^r (16e)^{r/t} \mathscr{S}_r$.

Skorokhod's inequality for multi-indexed sums reads as follows.

Theorem 2.14 *Let $\{X(\mathbf{n})\}$ be a field of independent random variables. Let $\mathfrak{L} \subset \mathbf{N}^d$ be an arbitrary finite set and $r \ge 2$ be a positive integer, $1 \le t \le 2$. Also let $\mathsf{E}[X(\mathbf{k})] = 0$ and $|X(\mathbf{k})| \le b$ almost surely for some non-random constant $b > 0$ and for all $\mathbf{k} \in \mathfrak{L}$. If $\mathsf{E}\left[\left|\sum_{\mathbf{k}\in\mathfrak{L}} X(\mathbf{k})\right|^t\right] \ge 1/8e$, then*

$$\mathsf{E}\left[\left|\sum_{\mathbf{k}\in\mathfrak{L}} X(\mathbf{k})\right|^r\right] \le \mathscr{S}_{r,t} \left(\sum_{\mathbf{k}\in\mathfrak{L}} \mathsf{E}\left[|X(\mathbf{k})|^t\right]\right)^{r/t}, \tag{2.17}$$

where $\mathscr{S}_{r,t}$ is the universal constant defined above.

Remark 2.3 It may appear that the assumption that the terms in the sum are bounded restricts the possible applications of Skorokhod's inequality. This is not the case for problems related to the almost sure asymptotic behavior of normalized sums of independent random variables, since a common trick in the proofs of such results is to truncate the random variables so that after a point we are dealing with bounded terms and Skorokhod's inequality applies.

Finally we note that the restriction $\mathsf{E}\left[|S_n|^r\right] \ge 1/8e$ is not critical at all. Instead one can assume that

$$\theta \overset{\text{def}}{=} \inf_n \mathsf{E}\left[|S_n|^r\right] > 0.$$

In this case, the constant $\mathscr{S}_{r,t}$ in (2.17) changes and depends on θ.

2.2 Maximal Inequalities for Distributions of Multi-Indexed Sums

2.2.1 A Generalization of Petrov's Inequality

Let $0 < q < 1$. Recall that a number $\varkappa_q = \varkappa_q(X)$ is called a *q-quantile* of a random variable X if

$$\mathsf{P}(X \le \varkappa_q) \ge q, \quad \mathsf{P}(X \ge \varkappa_q) \ge 1 - q.$$

A q-quantile is called a *median* if $q = \frac{1}{2}$.

Lemma 2.1 describes some key properties of quantiles.

Lemma 2.1 *Every random variable either possesses a unique q-quantile or the set of its q-quantiles coincides with a closed interval in the real line.*

If $\varkappa_q$ is a q-quantile of a random variable X, then $-\varkappa_q$ is a $(1-q)$-quantile of the random variable $-X$.

The following is a known inequality for $d = 1$ due to V. V. Petrov.

Theorem 2.15 (V. V. Petrov) *Let $0 < q < 1$. If $X_1, \ldots, X_n$ are independent random variables, then*

$$\mathsf{P}\left(\max_{k \le n}[S_k - \varkappa_q(S_k - S_n)] \ge x\right) \le q^{-1}\mathsf{P}(S_n \ge x) \tag{2.18}$$

for all $x \in \mathbf{R}$.

We refer to (2.18) as *Petrov's inequality*.

Our aim in this section is to generalize inequality (2.18) to the case $d > 1$. Let $\mathbf{n} \in \mathbf{N}^d$ be a fixed multi-index. Given $\mathbf{k} \le \mathbf{n}$, introduce $d + 1$ multi-indices $\mathbf{k}_0, \ldots, \mathbf{k}_d$: the "end" members of this sequence are defined by

$$\mathbf{k}_0 = \mathbf{k}, \quad \mathbf{k}_d = \mathbf{n}.$$

The construction is complete if $d = 1$. Let $d \ge 2$. If $0 < j < d$ is given, then the first j coordinates of the multi-index $\mathbf{k}_j$ coincide with the corresponding coordinates of the element $\mathbf{n}$, while the last $d - j$ coordinates of $\mathbf{k}_j$ coincide with those of the multi-index $\mathbf{k}$:

$$\mathbf{k}_j = (n_1, \ldots, n_j, k_{j+1}, \ldots, k_d), \quad 1 \le j \le d - 1. \tag{2.19}$$

The definition of the multi-indices $\mathbf{k}_1$ and $\mathbf{k}_{d-1}$ include the expressions "$n_1, \ldots, n_1$" and "$k_d, \ldots, k_d$", which are to be understood as just "n_1" and "k_d", respectively.

The procedure of creating a multi-index $\mathbf{k}_j$ from multi-indices $\mathbf{k} \le \mathbf{n}$ can be repeated several times. For example, $\left(\mathbf{k}_0\right)_1 = (k_1, k_2, \ldots, k_d)_1 = (n_1, k_2, \ldots, k_d)$, $\left(\mathbf{k}_1\right)_1 = (n_1, k_2, \ldots, k_d)_1 = (n_1, k_2, \ldots, k_d) = \mathbf{k}_1$,

$$\left(\mathbf{k}_1\right)_2 = (n_1, k_2, \dots, k_d)_2 = (n_1, n_2, k_3, \dots, k_d)$$

etc.

Let $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be independent random variables. As usual, their sums are denoted by $\{S(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$. Put

$$\Gamma(\mathbf{k}) = \sum_{j=1}^{d} \varkappa_q(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)). \tag{2.20}$$

Note that the number $\Gamma(\mathbf{k})$ depends not only on $\mathbf{k}$, but also on the multi-index $\mathbf{n}$. In fact, $\Gamma(\mathbf{k}) = \Gamma(\mathbf{k}, \mathbf{n})$.

Theorem 2.16 *Let* $0 < q < 1$. *If* $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *are independent random variables, then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}}[S(\mathbf{k}) - \Gamma(\mathbf{k})] \ge x\right) \le q^{-d}\mathsf{P}(S(\mathbf{n}) \ge x) \tag{2.21}$$

for all $x \in \mathbf{R}$, *where one can choose arbitrary quantiles* $\varkappa_q(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j))$ *to form the number* $\Gamma(\mathbf{k})$.

Proof Inequality (2.21) coincides with Petrov's inequality (2.18) if $d = 1$. We prove (2.21) for $d \ge 2$ by induction on the dimension of the space $\mathbf{N}^d$. Assume that Theorem 2.16 holds for all spaces $\mathbf{N}^b$ with $b < d$ and let us prove it for $b = d$.

Let $\mathbf{n} = (n_1, \dots, n_d)$. Given $1 \le j \le d$, we recursively define the sequence of random variables $\{\mathscr{M}_j(\ell), \ell \le n_j\}$ and the random number I_j, namely if the random variables $\{\mathscr{M}_j(\ell), \ell \le n_j\}$ are already defined for a given $1 \le j \le d$, then the random number I_j is chosen such that

$$I_j = \begin{cases} \min\{\ell : \mathscr{M}_j(\ell) \ge x\}, & \text{if } \max_{\ell\le n_j} \mathscr{M}_j(\ell) \ge x, \\ 0, & \text{if } \max_{\ell\le n_j} \mathscr{M}_j(\ell) < x. \end{cases} \tag{2.22}$$

The sequences $\{\mathscr{M}_j(\ell), \ell \le n_j\}$ are defined as follows: for $j = 1$,

$$\mathscr{M}_1(\ell) = \max_{\substack{k_2\le n_2 \\ \dots \\ k_d\le n_d}} [S(\ell, k_2, \dots, k_d) - \Gamma(\ell, k_2, \dots, k_d)];$$

for $2 \le j < d$ (clearly, this case happens only if $d \ge 3$),

$$\begin{aligned} \mathscr{M}_j(\ell) = \max_{\substack{k_{j+1}\le n_{j+1} \\ \dots \\ k_d\le n_d}} &[S(I_1, \dots, I_{j-1}, \ell, k_{j+1}, \dots, k_d) \\ &- \Gamma(I_1, \dots, I_{j-1}, \ell, k_{j+1}, \dots, k_d)]; \end{aligned}$$

and finally for $j = d$,

$$\mathscr{M}_d(\ell) = S(I_1, \dots, I_{d-1}, \ell) - \Gamma(I_1, \dots, I_{d-1}, \ell).$$

In the definition of $\mathscr{M}_2(\ell)$ and $\mathscr{M}_{d-1}(\ell)$, the expressions "$I_1, \dots, I_1$" and "$k_d, \dots, k_d$" are to be understood as "I_1" and "k_d", respectively. The sequences $\{\mathscr{M}_j(\ell), \ell \le n_j\}$ are well defined if we agree that $\Gamma(\mathbf{k}) = 0$ and $S(\mathbf{k}) = 0$ in the case where at least one of the coordinates of the element $\mathbf{k}$ equals 0.

