Probability Theory and Stochastic Modelling

Volume 80

Editors-in-chief
Søren Asmussen, Aarhus, Denmark
Peter W. Glynn, Stanford, CA, USA
Yves Le Jan, Orsay, France

Advisory Board
Martin Hairer, Coventry, UK
Peter Jagers, Gothenburg, Sweden
Ioannis Karatzas, New York, NY, USA
Frank P. Kelly, Cambridge, UK
Andreas Kyprianou, Bath, UK
Bernt Øksendal, Oslo, Norway
George Papanicolaou, Stanford, CA, USA
Etienne Pardoux, Marseille, France
Edwin Perkins, Vancouver, BC, Canada
Halil Mete Soner, Zürich, Switzerland
The **Probability Theory and Stochastic Modelling** series is a merger and continuation of Springer’s two well established series Stochastic Modelling and Applied Probability and Probability and Its Applications series. It publishes research monographs that make a significant contribution to probability theory or an applications domain in which advanced probability methods are fundamental. Books in this series are expected to follow rigorous mathematical standards, while also displaying the expository quality necessary to make them useful and accessible to advanced students as well as researchers. The series covers all aspects of modern probability theory including

- Gaussian processes
- Markov processes
- Random fields, point processes and random sets
- Random matrices
- Statistical mechanics and random media
- Stochastic analysis

as well as applications that include (but are not restricted to):

- Branching processes and other models of population growth
- Communications and processing networks
- Computational methods in probability and stochastic processes, including simulation
- Genetics and other stochastic models in biology and the life sciences
- Information theory, signal processing, and image synthesis
- Mathematical economics and finance
- Statistical methods (e.g. empirical processes, MCMC)
- Statistics for stochastic processes
- Stochastic control
- Stochastic models in operations research and stochastic optimization
- Stochastic models in the physical sciences

More information about this series at http://www.springer.com/series/13205
Asymptotic Theory of Weakly Dependent Random Processes
Preface

These lecture notes are the second version of the book “Théorie asymptotique des processus aléatoires faiblement dépendants”, written in French. In the process of translation, some misprints and inaccuracies have been removed, some proofs rewritten in more detail, some recent references have been added and three new sections incorporated. However, the numeration of the initial sections remains unchanged. Below I give comments on the three new sections.

Section 1.7 gives another covariance inequality for strongly mixing sequences. This covariance inequality is an improvement of Inequality (7) of Exercise 8, obtained by the author at the end of 1991. Section 4.4 deals with the central limit theorem for triangular arrays. This section is based on a paper of the author on the Lindeberg method. In Sect. 5.6 a meaningful coupling lemma of Peligrad for unbounded random variables is stated and proved.

The mathematics of the initial version was completed in January 1999. Since that time, there has been a huge amount of new results. In particular, it is now clear that the notions of weak dependence used in this book are too restrictive for some applications, for instance in the case of Markov chains associated to some dynamical systems. Consequently some new notions of weak dependence have been introduced. We refer to Dedecker et al. (2007) for an introduction to these new notions of dependence and their associated coefficients of dependence, as well as some applications of these new techniques. We will not treat this much broader spectrum of dependence in this second edition. Finally, it is a pleasure to thank Marina Reizakis and all the staff at Springer who contributed towards the production of this book.

Versailles, France

September 2016

Emmanuel Rio
Ce cours est une extension d’un cours commun avec Paul Doukhan effectué de 1994 à 1996 dans le cadre du DEA de modélisation stochastique et de statistiques de la faculté d’Orsay et du minicours que j’ai donné dans le cadre du séminaire Paris-Berlin 1997 de statistique mathématique. Il s’adresse avant tout aux étudiants de troisième cycle ainsi qu’aux chercheurs désireux d’approfondir leurs connaissances sur la théorie de la addition des variables aléatoires faiblement dépendantes.

Alors que la théorie de l’addition des variables aléatoires indépendantes est désormais largement avancée, peu d’ouvrages traitent des suites de variables faiblement dépendantes. Pour ne pas nous restreindre aux chaînes de Markov, nous étudions ici les suites de variables faiblement dépendantes dites fortement mélangées au sens de Rosenblatt (1956) ou absolument régulières au sens de Volkonskii et Rozanov (1959). Ces suites englobent de nombreux modèles utilisés en statistique mathématique ou en économétrie, comme le montre Doukhan (1994) dans son ouvrage sur le mélange. Nous avons choisi de nous concentrer sur les inégalités de moments ou de moyennes déviations pour les sommes de variables faiblement dépendantes et sur leurs applications aux théorèmes limites.

Je voudrais remercier ici tous ceux qui m’ont aidé lors de la rédaction de cet ouvrage, en particulier Paul Doukhan et Abdelkader Mokkadem, pour toutes les connaissances qu’ils m’ont transmises ainsi que pour leurs conseils avisés. Une partie de leurs résultats est développée dans ces notes. Je voudrais également remercier Sana Louhichi et Jérôme Dedecker pour les modifications et les nombreuses améliorations qu’il mont suggérées.

