Asymptotics for Associated Random Variables

Bearbeitet von
Paulo Eduardo Oliveira

1. Auflage 2012. Buch. x, 194 S. Hardcover
ISBN 978 3 642 25531 1
Format (B x L): 15,5 x 23,5 cm
Gewicht: 473 g

Chapter 2
Inequalities

Abstract This chapter sets the basic tools to prove the asymptotic results that are to come in the following chapters. The first three sections are concerned with different types of inequalities on joint distributions of associated random variables and moments of sums. It is interesting that, although association is defined with a somewhat vague requirement, it is possible to recover versions for moment inequalities which are quite close to the independent case, thus paving the way to find asymptotic results that are also similar to the ones found in the independence framework. One of the key issues with association is the ability to control joint distributions from the marginal distributions using the covariance structure of the random variables. This is explored mainly in Sects. 2.5 and 2.6. The inequalities proved in these sections will provide the means to use the coupling technique, common to prove convergence results. Sect. 2.5 shows that at least the convergence in distribution is concerned with the covariance structure that completely describes the behaviour of associated variables. This chapter is a fundamental one for the remaining text.

2.1 Introduction

As usual, inequalities play an important role in the development of a theory, as much of the proving efforts are spent obtaining good estimates of suitable quantities. The first result on inequalities for associated random variables appeared in Lebowitz [55], controlling covariances of blocks of variables motivated by the need to control some Hamiltonians appearing in the Ising spin models with ferromagnetic interactions. These are essentially covariance inequalities on transformations defined with indicator functions. The natural development considering more general transformations appeared with the contributions by Newman [69, 70], where the main goal was, however, the extension of inequalities on characteristic functions to transformed associated random variables. These inequalities on characteristic functions were motivated by the study of central limit problems using the classical approach: decomposing sums into sums of blocks and trying to treat these as if they were independent. This led the work in Newman [69, 70] to one of the main tools in the association literature about convergence in distribution, Theorem 2.37 and its extension to transformed variables, Theorem 2.40. Moreover, inequality (2.26) showed the importance of covariances on the characterization of the dependence structure

P.E. Oliveira, *Asymptotics for Associated Random Variables*,
DOI 10.1007/978-3-642-25532-8_2,

of associated random variables, leading naturally to conditions on the decay rate of the covariances when dealing with central limit problems, invariance principles or statistical estimation issues. The meaningfulness of the coefficient $u(n)$, introduced by Cox and Grimmett [25] (see Definition 2.13), is well justified by this characteristic functions inequality. As a sort of a side effect of the previous, it is worth mentioning the control of covariances of indicator functions by the covariances of the original associated variables, with a first version appearing in Yu [109] and later extended in Cai and Roussas [24]. This inequality, as expressed in Corollary 2.36, has shown to be of significant importance in the analysis of invariance principles and also when studying the behaviour of statistical estimators. The interest on extending covariance inequalities was also developed into another direction, controlling the covariance of not necessarily monotone transformations of the variables, with upper bounds depending on the derivatives of these transformations. A general methodology for approaching the control of such covariances really follows from a few concepts introduced in Newman [69, 70], although explicit results only appeared somewhat later, Bulinski [22]. The control of moments for partial sums was first treated in Birkel [13], where the usual $n^{r/2}$ bounds were proved under suitable decay rates on the pairwise covariances, expressed through a convenient coefficient, to be introduced below in Definition 2.13, and were later extended by Shao and Yu [94]. These tools were used by Masry [65] to prove almost optimal convergence rates for density estimators based on associated samples.

Finally, exponential inequalities have been an important tool for studying convergence rates, especially for laws of large numbers or large deviations. The first such inequality was proved by Prakasa Rao [83], who obtained an upper bound that is too weak to characterize convergence results. A more useful exponential inequality for bounded associated variables appeared in Ioannides and Roussas [48] and was used by the authors to prove the first convergence rates for Strong Laws of Large Numbers. This inequality was also used in statistical estimation to characterize convergence rates when approximating distribution functions in Azevedo and Oliveira [3] or Henriques and Oliveira [41] or for density estimators [42]. The approach was based on a block decomposition of partial sums, as for the proofs of Central Limit Theorems. This was later extended, dropping the boundedness of the variables by using truncation, by Oliveira [75]. Curiously, what seems to be the weak point of this approach was the treatment of the coupling independent variables used in the proof. This aspect was further improved by Sung [98] and Xing, Yang and Liu [106], always with the motivation of improving the convergence rates for the laws of large numbers in mind, who were able to obtain exponential inequalities that lead to convergence rates in the law of large numbers arbitrarily close to the rate for independent variables.

2.2 Block and Tail Inequalities

In this section we prove two simple inequalities whose proofs use essentially the same arguments. The first one concerns joint distributions and pairwise joint distributions, separating the variables, and was proved by Lebowitz [55]. This inequality can be thought of as a first version of the control of the distance between joint distributions and the product of marginal distributions that would be found in the case of independent random variables. This idea will be extended into a finer and more clear statement in Sect. 2.5, where this distance is truly controlled in terms of the covariances of the associated variables. The second inequality concerns distribution functions and tail probabilities, proving what one could expect to find, given the definition of association of random variables, that is, that joint probabilities tend to increase with respect to the product of the corresponding marginal ones. This is a direct consequence of the fact that covariances are nonnegative, which is one of the sources of difficulties when treating associated variables: covariances of moments of sums tend to increase with respect to what could be found for independent variables, so the inequalities seem to somehow go into a wrong direction.

Given $A, B \subset \{1, \dots, n\}$ and $x_1, \dots, x_n \in \mathbb{R}$, define

$$\begin{aligned} H_{A,B}(x_1, \dots, x_n) &= \mathbf{P}(X_i > x_i, i \in A \cup B) \\ &\quad - \mathbf{P}(X_j > x_j, j \in A)\mathbf{P}(X_k > x_k, k \in B), \\ H_{j,k}(x_1, \dots, x_n) &= H_{\{j\},\{k\}}(x_1, \dots, x_n). \end{aligned}$$

We will write, for simplicity, just $H_{A,B}$ or $H_{j,k}$, unless confusion arises. Notice that $H_{j,k} = \mathrm{Cov}(\mathbb{I}_{(x_j,+\infty)}(X_j), \mathbb{I}_{(x_k,+\infty)}(X_k)) \geq 0$, as the functions $\mathbb{I}_{(x,+\infty)}(u)$ are, for each fixed x, nondecreasing in u. Moreover, as already remarked before (see page 3),

$$\begin{aligned} H_{1,2}(x_1, x_2) &= \mathbf{P}(X_1 > x_1, X_2 > x_2) - \mathbf{P}(X_1 > x_1)\mathbf{P}(X_2 > x_2) \\ &= \mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) - \mathbf{P}(X_1 \leq x_1)\mathbf{P}(X_2 \leq x_2) \\ &= \mathrm{Cov}\big(\mathbb{I}_{(-\infty,x_1]}(X_1), \mathbb{I}_{(-\infty,x_2]}(X_2)\big). \end{aligned}$$

Theorem 2.1 (Lebowitz inequality) *Let $X_1, \dots, X_n$ be associated random variables, and $A, B \subset \{1, \dots, n\}$. Then $0 \leq H_{A,B} \leq \sum_{j\in A,k\in B} H_{j,k}$.*

Proof Recall the definition of $\mathbb{I}_A = \prod_{i\in A} \mathbb{I}_{(x_i,+\infty)}(X_i)$ (see page 22) and put $S_A = \sum_{i\in A} \mathbb{I}_{(x_i,+\infty)}(X_i)$, and analogously for $\mathbb{I}_B$ and S_B. Then, obviously,

$$H_{A,B} = \mathrm{Cov}(\mathbb{I}_A, \mathbb{I}_B), \qquad \sum_{j\in A,k\in B} H_{j,k} = \mathrm{Cov}(S_A, S_B)$$

and

$$\mathrm{Cov}(S_A, S_B) = \mathrm{Cov}(S_A - \mathbb{I}_A, S_B) + \mathrm{Cov}(\mathbb{I}_A, S_B - \mathbb{I}_B) + \mathrm{Cov}(\mathbb{I}_A, \mathbb{I}_B).$$

All the $\mathbb{I}_A$, $\mathbb{I}_B$, S_A and S_B are nondecreasing transformations of the $X_1, \dots, X_n$, thus are associated, according to Theorem 1.15. It follows that $H_{A,B} = \mathrm{Cov}(\mathbb{I}_A, \mathbb{I}_B) \geq 0$. Let us now fix $j \in A$. Then

$$S_A - \mathbb{I}_A = \sum_{\substack{\ell\in A\\ \ell\neq j}} \mathbb{I}_\ell + \mathbb{I}_j\Bigg(1 - \prod_{\substack{\ell\in A\\ \ell\neq j}} \mathbb{I}_\ell\Bigg).$$

The first term on the right does not depend on X_j, while the second is the product of a nondecreasing function of X_j, $\mathbb{I}_j$, by a nonnegative factor that does not depend on X_j. Thus, it follows that $S_A - \mathbb{I}_A$ is a nondecreasing function of X_j. Repeating this argument for each choice of $j \in A$ and each $k \in B$, it follows that $S_A - \mathbb{I}_A$ and $S_B - \mathbb{I}_B$ are nondecreasing in each variable they depend on. Thus, due to the association, $\mathrm{Cov}(S_A - \mathbb{I}_A, S_B) \geq 0$ and $\mathrm{Cov}(\mathbb{I}_A, S_B - \mathbb{I}_B) \geq 0$, so the theorem follows. □

This result implies immediately a very simple characterization of independence between associated variables completely described in terms of the covariances, generalizing Theorem 1.17.

Corollary 2.2 *Let $X_1, \ldots, X_n$ be associated random variables, and $A, B \subset \{1, \ldots, n\}$. Then the random variables X_i, $i \in A$, are jointly independent of X_j, $j \in B$, if and only if $\mathrm{Cov}(X_i, X_j) = 0$ for every $i \in A$ and $j \in B$.*

Theorem 2.1 above gives an upper bound for a term that may be thought of as a joint covariance using pairwise covariances. Next we prove a similar result but concerning directly the joint distributions. In this case, association implies a lower bound for the joint distributions.

Theorem 2.3 *Let $X_1, \ldots, X_n$ be associated random variables, and $x_1, \ldots, x_n \in \mathbb{R}$. For every $A \subset \{1, \ldots, n\}$,*

(a) $\mathbf{P}(X_i > x_i, i \in A) \geq \prod_{i \in A} \mathbf{P}(X_i > x_i)$.
(b) $\mathbf{P}(X_i \leq x_i, i \in A) \geq \prod_{i \in A} \mathbf{P}(X_i \leq x_i)$.

Proof (a) may be rewritten as $\mathrm{E}(\prod_{i \in A} \mathbb{I}_{(x_i, +\infty)}(X_i)) \geq \prod_{i \in A} \mathrm{E}(\mathbb{I}_{(x_i, +\infty)}(X_i))$. By permutating the random variables, which does not affect association, we may assume, without loss of generality, that $A = \{1, \ldots, k\}$ for some $k \leq n$. Then,

$$\begin{aligned}
&\mathrm{Cov}\left(\mathbb{I}_{(x_k, +\infty)}(X_k), \prod_{i=1}^{k-1} \mathbb{I}_{(x_i, +\infty)}(X_i)\right) \\
&\quad = \mathrm{E}\left(\mathbb{I}_{(x_k, +\infty)}(X_k) \prod_{i=1}^{k-1} \mathbb{I}_{(x_i, +\infty)}(X_i)\right) - \mathrm{E}\big(\mathbb{I}_{(x_k, +\infty)}(X_k)\big)\mathrm{E}\left(\prod_{i=1}^{k-1} \mathbb{I}_{(x_i, +\infty)}(X_i)\right) \\
&\quad \geq 0.
\end{aligned}$$

Iterating now this argument, (a) follows. As for (b), apply the same argument to the decreasing transformations $\mathbb{I}_{(-\infty, x_i]}(X_i) = 1 - \mathbb{I}_{(x_i, +\infty)}(X_i)$. □

A useful bound for $H_{j,k}$, in terms of the covariances of the original random variables will be proved later in Corollary 2.36.

The inequality in Theorem 2.1 concerns a special kind of nondecreasing transformations. In fact, the same is still true for the inequalities in Theorem 2.3. However, it is possible to go beyond nondecreasing and even nonmonotone transformations of associated variables if these functions are dominated by nondecreasing ones, in a convenient sense as introduced by Newman [69].

Definition 2.4 Let $f, g : \mathbb{R}^n \longrightarrow \mathbb{C}$, where $n \in \mathbb{N}$. We write $f \preceq g$ if $g - \mathrm{Re}(e^{i\alpha} f)$ is coordinatewise nondecreasing for every $\alpha \in \mathbb{R}$.

Remark 2.5 Notice that as $g = \frac{1}{2}[(g - \mathrm{Re}(f)) + (g - \mathrm{Re}(-f))]$, if $f \preceq g$, then g is real-valued and coordinatewise nondecreasing.