Then we introduce the random events $\{\mathscr{I}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ and $\{E(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ as follows:

$$\begin{aligned} \mathscr{I}(\mathbf{k}) &= \{\omega : I_1 = k_1, \dots, I_d = k_d\}, \\ E(\mathbf{k}) &= \{\omega : S(\mathbf{k}_1) - S(\mathbf{k}_0) - \varkappa_{1-q}(S(\mathbf{k}_1) - S(\mathbf{k}_0)) \ge 0\}. \end{aligned}$$

The events $\mathscr{I}(\mathbf{k}), \mathbf{k} \le \mathbf{n}$, are disjoint, that is $\mathscr{I}(\mathbf{k}) \cap \mathscr{I}(\mathbf{m}) = \varnothing$ if $\mathbf{k} \ne \mathbf{m}$. Moreover

$$\bigcup_{\mathbf{k} \le \mathbf{n}} \mathscr{I}(\mathbf{k}) = \left\{ \max_{\mathbf{k} \le \mathbf{n}} [S(\mathbf{k}) - \Gamma(\mathbf{k})] \ge x \right\}.$$

Note that the random event $\mathscr{I}(\mathbf{k})$ for every $\mathbf{k} \le \mathbf{n}$ is defined in terms of random variables $X(\mathbf{m})$ whose indices $\mathbf{m} = (m_1, \dots, m_d)$ are such that $m_1 \le k_1$, while the random event $E(\mathbf{k})$ is expressed in terms of the random variables $X(\mathbf{m})$ whose indices are such that $m_1 > k_1$. Thus $\mathscr{I}(\mathbf{k})$ and $E(\mathbf{k})$ are pairwise independent for all $\mathbf{k} \le \mathbf{n}$. Moreover, the definition of a $(1-q)$-quantile implies that

$$\mathsf{P}(E(\mathbf{k})) \ge q \quad \text{for all} \quad \mathbf{k} \le \mathbf{n}.$$

Thus

$$q\mathsf{P}\left(\max_{\mathbf{k} \le \mathbf{n}} [S(\mathbf{k}) - \Gamma(\mathbf{k})] \ge x \right) = q \sum_{\mathbf{k} \le \mathbf{n}} \mathsf{P}(\mathscr{I}(\mathbf{k})) \le \sum_{\mathbf{k} \le \mathbf{n}} \mathsf{P}(\mathscr{I}(\mathbf{k}) \cap E(\mathbf{k})).$$

For $\omega \in \mathscr{I}(\mathbf{k})$, we have $S(\mathbf{k}) \ge \Gamma(\mathbf{k}) + x$. Recall $\mathbf{k}_0 = \mathbf{k}$ and thus

$$S(\mathbf{k}_1) - \varkappa_{1-q}(S(\mathbf{k}_1) - S(\mathbf{k}_0)) \ge \Gamma(\mathbf{k}) + x$$

for $\omega \in \mathscr{I}(\mathbf{k}) \cap E(\mathbf{k})$. Since

$$\varkappa_{1-q}(S(\mathbf{k}_1) - S(\mathbf{k}_0)) + \Gamma(\mathbf{k}) = \Gamma(\mathbf{k}_1),$$

we get $S(\mathbf{k}_1) - \Gamma(\mathbf{k}_1) \ge x$ for $\omega \in \mathscr{I}(\mathbf{k}) \cap E(\mathbf{k})$. Therefore

$$
\begin{aligned}
q\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}}[S(\mathbf{k})-\Gamma(\mathbf{k})]\ge x\right)
&\le \sum_{\mathbf{k}\le\mathbf{n}} \mathsf{P}(\mathscr{I}(\mathbf{k}), S(\mathbf{k}_1)-\Gamma(\mathbf{k}_1)\ge x)\\
&\le \sum_{\mathbf{k}\le\mathbf{n}} \mathsf{P}\left(\mathscr{I}(\mathbf{k}), \max_{\substack{m_2\le n_2\\ \cdots\\ m_d\le n_d}}[S(\mathbf{m}_1)-\Gamma(\mathbf{m}_1)]\ge x\right)\\
&\le \mathsf{P}\left(\max_{\substack{k_2\le n_2\\ \cdots\\ k_d\le n_d}}[S(\mathbf{k}_1)-\Gamma(\mathbf{k}_1)]\ge x\right).
\end{aligned}
\tag{2.23}
$$

Now we use the inductive assumption. Put

$$
X'(\widetilde{\mathbf{k}})=\sum_{\ell=1}^{n_d} X(\ell,\widetilde{\mathbf{k}}),\quad S'(\widetilde{\mathbf{k}})=\sum_{\widetilde{\mathbf{m}}\le\widetilde{\mathbf{k}}} X'(\widetilde{\mathbf{m}}),
$$

where $\widetilde{\mathbf{k}}=(k_2,\dots,k_d)$ and $\widetilde{\mathbf{m}}=(m_2,\dots,m_d)$. The random variables $\{X'(\widetilde{\mathbf{k}}),\widetilde{\mathbf{k}}\le\widetilde{\mathbf{n}}\}$ are independent. Using the sums $S'(\widetilde{\mathbf{k}})$, we define the numbers $\Gamma'(\widetilde{\mathbf{k}})$ according to (2.20) but for elements $\widetilde{\mathbf{k}}$ of the space $\mathbf{N}^{d-1}$ rather than for elements of the space $\mathbf{N}^d$. By the inductive assumption,

$$
\mathsf{P}\left(\max_{\widetilde{\mathbf{k}}\le\widetilde{\mathbf{n}}}[S'(\widetilde{\mathbf{k}})-\Gamma'(\widetilde{\mathbf{k}})]\ge x\right)\le q^{-(d-1)}\mathsf{P}(S'(\widetilde{\mathbf{n}})\ge x).
$$

It is clear that $S'(\widetilde{\mathbf{k}})=S(\mathbf{k}_1)$, $\Gamma'(\widetilde{\mathbf{k}})=\Gamma(\mathbf{k}_1)$, and $S'(\widetilde{\mathbf{n}})=S(\mathbf{n})$. Thus (2.23) implies

$$
q\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}}[S(\mathbf{k})-\Gamma(\mathbf{k})]\ge x\right)\le q^{-(d-1)}\mathsf{P}(S(\mathbf{n})\ge x)
$$

and this completes the proof. □

Corollary 2.1 *Assume that there are a constant $c\ge 0$ and q-quantiles of the random variables $S(\mathbf{k}_{j-1})-S(\mathbf{k}_j)$ such that*

$$
\varkappa_q(S(\mathbf{k}_{j-1})-S(\mathbf{k}_j))\le c\quad \textit{for all}\quad \mathbf{k}\le\mathbf{n}\ \textit{and}\ 1\le j\le d.
$$

Then

$$
\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k})\ge x\right)\le q^{-d}\mathsf{P}(S(\mathbf{n})\ge x-dc)
\tag{2.24}
$$

for all $x\in\mathbf{R}$.

Proof Since $\Gamma(\mathbf{k}) \le dc$, we get

$$\left\{\omega\colon \max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k}) \ge x\right\} \subseteq \left\{\omega\colon \max_{\mathbf{k}\le\mathbf{n}}[S(\mathbf{k}) - \Gamma(\mathbf{k})] \ge x - dc\right\}$$

for all $x \in \mathbf{R}$. Thus inequality (2.24) follows from (2.21). □

Corollary 2.2 *If for some constants $c \ge 0$ and $q > 0$*

$$\mathsf{P}(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j) \le c) \ge q \quad \text{for all } j = 1, \dots, d \text{ and } \mathbf{k} \le \mathbf{n},$$

then inequality (2.24) holds for all $x \in \mathbf{R}$.

To prove Corollary 2.2 we need the following auxiliary results.

Lemma 2.2 *Let $0 < q < 1$. If $\mathsf{P}(X \le a) \ge q$ for some $a \in \mathbf{R}$, then there exists a quantile $\varkappa_q(X)$ such that $\varkappa_q(X) \le a$. If $\mathsf{P}(X \ge b) \ge q$, then there exists a quantile $\varkappa_{1-q}(X)$ such that $\varkappa_{1-q}(X) \ge b$.*

Proof of Lemma 2.2 We prove only the first assertion of the lemma (the second one is proved similarly). Let $\mathsf{P}(X \le a) \ge q$. Then $\mathsf{P}(X > a) \le 1 - q$. Let $\varkappa_q^*(X)$ be the minimal q-quantile of the random variable X. We have

$$\mathsf{P}(X \ge \varkappa_q^*(X)) \ge 1 - q \ge \mathsf{P}(X > a).$$

This implies that either $\varkappa_q^*(X) \le a$ or $\mathsf{P}(a < X < \varkappa_q^*(X)) = 0$. In the latter case every number from the interval $(a, \varkappa_q^*(X))$ is a q-quantile of the random variable X, that is $\varkappa_q^*(X)$ is not the minimal q-quantile. Thus $\varkappa_q^*(X) \le a$. □

Proof of Corollary 2.2 Lemma 2.2 implies that there exist q-quantiles such that $\varkappa_q(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)) \le c$ for all $\mathbf{k} \le \mathbf{n}$ and $1 \le j \le d$. Thus Corollary 2.2 follows from Corollary 2.1 with the same quantiles. □

2.2.2 Lévy's Inequality

Lévy's inequality is a particular case of inequality (2.18) corresponding to $q = \frac{1}{2}$, that is to medians. Put

$$M(\mathbf{k}) = \sum_{j=1}^{d} \operatorname{med}(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)),$$

where $\operatorname{med}(\xi)$ is a median of a random variable ξ.

Corollary 2.3 *Let* $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *be independent random variables. Then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} [S(\mathbf{k}) - M(\mathbf{k})] \ge x \right) \le 2^d \mathsf{P}(S(\mathbf{n}) \ge x) \tag{2.25}$$

for all $x \in \mathbf{R}$.

Proof Inequality (2.25) follows from inequality (2.21) with $q = \frac{1}{2}$. □

Corollary 2.4 *Let* $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *be independent symmetric random variables. Then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k}) \ge x \right) \le 2^d \mathsf{P}(S(\mathbf{n}) \ge x), \tag{2.26}$$

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{n})| \ge x \right) \le 2^d \mathsf{P}(|S(\mathbf{n})| \ge x) \tag{2.27}$$

for all $x \in \mathbf{R}$.