Orsay, France
Janvier 2000

Emmanuel Rio
## Contents

1 The Variance of Partial Sums ........................................... 1  
1.1 Introduction .................................................. 1  
1.2 Stationary Processes .......................................... 1  
1.3 A Covariance Inequality Under Strong Mixing ................ 3  
1.4 Variance of Partial Sums of a Strongly Mixing Sequence .... 8  
1.5 Applications to Density Estimation ........................... 15  
1.6 A Covariance Inequality Under Absolute Regularity ........... 22  
1.7 Other Covariance Inequalities Under Strong Mixing * .......... 26  
Exercises ....................................................... 30  

2 Algebraic Moments, Elementary Exponential Inequalities ....... 33  
2.1 Introduction .................................................. 33  
2.2 An Upper Bound for the Fourth Moment of Sums .............. 33  
2.3 Even Algebraic Moments .................................... 37  
2.4 Exponential Inequalities .................................... 40  
2.5 New Moment Inequalities ................................... 45  
Exercises ....................................................... 48  

3 Maximal Inequalities and Strong Laws ............................ 51  
3.1 Introduction .................................................. 51  
3.2 An Extension of the Maximal Inequality of Kolmogorov ...... 51  
3.3 Rates of Convergence in the Strong Law of Large Numbers ... 54  
Exercises ....................................................... 61  

4 Central Limit Theorems .............................................. 65  
4.1 Introduction .................................................. 65  
4.2 A Central Limit Theorem for Strongly Mixing and Stationary  
            Sequences ........................................... 65  
4.3 A Functional Central Limit Theorem for the Partial  
            Sum Process ........................................... 68
4.4 A Central Limit Theorem for Strongly Mixing Triangular Arrays* .......................................................... 70
Exercises ........................................................................ 84

5 Mixing and Coupling .................................................. 89
5.1 Introduction .................................................................. 89
5.2 A Coupling Lemma for Real-Valued Random Variables .... 89
5.3 A Coupling Lemma for $\beta$-Mixing Random Variables .... 91
5.4 Comparison of $\alpha$-Mixing and $\beta$-Mixing Coefficients for Finite $\sigma$-Fields .................................................. 95
5.5 Maximal Coupling and Absolutely Regular Sequences ...... 97
5.6 An Extension of Lemma 5.2 to Unbounded Random Variables * ................................................................. 97
Exercises ........................................................................ 99

6 Fuk–Nagaev Inequalities, Applications .......................... 101
6.1 Introduction .................................................................. 101
6.2 Exponential Inequalities for Partial Sums ..................... 101
6.3 Fuk–Nagaev Inequalities for Partial Sums ..................... 103
6.4 Application to Moment Inequalities ........................... 106
6.5 Application to the Bounded Law of the Iterated Logarithm . 109
Exercises ........................................................................ 110

7 Empirical Distribution Functions ................................. 113
7.1 Introduction .................................................................. 113
7.2 An Elementary Estimate ............................................. 114
7.3 Functional Central Limit Theorems .............................. 116
7.4 A Functional Central Limit Theorem for the Empirical Distribution Function .............................................. 118
7.5 Multivariate Distribution Functions ............................ 121

8 Empirical Processes Indexed by Classes of Functions ....... 125
8.1 Introduction .................................................................. 125
8.2 Classes of Regular Functions ...................................... 126
8.3 Maximal Coupling and Entropy with Bracketing .......... 130
Exercises ........................................................................ 145

9 Irreducible Markov Chains ........................................... 149
9.1 Introduction .................................................................. 149
9.2 Irreducible Markov Chains ......................................... 150
9.3 The Renewal Process of an Irreducible Chain ............... 152
9.4 Mixing Properties of Positively Recurrent Markov Chains:
An Example ..................................................................... 155
9.5 Small Sets, Absolute Regularity and Strong Mixing ....... 159
Contents

9.6 Rates of Strong Mixing and Rates of Ergodicity of Irreducible Chains .............................................. 164
9.7 On the Optimality of Theorem 4.2 ........................................... 167
Exercises ........................................................................... 169

Annex A: Young Duality and Orlicz Spaces ......................... 171
Annex B: Exponential Inequalities for Sums of Independent Random Variables ............................................. 175
Annex C: Upper Bounds for the Weighted Moments ............. 183
Annex D: Two Versions of a Lemma of Pisier ...................... 187
Annex E: Classical Results on Measurability ....................... 191
Annex F: The Conditional Quantile Transformation ............... 193
Annex G: Technical Tools .................................................. 195
References ....................................................................... 197
Index ............................................................................. 203
Notation