Remark 2.6 If f is a real-valued function, then it is obvious, by choosing $\alpha = \pi$ or $\alpha = 0$, that $f \preceq g$ if and only if both $g + f$ and $g - f$ are nondecreasing.

We first state a result allowing to deal with characteristic functions through the relation "$\preceq$".

Proposition 2.7 *Let f and g be functions defined on $\mathbb{R}^n$. Assume f is real-valued and $f \preceq g$. Let ρ be a complex-valued function defined on $\mathbb{R}$ such that, for every $u, v \in \mathbb{R}$, $|\rho(u) - \rho(v)| \leq |u - v|$. Then $\rho \circ f \preceq g$.*

Proof We need to prove that $g - \mathrm{Re}(e^{i\alpha} \rho \circ f)$ is nondecreasing. Let $t, s \in \mathbb{R}^n$ be such that $t \geq s$ (in the coordinatewise sense, that is, $t_j \geq s_j$, $j = 1, \ldots, n$). Then, taking into account the assumption on ρ, we have

$$\begin{aligned} &\left|\mathrm{Re}\big(e^{i\alpha} \rho\big(f(t)\big)\big) - \mathrm{Re}\big(e^{i\alpha} \rho\big(f(s)\big)\big)\right| \\ &\quad \leq \left|e^{i\alpha} \rho\big(f(t)\big) - e^{i\alpha} \rho\big(f(s)\big)\right| = \left|\rho\big(f(t)\big) - \rho\big(f(s)\big)\right| \leq \left|f(t) - f(s)\right|. \end{aligned}$$

Now, as $f \preceq g$, both $g + f$ and $g - f$ are nondecreasing. So, if $f(t) - f(s) > 0$, use the first to find $|f(t) - f(s)| = f(t) - f(s) \leq g(t) - g(s)$, and in case $f(t) - f(s) < 0$, use the later to find $|f(t) - f(s)| = f(s) - f(f) \leq g(t) - g(s)$. That is, we have in either case $|f(t) - f(s)| \leq g(t) - g(s)$. Finally, as g is nondecreasing,

$$\begin{aligned} &\left|\big(g(t) + \mathrm{Re}\big(e^{i\alpha} \rho\big(f(t)\big)\big)\big) - \big(g(s) + \mathrm{Re}\big(e^{i\alpha} \rho\big(f(s)\big)\big)\big)\right| \\ &\quad \geq g(t) - g(s) - \left|\mathrm{Re}\big(e^{i\alpha} \rho\big(f(t)\big)\big) - \mathrm{Re}\big(e^{i\alpha} \rho\big(f(s)\big)\big)\right| \geq 0. \end{aligned}$$

□

Notice that in the previous result we can choose $\rho(u) = e^{iu}$.

Lemma 2.8 *Let X_n, $n \in \mathbb{N}$, be associated random variables. Let f_1, f_2, g_1, g_2 be functions defined on $\mathbb{R}^n$ for some $n \in \mathbb{N}$, such that $f_1 \preceq g_1$ and $f_2 \preceq g_2$. Then*

$$\begin{aligned} &\left|\mathrm{Cov}\big(f_1(X_1, \ldots, X_n), f_2(X_1, \ldots, X_n)\big)\right| \\ &\quad \leq 2\left|\mathrm{Cov}\big(g_1(X_1, \ldots, X_n), g_2(X_1, \ldots, X_n)\big)\right|. \end{aligned} \tag{2.1}$$

Proof Assume that f_1, f_2, g_1, g_2 are real-valued functions. Then, it is enough to prove

$$\begin{aligned} &\left|\mathrm{Cov}\big(g_1(X_1, \ldots, X_n), g_2(X_1, \ldots, X_n)\big)\right| - \mathrm{Cov}\big(f_1(X_1, \ldots, X_n), h(X_1, \ldots, X_n)\big) \\ &\quad \geq 0 \end{aligned}$$

both for $h = f_2$ and $h = -f_2$. Now, according to Remark 2.6, $g_1 + f_1$, $g_1 - f_1$, $g_2 + f_2$ and $g_2 - f_2$ are all nondecreasing functions. Thus, for both considered choices of h, g_2+h and g_2-h are nondecreasing. Notice further that, according to Remark 2.5, g_1 and g_2 are nondecreasing functions. Thus, taking into account the association of the random variables, we have

$$\begin{aligned}
&\big|\mathrm{Cov}\big(g_1(X_1,\dots,X_n),g_2(X_1,\dots,X_n)\big)\big| \\
&\quad -\mathrm{Cov}\big(f_1(X_1,\dots,X_n),h(X_1,\dots,X_n)\big) \\
&=\mathrm{Cov}\big(g_1(X_1,\dots,X_n),g_2(X_1,\dots,X_n)\big) \\
&\quad -\mathrm{Cov}\big(f_1(X_1,\dots,X_n),h(X_1,\dots,X_n)\big) \\
&=\frac{1}{2}\big[\mathrm{Cov}\big(g_1(X_1,\dots,X_n)+f_1(X_1,\dots,X_n), \\
&\qquad g_2(X_1,\dots,X_n)-h(X_1,\dots,X_n)\big) \\
&\quad +\mathrm{Cov}\big(g_1(X_1,\dots,X_n)-f_1(X_1,\dots,X_n), \\
&\qquad g_2(X_1,\dots,X_n)+h(X_1,\dots,X_n)\big)\big] \\
&\geq 0,
\end{aligned}$$

again due to the association of the random variables for the final step. If f_1 and f_2 are complex-valued functions, separate them into the real and imaginary parts and apply twice the previous upper bound. □

Remark 2.9 Notice that, if one of the functions f_1 or f_2 is real-valued, we may drop the coefficient 2 in (2.1). This inequality, for real functions, has appeared in Birkel [14], while the extension to complex-valued functions was considered in Newman [69].

These inequalities give us a means to prove an extension of Theorem 2.1 considering smooth but not necessarily monotone transformations of the random variables.

Theorem 2.10 (Bulinsky inequality) *Let X_n, $n\in\mathbb{N}$, be associated random variables. Assume that $A,B\subset\mathbb{N}$ are two finite sets and that f_1 and f_2 are real-valued functions defined on $\mathbb{R}^{|A|}$ and $\mathbb{R}^{|B|}$, respectively, partially differentiable with bounded first-order partial derivatives. Then*

$$\big|\mathrm{Cov}\big(f_1(X_i,i\in A),f_2(X_j,j\in B)\big)\big|\leq\sum_{i\in A,j\in B}\left\|\frac{\partial f_1}{\partial t_i}\right\|_\infty\left\|\frac{\partial f_2}{\partial t_j}\right\|_\infty\mathrm{Cov}(X_i,X_j). \tag{2.2}$$

Proof Define the following functions:

$$g_1(s_1,\dots,s_{|A|})=\sum_{i\in A}\left\|\frac{\partial f}{\partial t_i}\right\|_\infty s_i \quad\text{and}\quad g_2(s_1,\dots,s_{|B|})=\sum_{j\in B}\left\|\frac{\partial g}{\partial t_j}\right\|_\infty s_j.$$

Then g_1-f_1, g_1+f_1, g_2-f_2 and g_2+f_2 are nondecreasing functions, that is, $f_1\preceq g_1$ and $f_2\preceq g_2$. Then, applying Lemma 2.8 and taking into account Remark 2.9, the theorem follows immediately. □

A useful consequence of Theorem 2.10 gives an upper bound when considering transformed associated random variables.

Corollary 2.11 *Let* X_n, $n \in \mathbb{N}$, *be associated random variables. Assume that* $A, B \subset \mathbb{N}$ *are two finite sets and that* h *is a bounded real-valued function defined on* $\mathbb{R}$ *with bounded first-order derivative. Then*

$$\left|\mathrm{Cov}\left(\prod_{i\in A} h(X_i), \prod_{j\in B} h(X_j)\right)\right| \le \|h\|_\infty^{a+b-1} \|h'\|_\infty^2 \sum_{i\in A, j\in B} \mathrm{Cov}(X_i, X_j).$$

Proof Apply Theorem 2.10 to $f_1(s_1, \ldots, s_{|A|}) = h(s_1) \cdots h(s_{|A|})$ and $f_2(s_1, \ldots, s_{|B|}) = h(s_1) \cdots h(s_{|B|})$. □

Remark 2.12 Notice that the inequalities proved in Theorem 2.10 and Corollary 2.11 makes the control of dependence through the pairwise covariances.

Following the above remark, we close the section introducing an essential coefficient, firstly used by Cox and Grimmett [25], for the control of the dependence for associated variables.

Definition 2.13 Let X_n, $n \in \mathbb{N}$, be a sequence of random variables. Denote

$$u(n) = \sup_{k\in\mathbb{N}} \sum_{j:|j-k|\ge n} \mathrm{Cov}(X_j, X_k).$$

Remark 2.14 Notice that if we assume the random variables to be stationary, then

$$u(n) = \mathrm{E}X_1 + 2 \sum_{j=n+1}^{\infty} \mathrm{Cov}(X_1, X_j).$$

One can recognize this expression as the asymptotic variance in central limit theorems for dependent variables if we choose $n = 0$.

2.3 Moment Inequalities

Moment inequalities play a central role in proving asymptotic results for sums of random variables. As is well known, for independent random variables, the growth of $\mathrm{E}S_n^r$ is controlled by $n^{r/2}$. We find in this section sufficient conditions for this growth rate to hold for associated variables. The first such result was proved by Birkel [13] and later extended by Shao and Yu [94], obtaining the control for sums of transformations of associated variables, as described in Theorem 2.18 below. The route for the proof of both these versions is much alike, although the technicalities are somewhat different: find an appropriate control of covariances of powers of sums and use an induction argument to proceed. So, let us go into the first step for proving the main moment inequality, the control of covariances between sums of transformed variables.

Lemma 2.15 *Let $2 < p < r \le +\infty$, and f be an absolutely continuous function such that $\sup_{x\in\mathbb{R}} |f'(x)| \le C_0$. Assume that the random variables X_n, $n \in \mathbb{N}$, are associated with $\|f(X_n)\|_r < \infty$, $n \in \mathbb{N}$. Let A and B be two finite subsets of $\mathbb{N}$. Then*

$$\left|\mathrm{Cov}\left(\left|\sum_{i\in A} f(X_i)\right|, \left|\sum_{j\in B} f(X_j)\right|^{p-1}\right)\right|$$
$$\le p\left(\mathrm{E}\left|\sum_{j\in B} f(X_j)\right|^p\right)^{(r-1)(p-2)/r_p} \left(\sum_{i\in A}\left\| f(X_j)\right\|_r\right)^{r(p-2)/r_p}$$
$$\times\left(C_0^2 \sum_{i\in A, j\in B} \mathrm{Cov}(X_i, X_j)\right)^{(r-p)/r_p}, \tag{2.3}$$

where $r_p = r(p-1) - p$.

Proof Let $C_1 > 0$ be fixed and denote as usual by $|A|$ the cardinality of a set A. Define $g(t_1, \ldots, t_{|A|}) = |\sum_{i=1}^{|A|} f(t_i)|$ and

$$h(t_1, \ldots, t_{|B|}) = \left|\sum_{j=1}^{|B|} f(t_j)\right|^{p-1} \mathbb{I}_{\{|\sum_j f(t_j)| \le C_1\}} + C_1^{p-1} \mathbb{I}_{\{|\sum_j f(t_j)| > C_1\}}.$$

It is easily verified that

$$\left\|\frac{\partial g}{\partial t_i}\right\|_\infty \le C_1 \quad \text{and} \quad \left\|\frac{\partial h}{\partial t_j}\right\|_\infty \le (p-1) C_1^{p-2} C_0$$

(in fact, h is not differentiable at every point, but, as f has a bounded derivative, one may arbitrarily approximate h by a differentiable function and then take limits), and thus, from Theorem 2.10 it follows that

$$\left|\mathrm{Cov}\big(g(X_i, i \in A), h(X_j, j \in B)\big)\right| \le (p-1) C_1^{p-2} C_0^2 \sum_{i\in A, j\in B} \mathrm{Cov}(X_i, X_j).$$

To complete the proof, we now find an upper bound for

$$\left|\mathrm{Cov}\left(g(X_i, i \in A), \left|\sum_{j\in B} f(X_j)\right|^{p-1} - h(X_j, j \in B)\right)\right|$$
$$= \left|\mathrm{Cov}\left(g(X_i, i \in A), \left(\left|\sum_{j\in B} f(X_j)\right|^{p-1} - C_1^{p-1}\right) \mathbb{I}_{\{|\sum_{j\in B} f(X_j)| > C_1\}}\right)\right|.$$

By rewriting this expression in terms of mathematical expectations, this covariance is obviously less than or equal to

$$\max\left[\sum_{i\in A} \mathrm{E}\left(\left|f(X_i)\right| \left|\sum_{j\in B} f(X_j)\right|^{p-1} \mathbb{I}_{\{|\sum_{j\in B} f(X_j)| > C_1\}}\right),\right.$$
$$\left.\sum_{i\in A} \mathrm{E}\left|f(X_i)\right| \mathrm{E}\left(\left|\sum_{j\in B} f(X_j)\right|^{p-1} \mathbb{I}_{\{|\sum_{j\in B} f(X_j)| > C_1\}}\right)\right].$$

Applying the Hölder inequality to both terms, the above is still less than or equal to

$$\sum_{i\in A}\|f(X_i)\|_r\left(\mathrm{E}\left(\left|\sum_{j\in B}f(X_j)\right|^{(p-1)r/(r-1)}\mathbb{I}_{\{|\sum_{j\in B}f(X_j)|>C_1\}}\right)\right)^{(r-1)/r}.$$

Finally, using again the Hölder inequality followed by the Markov inequality, we find

$$\left|\mathrm{Cov}\left(g(X_i,i\in A),\left|\sum_{j\in B}f(X_j)\right|^{p-1}-h(X_j,j\in B)\right)\right|$$
$$\le C_1^{-(r-p)/r}\sum_{i\in A}\|f(X_i)\|_r\left(\mathrm{E}\left|\sum_{j\in B}f(X_j)\right|^{p}\right)^{(r-1)/r}.$$

The lemma now follows by choosing

$$C_1=\left(\frac{\sum_{i\in A}\|f(X_i)\|_r\mathrm{E}(|\sum_{j\in B}f(X_j)|^p)^{(r-1)/r}}{C_0^2\sum_{i\in A,j\in B}\mathrm{Cov}(X_i,X_j)}\right)^{r/r_p}.$$ □

Remark 2.16 Notice that, in the proof of this inequality, the association of the random variables is only used through Bulinsky's inequality.