Proof One of the medians $\mu(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j))$ equals zero in the case of symmetric random variables $\{X(\mathbf{k})\}$ for all $j = 1, \dots, d$ and $\mathbf{k} \le \mathbf{n}$. Thus inequality (2.26) follows directly from (2.25) where one takes the zero medians. Applying (2.26) to the random variables $\{-X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ we obtain (2.27). □

Remark 2.4 Corollary 2.4 generalizes Lévy's inequality (2.4) to the case $d > 1$. The constant 2^d on the right-hand side is optimal for $d = 1$. This follows by considering the case of Bernoulli random variables $X_1, \dots, X_n$ and non-integer x, since both sides of inequality (2.4) coincide in this case. The question of whether or not the constant 2^d is optimal for $d > 1$ is not as simple, however the answer is positive in this case, too.

According to Wichura's invariance principle [419]

$$\frac{1}{n^{d/2}} S([nt_1], \dots, [nt_d]) \quad \Longrightarrow \quad W(t_1, \dots, t_d),$$

in the space $D[0,1]^d$, where W is the so-called Chentsov–Yeh random field. In particular,

$$\lim_{n\to\infty} \mathsf{P}\left(\frac{1}{n^{d/2}} \max_{\mathbf{k}\le(n,\dots,n)} S(\mathbf{k}) \ge x \right) = \mathsf{P}\left(\sup_{t_1,\dots,t_d \in [0,1]^d} W(t_1, \dots, t_d) \ge x \right).$$

It is shown in Chap. 2 of [341] that

$$\mathsf{P}\left(\sup_{t_1,\dots,t_d \in [0,1]^d} W(t_1, \dots, t_d) \ge x \right) \sim 2^d \mathsf{P}(W(1, \dots, 1) \ge x) \qquad \text{as } x \to \infty.$$

This means that the inequality

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{n}) \ge x\right) \le c\mathsf{P}(S(\mathbf{n}) \ge x)$$

does not hold for any $c < 2^d$ if x and $n_1, \dots, n_d$ are sufficiently large.

On the other hand it is still unknown whether or not the latter inequality with $c < 2^d$ may hold for small x, or small $n_1, \dots, n_d$, or in the case where the second moments do not exist.

2.2.3 Ottaviani's Inequality

One of the achievements of probability theory in the first half of the 20th century is Lévy's theorem on the equivalence of almost sure convergence and convergence in probability of a series of independent random variables (Theorem 5.2 is an analog of Lévy's theorem for multi-indexed sums). One of the proofs of Levy's theorem is based on the following result, known as *Ottaviani's inequality*.

Theorem 2.17 (G. Ottaviani) *If*

$$\mathsf{P}(|S_k - S_n| \ge \varepsilon) \le \alpha, \qquad \textit{for all} \quad k = 0, \dots, n-1,$$

for some $\varepsilon > 0$ *and* $0 < \alpha < 1$, *then*

$$\mathsf{P}\left(\max_{k\le n} |S_k| \ge 2\varepsilon\right) \le \frac{\alpha}{1-\alpha}. \tag{2.28}$$

One can prove a generalization of Ottaviani's inequality for $d > 1$ with the help of inequality (2.21).

Corollary 2.5 *Assume that*

$$\mathsf{P}(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j) > \varepsilon) \le \alpha \qquad \textit{for all} \quad j = 1, \dots, d \;\; \textit{and} \quad \mathbf{k} \le \mathbf{n}, \tag{2.29}$$

for some $\varepsilon > 0$ *and* $0 < \alpha < 1$, *where the multi-indices* $\mathbf{k}_j$ *are defined by equality* (2.19). *Then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k}) \ge (d+1)\varepsilon\right) \le (1-\alpha)^{-d}\mathsf{P}(S(\mathbf{n}) \ge \varepsilon). \tag{2.30}$$

Moreover, if

$$\mathsf{P}(|S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)| > \varepsilon) \le \alpha \quad \textit{for all}\ j = 1, \dots, d \;\; \textit{and} \quad \mathbf{k} \le \mathbf{n},$$

then

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})| \ge (d+1)\varepsilon \right) \le (1-\alpha)^{-d}\mathsf{P}(|S(\mathbf{n})| \ge \varepsilon). \tag{2.31}$$

Proof The assumption of the corollary implies that

$$\mathsf{P}(S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j) \le \varepsilon) \ge 1-\alpha \qquad \text{for all} \quad j = 1, \dots, d \ \text{ and } \quad \mathbf{k} \le \mathbf{n}.$$

Thus Corollary 2.2 with $c = \varepsilon$, $x = \varepsilon$, and $q = 1-\alpha$ implies inequality (2.30).

The same reasoning for random variables $-X(\mathbf{k})$ yields

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} [-S(\mathbf{k})] \ge (d+1)\varepsilon \right) \le (1-\alpha)^{-d}\mathsf{P}(-S(\mathbf{n}) \ge \varepsilon).$$

Since

$$\left\{\omega\colon \max_{\mathbf{k}\in A} S(\mathbf{k}) \ge z\right\} \cup \left\{\omega\colon \max_{\mathbf{k}\in A} [-S(\mathbf{k})] \ge z\right\} = \left\{\omega\colon \max_{\mathbf{k}\in A} |S(\mathbf{k})| \ge z\right\}$$

for all $A \subseteq \mathbf{N}^d$ and $z > 0$, inequality (2.31) is also proved. □

The most useful variant of Corollary 2.5 is given below.

Corollary 2.6 *Assume that condition* (2.29) *holds for some* $\varepsilon > 0$ *and* $0 < \alpha < \frac{1}{d}$. *Then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})| \ge (d+1)\varepsilon \right) \le \frac{\alpha}{1-d\alpha}. \tag{2.32}$$

Proof Corollary 2.6 follows from Corollary 2.5 in view of the Bernoulli inequality: $1 - dt \le (1-t)^d$, $0 \le t \le 1$. □

2.2.4 Kolmogorov's Inequality for Probabilities

The proof of the law of the iterated logarithm for $d = 1$ is based on the so-called *Kolmogorov inequality*.

Theorem 2.18 (A. N. Kolmogorov) *If* $\mathsf{E}[X_i] = 0$ *and* $\mathsf{E}\left[X_i^2\right] < \infty$ *for all* $i = 1, \dots, n$, *then*

$$\mathsf{P}\left(\max_{k\le n} S_n \ge x \right) \le 2\mathsf{P}\left(S_n \ge x - \sqrt{2B_n} \right) \tag{2.33}$$

for all $x \in \mathbf{R}$, *where* $B_n = \mathsf{E}\left[S_n^2\right]$.

An analogous inequality for the case $d > 1$ can be derived from Corollary 2.2.

Corollary 2.7 *Let* $\mathsf{E}[X(\mathbf{k})] = 0$ *and* $\mathsf{E}\left[X^2(\mathbf{k})\right] = 0$ *for all* $\mathbf{k} \le \mathbf{n}$. *Put* $B(\mathbf{n}) = \mathsf{E}\left[S^2(\mathbf{n})\right]$. *Then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k}) \ge x\right) \le 2^d \mathsf{P}\left(S(\mathbf{n}) \ge x - d\sqrt{2B(\mathbf{n})}\right), \tag{2.34}$$

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})| \ge x\right) \le 2^d \mathsf{P}\left(|S(\mathbf{n})| \ge x - d\sqrt{2B(\mathbf{n})}\right) \tag{2.35}$$

for all $x \in \mathbf{R}$.

An even more general result can be proved for which Corollary 2.7 is the particular case corresponding to $q = \frac{1}{2}$ and $r = 2$.

Corollary 2.8 *Let* $\mathbf{n} \in \mathbf{N}^d$, $0 < q < 1$, *and* $r > 0$. *Assume that*

$$\mathsf{E}\left[|X(\mathbf{k})|^r\right] < \infty \quad \textit{for all } \mathbf{k} \le \mathbf{n}.$$

Put

$$M_r(\mathbf{n}) = \sum_{\mathbf{k}\le\mathbf{n}} \mathsf{E}\left[|X(\mathbf{k})|^r\right].$$

If $0 < r \le 1$, *then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k}) \ge x\right) \le q^{-d} \mathsf{P}\left(S(\mathbf{n}) \ge x - d(1-q)^{-1/r} M_r^{1/r}(\mathbf{n})\right) \tag{2.36}$$

for all $x \in \mathbf{R}$.

If $1 \le r \le 2$ *and* $\mathsf{E}[X(\mathbf{k})] = 0$ *for all* $\mathbf{k} \le \mathbf{n}$, *then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k}) \ge x\right) \le q^{-d} \mathsf{P}\left(S(\mathbf{n}) \ge x - d(1-q)^{-1/r} \left(c_r M_r(\mathbf{n})\right)^{1/r}\right) \tag{2.37}$$

for all $x \in \mathbf{R}$, *where*

$$c_r = \begin{cases} 1, & \textit{for } r = 1 \textit{ or } r = 2, \\ 2, & \textit{for } 1 < r < 2. \end{cases}$$

Finally, if $r \ge 1$, *then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} S(\mathbf{k}) \ge x\right) \le q^{-d} \mathsf{P}\left(S(\mathbf{n}) \ge x - d(1-q)^{-1/r} \left(\mathsf{E}\left[|S(\mathbf{n})|^r\right]\right)^{1/r}\right) \tag{2.38}$$

for all $x \in \mathbf{R}$.