\( a \wedge b \) \quad \min(a, b)
\( a \vee b \) \quad \max(a, b)
\( \xi, x \) \quad Euclidean scalar product of \( \xi \) and \( x \)
\( x^+, x_+ \) \quad For a real \( x \), the number \( \max(0, x) \)
\( x^-, x_- \) \quad For a real \( x \), the number \( \max(0, -x) \)
\( [x] \) \quad For a real \( x \), the integer part of \( x \)
\( f(x - 0) \) \quad For a function \( f \), the left limit of \( f \) at point \( x \)
\( f(x + 0) \) \quad For a function \( f \), the right limit of \( f \) at point \( x \)
\( f^{-1} \) \quad The generalized inverse function of \( f \)
\( \alpha^{-1} \) \quad The function defined by \( \alpha^{-1}(u) = \sum_{i \in \mathbb{N}} \mathbb{I}_{u < \alpha_i} \)
\( Q_X \) \quad For a random variable \( X \), the inverse of \( t \to \mathbb{P}(|X| > t) \)
\( \mathbb{E}(X) \) \quad For an integrable random variable \( X \), the expectation of \( X \)
\( B^c \) \quad The complement of \( B \)
\( \mathbb{I}_B \) \quad The indicator function of \( B \)
\( \mathbb{E}(X|A) \) \quad Conditional expectation of \( X \) conditionally to the \( \sigma \)-field \( A \)
\( \mathbb{P}(B|A) \) \quad Conditional expectation of \( \mathbb{I}_B \) conditionally to \( A \)
\( \text{Var} X \) \quad The variance of \( X \)
\( \text{Cov}(X, Y) \) \quad Covariance between \( X \) and \( Y \)
\( \|\mu\| \) \quad Measure of total variation associated to the measure \( \mu \)
\( \|\mu\| \) \quad Total variation of the measure \( \mu \)
\( \mu \otimes \nu \) \quad Tensor product of the measures \( \mu \) and \( \nu \)
\( \mathcal{A} \otimes \mathcal{B} \) \quad Tensor product of the \( \sigma \)-fields \( \mathcal{A} \) and \( \mathcal{B} \)
\( \mathcal{A} \vee \mathcal{B} \) \quad The \( \sigma \)-field generated by \( \mathcal{A} \cup \mathcal{B} \)
\( \sigma(X) \) \quad \( \sigma \)-field generated by \( X \)
\( L^r \) \quad For \( r \geq 1 \), the space of random variables \( X \) such that \( \mathbb{E}(|X|^r) < \infty \)
\( \|X\|_r \) \quad For \( r \geq 1 \), the usual norm on \( L^r \)
\( L^\infty \) \quad The space of almost surely bounded random variables
\( \|X\|_\infty \) \quad The usual norm on \( L^\infty \)
$L'(P)$  For a $P$ law on $\mathcal{X}$, the space of functions $f$ such that $\int_{\mathcal{X}} |f|' dP < \infty$

$L^\phi$  The Orlicz space associated to a convex function $\phi$

$\| \cdot \|_\phi$  The usual norm on $L^\phi$

$\phi^*$  For a function $\phi$, the Young dual of $\phi$
Introduction

This book is a translation and a second version of “Théorie asymptotique des processus aléatoires faiblement dépendants”, published in 2000 by Springer. It is devoted to inequalities and limit theorems for weakly dependent sequences. Our aim is to give efficient technical tools to Mathematicians or Statisticians who are interested in weak dependence. We will essentially consider classical notions of weak dependence, called mixing conditions. Sometimes we will give more general results. Nevertheless, the strong mixing coefficients of Rosenblatt (1956) will be used in most of the results.

Here the strong mixing coefficient between two $\sigma$-fields $\mathcal{A}$ and $\mathcal{B}$ is defined by

$$\alpha(\mathcal{A}, \mathcal{B}) = 2 \sup \{ \mathbb{P}(A \cap B) - \mathbb{P}(A) \mathbb{P}(B) : (A, B) \in \mathcal{A} \times \mathcal{B} \}.$$  

This coefficient is equal to the strong mixing coefficient of Rosenblatt (1956), up to the multiplicative factor 2. This coefficient is a measure of the dependence between $\mathcal{A}$ and $\mathcal{B}$. For example, $\alpha(\mathcal{A}, \mathcal{B}) = 0$ if and only if $\mathcal{A}$ and $\mathcal{B}$ are independent. For a sequence $(X_i)_{i \in \mathbb{Z}}$ of random variables in some Polish space $\mathcal{X}$, let $\mathcal{F}_k = \sigma(X_i : i \leq k)$ and $\mathcal{G}_l = \sigma(X_i : i \geq l)$. The strong mixing coefficients $(\alpha_n)_{n \geq 0}$ of the sequence $(X_i)_{i \in \mathbb{Z}}$ are defined by

$$\alpha_0 = 1/2 \text{ and } \alpha_n = \sup_{k \in \mathbb{Z}} \alpha(\mathcal{F}_k, \mathcal{G}_{k+n}) \text{ for any } n > 0.$$  