The previous lemma is essential for the proof of the main inequality in this section, as done in Shao and Yu [94]. We still need a technical lemma to achieve this extension.

Lemma 2.17 *Let* $\alpha,\beta\in(0,1)$ *and* $x,a,b,c\ge 0$. *If* $x\le a+bx^{1-\alpha}+cx^{1-\beta}$, *then* $x\le 2a+(4b)^{1/\alpha}+(4c)^{1/\beta}$.

Proof As

$$s^\theta t^{1-\theta}\le s+t,\qquad s,t\ge 0,\theta\in[0,1],\tag{2.4}$$

it follows that $bx^{1-\alpha}=((4b)^{(1-\alpha)/\alpha}b)^\alpha(\frac{x}{4})^{1-\alpha}\le 4^{(1-\alpha)/\alpha}b^{1/\alpha}+\frac{x}{4}$. Analogously, $cx^{1-\beta}\le 4^{(1-\beta)/\beta}b^{1/\beta}+\frac{x}{4}$. Using these bounds on the assumption, the lemma follows readily. □

We may now, following Shao and Yu [94], prove a moment bound for partial sums requiring a suitable decay rate on the covariance structure, expressed using the coefficient $u(n)$ introduced in Definition 2.13.

Theorem 2.18 *Let* $2<p<r\le+\infty$, *and* f *be an absolutely continuous function such that* $\sup_{x\in\mathbb{R}}|f'(x)|\le C_0$. *Assume that the random variables* X_n, $n\in\mathbb{N}$, *are associated*, $\mathrm{E}f(X_n)=0$, $\|f(X_n)\|_r<\infty$, $n\in\mathbb{N}$, *and* $u(n)\le C_1n^{-\theta}$ *for some* $C_1>0$ *and* $\theta>0$. *Then, for each* $\varepsilon>0$, *there exists* K, *depending on* ε, r, p *and* θ, *such that*

$$
\begin{aligned}
&\mathrm{E}\left|\sum_{i=1}^{n} f(X_i)\right|^p \\
&\leq K\left[n^{1+\varepsilon}\max_{i\leq n}\mathrm{E}\big(\big|f(X_i)\big|^p\big)+\left(n\max_{i\leq n}\sum_{j=1}^{n}\big|\mathrm{Cov}\big(f(X_i),f(X_j)\big)\big|\right)^{p/2}\right. \\
&\quad + n^{(r(p-1)-p-\theta(r-p))/(r-2)\vee(1+\varepsilon)}\max_{i\leq n}\|f(X_i)\|_r^{r(p-2)/(r-2)} \\
&\quad \left.\times\big(C_0^2C_1\big)^{(r-p)/(r-2)}\right]. \qquad (2.5)
\end{aligned}
$$

Proof We shall prove the theorem by induction on the number of terms in the summation on the left side of (2.5). Notice that (2.5) is obvious for $n = 1$. So, assume that the theorem is true for each $k < n$. Denote, for each $n \in \mathbb{N}$, $T_n = \sum_{i=1}^{n} f(X_i)$ and $r_p = r(p-1) - p$, as in Lemma 2.15. Let $a \in (0, \frac{1}{2})$ be fixed, $m = [na] + 1$, and denote $k_n = [\frac{n}{2m}] + 1$. Now, decompose T_n into the sum of several blocks of length m:

$$
\xi_\ell = \sum_{j=2(\ell-1)m+1}^{n\wedge(2\ell-1)m} f(X_j) \quad \text{and} \quad \eta_\ell = \sum_{j=(2\ell-m)+1}^{n\wedge 2\ell m} f(X_j), \quad \ell = 1,\ldots,k_n.
$$

Further, define the sums of alternating blocks:

$$
T_{1,n} = \sum_{\ell=1}^{k_n}\xi_\ell \quad \text{and} \quad T_{2,n} = \sum_{\ell=1}^{k_n}\eta_\ell.
$$

It is obvious, using the binomial inequality, that $\mathrm{E}|T_n|^p \leq 2^{p-1}(\mathrm{E}|T_{1,n}|^p + \mathrm{E}|T_{2,n}|^p)$. We will concentrate on finding an upper bound for $\mathrm{E}|T_{1,n}|^p$, as the other mathematical expectation is analogous. Using again the binomial inequality, we find

$$
\begin{aligned}
\mathrm{E}|T_{1,n}|^p &= \mathrm{E}\left(\sum_{\ell=1}^{k_n}|\xi_\ell||\xi_\ell + T_{1,n} - \xi_\ell|^{p-1}\right) \\
&\leq 2^{p-2}\sum_{\ell=1}^{k_n}\mathrm{E}\big(|\xi_\ell|\big(|\xi_\ell|^{p-1} + |T_{1,n} - \xi_\ell|^{p-1}\big)\big) \\
&\leq 2^{p-2}\sum_{\ell=1}^{k_n}\mathrm{E}|\xi_\ell|^p + 2^{p-2}\sum_{\ell=1}^{k_n}\mathrm{E}\big(|\xi_\ell||T_{1,n} - \xi_\ell|^{p-1}\big). \qquad (2.6)
\end{aligned}
$$

Denote the first summation above by A_1 and the second by A_2. The association of the random variables enables the control of A_2 as, taking into account Lemma 2.15, it follows that

$$A_2 \le \sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell| \mathrm{E}|T_{1,n} - \xi_\ell|^{p-1} + \sum_{\ell=1}^{k_n} p\big(\mathrm{E}|T_{1,n} - \xi_\ell|^p\big)^{(r-1)(p-2)/r_p} \Big(m \max_{j\le m} \big\| f(X_j)\big\|_r\Big)^{r(p-2)/r_p} \times \big(C_0^2 m u(m)\big)^{(r-p)/r_p}. \tag{2.7}$$

Using the binomial and Hölder inequalities, we get that the first term above is less than or equal to

$$2^{p-2} \sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell| \mathrm{E}\big(|\xi_\ell|^{p-1} + |T_{1,n}|^{p-1}\big) \le 2^{p-2}\Bigg[\sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell|^p + \sum_{\ell=1}^{k_n} \big(\mathrm{E}|T_{1,n}|^p\big)^{(p-1)/p} \mathrm{E}|\xi_\ell|\Bigg].$$

Put $C_2 = \max_{j\le m} \| f(X_j)\|_r^{r(p-2)/r_p} (C_0^2 u(m))^{(r-p)/r_p}$. Then, using again the binomial inequality on $\mathrm{E}|T_{1,n} - \xi_\ell|^p$, we obtain

$$A_2 \le 2^{p-2} A_1 + 2^{p-2}\big(\mathrm{E}|T_{1,n}|^p\big)^{(p-1)/p} \sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell| + p2^{p-1} C_2 m \Bigg(\sum_{\ell=1}^{k_n} \big(\mathrm{E}|\xi_\ell|^p\big)^{1-(r-2)/r_p} + k_n\big(\mathrm{E}|T_{1,n}|^p\big)^{1-(r-2)/r_p}\Bigg). \tag{2.8}$$

Using (2.4), we easily see that

$$C_2 m \sum_{\ell=1}^{k_n} \big(\mathrm{E}|\xi_\ell|^p\big)^{1-(r-2)/r_p} \le \sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell|^p + k_n (C_2 m)^{r_p/(r-2)},$$

so, replacing in (2.8), we have

$$A_2 \le 2^{p-2}(1+2p) A_1 + 2^{p-2}\big(\mathrm{E}|T_{1,n}|^p\big)^{(p-1)/p} \sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell| + p2^{p-1}(C_2 m k_n)^{r_p/(r-2)} + p2^{p-1} C_2 m k_n \big(\mathrm{E}|T_{1,n}|^p\big)^{1-(r-2)/r_p}.$$

Insert now this into (2.6) and use Lemma 2.17, to find that

$$\mathrm{E}|T_{1,n}|^p \le 2^{p-2}\big(1 + 2^{p-2}(1+2p)\big) A_1 + p2^{2p-3}(C_1 m k_n)^{r_p/(r-2)} + 2^{2(p-2)}\big(\mathrm{E}|T_{1,n}|^p\big)^{(p-1)/p} \sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell| + p2^{2p-3} C_2 m k_n \big(\mathrm{E}|T_{1,n}|^p\big)^{1-(r-2)/r_p}$$

$$\leq 2^{p-1}\big(1+2^{p-2}(1+2p)\big)A_1 + 2^{2p(p-1)}\left(\sum_{\ell=1}^{k_n} \mathrm{E}|\xi_\ell|\right)^p + 3p2^{2(p-1)}(C_2 m k_n)^{r_p/(r-2)}.$$

To simplify the writing of the expressions, denote

$$a_p = 2^{p-1}\big(1+2^{p-2}(1+2p)\big)$$

and

$$b_p = 3p2^{2(p-1)} \max_{j\leq m}\|f(X_j)\|_r^{r(p-2)/r_p}\big(C_0^2 C_1\big)^{(r-p)/(r-2)}.$$

Then, recalling the assumption on $u(n)$, we have

$$\mathrm{E}|T_{1,n}|^p \leq a_p A_1 + 2^{2p(p-1)}\left(\sum_{\ell=1}^{k_n}\big(\mathrm{E}\xi_\ell^2\big)^{1/2}\right)^p + b_p (m k_n)^{r_p/(r-2)} m^{\theta(p-r)/(r-2)}.$$

Denote $v_n = \max_{i\leq n}\sum_{j=1}^n |\mathrm{Cov}(f(X_i), f(X_j))|$, so that $\mathrm{E}\xi_\ell^2 \leq m v_m \leq m v_n$. Hence,

$$\mathrm{E}|T_{1,n}|^p \leq a_p A_1 + 2^{2p(p-1)} k_n^p (m v_n)^{p/2} + b_p (m k_n)^{r_p/(r-2)} m^{\theta(p-r)/(r-2)}. \tag{2.9}$$

Of course, $T_{2,n}$ verifies an analogous inequality. Therefore,

$$\mathrm{E}|T_n|^p \leq 2^{p-1} a_p \left(\sum_{\ell=1}^{k_n}\mathrm{E}|\xi_\ell|^p + \sum_{\ell=1}^{k_n}\mathrm{E}|\eta_\ell|^p\right) + 2^{(2p+1)(p-1)+1} k_n^p (m v_n)^{p/2} + 2^p b_p (m k_n)^{r_p/(r-2)} m^{\theta(p-r)/(r-2)}.$$

We may now use the induction hypothesis to bound the summations $\sum_{\ell=1}^{k_n}\mathrm{E}|\xi_\ell|^p$ and $\sum_{\ell=1}^{k_n}\mathrm{E}|\eta_\ell|^p$, so it follows that

$$\begin{aligned}\mathrm{E}|T_n|^p \leq\ & 2^p a_p k_n K\Big(m^{1+\varepsilon}\max_{j\leq n}\mathrm{E}\big|f(X_j)\big|^p + (m v_n)^{p/2} \\ &+ 2 b_p m^{(r_p+\theta(p-r))/(r-2)\vee(1+\varepsilon)}\Big) \\ &+ 2^{(2p+1)(p-1)+1} k_n^p (m v_n)^{p/2} + 2^p b_p (m k_n)^{r_p/(r-2)} m^{\theta(p-r)/(r-2)}.\end{aligned}$$

Choosing $a = (2^{p-1}a_p)^{-1/\varepsilon}$ and

$$K = \max\left(\frac{2^{(2p+1)(p-1)+1} a^{p(\varepsilon-1)/2}}{1-a^{(p/2-1)(1+\varepsilon)}}, \frac{2^p b_p a^{(1+\varepsilon)r_p/(r-2)}}{1-2b_p a^{(1+\varepsilon)(p(r-1)/(r-2)-2)}}\right),$$

we get inequality (2.5), so the proof of the theorem is concluded. □

The inequality just proved in Theorem 2.18 above plays an important role in the study of convergence in distribution of empirical processes, allowing the control of the moments of increments needed to prove the tightness of the empirical process (refer to Sect. 5.4). An extension to the multivariate case, with an application to density estimation may be found in Masry [65]. A version of this inequality for LPQD random variables has been proved in Louhichi [61].