Proof For any $t > 0$,

$$\mathsf{P}\left(|S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)| \geq t(1-q)^{-1/r}\right) \leq (1-q)\frac{\mathsf{E}\left[|S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)|^r\right]}{t^r} \quad (2.39)$$

by the Chebyshev–Markov inequality for all $j = 1, \ldots, d$ and $\mathbf{k} \leq \mathbf{n}$. To prove (2.36) we choose $t = (M_r(\mathbf{n}))^{1/r}$ and use the inequality

$$\mathsf{E}\left[|S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)|^r\right] \leq M_r(\mathbf{n})$$

that results from $|x + y|^r \leq |x|^r + |y|^r$ for $0 < r \leq 1$.

To prove (2.37) we choose $t = (c_r M_r(\mathbf{n}))^{1/r}$ and use the inequality

$$\mathsf{E}\left[|S(\mathbf{k}_{j-1}) - S(\mathbf{k}_j)|^r\right] \leq c_r M_r(\mathbf{n})$$

that follows from the Bahr–Esseen inequality (2.8).

Finally, to prove (2.38) we choose $t = \left(\mathsf{E}\left[|S(\mathbf{n})|^r\right]\right)^{1/r}$ and use the inequality

$$\mathsf{E}\left[|S(\mathbf{k}_j) - S(\mathbf{k}_{j-1})|^r\right] \leq \mathsf{E}\left[|S(\mathbf{n})|\right]^r,$$

which is a consequence of the martingale property of the sequence $\zeta_j = S(k_j)$.

In either case,

$$\mathsf{P}\left(S(\mathbf{k}_j) - S(\mathbf{k}_{j-1}) \geq -t(1-q)^{-1/r}\right) \geq q$$

and thus Corollary 2.8 follows from Corollary 2.2. □

2.3 Maximal Inequalities

2.3.1 A Generalization of Kolmogorov's Inequality for Moments

An analog of inequality (2.1) for $d > 1$ can be proved by using the results of the preceding section. Moreover, this method allows us to obtain upper bounds not only for probabilities but also for moments as in Theorem 2.4.

Corollary 2.9 *Let* $\mathbf{n} \in \mathbf{N}^d$ *and* $r > 0$. *If* $\{X(\mathbf{k}), \mathbf{k} \leq \mathbf{n}\}$ *are independent symmetric random variables such that* $\mathsf{E}\left[|X(\mathbf{k})|^r\right] < \infty$ *for all* $\mathbf{k} \leq \mathbf{n}$, *then*

$$\mathsf{E}\left[\max_{\mathbf{k}\leq\mathbf{n}} |S(\mathbf{k})|^r\right] \leq 2^d \mathsf{E}\left[|S(\mathbf{n})|^r\right]. \quad (2.40)$$

Proof First we multiply inequality (2.27) by rx^{r-1} and integrate the result in the interval $[0, \infty)$. Using the representation for the moment of the random variable ξ

$$\mathsf{E}\left[|\xi|^r\right] = r\int_0^\infty x^{r-1}\mathsf{P}(|\xi| \geq x)\,dx$$

(see Sect. A.19, we prove inequality (2.40)). □

Corollary 2.10 *Let* $\mathbf{n} \in \mathbf{N}^d$ *and* $r > 0$. *Let* $\{X(\mathbf{k}), \mathbf{k} \leq \mathbf{n}\}$ *be independent random variables such that* $\mathsf{E}\left[|X(\mathbf{k})|^r\right] < \infty$ *for all* $\mathbf{k} \leq \mathbf{n}$. *Then*

$$\mathsf{E}\left[\max_{\mathbf{k}\leq\mathbf{n}} |S(\mathbf{k})|^r\right] \leq \sum_{\mathbf{k}\leq\mathbf{n}} \mathsf{E}\left[|X(\mathbf{k})|^r\right] \tag{2.41}$$

in the case when $0 < r \leq 1$.

If $r \geq 1$ *and* $\mathsf{E}\,[X(\mathbf{k})] = 0$ *for all* $\mathbf{k} \leq \mathbf{n}$, *then*

$$\mathsf{E}\left[\max_{\mathbf{k}\leq\mathbf{n}} |S(\mathbf{k})|^r\right] \leq 2^{d+r}\mathsf{E}\left[|S(\mathbf{n})|^r\right]. \tag{2.42}$$

Moreover, under the same assumptions

$$\mathsf{E}\left[\max_{\mathbf{k}\leq\mathbf{n}} |S(\mathbf{k})|^r\right] \leq 2^{d+r+1}\sum_{\mathbf{k}\leq\mathbf{n}} \mathsf{E}\left[|X(\mathbf{k})|^r\right] \tag{2.43}$$

for $1 \leq r \leq 2$ *and*

$$\mathsf{E}\left[\max_{\mathbf{k}\leq\mathbf{n}} |S(\mathbf{k})|^r\right] \leq 2^{d+r} J_r |\mathbf{n}|^{-1+r/2}\sum_{\mathbf{k}\leq\mathbf{n}} \mathsf{E}\left[|X(\mathbf{k})|^r\right] \tag{2.44}$$

for $r \geq 2$, *where* J_r *is the Dharmadhikari–Jogdeo constant* (2.9).

Remark 2.5 As seen from the proof below, the constant on the right-hand side of (2.43) equals 2^{d+2} if $r = 2$. In fact, this constant can be decreased up to 2^{d+1} in this case.

Proof The inequality (2.41) for $0 < r \leq 1$ follows from the bound

$$|S(\mathbf{k})|^r \leq \sum_{\mathbf{m}\leq\mathbf{n}} \mathsf{E}\left[|S(\mathbf{m})|^r\right] \qquad \text{for all } \mathbf{k} \leq \mathbf{n}.$$

To prove inequality (2.42) for $r \geq 1$, we first apply inequality (2.40) for symmetrized $\{S^{(s)}(\mathbf{k}), \mathbf{k} \leq \mathbf{n}\}$:

$$\mathsf{E}\left[\max_{\mathbf{k}\leq\mathbf{n}} |S^{(s)}(\mathbf{k})|^r\right] \leq 2^d\mathsf{E}\left[|S^{(s)}(\mathbf{n})|^r\right].$$

Then we estimate the right-hand side by the Hölder inequality

$$\mathsf{E}\left[|S^{(s)}(\mathbf{n})|^r\right] \le 2^r \mathsf{E}\left[|S(\mathbf{n})|^r\right].$$

Note that if $r = 2$, then, in fact, the constant on the right-hand side equals 2.

To complete the proof of inequality (2.42), we denote by $\mathfrak{F}$ the σ-algebra generated by the random variables $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$. Then we deduce from Jensen's inequality for conditional expectations that

$$\begin{aligned}\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} |S^{(s)}(\mathbf{k})|^r\right] &= \mathsf{E}\left[\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} |S^{(s)}(\mathbf{k})|^r / \mathfrak{F}\right]\right] \ge \mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} \mathsf{E}\left[|S^{(s)}(\mathbf{k})|^r / \mathfrak{F}\right]\right] \\ &\ge \mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} \left|\mathsf{E}\left[S^{(s)}(\mathbf{k}) / \mathfrak{F}\right]\right|^r\right] = \mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})|^r\right],\end{aligned}$$

since $\mathsf{E}\left[S(\mathbf{k})\right] = 0$.

Finally, to prove inequalities (2.43) and (2.44), we use (2.42) and apply the Bahr–Esseen inequality (2.8) if $1 \le r \le 2$ or the Dharmadhikari–Jogdeo inequality (2.9) if $r \ge 2$. □

Kolmogorov's inequality (2.1) follows from Corollary 2.10 with $r = 2$ and Chebyshev's inequality.

Corollary 2.11 *Let $\mathbf{n} \in \mathbf{N}^d$ and let $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be independent random variables such that $\mathsf{E}\left[X(\mathbf{k})\right] = 0$ and $\mathsf{E}\left[X^2(\mathbf{k})\right] < \infty$ for all $\mathbf{k} \le \mathbf{n}$. Then*

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})| \ge x\right) \le \frac{2^{d+1}}{x^2} \sum_{\mathbf{k}\le\mathbf{n}} \mathsf{E}\left[X^2(\mathbf{k})\right] \tag{2.45}$$

for all $x > 0$.

Further bounds for the expectation of the maximum of sums of independent random variables can be derived from inequality (2.42). For example, the following result follows from Rosenthal's inequality (2.12).

Corollary 2.12 *Let $\mathbf{n} \in \mathbf{N}^d$, $r \ge 2$, and let $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be independent random variables such that $\mathsf{E}\left[X(\mathbf{k})\right] = 0$ and $\mathsf{E}\left[|X(\mathbf{k})|^r\right] < \infty$ for all $\mathbf{k} \le \mathbf{n}$. Then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})|^r\right] \le 2^{d+r} R_r \left[\sum_{\mathbf{k}\le\mathbf{n}} \mathsf{E}\left[|X(\mathbf{k})|^r\right] + (\mathsf{var}\left[S(\mathbf{n})\right])^{r/2}\right], \tag{2.46}$$

where R_r is the Rosenthal constant.

2.4 A Generalization of the Hájek–Rényi Inequality

Two methods can be used to establish a generalization of the Hájek–Rényi inequality (2.2) for $d > 1$. The first allows us to obtain results for normalizing fields with non-negative increments (see Definition A.11). The second method works for a wider class of monotone normalizing fields and has the advantage that it leads to better constants in the bounds. We show how the first method works in the proof of the Hájek–Rényi version of Kolmogorov's inequality.