If we assume that $u(n)$ decreases fast enough, we may be more explicit about the growth rate of the third term in (2.5).

Corollary 2.19 *Under the assumptions of Theorem 2.18, if $\theta \geq \frac{r(p-2)}{2(r-p)}$, then, for each $\varepsilon > 0$, there exists K, depending on ε, r, p and θ, such that*

$$\begin{aligned} &\mathrm{E}\left|\sum_{i=1}^{n} f(X_i)\right|^p \\ &\quad \leq K\Bigg[n^{1+\varepsilon}\max_{i\leq n}\mathrm{E}\big(\big|f(X_i)\big|^p\big) + \Bigg(n\max_{i\leq n}\sum_{j=1}^{n}\big|\mathrm{Cov}\big(f(X_i), f(X_j)\big)\big|\Bigg)^{p/2} \\ &\qquad + n^{p/2}\max_{i\leq n}\big\|f(X_i)\big\|_r^{r(p-2)/(r-2)}\big(C_0^2C_1\big)^{(r-p)/(r-2)}\Bigg]. \end{aligned} \tag{2.10}$$

The following result is an immediate consequence of Corollary 2.19, choosing f as the identity function and $\varepsilon = \frac{p-2}{2}$, to obtain the $n^{r/2}$ growth rate for the rth moment of partial sums as for independent random variables.

Corollary 2.20 *Let $2 < p < r \leq \infty$, and X_n, $n \in \mathbb{N}$, be centred and associated random variables satisfying $u(n) \leq C_1 n^{-\theta}$ for some $C_1 > 0$, with $\theta \geq \frac{r(p-2)}{2(r-p)}$, and $\|X_n\|_r < \infty$ for $n \geq 1$. Then, there exists a constant $K = K(p,r)$ such that, for all $n \geq 1$,*

$$\begin{aligned} \mathrm{E}\left|\sum_{i=1}^{n} X_i\right|^p \leq K n^{p/2}\Bigg[&\max_{i\leq n}\mathrm{E}|X_i|^p + \Bigg(\max_{i\leq n}\sum_{j=1}^{n}\mathrm{Cov}(X_i, X_j)\Bigg)^{p/2} \\ &+ \max_{i\leq n}\|X_i\|_r^{r(p-2)/(r-2)} C_1^{(r-p)/(r-2)}\Bigg]. \end{aligned} \tag{2.11}$$

This corollary is essentially a version of the result proved by Birkel [13] that we state next for convenience later when studying invariance principles (see Sect. 5.4).

Corollary 2.21 *Let $2 \leq p < r < \infty$, and X_n, $n \in \mathbb{N}$, be centred and associated random variables such that $u(n) \leq C_1 n^{-\theta}$ for some $C_1 > 0$ and $\theta > 0$, and $\sup_{n\in\mathbb{N}} \mathrm{E}|X_n|^{r+\eta} < \infty$ for some $\eta > 0$. Then, writing $S_n = X_1 + \cdots + X_n$, there exists a constant $K > 0$ such that*

$$\sup_{m\in\mathbb{N}} \mathrm{E}\big(|S_{m+n} - S_m|^p\big) \leq K n^{p/2}. \tag{2.12}$$

Proof Just remark that, in Corollary 2.20, the constant K in (2.11) does not depend on n and the expression inside the large square brackets is bounded above by $\sup_n \mathrm{E}|X_n|^p + (u(0))^{p/2} + \sup_n \|X_i\|_r^{r(p-2)/(r-2)} C^{(r-p)/(r-2)}$, which is also independent from n. Of course, in (2.11) one does not have to start the summations at $i = 1$. □

2.4 Maximal Inequalities

An usual, the way to prove functional central limit theorems is based on suitable maximal inequalities. In fact, such inequalities are at the base of most of the arguments needed to prove the tightness of the sequences of random functions in the most popular spaces of continuous or càdlàg functions as described, for example, in Billingsley [10]. There is a huge literature on such problems that is directly inspired on this approach. Thus, we are interested on extending such inequalities to associated random variables. An application of the following results to proving functional central limit theorems will be treated later on Sect. 5.3.

The first maximal inequality appeared in Newman and Wright [71], controlling the second-order moments of maxima. We will then discuss some extensions to higher-order moments and conclude this section considering the case where we only have moments of order strictly smaller than 2.

Throughout this section we will be referring to the maxima of partial sums, so we introduce the following notation: for each $n \in \mathbb{N}$, denote $M_n = \max(S_1, \ldots, S_n)$ and $M_n^* = \max(|S_1|, \ldots, |S_n|)$.

Theorem 2.22 *Let $X_n, n \in \mathbb{N}$, be centred, square-integrable and associated random variables. Then $\mathrm{E}M_n^2 \leq \mathrm{Var}(S_n) = \mathrm{E}S_n^2$ for every $n \in \mathbb{N}$.*

Proof The inequality is trivial for $n = 1$. The proof will now be completed by an induction argument. Thus, let us assume that the result holds for the maxima of partial sums involving $n - 1$ variables and define, for each $n \in \mathbb{N}$,

$$\begin{aligned}
K_n &= \min(X_2 + \cdots + X_n, X_3 + \cdots + X_n, \ldots, X_n, 0), \\
L_n &= \max(X_2, X_2 + X_3, \ldots, X_2 + \cdots + X_n), \\
J_n &= \max(0, L_n).
\end{aligned}$$

Notice that all these variables depend only on $X_2, \ldots, X_n$. It is clear that $K_n = X_2 + \cdots + X_n - J_n$, so, as $M_n = X_1 + J_n$,

$$\begin{aligned}
\mathrm{E}M_n^2 &= \mathrm{Var}(X_1) + 2\,\mathrm{Cov}(X_1, J_n) + \mathrm{E}J_n^2 \\
&= \mathrm{Var}(X_1) + 2\,\mathrm{Cov}(X_1, X_2 + \cdots + X_n) - 2\,\mathrm{Cov}(X_1, K_n) + \mathrm{E}J_n^2.
\end{aligned}$$

The random variables K_n are increasing transformations of the original variables, so according to Theorem 1.15, they are associated and associated with X_1, and thus $\mathrm{Cov}(X_1, K_n) \geq 0$. Moreover, $J_n^2 \leq L_n^2$, so it follows that

$$\mathrm{E}M_n^2 \leq \mathrm{Var}(X_1) + 2\,\mathrm{Cov}(X_1, X_2 + \cdots + X_n) + \mathrm{E}L_n^2.$$

Finally, by the induction hypothesis, $\mathrm{E}L_n^2 \leq \mathrm{Var}(X_2 + \cdots + X_n)$, so the theorem follows. □

Given $j, n \in \mathbb{N}$, define

$$T_{j,n} = \begin{cases} j\text{th largest among } (S_1, \ldots, S_n) & \text{if } j \leq n, \\ \min(S_1, \ldots, S_n) & \text{if } j > n. \end{cases}$$

Obviously, $T_{n,n} = \min(S_1, \ldots, S_n)$ and $T_{1,n} = \max(S_1, \ldots, S_n)$. The following gives a general inequality.

Lemma 2.23 *Let X_n, $n \in \mathbb{N}$, be associated random variables, and m a nondecreasing function such that $m(0) = 0$. Then, for all $j, n \in \mathbb{N}$,*

$$\mathrm{E} \int_0^{T_{j,n}} u m(du) \leq \mathrm{E}\big(S_n m(T_{j,n})\big). \tag{2.13}$$

Thus, for every $c > 0$,

$$\lambda \mathbf{P}(T_{j,n} \geq c) \leq \int_{\{T_{j,n} \geq c\}} S_n \, d\mathbf{P}. \tag{2.14}$$

Proof Write

$$m(T_{j,n}) S_n = \sum_{k=0}^{n-1} S_{k+1}\big(m(T_{j,k+1}) - m(T_{j,k})\big) + \sum_{k=1}^{n-1} (S_{k+1} - S_k) m(T_{j,k}). \tag{2.15}$$

By the definition of $T_{j,n}$, if $k < j$, we have either $T_{j,k} = T_{j,k+1}$ or $S_{k+1} = Y_{j,k+1}$. Analogously, if $k \geq j$, we have either $T_{j,k} = T_{j,k+1}$ or $S_{k+1} \geq Y_{j,k+1}$. Thus, for every $k \geq 1$,

$$S_{k+1}\big(m(T_{j,k+1}) - m(T_{j,k})\big) \geq T_{j,k+1}\big(m(T_{j,k+1}) - m(T_{j,k})\big) \geq \int_{T_{j,k}}^{T_{j,k+1}} u m(du).$$

Now, take expectations, sum these terms, and recall that $T_{j,n} = S_n$ and that, due to the association of the random variables and that S_n are nondecreasing transformations, all the terms in the second summation of (2.15) are nonnegative, so (2.13) follows. Finally, to prove (2.14), choose $m(u) = \mathbb{I}_{[c,+\infty)}(u)$ and apply (2.13). □

Remark 2.24 The previous lemma was proved by Newman and Wright [72] for a somewhat more general framework. In fact, all that is used in the proof is just that $\mathrm{Cov}(X_{n+1}, m(S_1, \ldots, S_n)) \geq 0$. A sequence of random variables verifying this condition was called in Newman and Wright [72] a demimartingale.

We may now prove an extended version of the inequality in Theorem 2.22.

Theorem 2.25 *Let X_n, $n \in \mathbb{N}$, be centred and associated random variables. Then $\mathrm{E}T_{j,n}^2 \leq \mathrm{E}S_n^2$.*

Proof Define the random variables $Z_1 = 0$ and $Z_k = \sum_{i=n-k+2}^{n} X_i$, $k = 2, 3, \ldots, n+1$, and, for $j \le n$, $Z_{j,n}$ the jth largest among $(Z_1, \ldots, Z_n)$. Then, from Theorem 2.23 with $m(u) = u$ we have

$$\frac{1}{2}\mathrm{E}Z_{n-j+1,n}^2 \le \mathrm{E}(Z_n T_{n-j+1,n}) \le \mathrm{E}(Z_{n+1} T_{n-j+1,n}),$$

so $\mathrm{E}(Z_{n+1} - T_{n-j+1,n})^2 \le \mathrm{E}Z_{n+1}^2$, which is equivalent to the statement of this theorem. □

We may improve the upper bound for $\mathrm{E}M_n^2$ if we assume a more precise convergence decrease rate on the covariance of the random variables. The following result appeared much later in the literature (Yang, Su and Yu [108]) and was motivated by the search for convenient characterizations of the convergence rate in Strong Laws of Large Numbers.