2.4.1 First Method

Let $\mathbf{n} \in \mathbf{N}^d$ and let a field $\{b(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ have non-negative increments. Consider a field $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ of independent random variables such that $\mathsf{E}[X(\mathbf{k})] = 0$ and $\mathsf{E}\left[X^2(\mathbf{k})\right] < \infty$ for all $\mathbf{k} \le \mathbf{n}$. According to Proposition A.8,

$$\max_{\mathbf{k}\le\mathbf{n}} \left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right| \le 2^d \max_{\mathbf{k}\le\mathbf{n}} \left|\sum_{\mathbf{m}\le\mathbf{k}} \frac{X(\mathbf{m})}{b(\mathbf{m})}\right|. \tag{2.47}$$

Now we square both sides of the latter inequality and pass to the mathematical expectation. Then we apply inequality (2.43) with $r = 2$ to the right-hand side:

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} \left|\sum_{\mathbf{m}\le\mathbf{k}} \frac{X(\mathbf{m})}{b(\mathbf{m})}\right|^2\right] \le 2^{d+1} \sum_{\mathbf{k}\le\mathbf{n}} \frac{\mathsf{E}\left[X^2(\mathbf{k})\right]}{b^2(\mathbf{k})}. \tag{2.48}$$

Combining all the above results with the Chebyshev inequality we get, for $x > 0$,

$$\mathsf{P}\left(\max_{\mathbf{k}\le\mathbf{n}} \left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right| \ge x\right) \le \frac{2^{2d+1}}{x^2} \sum_{\mathbf{k}\le\mathbf{n}} \frac{\mathsf{E}\left[X^2(\mathbf{k})\right]}{b^2(\mathbf{k})}.$$

This is an analog of inequality(2.2) for $d > 1$.

Inequality (2.47) holds in the general case, that is not only for independent random variables. On the other hand, the second step of the proof of inequality (2.48) requires the property of independence.

2.4.2 Second Method

The second method allows us to study monotone normalizing fields $\{b(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ (see Definition A.11). Recall that the property of monotonicity of a field is less restrictive than the property of non-negativity of its increments (see Proposition A.7). Moreover, the second method also works for dependent random variables.

Let $\mathbf{n} \in \mathbf{N}^d$, $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be random variables, and let $\{b(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be an increasing field of positive real numbers such that $b(\mathbf{1}) > 0$. Let $c > 1$ and let the sets A_t, $t \ge 0$, be defined as follows:

$$A_t = \{\mathbf{k} \le \mathbf{n} : b(\mathbf{k}) \le c^t b(\mathbf{1})\}. \tag{2.49}$$

It is clear that $A_0 = \{\mathbf{k} : b(\mathbf{k}) = b(\mathbf{1})\}$.

Let $r > 0$ and let $\{\lambda(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be a field of non-negative numbers for which

$$\mathsf{E}\left[\max_{\mathbf{k}\in A_t} |S(\mathbf{k})|\right]^r \le \sum_{\mathbf{k}\in A_t} \lambda(\mathbf{k}) \qquad \text{for all } t \ge 0. \tag{2.50}$$

A field $\{\lambda(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ satisfying condition (2.50) exists for every field of random variables $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$. Indeed, put

$$g(t) = \mathsf{E}\left[\max_{\mathbf{k}\in A_t} |S(\mathbf{k})|^r\right]$$

and, for every $t \ge 0$, choose an arbitrary point $\mathbf{k}_t \in A_t \setminus A_{t-1}$ if the set $A_t \setminus A_{t-1}$ is non-empty. It is obvious that the field

$$\lambda(\mathbf{k}) = \begin{cases} g(t), & \text{if } \mathbf{k} = \mathbf{k}_t, \\ 0, & \text{if } \mathbf{k} \ne \mathbf{k}_t \end{cases}$$

satisfies condition (2.50).

Note that the choice of $\{\lambda(\mathbf{k})\}$ described above is not unique. Moreover, this choice is not optimal in many cases. Below we will consider other fields $\{\lambda(\mathbf{k})\}$ as well.

Theorem 2.19 *Let* $\mathbf{n} \in \mathbf{N}^d$, $r > 0$, *and let* $\{b(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *be an increasing field of positive real numbers. If condition* (2.50) *holds for some* $c > 1$ *and some non-negative numbers* $\{\lambda(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$, *then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} \left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le \frac{c^{2r}}{c^r - 1} \sum_{\mathbf{k}\le\mathbf{n}} \frac{\lambda(\mathbf{k})}{b^r(\mathbf{k})}. \tag{2.51}$$

Proof Put $B_0 = A_0$ and $B_t = A_t \setminus A_{t-1}$ for $t \ge 1$, where the sets A_t are defined in (2.49). Note that the sets B_t are empty for sufficiently large t. Moreover,

$$\bigcup_{i=0}^{t} B_i = A_t, \qquad \bigcup_{t=0}^{\infty} B_t = \{\mathbf{k} : \mathbf{k} \le \mathbf{n}\}.$$

It is clear that

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le \sum_{t=0}^{\infty}\mathsf{E}\left[\max_{\mathbf{k}\in B_t}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le \left(\frac{c}{b_1}\right)^r \sum_{t=0}^{\infty} c^{-tr}\mathsf{E}\left[\max_{\mathbf{k}\in A_t}|S(\mathbf{k})|^r\right],$$

where $b_1 = b(\mathbf{1})$. The assumptions of the theorem imply that

$$\sum_{t=0}^{\infty} c^{-tr}\mathsf{E}\left[\max_{\mathbf{k}\in A_t}|S(\mathbf{k})|^r\right] \le \sum_{t=0}^{\infty} c^{-tr}\sum_{i=0}^{t}\sum_{\mathbf{k}\in B_i}\lambda(\mathbf{k}) \le \sum_{i=0}^{\infty}\sum_{\mathbf{k}\in B_i}\lambda(\mathbf{k})\sum_{t=i}^{\infty}c^{-tr}$$
$$= \frac{c^r}{c^r-1}\sum_{i=0}^{\infty}c^{-ir}\sum_{\mathbf{k}\in B_i}\lambda(\mathbf{k}) \le b_1^r\frac{c^r}{c^r-1}\sum_{\mathbf{k}\le\mathbf{n}}\frac{\lambda(\mathbf{k})}{b^r(\mathbf{k})}.$$

The latter two bounds complete the proof of the theorem. □

Corollary 2.13 *Let* $\mathbf{n}\in\mathbf{N}^d$, $r>0$, *and let* $\{b(\mathbf{k}), \mathbf{k}\le\mathbf{n}\}$ *be an increasing field of positive numbers. If condition* (2.50) *holds for some non-negative numbers* $\lambda(\mathbf{k})$, $\mathbf{k}\le\mathbf{n}$, *and all* $c>1$, *then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le 4\sum_{\mathbf{k}\le\mathbf{n}}\frac{\lambda(\mathbf{k})}{b^r(\mathbf{k})}. \tag{2.52}$$

Proof Corollary 2.13 follows from Theorem 2.19, since

$$\min_{c>1}\frac{c^{2r}}{c^r-1} = 4.$$

□

The assumptions of Corollary 2.13 can easily be checked for independent random variables.

Corollary 2.14 *Let* $\mathbf{n}\in\mathbf{N}^d$ *and* $r>0$. *Let* $\{X(\mathbf{k}), \mathbf{k}\le\mathbf{n}\}$ *be a field of independent random variables such that* $\mathsf{E}\left[|X(\mathbf{k})|^r\right]<\infty$ *for all* $\mathbf{k}\le\mathbf{n}$. *Let* $\{b(\mathbf{k}), \mathbf{k}\le\mathbf{n}\}$ *be an increasing field of positive real numbers. Then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le 4\sum_{\mathbf{k}\le\mathbf{n}}\frac{\mathsf{E}\left[|X(\mathbf{k})|^r\right]}{b^r(\mathbf{k})} \tag{2.53}$$

if $0<r\le 1$.

If $1\le r\le 2$ *and* $\mathsf{E}[X(\mathbf{k})]=0$ *for all* $\mathbf{k}\le\mathbf{n}$, *then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le 2^{d+r+3}\sum_{\mathbf{k}\le\mathbf{n}}\frac{\mathsf{E}\left[|X(\mathbf{k})|^r\right]}{b^r(\mathbf{k})} \tag{2.54}$$

(*the constant on the right-hand side can be decreased up to* 2^{d+3} *if* $r = 2$).

Finally, if $r \ge 2$ *and* $\mathsf{E}[X(\mathbf{k})] = 0$ *for all* $\mathbf{k} \le \mathbf{n}$, *then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le 2^{d+r+2} J_r |\mathbf{n}|^{-1+r/2} \sum_{\mathbf{k}\le\mathbf{n}} \frac{\mathsf{E}\left[|X(\mathbf{k})|^r\right]}{b^r(\mathbf{k})}, \tag{2.55}$$

where J_r *is the Dharmadhikari–Jogdeo constant* (2.9).