Theorem 2.26 *Let X_n, $n \in \mathbb{N}$, be centred, square-integrable and associated random variables such that*

$$\sum_{i=1}^{\infty} u^{1/2}\big(2^i\big) < \infty. \tag{2.16}$$

Then, there exists a positive constant C such that

$$\mathrm{E}M_n^2 \le Cn\Big(\max_{k \le n} \mathrm{E}X_k^2 + 1\Big). \tag{2.17}$$

Proof For each $n \in \mathbb{N}$, define the sequence of random variables $Y_{i,n} = X_i \mathbb{I}_{[1,n]}(i)$, $i = 1, 2, \ldots$. These random variables are obtained as nondecreasing transformations of the original ones and thus are associated. Consider, on the sequel, n fixed. Given $j, k \in \mathbb{N}$, define $S_j(k) = Y_{j+1,n} + \cdots + Y_{j+k,n}$ and, to deal with the nonstationarity, $s_k = \sup_{j \in \mathbb{N}} \|S_j(k)\|_2$, where $\|S_j(k)\|_2 = (\mathrm{E}(S_j(k)))^{1/2}$ is the L^2 norm of $S_j(k)$. Then, obviously,

$$\begin{aligned}
\big\|S_j(2k)\big\|_2 &\le \big\|S_j(k) + S_{j+k+[k^{1/3}]}(k)\big\|_2 + \big\|S_{j+k}\big(\big[k^{1/3}\big]\big)\big\|_2 + \big\|S_{j+2k}\big(\big[k^{1/3}\big]\big)\big\|_2 \\
&\le \big\|S_j(k) + S_{j+k+[k^{1/3}]}(k)\big\|_2 + 2k^{1/3} \sup_{i \in \mathbb{N}} \|Y_{i,n}\|_2 \\
&= \big\|S_j(k) + S_{j+k+[k^{1/3}]}(k)\big\|_2 + 2k^{1/3} \max_{i \le n} \|X_i\|_2.
\end{aligned}$$

Now we use the association of the random variables, implying that the covariances are nonnegative, so that

$$\begin{aligned}
\big\|S_j(k) + S_{j+k+[k^{1/3}]}(k)\big\|_2^2 &= \mathrm{E}\big(S_j(k) + S_{j+k+[k^{1/3}]}(k)\big)^2 \\
&\le 2s_k^2 + 2\sum_{i=j+1}^{j+k} \sum_{\ell=j+k+[k^{1/3}]+1}^{\infty} \mathrm{Cov}(Y_{i,n}, Y_{\ell,n}) \\
&\le 2s_k^2 + 2\sum_{i=j+1}^{j+k} u\big(\big[k^{1/3}\big]\big) = 2s_k^2 + 2ku\big(\big[k^{1/3}\big]\big).
\end{aligned}$$

Thus, inserting this in the previous majorization, it follows

$$s_{2k} \le \sqrt{2} s_k + \sqrt{2ku\big(\big[k^{1/3}\big]\big)} + 2k^{1/3} s_1.$$

We now use recursively the previous bound to find

$$\begin{aligned} s_{2^r} &\le 2^{r/2} s_1 + 2s_1 \sum_{i=0}^{r-1} 2^{(r-1-i)/2+i/3} + 2^{r/2} \sum_{i=0}^{r-1} \sqrt{u\big(\big[2^{i/3}\big]\big)} \\ &\le 2^{r/2} s_1 + 2^{(r+1)/2} s_1 \sum_{i=0}^{\infty} 2^{-i/6} + 2^{r/2} \sum_{i=0}^{\infty} \sum_{j=3i}^{3i+2} \sqrt{u\big(\big[2^{j/3}\big]\big)} \\ &\le 14 \times 2^{r/2} s_1 + 3 \times 2^{r/2} \sum_{i=0}^{\infty} \sqrt{u\big(\big[2^{i}\big]\big)} \\ &\le C 2^{r/2}(s_1 + 1), \end{aligned}$$

where $C = \max(14, 3\sum_{i=0}^{\infty} \sqrt{u([2^i])})$. Assume now that $2^r \le k < 2^{r+1}$. Then, due to the association,

$$\mathrm{E}S_k^2 \le \mathrm{E}\big(S_0^2\big(2^{r+1}\big)\big) \le s_{2^{2+1}}^2 \le C2^{(r+1)/2}\big(s_1^2+1\big) \le 2C2^{2/r}\big(s_1^2+1\big),$$

so, from Theorem 2.22 the result follows. □

Assumption (2.16) that has been used for the first time in Yang [107] is a rather mild one. In fact, (2.16) is verified if $u(n)$ is of order $(\log n)^{-2}(\log\log n)^{-3}$, much weaker than the typical hypothesis used in Sect. 2.3, where a polynomial decrease rate was often assumed.

The result above assumes the existence of second-order moments, but, being a statement that does not require stationarity, these moments may be unbounded. Of course, the case where the second-order moments are bounded is included in the framework of Theorem 2.26. As could be expected, it is possible to prove a version of the upper bound better adapted to this situation, which will be explored later when dealing with truncated variables, allowing to explore the behaviour of the truncating sequence.

Theorem 2.27 *Let X_n, $n \in \mathbb{N}$, be centred and associated random variables such that* $\sup_{n\in\mathbb{N}} \mathrm{E}X_n^2 < \infty$ *and*

$$K = \sup_{j\in\mathbb{N}} \sum_{k:k-j>1}^{\infty} \mathrm{Cov}^{1/2}(X_j, X_k) < \infty. \tag{2.18}$$

Then, for every $n \in \mathbb{N}$,

$$\mathrm{E}M_n^2 \le 2n \sup_{n\in\mathbb{N}} \mathrm{E}X_n^2 + 4nK\Big(\sup_{n\in\mathbb{N}} \mathrm{E}X_n^2\Big)^{1/2}.$$

Proof First remark that, obviously, for every $j \le n$,

$$S_j^2 \le \max\Bigl(\Bigl(\min_{k\le n} S_k\Bigr)^2, \Bigl(\max_{k\le n} S_k\Bigr)^2\Bigr) \le \Bigl(\min_{k\le n} S_k\Bigr)^2 + \Bigl(\max_{k\le n} S_k\Bigr)^2,$$

so it follows that $M_n^2 \le (\min_{k\le n} S_k)^2 + (\max_{k\le n} S_k)^2 \le 2(\max_{k\le n} S_k)^2$. Thus, applying Theorem 2.25, we have that $\mathrm{E}(\max_{k\le n} S_k)^2 \le \mathrm{E}S_n^2$, so

$$\mathrm{E}\Bigl(\max_{k\le n} S_k^2\Bigr) \le 2\mathrm{E}S_n^2 = 2\sum_{j,k=1}^{n} \mathrm{Cov}(X_j, X_k).$$

Define $K_1 = \sup_{n\in\mathbb{N}} \mathrm{E}X_n^2$. Then, it is obvious that

$$\mathrm{Cov}(X_j, X_k) \le \bigl(\mathrm{E}X_j^2 \mathrm{E}X_k^2\bigr)^{1/2} \le K_1$$

and

$$\mathrm{Cov}(X_j, X_k) \le \bigl(K_1 \,\mathrm{Cov}(X_j, X_k)\bigr)^{1/2}.$$

Thus,

$$\sum_{j,k=1}^{n} \mathrm{Cov}(X_j, X_k) = \sum_{j=1}^{n} \mathrm{E}X_j^2 + 2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} \mathrm{Cov}(X_j, X_k) \le K_1 n + 2K_1^{1/2} nK,$$

so the result follows. □

We now prove an extension to associated random variables of a well-known maximal inequality under independence.

Theorem 2.28 *Let X_n, $n \in \mathbb{N}$, be centred, square-integrable and associated random variables. Then, for all $\lambda > 0$ and $n \in \mathbb{N}$,*

$$\mathbf{P}\bigl(M_n^* \ge \lambda s_n\bigr) \le 2\mathbf{P}\bigl(|S_n| \ge (\lambda - \sqrt{2})s_n\bigr). \tag{2.19}$$

Proof Define $M_n^+ = \max(0, S_1, \ldots, S_n)$. Given real numbers $x < y$, we have

$$\begin{aligned}
&\mathbf{P}\bigl(M_n^+ \ge y\bigr)\\
&\quad \le \mathbf{P}(S_n \ge x) + \mathbf{P}\bigl(M_{n-1}^+ \ge y, M_{n-1}^+ - S_n > y - x\bigr)\\
&\quad \le \mathbf{P}(S_n \ge x) + \mathbf{P}\bigl(S_{n-1}^+ \ge y\bigr)\mathbf{P}\bigl(M_{n-1}^+ - S_n > y - x\bigr)\\
&\quad \le \mathbf{P}(S_n \ge x) + \mathbf{P}\bigl(M_n^+ \ge y\bigr)\frac{\mathrm{E}(M_{n-1}^+ - S_n)^2}{(y-x)^2},
\end{aligned}$$

using Theorem 2.3, as M_{n-1}^+ and $S_n - M_{n-1}^+$ are associated. The mathematical expectation above may be rewritten as

$$\mathrm{E}\bigl(M_{n-1}^+ - S_n\bigr)^2 = \mathrm{E}\bigl[\max(X_n, X_n + X_{n-1}, \ldots, X_n + \cdots + X_1)^2\bigr] \le \mathrm{E}S_n^2,$$

taking into account Theorem 2.22. Now, if $(y-x)^2 \ge s_n^2 = \mathrm{E}S_n^2$, it follows then that

$$\mathbf{P}\bigl(M_n^+ \ge y\bigr) \le \frac{(y-x)^2}{(y-x)^2 - s_n^2}\mathbf{P}(S_n \ge x). \tag{2.20}$$

Repeating these arguments with $M_n^- = \max(0, -S_1, \ldots, -S_n)$ and adding to (2.20), we find, whenever $y - x \geq \sqrt{2}s_n$,

$$\mathbf{P}\big(M_n^* \geq y\big) \leq 2\mathbf{P}\big(|S_n| \geq x\big). \tag{2.21}$$

Finally, choosing $y = \lambda s_n$ and $x = (\lambda - \sqrt{2})s_n$, it follows that

$$\mathbf{P}\big(M_n^* \geq \lambda s_n\big) \leq 2\mathbf{P}\big(|S_n| \geq (\lambda - \sqrt{2})s_n\big).$$

□

Notice that (2.19) has exactly the same form as in the independent case (see Sect. 10 in Billingsley [10], for example).

The previous theorem implies, in a very simple way, a maximal inequality for higher-order moments, thus extending Theorem 2.22.

Corollary 2.29 *Let X_n, $n \in \mathbb{N}$, be centred and associated random variables with finite moments of order $p \geq 2$. Then $\mathrm{E}M_n^p \leq \sqrt{2}(\mathrm{E}|S_n|^p)^{1/p} + 2\mathrm{E}S_n^p$.*

Proof Recall that $M_n^+ = \max(0, S_1, \ldots, S_n)$. As M_n^+ is a nonnegative variable, by (2.19) and Hölder's inequality, it follows that, with $s_n^2 = \mathrm{E}S_n^2$,

$$\begin{aligned}
&\mathrm{E}\big(M_n^+\big)^p \\
&\quad= \int_0^{+\infty} \mathbf{P}\big(\big(M_n^+\big)^p > y\big)\,dy \leq \sqrt{2}s_n + 2\int_{\sqrt{2}s_n}^{+\infty} \mathbf{P}\big(|S_n|^p > y - \sqrt{2}s_n\big)\,dy \\
&\quad\leq \sqrt{2}s_n + 2\mathrm{E}S_n^p \leq \sqrt{2}\big(\mathrm{E}|S_n|^p\big)^{1/p} + 2\mathrm{E}S_n^p.
\end{aligned}$$

Obviously, the same applies to $M_n^- = \max(0, -S_1, \ldots, -S_n)$, so the result follows. □

If we only have moments of order smaller than 2, it is still possible to prove bounds for the corresponding moments of maxima. These will be useful later on, when analysing extensions of Strong Laws of Large Numbers. In this case we need a preparatory lemma providing control on the tail of M_n. In order to state the result, we need to introduce some extra notation. Recall the definition of $H_{j,k}$ (see page 3): $H_{j,k}(x, y) = \mathbf{P}(X_j > x, X_k > y) - \mathbf{P}(X_j > x)\mathbf{P}(X_k > y)$. Next, given $v > 0$, define $g_v(u) = \max(\min(u, v), -v)$ and

$$G_{j,k}(v) = \mathrm{Cov}\big(g_v(X_j), g_v(X_k)\big) = \int\int_{[-v,v]^2} H_{j,k}(x, y)\,dx\,dy.$$

It is obvious that $G_{j,k}(+\infty) = \mathrm{Cov}(X_j, X_k)$. Moreover, as each g_v is an increasing function bounded by v^2, it follows, assuming that the random variables X_n are associated, that $0 \leq G_v(x, y) \leq v^2$.

Lemma 2.30 *Let X_n, $n \in \mathbb{N}$, be associated random variables. Assume that there exists a nonnegative random variable Y such that, for every $x > 0$, $\sup_{n\in\mathbb{N}} \mathbf{P}(|X_n| > x) \le \mathbf{P}(Y > x)$. Then, for every $x, m > 0$,*

$$\mathbf{P}(M_n > x) \le \frac{4n}{x^2}\mathrm{E}\big(Y^2 \mathbb{I}_{\{Y\le m\}}\big) + \frac{4n}{x}\mathrm{E}(Y\mathbb{I}_{\{Y>m\}}) + \frac{4nm^2}{x^2}\mathbf{P}(Y>m) + \frac{8}{x^2}\sum_{\substack{j,k=1\\ j\ne k}}^{n} G_{j,k}(m).$$

Proof Let $m > 0$ be fixed and define, for each $j \in \mathbb{N}$,

$$X_{1,j} = g_m(X_j), \qquad S_{1,n} = \sum_{j=1}^{n}(X_{1,j} - \mathrm{E}X_{1,j}),$$
$$M_{1,n} = \max_{k\le n} S_{1,k}, \qquad X_{2,j} = X_j - X_{1,j}.$$

As, obviously,

$$M_n \le M_{1,n} + \sum_{j=1}^{n}\big(|X_{2,j}| + \mathrm{E}|X_{2,j}|\big),$$

it follows from Markov's inequality that

$$\mathbf{P}(M_n > x) \le \mathbf{P}\left(M_{1,n} > \frac{x}{2}\right) + \frac{4}{x}\sum_{j=1}^{n}\mathrm{E}|X_{2,j}| \le \frac{4}{x^2}\mathrm{E}M_{1,n}^2 + \frac{4}{x}\sum_{j=1}^{n}\mathrm{E}|X_{2,j}|.$$

The random variables $X_{1,n}$ are associated, so we may apply Theorem 2.22 to bound the first term on the right above, to find

$$\begin{aligned}\mathbf{P}(M_n > x) &\le \frac{4}{x^2}\mathrm{E}S_{1,n}^2 + \frac{4}{x}\sum_{j=1}^{n}\mathrm{E}|X_{2,j}| \\ &\le \frac{4}{x^2}\sum_{j=1}^{n}\mathrm{E}X_{1,j}^2 + \frac{8}{x^2}\sum_{\substack{j,k=1\\ j\ne k}}^{n}\mathrm{Cov}(X_{1,j}, X_{1,k}) + \frac{4}{x}\sum_{j=1}^{n}\mathrm{E}|X_{2,j}|.\end{aligned}$$

Finally, notice that $\mathrm{E}X_{1,j}^2 = \int \mathbf{P}(X_{1,j}^2 > x)\,dx \le \mathrm{E}(Y^2\mathbb{I}_{\{Y\le m\}}) + m^2\mathbf{P}(Y > m)$ and, analogously, $\mathrm{E}|X_{2,j}| \le \mathrm{E}((|X_j| - m)\mathbb{I}_{\{|X_j|>m\}}) \le \mathrm{E}(Y\mathbb{I}_{\{Y>m\}})$. □

Now we may prove the announced maximal inequality.