Proof We check the assumptions of Corollary 2.13. For a fixed $c > 1$, define the sets A_t according to (2.49). Now fix $t \ge 0$ and define the random variables

$$X'(\mathbf{k}) = \begin{cases} X(\mathbf{k}), & \mathbf{k} \in A_t, \\ 0, & \mathbf{k} \notin A_t, \end{cases}$$

their sums $S'(\mathbf{k}) = \sum_{\mathbf{m}\le\mathbf{k}} X'(\mathbf{m})$, $\mathbf{m} \le \mathbf{n}$, and the field

$$\lambda(\mathbf{k}) = \begin{cases} \mathsf{E}\left[|X(\mathbf{k})|^r\right], & 0 < r \le 1, \\ 2^{d+1+r}\mathsf{E}\left[|X(\mathbf{k})|^r\right], & 1 < r < 2, \\ 2^{d+1}\mathsf{E}\left[|X(\mathbf{k})|^2\right], & r = 2, \\ 2^{d+r} J_r |\mathbf{n}|^{-1+r/2}\mathsf{E}\left[|X(\mathbf{k})|^r\right], & r > 2. \end{cases}$$

Note that $S(\mathbf{k}) = S'(\mathbf{k})$ for $\mathbf{k} \in A_t$ and thus

$$\max_{\mathbf{k}\in A_t}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right| \le \max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S'(\mathbf{k})}{b(\mathbf{k})}\right|.$$

Applying Corollary 2.10 to the sum $S'(\mathbf{k})$ we prove condition (2.50). Now Corollary 2.14 follows from Corollary 2.13. □

Remark 2.6 Even in the case when $d = 1$, the method described above allows us to obtain new results. One example of this kind is based on the monotonicity of the sequence $\mathsf{E}\left[|S_n|^r\right]$ if the random variables $\{X_k\}$ are independent. Let $r > 1$ and put $S_0 = 0$ and

$$\lambda_k = \left(\frac{r}{r-1}\right)^r \left(\mathsf{E}\left[|S_k|^r\right] - \mathsf{E}\left[|S_{k-1}|^r\right]\right), \qquad k \ge 1.$$

Then Doob's inequality for martingales implies that

$$\mathsf{E}\left[\max_{k\le n}|S_k|^r\right] \le \left(\frac{r}{r-1}\right)^r \mathsf{E}\left[|S_n|^r\right] = \sum_{k\le n}\lambda_k.$$

Thus Corollary 2.13 yields

$$\mathsf{E}\left[\max_{k\le n}|S_k|^r\right]\le 4\left(\frac{r}{r-1}\right)^r\sum_{1\le k\le n}\frac{\mathsf{E}\left[|S_k|^r\right]-\mathsf{E}\left[|S_{k-1}|^r\right]}{b_k^r} \tag{2.56}$$

for all monotone sequences $\{b_k, k\le n\}$. The right-hand side of inequality (2.56) for $r=2$ coincides, up to a multiplicative factor, with that of the Hájek–Rényi inequality. Note that (2.56) is more general, since it holds for moments, while the original Hájek–Rényi inequality is valid only for probabilities. For other r, inequality (2.56) may even be a better estimate of the probabilities involved in the Hájek–Rényi inequality.

Example 2.1 Let, for example, $r=4$, $b_k=k$, and let $X_1,\dots,X_n$ be Gaussian random variables such that

$$\mathsf{E}\left[X_k\right]=0,\qquad \mathsf{E}\left[S_k^2\right]=\frac{k^2}{\ln(3k)}.$$

The Hájek–Rényi inequality provides the bound

$$\mathsf{P}\left(\max_{k\le n}\left|\frac{S_k}{k}\right|\ge x\right)\le\frac{1}{x^2}\sum_{k=1}^{n}\frac{\mathsf{E}\left[X_k^2\right]}{k^2}\asymp\ln\ln n. \tag{2.57}$$

On the other hand,

$$\mathsf{E}\left[S_k^4\right]-\mathsf{E}\left[S_{k-1}^4\right]\asymp\frac{k^3}{\ln^2 k}$$

and inequality (2.56) for $r=4$ means that

$$\mathsf{P}\left(\max_{k\le n}\left|\frac{S_k}{k}\right|\ge x\right)\le\frac{16}{x^2}\sum_{k=1}^{n}\frac{\mathsf{E}\left[S_k^4\right]-\mathsf{E}\left[S_{k-1}^4\right]}{k^4}\asymp 1.$$

Clearly the latter bound is asymptotically better than the bound in (2.57).

2.4.3 The Hájek–Rényi Inequality for some Classes of Dependent Random Variables

Just to demonstrate how general the second method is, we provide below some maximal inequalities for the three particular cases of orthogonal, martingale, and homogeneous random variables.

2.4.3.1 The Hájek–Rényi Inequality for Orthogonal Random Variables

Throughout this section, log stands for $\log_2$.

A field $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ of random variables is called *orthogonal* if $\mathsf{E}\left[X^2(\mathbf{k})\right] < \infty$ for all $\mathbf{k} \le \mathbf{n}$ and

$$\mathsf{E}\left[X(\mathbf{k})X(\mathbf{l})\right] = 0 \qquad \text{for all } \mathbf{k} \ne \mathbf{l}.$$

Members of an orthogonal field are called *orthogonal random variables*.

Corollary 2.15 *Let* $\mathbf{n} \in \mathbf{N}^d$ *and let* $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *be orthogonal random variables. Assume that* $\{b(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *is an increasing field of positive real numbers. Then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^2\right] \le 4\,(\log(2n_1))^2 \ldots (\log(2n_d))^2 \sum_{\mathbf{k}\le\mathbf{n}} \frac{\mathsf{E}\left[X^2(\mathbf{k})\right]}{b^2(\mathbf{k})}. \tag{2.58}$$

Proof Moricz [318] proved the following analog of the Menschoff–Rademacher maximal inequality for orthogonal random variables in the case of $d > 1$:

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} S^2(\mathbf{k})\right] \le (\log(2n_1))^2 \ldots (\log(2n_d))^2 \sum_{\mathbf{k}\le\mathbf{n}} \mathsf{E}\left[X^2(\mathbf{k})\right].$$

The rest of the proof is the same as that of Corollary 2.14. □

2.4.3.2 The Hájek–Rényi Inequality for Martingales

Cairoli [43] established several examples showing that some classical inequalities for maximums of submartingales with discrete time are not valid in the case of *submartingales with multi-dimensional time*. Below we show that, despite his examples, some other inequalities can be generalized to the case of $d > 1$, too.

Definition 2.1 Let $\{\mathfrak{F}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be a family of σ-algebras such that

$$\mathfrak{F}(\mathbf{k}) \subseteq \mathfrak{F}(\mathbf{l}) \qquad \text{for all } \mathbf{k} \le \mathbf{n}.$$

A field of random variables $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ is called a *martingale difference* with respect to $\{\mathfrak{F}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ if $X(\mathbf{k})$ is a $\mathfrak{F}(\mathbf{k})$-measurable random variable for all $\mathbf{k} \le \mathbf{n}$ and

$$\mathsf{E}\left[S(\mathbf{k})/\mathfrak{F}(\mathbf{l})\right] = S(\min\{k_1, l_1\}, \ldots, \min\{k_d, l_d\}) \qquad \text{for all } \mathbf{k}, \mathbf{l} \le \mathbf{n}.$$

The field $\{S(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ is called a *martingale* with respect to $\{\mathfrak{F}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$.

Clearly, when $d = 1$, Definition 2.1 defines a usual martingale sequence with respect to an increasing sequence of σ-algebras.

Let $r > 1$. Cairoli [43] proved the following analog of Doob's inequality for $d > 1$:

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})|^r\right] \le \left(\frac{r}{r-1}\right)^{dr} \mathsf{E}\left[|S(\mathbf{n})|^r\right]. \tag{2.59}$$

The right-hand side of this inequality is not always a field with non-negative increments for $d > 1$.

Indeed, this is the case if, for example, $r = 2$. But, in general, this inequality cannot be used to establish a Hájek–Rényi type inequality for martingales if $d > 1$. On the other hand, $\mathsf{E}\left[|S(\mathbf{n})|^r\right]$ can be estimated with the help of the Bahr–Esseen or Dharmadhikari–Jogdeo inequality and the result can already be used to derive a Hájek–Rényi inequality.

Since

$$\zeta_i = \sum_{k_1=1}^{n_1} \cdots \sum_{k_{d-1}=1}^{n_{d-1}} \sum_{k_d=1}^{i} X(\mathbf{k})$$

is a usual martingale with respect to the σ-algebras $\mathfrak{F}(n_1, \ldots, n_{d-1}, i)$, $i \le n_d$, we use induction with respect to the dimension of the spaces $\mathbf{N}^d$ and prove that the Bahr–Esseen inequality [16] for martingales with one-dimensional indices implies that

$$\mathsf{E}\left[|S(\mathbf{n})|^r\right] \le 2^d \sum_{\mathbf{k}\le\mathbf{n}} \mathsf{E}\left[|X(\mathbf{k})|^r\right] \tag{2.60}$$

if $1 \le r \le 2$, while the Dharmadhikari–Jogdeo inequality [73] for usual martingales implies that

$$\mathsf{E}\left[|S(\mathbf{n})|^r\right] \le J_r^d |\mathbf{n}|^{-1+r/2} \sum_{\mathbf{k}\le\mathbf{n}} \mathsf{E}\left[|X(\mathbf{k})|^r\right] \tag{2.61}$$

if $r > 2$.