Theorem 2.31 *Let X_n, $n \in \mathbb{N}$, be associated random variables. Assume that there exists a nonnegative random variable Y such that, for some $p \in (1,2)$, $\mathrm{E}Y^p < \infty$*

and, for every $x > 0$, $\sup_{n\in\mathbb{N}} \mathbf{P}(|X_n| > x) \le \mathbf{P}(Y > x)$. *Then, there exists a constant* $c_p > 0$, *depending only on* p, *such that*

$$\mathrm{E}\Big(\max_{k\le n} |S_k|^p\Big) \le c_p\left(n\mathrm{E}Y^p + \sum_{\substack{j,k=1\\ j\ne k}}^{n} \int_0^\infty x^{p-3} G_{j,k}(x)\,dx\right). \tag{2.22}$$

Proof Applying Lemma 2.30 with $m = x$ and taking into account that $p < 2$, we have

$$\begin{aligned}
\mathrm{E}M_n^p &= p\int_0^\infty x^{p-1}\mathbf{P}(M_n > x)\,dx \\
&\le p\int_0^\infty x^{p-1}\left(\frac{4n}{x^2}\mathrm{E}\big(Y^2\mathbb{I}_{\{Y\le x\}}\big) + \frac{4n}{x}\mathrm{E}(Y\mathbb{I}_{\{Y>x\}}) + 4n\mathbf{P}(Y > x)\right)dx \\
&\quad + p\int_0^\infty x^{p-1}\frac{8}{x^2}\sum_{\substack{j,k=1\\ j\ne k}}^{n} G_{j,k}(x)\,dx \\
&\le 8n\int_0^\infty x^{p-3}\mathrm{E}\big(Y^2\mathbb{I}_{\{Y\le x\}}\big) + x^{p-2}\mathrm{E}(Y\mathbb{I}_{\{Y>x\}}) + x^{p-1}\mathbf{P}(Y \ge x)\,dx \\
&\quad + 8\int_0^\infty x^{p-3}\sum_{\substack{j,k=1\\ j\ne k}}^{n} G_{j,k}(x)\,dx \\
&\le 8n\left(\frac{1}{2-p} + \frac{1}{p-1} + \frac{1}{p}\right)\mathrm{E}Y^p + 8\int_0^\infty x^{p-3}\sum_{\substack{j,k=1\\ j\ne k}}^{n} G_{j,k}(x)\,dx
\end{aligned}$$

and choose $c_p = 8\max(\frac{1}{2-p} + \frac{1}{p-1} + \frac{1}{p}, 1)$. Finally, notice that one may replace the variables X_n by $-X_n$, keeping the association, so the previous inequality also applies, thus proving the theorem. □

2.5 Characteristic Functions

This section presents a few inequalities, first proved by Newman [68] (see also Newman [70] for a more complete presentation), controlling the distance between joint distributions and the product of marginal distributions, based on characteristic functions. As it will be shown, this distance is completely characterized by the covariances of the variables, and thus, when seeking for asymptotic results, it becomes natural to look for assumptions on the covariance structure. These inequalities will play a major role in proving central limit theorems for associated variables. An analogous inequality concerning moment generating functions is also proved. This later inequality, first used by Dewan and Prakasa Rao [30] in a different context, will be a useful tool for proving exponential inequalities of the next section and to derive convergence rates.

We start by setting some notation.

Definition 2.32 Given the random variable X, we denote its *characteristic function* by

$$\varphi_X(u) = \mathrm{E}e^{iuX}, \quad u \in \mathbb{R}.$$

Given random variables $X_1, \ldots, X_n$, we denote their *joint characteristic function* by

$$\varphi_{(X_1,\ldots,X_n)}(u_1, \ldots, u_n) = \mathrm{E}e^{i(u_1X_1+\cdots+u_nX_n)}, \quad u_1, \ldots, u_n \in \mathbb{R}.$$

Given a set $A \subset \{1, \ldots, n\}$, we denote

$$\varphi_A(u_1, \ldots, u_n) = \mathrm{E}e^{i\sum_{j\in A} u_j X_j}.$$

If $A = \{j\}$, we denote, for simplicity, φ_A by φ_j.

We first prove the inequality that describes the control of the distance between joint distributions and the product of marginal distributions for two random variables. The proof of the main result in this section, Theorem 2.37, is built on this version proceeding by induction.

Lemma 2.33 *Let X and Y be associated random variables. Then, for every $u, v \in \mathbb{R}$,*

$$\left|\mathrm{E}e^{iuX+ivY} - \mathrm{E}e^{iuX}\mathrm{E}e^{ivY}\right| \le |uv| \,\mathrm{Cov}(X, Y). \tag{2.23}$$

Proof The expression inside the absolute value may be rewritten as $\mathrm{Cov}(e^{iuX}, e^{ivY})$. Now, representing by $\mathbf{P}_X$, $\mathbf{P}_Y$ and $\mathbf{P}_{(X,Y)}$ the distributions of X, Y and (X, Y), respectively, and recalling that $H(s,t) = \mathbf{P}(X > s, Y > t) - \mathbf{P}(X > s)\mathbf{P}(Y > t)$ and integrating by parts (see Theorem C.4), we find

$$\begin{aligned}\mathrm{Cov}\big(e^{iuX}, e^{ivY}\big) &= \int\int e^{ius+ivt}(\mathbf{P}_{(X,Y)} - \mathbf{P}_X \otimes \mathbf{P}_Y)(ds\,dt)\\ &= \int\int \frac{\partial^2}{\partial s\,\partial t} e^{ius+ivt} H(s,t)\,ds\,dt\\ &= \int\int iue^{ius} ive^{ivt} H(s,t)\,ds\,dt.\end{aligned}$$

Noticing that due to the association of the variables, H is a nonnegative function, the lemma follows immediately from

$$\left|\mathrm{Cov}\big(e^{iuX}, e^{ivY}\big)\right| \le |uv| \int\int H(s,t)\,ds\,dt = |uv|\,\mathrm{Cov}(X, Y),$$

using Hoeffding's formula (1.2). □

Remark 2.34 Notice that the proof depends only on the fact that H has constant sign. Thus, the previous result holds for positively dependent variables or for negatively dependent variables taking, in this later case, the absolute value of the covariance on the left-hand side of (2.23). Of course, for general random variables, the upper bound would be

$$|uv| \int \int \big|H(s,t)\big| \, ds \, dt.$$

Before stating an extension of this lemma, it is useful to relate the covariances H with the covariances between the original random variables. This is a direct consequence of the two-dimensional version of the classical Berry–Esséen inequalities, given in Theorem A.2.

Lemma 2.35 *Let X and Y be associated random variables with absolutely continuous distributions. Assume that the marginal densities f_X and f_Y are bounded by M. Then, for every $T > 0$,*

$$H(x,y) = \mathrm{Cov}\big(\mathbb{I}_{(-\infty,x]}(X), \mathbb{I}_{(-\infty,y]}(Y)\big) \le M^* \left(T^2 \,\mathrm{Cov}(X,Y) + \frac{1}{T}\right), \tag{2.24}$$

where $M^* = \max(\frac{2}{\pi^2}, 45M)$.

Proof Using Corollary A.3 together with (2.23), the lemma follows immediately. □

Optimizing the choice of T on the previous result, we find the following important inequality.

Corollary 2.36 *Under the same assumptions as in Lemma* 2.35, *if* $\mathrm{Cov}(X,Y) > 0$, *we have that*

$$\mathrm{Cov}\big(\mathbb{I}_{(-\infty,x]}(X), \mathbb{I}_{(-\infty,y]}(Y)\big) \le \frac{1}{M^*} \mathrm{Cov}^{1/3}(X,Y). \tag{2.25}$$

Proof In (2.24), choose $T = (2\,\mathrm{Cov}(X,Y))^{-1/3}$, and the inequality follows. □

This inequality plays an important role while studying invariance principles and the asymptotics for the density and regression estimators that depend on transformations using indicator functions on the sequence of random variables. In fact, (2.25) enables the control of covariances between indicator functions using covariances between the original random variables. Thus, it gives a way to obtain sufficient conditions expressed in terms of the initial variables.

Now we extend Lemma 2.33 to any number of associated variables.

Theorem 2.37 (Newman inequality) *Let X_n, $n \in \mathbb{N}$, be associated variables. Then, for all $n \in \mathbb{N}$ and $u_1, \ldots, u_n \in \mathbb{R}$,*

$$\left| \varphi_{(X_1,\ldots,X_n)}(u_1,\ldots,u_n) - \prod_{j=1}^{n} \varphi_j(u_j) \right| \leq \frac{1}{2} \sum_{\substack{j,k=1 \\ j\neq k}}^{n} |u_j u_k| \operatorname{Cov}(X_j, X_k). \qquad (2.26)$$

Proof When $n = 2$, inequality (2.26) reduces to (2.23), so we may proceed by induction on n to prove this theorem. Assume then that (2.26) holds whenever there are only $n-1$ variables involved. To prove the inequality for n variables, split the set $\{1,\ldots,n\}$ in the following way:

(a) if all the $u_1,\ldots,u_n$ have the same sign, take $A = \{1,\ldots,n-1\}$ and $B = \{n\}$;
(b) if not all the $u_1,\ldots,u_n$ have the same sign, take $A = \{j \in \{1,\ldots,n\} : u_j > 0\}$ and $B = \{1,\ldots,n\} \setminus A$.

Define now the variables $U = \sum_{j\in A} |u_j| X_j$ and $V = \sum_{j \in B} |u_j| X_j$. Notice that these variables are increasing transformations of the X_n's, so they are still associated. Moreover, we can write

$$\varphi_{(X_1,\ldots,X_n)}(u_1,\ldots,u_n) = \mathrm{E} e^{i(U-V)} = \varphi_{U-V}(1),$$
$$\varphi_A(u_j, j\in A) = \varphi_U(1) \quad \text{and} \quad \varphi_B(u_j, j \in B) = \varphi_V(-1).$$

Then,

$$\begin{aligned}
&\left| \varphi_{(X_1,\ldots,X_n)}(u_1,\ldots,u_n) - \prod_{j=1}^{n} \varphi_j(u_j) \right| \\
&\quad \leq \left| \varphi_{U-V}(1) - \varphi_U(1)\varphi_V(-1) \right| + \left|\varphi_U(1)\right| \left| \varphi_V(-1) - \prod_{j\in B} \varphi_j(u_j) \right| \\
&\qquad + \left| \prod_{j\in B} \varphi_j(u_j) \right| \left| \varphi_U(1) - \prod_{j \in A} \varphi_j(u_j) \right|. \qquad (2.27)
\end{aligned}$$

Of course, characteristics functions have absolute values bounded by 1, so the second and third terms may be bounded using the induction hypothesis. It remains to bound the first term on (2.27): as U and V are associated, we refer to Lemma 2.33 and, with respect to the notation used in this lemma, choose, for the case (a), $u = v = 1$ if the u_j are positive or $u = v = -1$ if the u_j are negative, and for the case (b), $u = 1$ and $v = -1$. Applying now Lemma 2.33 and the induction hypothesis, we immediately get (2.27). □

Remark 2.38 A close look at the proof shows that the association assumption can be weakened. In fact, what is used throughout the proof is the fact that linear combinations with nonnegative coefficients (notice that we multiply by -1 the coefficients that are negative) of the random variables have nonnegative covariance. Thus, the previous result holds if the random variables are LPQD (see Definition 1.58).