Corollary 2.16 *Let* $\mathbf{n} \in \mathbf{N}^d$, $r > 1$, *and let* $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *be a martingale difference with respect to a family of increasing* σ*-algebras. Let* $\{b(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *be an increasing field of positive numbers. Then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} \left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le 4\left(\frac{r}{r-1}\right)^{dr} \sum_{\mathbf{k}\le\mathbf{n}} \frac{\mathsf{E}\left[|X(\mathbf{k})|^r\right]}{b^r(\mathbf{k})} \tag{2.62}$$

if $0 < r \le 1$;

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} \left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le 2^{d+2}\left(\frac{r}{r-1}\right)^{dr} \sum_{\mathbf{k}\le\mathbf{n}} \frac{\mathsf{E}\left[|X(\mathbf{k})|^r\right]}{b^r(\mathbf{k})} \tag{2.63}$$

if $1 < r \le 2$; *and*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}}\left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^r\right] \le 4J_r^d\left(\frac{r}{r-1}\right)^{dr}|\mathbf{n}|^{-1+r/2}\sum_{\mathbf{k}\le\mathbf{n}}\frac{\mathsf{E}\left[|X(\mathbf{k})|^r\right]}{b^r(\mathbf{k})} \tag{2.64}$$

if $r > 2$, *where* J_r *is the Dharmadhikari–Jogdeo constant* (2.9).

Proof Let $c > 1$. The proof of Corollary 2.16 for $0 < r \le 1$ coincides with that of Corollary 2.14.

Now let $r > 1$. Define the sets A_t, $t \ge 0$, according to the rule (2.49). Then we apply Lemma 2.4 given below for all $\mathbf{k}_0 \notin A_t$ and prove that the field

$$X'(\mathbf{k}) = \begin{cases} X(\mathbf{k}), & \mathbf{k} \in A_t, \\ 0, & \mathbf{k} \notin A_t, \end{cases}$$

is a martingale difference with respect to the σ-algebras $\{\mathfrak{F}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$. Now we use inequalities (2.60) and (2.61) for random variables $\{X'(\mathbf{k})\}$. As a result, we prove that condition (2.50) holds for $\{X(\mathbf{k})\}$ with the corresponding numbers $\{\lambda(\mathbf{k})\}$. Inequality (2.63) follows from (2.59), Corollary 2.13, and the Bahr–Esseen inequality (2.60), while (2.64) follows from (2.59), Corollary 2.13, and the Dharmadhikari–Jogdeo inequality (2.61). □

Lemma 2.3 *Let* $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ *be a martingale difference with respect to a family of increasing* σ*-algebras* $\{\mathfrak{F}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$. *Then, for all* $\mathbf{k}, \mathbf{l} \le \mathbf{n}$,

$$\mathsf{E}\left[X(\mathbf{k})/\mathfrak{F}(\mathbf{l})\right] = \begin{cases} X(\mathbf{k}), & \mathbf{k} \le \mathbf{l}, \\ 0, & \mathbf{k} \not\le \mathbf{l}. \end{cases} \tag{2.65}$$

Proof Let $\mathbf{k} = (k_1, \dots, k_d)$, $\mathbf{l} = (l_1, \dots, l_d)$. As usual, we also let $S(\mathbf{l}) = \sum_{\mathbf{k}\le\mathbf{l}} X(\mathbf{k})$. Then we use the representation

$$X(\mathbf{k}) = \sum_{\boldsymbol{\varepsilon}\in\mathscr{E}_d}(-1)^{\varepsilon_1+\dots+\varepsilon_d}S(\mathbf{k}-\boldsymbol{\varepsilon})$$

(see Proposition A.5), where $\boldsymbol{\varepsilon}$ is a vector $(\varepsilon_1, \dots, \varepsilon_d)$ and each of its coordinates ε_i is either equal to 0 or 1. We deduce from this representation that

$$\mathsf{E}\left[X(\mathbf{k})/\mathfrak{F}(\mathbf{l})\right] = \sum_{\boldsymbol{\varepsilon}\in\mathscr{E}_d}(-1)^{\varepsilon_1+\dots+\varepsilon_d}\mathsf{E}\left[S(\mathbf{k}-\boldsymbol{\varepsilon})/\mathfrak{F}(\mathbf{l})\right].$$

If $\mathbf{k} \le \mathbf{l}$, then $\mathsf{E}\left[S(\mathbf{k}-\boldsymbol{\varepsilon})/\mathfrak{F}(\mathbf{l})\right] = S(\mathbf{k}-\boldsymbol{\varepsilon})$ for all $\boldsymbol{\varepsilon} \in \mathscr{E}_d$ and thus (2.65) holds.

Otherwise, that is if $\mathbf{k} \not\le \mathbf{l}$, then there is a non-empty set $J \subseteq \{1, \dots, d\}$ such that $k_j > l_j$ for $j \in J$ and $k_j \le l_j$ for $j \notin J$. Note that if $J = \{1, \dots, d\}$, then there is no $j \in J$ for which $k_j \le l_j$ and thus the latter condition does not hold at all.

First we consider the case of $J = \{1, \dots, d\}$, that is, the case of $k_1 > l_1, \dots, k_d > l_d$. Then $\mathbf{l} \le \mathbf{k} - \boldsymbol{\varepsilon}$, whence $\mathsf{E}\left[S(\mathbf{k}-\boldsymbol{\varepsilon}) / \mathfrak{F}(\mathbf{l})\right] = S(\mathbf{l})$ by Definition 2.1. Hence

$$\mathsf{E}\left[X(\mathbf{k}) / \mathfrak{F}(\mathbf{l})\right] = S(\mathbf{l}) \sum_{\boldsymbol{\varepsilon} \in \mathscr{E}_d} (-1)^{\varepsilon_1 + \dots + \varepsilon_d} = 0.$$

Now we consider the case of a set $J = \{1, \dots, c\}$, $1 \le c < d$, that is, the case where $k_1 > l_1, \dots, k_c > l_c$ and $k_{c+1} \le l_{c+1}, \dots, k_d \le l_d$. Then

$$\mathsf{E}\left[S(\mathbf{k}-\boldsymbol{\varepsilon}) / \mathfrak{F}(\mathbf{l})\right] = S(l_1, \dots, l_c, k_{c+1} - \varepsilon_{c+1}, \dots, k_d - \varepsilon_d),$$

whence

$$\mathsf{E}\left[X(\mathbf{k}) / \mathfrak{F}(\mathbf{l})\right] = \sum_{\mathscr{E}_{1,\dots,c}} (-1)^{\varepsilon_1 + \dots + \varepsilon_c} \sum_{\mathscr{E}_{c+1,\dots,d}} (-1)^{\varepsilon_{c+1} + \dots + \varepsilon_d} S(l_1, \dots, l_c, k_{c+1} - \varepsilon_{c+1}, \dots, k_d - \varepsilon_d) = 0,$$

where the sums are formed with respect to the sets $\mathscr{E}_{1,\dots,c}$ and $\mathscr{E}_{c+1,\dots,d}$ consisting of those elements $(\varepsilon_1, \dots, \varepsilon_c)$ and $(\varepsilon_{c+1}, \dots, \varepsilon_d)$, respectively, whose coordinates equal either 0 or 1.

The case of a general set J is considered similarly to the case of $J = \{1, \dots, c\}$ in view of symmetry. The lemma is completely proved. □

Lemma 2.4 *Let $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be a martingale difference with respect to a family of increasing σ-algebras $\{\mathfrak{F}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ and let $\mathbf{k}_0 \le \mathbf{n}$ be fixed. Put $S(\mathbf{k}) = \sum_{\mathbf{l} \le \mathbf{k}} X(\mathbf{l})$ and*

$$X'(\mathbf{k}) = \begin{cases} X(\mathbf{k}), & \mathbf{k} \ne \mathbf{k}_0, \\ 0, & \mathbf{k} = \mathbf{k}_0, \end{cases} \qquad S'(\mathbf{m}) = \sum_{\mathbf{k} \le \mathbf{m}} X'(\mathbf{k}), \quad \mathbf{m} \le \mathbf{n}.$$

Then $\{S'(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ is a martingale with respect to the σ-algebras $\{\mathfrak{F}(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$.

Proof It is clear that

$$S'(\mathbf{k}) = \begin{cases} S(\mathbf{k}), & \mathbf{k}_0 \not\le \mathbf{k}, \\ S(\mathbf{k}) - X(\mathbf{k}_0), & \mathbf{k}_0 \le \mathbf{k}. \end{cases}$$

If $\mathbf{k}_0 \not\le \min\{\mathbf{k}, \mathbf{l}\}$ and $\mathbf{k}_0 \not\le \mathbf{k}$, then

$$\mathsf{E}\left[S'(\mathbf{k}) / \mathfrak{F}(\mathbf{l})\right] = \mathsf{E}\left[S(\mathbf{k}) / \mathfrak{F}(\mathbf{l})\right] = S(\min\{\mathbf{k}, \mathbf{l}\}) = S'(\min\{\mathbf{k}, \mathbf{l}\}).$$

On the other hand, if $\mathbf{k}_0 \le \min\{\mathbf{k}, \mathbf{l}\}$, then $\mathbf{k}_0 \le \mathbf{k}$ and

$$\mathsf{E}\left[S'(\mathbf{k}) / \mathfrak{F}(\mathbf{l})\right] = \mathsf{E}\left[S(\mathbf{k}) - X(\mathbf{k}_0) / \mathfrak{F}(\mathbf{l})\right] = S(\min\{\mathbf{k}, \mathbf{l}\}) - X(\mathbf{k}_0) = S'(\min\{\mathbf{k}, \mathbf{l}\})$$

by Lemma 2.3. Finally, if $\mathbf{k}_0 \not\leq \min\{\mathbf{k}, \mathbf{l}\}$ and $\mathbf{k}_0 \leq \mathbf{k}$, then $\mathbf{k}_0 \not\leq \mathbf{l}$ and thus

$$\mathsf{E}\left[S'(\mathbf{k})/\mathfrak{F}(\mathbf{l})\right] = \mathsf{E}\left[S(\mathbf{k}) - X(\mathbf{k}_0)/\mathfrak{F}(\mathbf{l})\right] = S(\min\{\mathbf{k}, \mathbf{l}\}) = S'(\min\{\mathbf{k}, \mathbf{l}\})$$

again by Lemma 2.3. The proof is complete. □

2.4.3.3 The Hájek–Rényi Inequality for Quadratic Forms

A random quadratic form is a particular case of a martingale for $d = 2$.