Remark 2.39 Notice that, as far as what convergence in distribution is concerned, inequality (2.26) means that the covariance structure of the random variables completely determines the properties of the approximation of joint distributions to independence. This remark is at the heart of most of the results included in Chap. 4. It also justifies that, for random variables, it is natural to seek for assumptions on the behaviour of these covariances.

Next, we prove an extension of Theorem 2.37, allowing for applications going beyond associated random variables themselves, by considering suitable transformations of the variables. Of course, Theorem 2.37 still applies if we consider transformations of the initial variables that are either all increasing or all decreasing as, according to Theorem 1.15, such transformations keep the association. However, it is possible to prove a version of (2.26) for nonmonotone transformations of associated variables, using Lemma 2.8, if these functions are dominated by nondecreasing ones, as described by the relation "$\preceq$" introduced in Definition 2.4.

Theorem 2.40 *Let Y_n, $n \in \mathbb{N}$, be associated random variables. Assume that, for each $n \in \mathbb{N}$, f_n, g_n are real-valued functions such that $f_n \preceq g_n$, and denote $X_n = f_n(Y_1, Y_2, \ldots)$ and $X_n^* = g_n(Y_1, Y_2, \ldots)$. Then, for every $n \in \mathbb{N}$, given $A, B \subset \{1, \ldots, n\}$ and $u_1, \ldots, u_n \in \mathbb{R}$,*

$$\begin{aligned} &\left|\varphi_{A\cup B}(u_1, \ldots, u_n) - \varphi_A(u_1, \ldots, u_n)\varphi_B(u_1, \ldots, u_n)\right| \\ &\quad \le 2 \sum_{j\in A, k\in B} |u_j u_k| \left|\operatorname{Cov}\left(X_j^*, X_k^*\right)\right| \end{aligned} \tag{2.28}$$

and

$$\left|\varphi_{(X_1,\ldots,X_n)}(u_1, \ldots, u_n) - \prod_{j=1}^{n} \varphi_j(u_j)\right| \le 2 \sum_{\substack{j,k=1 \\ j\neq k}}^{n} |u_j u_k| \operatorname{Cov}\left(X_j^*, X_k^*\right). \tag{2.29}$$

Proof To prove (2.28), define $f_1(u_1, \ldots, u_n) = \exp(i \sum_{j\in A} u_j X_j)$ and $f_2(u_1, \ldots, u_n) = \exp(i \sum_{k\in B} u_k X_k)$. As $\sum_{j\in A} u_j f_j \preceq \sum_{j\in A} |u_j| g_j$, Proposition 2.7 applies, so (2.28) is an immediate consequence of Lemma 2.8.

To prove (2.29), argue by induction as in the proof of Theorem 2.37, decomposing the set $\{1, \ldots, n\}$ in exactly the same way and using decomposition (2.27). As previously, the second and third terms of this decomposition are controlled directly from the induction hypothesis. To control the first term, define $U = f_1(u_1, \ldots, u_n)$ and $V = f_s(u_1, \ldots, u_n)$, which, according to Proposition 2.7, verify $U \preceq \sum_{j\in A} |u_j| g_j$ and $V \preceq \sum_{k\in B} |u_k| g_k$, and apply again Lemma 2.8. □

A straightforward modification of the proof of Lemma 2.33 gives an upper bound in terms of moment generating functions. In fact, we may write

$$\begin{aligned} \operatorname{Cov}\left(e^{\lambda X}, e^{\lambda Y}\right) &= \int\int e^{\lambda(s+t)} (\mathbf{P}_{(X,Y)} - \mathbf{P}_X \otimes \mathbf{P}_Y)(ds\, dt) \\ &= \int\int \lambda^2 e^{\lambda(s+t)} H(s,t)\, ds\, dt. \end{aligned}$$

The following statement is now obvious.

Lemma 2.41 *Let X and Y be associated random variables such that $|X|, |Y| \leq C$ for some constant $C > 0$. Then, for every $\lambda \in \mathbb{R}$,*

$$\left|\mathrm{E}e^{\lambda(X+Y)} - \mathrm{E}e^{\lambda X}\mathrm{E}e^{\lambda Y}\right| \leq \lambda^2 e^{2\lambda C}\,\mathrm{Cov}(X,Y). \tag{2.30}$$

Taking into account that, for any $A \subset \{1,\dots,n\}$, $\mathrm{E}e^{\lambda\sum_{j\in A}X_j} \leq e^{\lambda|A|C}$, where $|A|$ is the number of elements in A, the extension for n random variables is immediate, following the arguments of the proof of Theorem 2.37.

Theorem 2.42 *Let X_n, $n \in \mathbb{N}$, be associated variables such that $|X_n| \leq C$ for some constant $C > 0$, not depending n. Then, for all $n \in \mathbb{N}$ and $\lambda \in \mathbb{R}$,*

$$\left|\mathrm{E}e^{\lambda\sum_{j=1}^n X_j} - \prod_{j=1}^n \mathrm{E}e^{\lambda X_j}\right| \leq \frac{\lambda^2}{2}e^{n\lambda C}\sum_{\substack{i,j=1\\ i\neq j}}^n \mathrm{Cov}(X_i,X_j). \tag{2.31}$$

2.6 Exponential Inequalities

One of the main tools used for characterizing convergence rates in strong laws has been convenient versions of the so-called Bernstein-type exponential inequalities. There exist several versions of such inequalities available in the literature for independent sequences of variables with assumptions of uniform boundedness or some, quite relaxed, control on their (centred or noncentred) moments. For associated random variables, a first exponential inequality was proved by Prakasa Rao [83], but it was too weak to be really useful. In fact, this inequality does not even recover the known results if one assumes the variables to be independent. A stronger inequality, effectively extending results from the independent case, was proved by Ioannides and Roussas [48]. The technique was based on decomposing S_n into the sum of convenient blocks in both cases. But, while Prakasa Rao [83] tried to control everything just using properties of the exponential function, Ioannides and Roussas [48] controlled the mathematical expectations by coupling the blocks by independent ones. The route of the proof consists then in controlling the distance between the original blocks and the coupling independent variables, achieved using an induction argument, and finding convenient bounds for the independent coupling terms. Their inequality was later extended in Oliveira [75], where the approximation to independence was controlled in a different way, based on Theorem 2.42, avoiding the induction argument and dropping some technical assumptions appearing in course of the proof proposed by Ioannides and Roussas [48]. One way to have some insight on how optimized the exponential inequality is was to use it to characterize convergence rates for Strong Laws of Large Numbers to see how close the optimal rates for independent variables remain. This will be explored later in Sect. 3.2. We should note at this point that the main term in such convergence rate characterizations is the one that controls the independent coupling terms. This led to some effort in improving the control of these independent like terms by Sun [98], Xing, Yang and Liu [106], Henriques and Oliveira [43] and Xing and Yang [105], but we defer this to Sect. 3.2. Nevertheless, in all these references, the way association was used to find the exponential inequality was essentially the same, and that is what we will be concentrating on this section.

Let us now introduce the notation to be used throughout this section. Let c_n, $n \in \mathbb{N}$, be a sequence of nonnegative real numbers such that $c_n \longrightarrow +\infty$ and, given the random variables X_n, $n \in \mathbb{N}$, define, for all $i, n \geq 1$,

$$\begin{aligned} X_{1,i,n} &= -c_n \mathbb{I}_{(-\infty,-c_n)}(X_i) + X_i \mathbb{I}_{[-c_n,c_n]}(X_i) + c_n \mathbb{I}_{(c_n,+\infty)}(X_i), \\ X_{2,i,n} &= (X_i - c_n)\mathbb{I}_{(c_n,+\infty)}(X_i), \qquad X_{3,i,n} = (X_i + c_n)\mathbb{I}_{(-\infty,-c_n)}(X_i). \end{aligned} \tag{2.32}$$

For each fixed $n \geq 1$, the variables $X_{1,1,n}, \ldots, X_{1,n,n}$ are uniformly bounded, and thus they may be treated using Theorem 2.42. Note that, for each fixed $n \geq 1$, all these variables are monotone transformations of the initial variables X_n. This implies that an association assumption is preserved by this construction.

The proof of an exponential inequality will use, besides the truncation introduced before, a convenient decomposition of the sums into blocks. This block decomposition is a means to an approximation to independence technique on the truncated variables. The tails will be treated directly using Laplace transforms.

Consider a sequence of natural numbers p_n such that, for each $n \geq 1$, $p_n < \frac{n}{2}$ and define r_n as the greatest integer less than or equal to $\frac{n}{2p_n}$. Define then, for $q = 1, 2, 3$ and $j = 1, \ldots, 2r_n$,

$$Y_{q,j,n} = \sum_{\ell=(j-1)p_n+1}^{jp_n} (X_{q,\ell,n} - \mathrm{E}X_{q,\ell,n}). \tag{2.33}$$

Finally, for all $q = 1, 2, 3$ and $n \geq 1$, define

$$\begin{aligned} Z_{q,n,od} &= \sum_{j=1}^{r_n} Y_{q,2j-1,n}, \qquad Z_{q,n,ev} = \sum_{j=1}^{r_n} Y_{q,2j,n}, \\ R_{q,n} &= \sum_{\ell=2r_n p_n+1}^{n} (X_{q,\ell,n} - \mathrm{E}X_{q,\ell,n}). \end{aligned} \tag{2.34}$$

The proof of the main result is now divided into the control of the bounded terms, corresponding to the index $q = 1$, and the control of the nonbounded terms that correspond to the indices $q = 2, 3$. The next result takes care of the approximation between the joint distribution of the blocks and what one would find if these blocks were independent.

Lemma 2.43 *Let X_n, $n \in \mathbb{N}$, be strictly stationary and associated random variables. Then, for every $\lambda > 0$,*

$$\left| \mathrm{E}e^{\lambda Z_{1,n,od}} - \prod_{j=1}^{r_n} \mathrm{E}e^{\lambda Y_{1,2j-1,n}} \right| \leq \frac{\lambda^2 n}{2} e^{\lambda n c_n} \sum_{j=p_n+2}^{(2r_n-1)p_n} \mathrm{Cov}(X_1, X_j), \tag{2.35}$$

and analogously for the term corresponding to $Z_{1,n,ev}$.

Proof Taking into account (2.34) and the fact that the variables defined in (2.32) are associated, we have, applying directly Theorem 2.42,

$$\left| \mathrm{E}e^{\lambda Z_{1,n,od}} - \prod_{j=1}^{r_n} \mathrm{E}e^{\lambda Y_{1,2j-1,n}} \right| \leq \lambda^2 r_n p_n e^{2\lambda r_n p_n c_n} \sum_{1\leq j<j'\leq r_n} \mathrm{Cov}(Y_{1,2j-1,n}, Y_{1,2j'-1,n}). \tag{2.36}$$

As $2r_n p_n \leq n$, we are left with the sum of the covariances to deal with. Using the stationarity of the variables, it follows that

$$\sum_{1\leq j<j'\leq r_n} \mathrm{Cov}(Y_{1,2j-1,n}, Y_{1,2j'-1,n}) = \sum_{j=1}^{r_n-1} (r_n - j)\,\mathrm{Cov}(Y_{1,1,n}, Y_{1,2j-1,n}).$$

A further invocation of the stationarity implies that

$$\begin{aligned} \mathrm{Cov}&(Y_{1,1,n}, Y_{1,2j-1,n}) \\ &= \sum_{\ell=0}^{p_n-1} (p_n - \ell)\,\mathrm{Cov}(X_{1,1,n}, X_{1,2jp_n+\ell+1,n}) \\ &\quad + \sum_{\ell=1}^{p_n-1} (p_n - \ell)\,\mathrm{Cov}(X_{1,\ell+1,n}, X_{1,2jp_n+1,n}) \\ &\leq p_n \sum_{\ell=(2j-1)p_n+2}^{(2j+1)p_n} \mathrm{Cov}(X_{1,1,n}, X_{1,\ell,n}). \end{aligned} \tag{2.37}$$

We now analyse the covariances using the Hoeffding formula (1.2):

$$\begin{aligned} \mathrm{Cov}(X_{1,i,n}, X_{1,j,n}) = \int_{\mathbb{R}^2} &\mathbf{P}(X_{1,i,n} > u, X_{1,j,n} > v) \\ &- \mathbf{P}(X_{1,i,n} > u)\mathbf{P}(X_{1,j,n} > v)\,du\,dv. \end{aligned} \tag{2.38}$$

If we take into account the truncation made in (2.32), it follows that the integrand function vanishes outside the square $[-c_n, c_n]^2$. Moreover, for $u, v \in [-c_n, c_n]$, we may replace, in the integrand, the variables $X_{1,i,n}$ and $X_{1,j,n}$ by X_i and X_j, respectively, so that

$$\begin{aligned} \mathrm{Cov}&(X_{1,i,n}, X_{1,j,n}) \\ &= \int_{[-c_n,c_n]^2} \mathbf{P}(X_i > u, X_j > v) - \mathbf{P}(X_i > u)\mathbf{P}(X_j > v)\,du\,dv \\ &\leq \int_{\mathbb{R}^2} \mathbf{P}(X_i > u, X_j > v) - \mathbf{P}(X_i > u)\mathbf{P}(X_j > v)\,du\,dv = \mathrm{Cov}(X_i, X_j), \end{aligned}$$

due to the nonnegativity of the latter integrand function, as follows from the association of the original variables. Inserting this into (2.36) and (2.37), the lemma follows. □

The next step is to find some convenient control on the variables that couple the blocks $Y_{1,2j-1,n}$. We first prove a small extension of the moment inequalities studied in Sect. 2.3, that is better suited for our present purposes.