Assume that $\{\xi_i, i \leq n\}$ are independent random variables such that $\mathsf{E}[\xi_i] = 0$ and $\mathsf{E}\left[\xi_i^2\right] < \infty$ for all $i \leq n$. Let $\{a_{ij}, i, j \leq n\}$ be some collection of real numbers. Put

$$X(i, j) = a_{ij}\left(\xi_i \xi_j - \mathsf{E}\left[\xi_i \xi_j\right]\right), \qquad i, j \leq n,$$
$$S(m_1, m_2) = \sum_{i_1=1}^{m_1} \sum_{i_2=1}^{m_2} X(i_1, i_2).$$

Then $\{X(i, j); (i, j) \leq (n, n)\}$ is a martingale-difference with respect to the family of σ-algebras

$$\mathfrak{F}(k, l) = \sigma\{X(i, j); (i, j) \leq (k, l)\} = \sigma(\xi_i, i \leq \max\{k, l\}).$$

Further, let $\{b(i, j); (i, j) \leq (n, n)\}$ be an increasing field of positive real numbers. Put $\sigma_i^2 = \mathsf{E}\left[\xi_i^2\right]$. Corollary 2.16 for $r = 2$ implies that

$$\mathsf{E}\left[\max_{(k,l)\leq(n_1,n_2)} \left|\frac{1}{b(k, l)} \sum_{i=1}^{k} \sum_{j=1}^{l} a_{ij}\left(\xi_i \xi_j - \mathsf{E}\left[\xi_i \xi_j\right]\right)\right|^2\right]$$
$$\leq 2^{2d+3} \mathsf{E}\left[S^2(n_1, n_2)\right] \leq 2^{2d+3}\left[\sum_{i=1}^{\min\{n_1,n_2\}} \frac{a_{ii}^2}{b^2(i, i)} \mathsf{E}\left[\left(\xi_i^2 - \sigma_i^2\right)^2\right] + \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{a_{ij}^2}{b^2(i, j)} \sigma_i^2 \sigma_j^2\right]$$

for all $n_1, n_2 \leq n$.

2.4.3.4 The Hájek–Rényi Inequality for Homogeneous Random Fields

A random field $\{X(\mathbf{k}), \mathbf{k} \leq \mathbf{n}\}$ is called *homogeneous* if there exists a function R depending on d arguments such that

$$\mathsf{E}\,[X(\mathbf{k})] = \text{const}, \qquad \mathsf{E}\,[X(\mathbf{k})X(\mathbf{l})] = R(\mathbf{k}-\mathbf{l}) \qquad \text{for all} \quad \mathbf{k}, \mathbf{l} \le \mathbf{n}.$$

In what follows we assume that $\mathsf{E}\,[X(\mathbf{k})] = 0$ for $\mathbf{k} \le \mathbf{n}$. It is easy to show that

$$\mathsf{E}\left[S^2(\mathbf{n})\right] \le |\mathbf{n}| \sum_{\mathbf{0}\le\mathbf{k}\le\mathbf{n}-\mathbf{1}} |R(\mathbf{k})|.$$

Since the right-hand side of the latter inequality is of a super-additive structure, Moricz's inequality [318] implies that

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} |S(\mathbf{k})|^2\right] \le (\log(2n_1))^2 \dots (\log(2n_d))^2\, |\mathbf{n}| \sum_{\mathbf{0}\le\mathbf{k}\le\mathbf{n}-\mathbf{1}} |R(\mathbf{k})|.$$

This bound together with Corollary 2.13 yields an analog of the Hájek–Rényi inequality for homogeneous fields.

Corollary 2.17 *Let $\mathbf{n} \in \mathbf{N}^d$ and let $\{X(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be a homogeneous field. Further let $\{b(\mathbf{k}), \mathbf{k} \le \mathbf{n}\}$ be an increasing field of positive real numbers. Then*

$$\mathsf{E}\left[\max_{\mathbf{k}\le\mathbf{n}} \left|\frac{S(\mathbf{k})}{b(\mathbf{k})}\right|^2\right] \le 4\,(\log(2n_1))^2 \dots (\log(2n_d))^2\, |\mathbf{n}| \sum_{\mathbf{0}\le\mathbf{k}\le\mathbf{n}-\mathbf{1}} \frac{|R(\mathbf{k})|}{b^2(\mathbf{k})}. \tag{2.66}$$

2.5 Comments

Theorem 2.1 was stated (without proof) by Bienayme in 1853; Chebyshev independently proved it in 1867. Theorem 2.2 was proved by Kolmogorov [254]; its generalization, Theorem 2.3, is mentioned by Hájek and Rényi [150]. Theorem 2.4 is a corollary of Doob's inequality [77] for martingales. Theorem 2.5 is obtained by Lévy [283] in his studies of the convergence of sums of independent random variables to the stable laws.

Section 2.1 The Bahr–Esseen inequality (2.6) is proved in [16]; the Dharmadhikari–Jogdeo inequality (2.7) in [73]; Rosenthal's inequality (2.11) in [360]; and the Marcinkiewicz–Zygmund inequality (2.11) in [301]. Theorem 2.15 is due to Skorokhod [371].

Section 2.2 Theorem 2.15 is due to Petrov [337]. Lemma 2.2 can be found, for example, in Petrov's monograph [338]. Theorem 2.16 is proved by Petrov [337] in the case $d = 1$ (see inequality (2.18)). Corollary 2.4 for $d = 1$ is proved by Lévy [283], see [215] for the case $d > 1$. Inequality (2.26) is mentioned by Zimmerman [433] in the case $d = 2$; the same inequality for $d > 1$ is given by Paranjape and Park [335] without proof.

Theorem 2.17 is proved by Ottaviani [333]; its analog for $d > 1$ (Corollary 2.6) is obtained in [210] using another method. A similar result is given by N. Etemadi in [89] for $d = 1$ and in [92] for $d > 1$.

The idea to use Corollary 2.2 to prove inequality (2.33) is proposed by Petrov [337] for $d = 1$. Corollary 2.8 is proved by Petrov [337] for the case of $d = 1$ and $r = 2$ (the general case is obtained in [215]). All other results of this section for $d > 1$ are proved in [215].

Section 2.3 Inequality (2.43) for $r = 2$ is due to Wichura [419] with the constant 4^d instead of 2^{d+2} on the right-hand side. All other results for $d > 1$ are obtained in [211].

Christofides [59] and Christofides and Serfling [62] describe an application of inequality (2.44) for generalized U-statistics. Inequality (2.45) is used by Shcherbakova [365] to study the almost sure limit properties of increments of random fields.

Section 2.4 The first method of proof for the Hájek–Rényi type inequalities for $d > 1$ is proposed by Shorack and Smythe [366]; the second method is presented in [211]. Etemadi [92] obtains an inequality of type (2.54) for fields with independent increments.

The relationship between inequalities like (2.2) and the strong law of large numbers is clear. It is shown in [98] and [234] that the strong law of large numbers is always a consequence of such an inequality (Khoshnevisan [205] discusses a similar idea for multi-indexed sums of independent identically distributed random variables). This approach was later used by many authors for various types of dependence: Wang et al. [411] study the so-called d-demimartingales; Wang and Jin [414] reduce the complete convergence of multi-indexed sums to the corresponding inequalities for weighted sums; Hung and Tien [173] show some applications of the Hájek–Rényi type inequalities to Banach-valued martingale fields.

The literature on Hájek–Rényi type inequalities for various dependence schemes is rather extensive. Below is a short list of relevant schemes: pairwise independent random variables (Etemadi [91]), pairwise NQ dependent random variables (Yang et al. [424]), linearly dependent NQD random variables (Hu et al. [170]), AANA sequences (Wang et al. [407]), sequences with φ-mixing (Wang et al. [409]) or with ψ-mixing (Yang et al. [410]), mixingales (Wang et al. [413]), demimartingales (Wang and Hu [406]).

Several generalizations of martingales are known in the case of sequences depending on multi-indices (Cairoli and Walsh [46], Wong and Zakai [421], Zakai [430]). Definition 2.1 was introduced by Shorack and Smythe [366] (they also prove inequality (2.59)). Below is a list of papers where Kolmogorov and Hájek–Rényi inequalities have been discovered for multi-parameter martingales: Li [285], Fazekas [95], Sung [390], He [157], and Christofides and Serfling [60, 61].

Inequalities for maximums of multi-indexed sums of random variables with different types of dependence have been proved by Kryzhanovskaya [264], Borodikhin [29], Lagodowski [271], Kurbanmuradov and Sabelfeld [269], Shashkin [364], and Bakhtin and Bulinskii [17].

Springer

http://www.springer.com/978-3-662-44387-3

Limit Theorems for Multi-Indexed Sums of Random Variables
Klesov, O.
2014, XVIII, 483 p. 2 illus., Hardcover
ISBN: 978-3-662-44387-3