Lemma 2.44 *Let $c > 0$ and $S_{1,n} = \sum_{i=1}^{n}(X_{1,i,n} - \mathrm{E}X_{1,i,n})$. Assume that the random variables X_n, $n \in \mathbb{N}$, are strictly stationary, associated and $u(0) < \infty$. Then $\mathrm{E}S_{1,n}^2 \leq 2nc_n^*$, where $c_n^* \geq c_n^2 + u(0)$.*

Proof Using the stationarity, we easily get that

$$\begin{aligned}\mathrm{E}S_{1,n}^2 &= n\,\mathrm{Var}(X_{1,1,n}) + 2\sum_{j=1}^{n-1}(n-j)\,\mathrm{Cov}(X_{1,1,n}, X_{1,j+1,n}) \\ &\leq 2nc_n^2 + 2nu(0) \leq 2nc_n^*,\end{aligned}$$

since $\mathrm{Cov}(X_{1,1,n}, X_{1,j+1,n}) \leq \mathrm{Cov}(X_1, X_{j+1})$ due to the association of the random variables, as mentioned in the proof of the previous lemma. □

Remark 2.45 Notice that we can assume that $c_n^* = 2c_n^2$, at least as $c_n \longrightarrow +\infty$, as is the case for our framework.

The previous inequality will help improving the control of the independent-like terms used in Ioannides and Roussas [48] and Oliveira [75]. An improvement based on the Hölder inequality appears in Sung [98], but the approach by Xing, Yan and Liu [106] that goes along the arguments to be used below produces a better upper bound.

Lemma 2.46 *Let $X_n, n \in \mathbb{N}$, be strictly stationary and associated random variables such that $u(0) < \infty$. If $0 < \lambda < \frac{1}{2c_n p_n}$, then*

$$\prod_{j=1}^{r_n} \mathrm{E}e^{\lambda Y_{1,2j-1,n}} \leq \exp\bigl(\lambda^2 nc_n^*\bigr),$$

and the same bound holds for $\prod_{j=1}^{r_n} \mathrm{E}e^{\lambda Y_{1,2j,n}}$.

Proof From the definition (2.33) it is obvious that $|Y_{1,2j-1,n}| \leq 2c_n p_n$. Using a Taylor expansion and Lemma 2.44, we get that, for each $j = 1, \ldots, r_n$,

$$\mathrm{E}e^{\lambda Y_{1,2j-1,n}} \leq 1 + \lambda^2 \mathrm{E}Y_{1,2j-1,n}^2 \sum_{k=2}^{\infty} \frac{(2c_n\lambda p_n)^{k-2}}{k!} \leq \exp\bigl(2\lambda^2 p_n c_n^*\bigr),$$

using the inequality $1 + x \leq e^x$ for ≥ 0 and taking into account the assumption on λ. Finally, recall that $2r_n p_n \leq n$. □

We may now prove an exponential inequality for the sum of odd indexed or even indexed terms.

Lemma 2.47 *Let X_n, $n \in \mathbb{N}$, be strictly stationary and associated random variables. Assume that $p_n > c_n > u(0)$ and*

$$\frac{n}{c_n^4}\exp\left(\frac{n}{4c_n}\right)u(p_n) \leq C_0 < \infty. \tag{2.39}$$

Then, for every $\varepsilon \in (0, \frac{c_n}{p_n})$,

$$\mathbf{P}\left(\frac{1}{n}|Z_{1,n,od}| > \varepsilon\right) \leq (1+32C_0)\exp\left(-\frac{n\varepsilon^2}{8c_n^2}\right), \tag{2.40}$$

and analogously for $Z_{1,n,ev}$.

Proof Applying Markov's inequality and using Lemma 2.43, we find that, for every $\lambda > 0$ small enough,

$$\mathbf{P}\left(\frac{1}{n}|Z_{1,n,od}| > \varepsilon\right) \leq \frac{\lambda^2 n}{2}\exp(\lambda n c_n - \lambda n\varepsilon)\sum_{j=p_n+2}^{(2r_n-1)p_n}\mathrm{Cov}(X_1, X_l) + \exp\left(\lambda^2 n c_n^* - \lambda n\varepsilon\right). \tag{2.41}$$

To optimize the exponent in the last term of the upper bound in (2.41), we choose $\lambda = \frac{\varepsilon}{2c_n^*}$, so that $\lambda^2 n c_n^* - \lambda n\varepsilon = -\frac{n\varepsilon^2}{4c_n^*}$. Notice that as $\varepsilon < \frac{c_n}{p_n}$, it follows that the requirement on λ of Lemma 2.46 is fulfilled. Replacing now λ in the first term of the upper bound and taking into account (2.39), we get that

$$\begin{aligned}
&\mathbf{P}\left(\frac{1}{n}|Z_{1,n,od}| > \varepsilon\right)\\
&\leq \frac{\varepsilon^2 n}{8(c_n^*)^2}\exp\left(\frac{nc_n}{2c_n^*} - \frac{n\varepsilon^2}{2c_n^*}\right)\sum_{j=p_n+2}^{(2r_n-1)p_n}\mathrm{Cov}(X_1, X_l) + \exp\left(-\frac{n\varepsilon^2}{4c_n^*}\right)\\
&\leq 32C_0\exp\left(-\frac{n\varepsilon^2}{2c_n^*}\right) + \exp\left(-\frac{n\varepsilon^2}{4c_n^*}\right)\\
&\leq (1+32C_0)\exp\left(-\frac{n\varepsilon^2}{4c_n^*}\right)\\
&= (1+32C_0)\exp\left(-\frac{n\varepsilon^2}{8c_n^2}\right).
\end{aligned}$$

□

Remark 2.48 Note that assumption (2.39), which involves the covariance structure on the previous lemma, is much stronger than (2.16), used to control second-order moments of maxima.

To complete the treatment of the bounded terms, it remains to control the sum corresponding to the indices after $2r_n p_n$, that is, $R_{1,n}$.

Lemma 2.49 *Let X_n, $n \in \mathbb{N}$, be strictly stationary associated random variables and assume that*

$$\frac{n}{c_n p_n} \longrightarrow +\infty. \tag{2.42}$$

Then, for n large enough and every $\varepsilon > 0$, we have $\mathbf{P}(|R_{1,n}| > n\varepsilon) = 0$.

Proof Recall the definition of $R_{1,n} = \sum_{\ell=2r_n p_n+1}^{n} (X_{1,\ell,n} - \mathrm{E}X_{1,\ell,n})$. Taking into account the construction of r_n and p_n, we get that $|R_{1,n}| \le 2(n - 2r_n p_n)c_n \le 4c_n p_n$. Now $\mathbf{P}(|R_{1,n}| > n\varepsilon) \le \mathbf{P}(\frac{4}{\varepsilon} > \frac{n}{c_n p_n})$ and, by (2.42), this is equal zero for n large enough. □

The variables $X_{2,i,n}$ and $X_{3,i,n}$ are associated but not bounded, even for fixed n. This means that Theorem 2.42 may not be applied to the sum of such terms. But, we may note that these variables depend only on the tails of the distributions of the original variables. So, by controlling the decrease rate of these tails we may prove an exponential inequality for sums of $X_{2,i,n}$ or $X_{3,i,n}$. A first upper bound using such an approach was obtained in Oliveira [75], where the association was not explored, and later improved by Xing, Yang and Liu [106], using explicitly the association of the random variables via the maximal inequality proved in Theorem 2.27.

Lemma 2.50 *Let X_n, $n \in \mathbb{N}$, be associated random variables such that (2.18) holds and there exist $M > 0$ and $\delta > 0$ such that*

$$\sup_{|t|\le\delta} \mathrm{E}e^{tX_1} \le M < +\infty. \tag{2.43}$$

Then, for $t \in (0, \delta]$ and $q = 2, 3$,

$$\mathbf{P}\left(\max_{1\le k\le n}\left|\sum_{i=1}^{k}(X_{q,i,n} - \mathrm{E}X_{q,i,n})\right| > n\varepsilon\right) \le \frac{2MCe^{-tc_n}}{nt^2\varepsilon^2} + \frac{(2M)^{1/2}Ce^{-tc_n/2}}{nt\varepsilon^2}, \tag{2.44}$$

where C is the constant introduced in Theorem 2.27.

Proof According to Theorem 2.27, using Markov's inequality, we have

$$\begin{aligned}
&\mathbf{P}\left(\max_{1\le k\le n}\left|\sum_{i=1}^{k}(X_{q,i,n} - \mathrm{E}X_{q,i,n})\right| > n\varepsilon\right) \\
&\quad\le \frac{1}{n^2\varepsilon^2}\mathrm{E}\left(\max_{1\le k\le n}\left|\sum_{i=1}^{k}(X_{q,i,n} - \mathrm{E}X_{q,i,n})\right|^2\right) \\
&\quad\le \frac{C}{n\varepsilon^2}\left(\sup_i \mathrm{E}X_{q,i,n}^2 + \left(\sup_i \mathrm{E}X_{q,i,n}^2\right)^{1/2}\right),
\end{aligned} \tag{2.45}$$

so we need to find upper bounds for $\mathrm{E}X_{q,i,n}^2$. Write $\overline{F}_i(x) = \mathbf{P}(X_i > x)$. Using again Markov's inequality, we get that, for $t \in (0, \delta)$, $\overline{F}_i(x) \le e^{-tx}\mathrm{E}e^{tX_i} \le Me^{-tx}$.

Writing the mathematical expectation as a Stieltjes integral and integrating by parts, we find

$$\mathrm{E}X_{2,i,n}^2 = -\int_{(c_n,+\infty)} (x-c_n)^2 \overline{F}_i(dx) = \int_{c_n}^{+\infty} 2(x-c_n)\overline{F}_i(x)\,dx \le 2M\frac{e^{-tc_n}}{t^2}.$$

Inserting this into (2.45), we get the result. □

Remark 2.51 Notice that in (2.44), if $c_n \longrightarrow +\infty$, the second term in (2.44) is the dominating one.

We can now collect all the previous upper bounds into the following result, stating an exponential inequality regardless of the boundedness of the variables.

Theorem 2.52 *Let X_n, $n \in \mathbb{N}$, be strictly stationary and associated random variables such that* (2.18), (2.39), (2.42) *and* (2.43) *hold. Then, for n large enough, $\varepsilon \in (0, \frac{c_n}{p_n})$ and $t \in (0, \delta]$,*

$$\mathbf{P}\left(\left|\sum_{i=1}^n X_i - \mathrm{E}X_i\right| > n\varepsilon\right) \le 2(1+32C_0)\exp\left(-\frac{n\varepsilon^2}{648c_n^2}\right) + \frac{36MCe^{tc_n}}{nt^2\varepsilon^2} + \frac{18(2M)^{1/2}Ce^{tc_n/2}}{nt\varepsilon^2}. \quad (2.46)$$

Proof Just write

$$\begin{aligned}
&\mathbf{P}\left(\left|\sum_{i=1}^n X_i - \mathrm{E}X_i\right| > n\varepsilon\right)\\
&\le \sum_{q=1}^3 \mathbf{P}\left(\left|\sum_{i=1}^n X_{q,i,n} - \mathrm{E}X_{q,i,n}\right| > \frac{n\varepsilon}{3}\right)\\
&\le \mathbf{P}\left(|Z_{1,n,od}| > \frac{n\varepsilon}{9}\right) + \mathbf{P}\left(|Z_{1,n,ev}| > \frac{n\varepsilon}{9}\right) + \mathbf{P}\left(|R_{1,n}| > \frac{n\varepsilon}{9}\right)\\
&\quad + \mathbf{P}\left(\left|\sum_{i=1}^n X_{2,i,n} - \mathrm{E}X_{2,i,n}\right| > \frac{n\varepsilon}{3}\right) + \mathbf{P}\left(\left|\sum_{i=1}^n X_{3,i,n} - \mathrm{E}X_{3,i,n}\right| > \frac{n\varepsilon}{3}\right)
\end{aligned}$$

and use (2.40) and (2.44) to conclude. □

These inequalities are not yet in an adequate form for characterizing convergence rates. This will be done in Sect. 3.2, essentially allowing ε to depend on n and identifying a convenient decrease rate such that the derived upper bound still defines a convergent series.

Springer

http://www.springer.com/978-3-642-25531-1

Asymptotics for Associated Random Variables
Oliveira, P.E.
2012, X, 194 p., Hardcover
ISBN: 978-3-642-25531-1