(1)

The sequence $(X_i)_{i \in \mathbb{Z}}$ is said to be strongly mixing in the sense of Rosenblatt (1956) if $\lim_{n \to \infty} \alpha_n = 0$. In the stationary case, this means that the $\sigma$-field $\mathcal{G}_n$ of the future after time $n$ is asymptotically independent of $\mathcal{F}_0$, which is the $\sigma$-field of the past before time 0. We refer to Bradley (1986, 2005) for other coefficients of weak dependence and the relations between them.

In these notes, we will mainly establish results for strongly mixing sequences or for absolutely regular sequences in the sense of Volkonskii and Rozanov (1959). Indeed, these notions of weak dependence are less restrictive than the notions of
\( \rho \)-mixing and uniform mixing in the sense of Ibragimov (1962). For example, in the case of autoregressive models with values in \( \mathbb{R}^d \) defined by the recursive equation
\[
X_{n+1} = f(X_n) + \varepsilon_{n+1},
\]
(2)
for some sequence of independent and identically distributed integrable innovations \((\varepsilon_n)_n\) with a positive continuous bounded density, the stationary sequence \((X_i)_{i \in \mathbb{Z}}\) solution of (2) is uniformly mixing in the sense of Ibragimov only if the function \(f\) is uniformly bounded over \( \mathbb{R}^d \). This condition is too restrictive for the applications. By contrast, the stationary solution of (2) is strongly mixing with a geometric rate of strong mixing as soon as there exists \( M > 0, s > 0 \) and \( \rho < 1 \) such that
\[
\mathbb{E}(|f(x) + \varepsilon_0|^s) \leq \rho |x|^s \quad \text{for } x > M \quad \text{and} \quad \sup_{|x| \leq M} \mathbb{E}(|f(x) + \varepsilon_0|^s) < \infty. \quad (3)
\]

We refer to Doukhan and Ghindès (1983) and to Mokkadem (1985) for more about the model (2), and to Doukhan (1994) for other examples of Markov models satisfying mixing conditions. Although the notions of strong mixing or absolute regularity are less restrictive than the notions of \( \rho \)-mixing and uniform mixing, they are adequate for the applications. For example, Viennet (1997) obtains optimal results for linear estimators of the density in the case of absolutely regular sequences.

We now summarize the contents of these lecture notes. Our main tools are covariance inequalities for random variables satisfying mixing conditions and coupling results which are similar to the coupling theorems of Berbee (1979) and Goldstein (1979). Chapters 1–4 are devoted to covariance inequalities, moment inequalities and classical limit theorems. Chapters 5–8 mainly use coupling techniques. The coupling techniques are applied to the law of the iterated logarithm for partial sums in Chap. 6 and then to empirical processes in Chaps. 7 and 8.

In Chap. 1, we give covariance inequalities for random variables satisfying a strong mixing condition or an absolute regularity condition. Let us recall Ibragimov’s (1962) covariance inequality for bounded random variables: if \( X \) and \( Y \) are uniformly bounded real-valued random variables, then
\[
|\text{Cov}(X, Y)| \leq 2\alpha(\sigma(X), \sigma(Y)) \|X\|_\infty \|Y\|_\infty, \quad (4)
\]
where \( \sigma(X) \) and \( \sigma(Y) \) denote the \( \sigma \)-fields generated by \( X \) and \( Y \), respectively. We give extensions of (4) to unbounded random variables. We then apply these covariance inequalities to get estimates of the variance of partial sums. In the dependent case, the variance of the sum may be much larger than the sum of variances. We refer to Bradley (1997) for lower bounds for the variance of partial sums in the strong mixing case. Nevertheless, adequate applications of the variance estimates still provide efficient results. For example, we give in Sects. 1.5 and 1.6 some impressive applications of mixing conditions to density estimation. In Sect. 1.7*
we give other covariance inequalities (the * indicates that the section is new to this edition).

Chapter 2 is devoted to the applications of covariance inequalities to moment inequalities for partial sums. In Sects. 2.2 and 2.3, we apply the covariance inequalities of Chap. 1 to algebraic moments of sums. Our methods are similar to the methods proposed by Doukhan and Portal (1983, 1987). They lead to Rosenthal type inequalities. In Sects. 2.4 and 2.5, we prove Marcinkiewicz type moment inequalities for the absolute moments of order $p > 2$, and we give a way to derive exponential inequalities from these results. In Chap. 3 we give extensions of the maximal inequalities of Doob and Kolmogorov to dependent sequences. These maximal inequalities are then used to obtain Baum–Katz type laws of large numbers, and consequently rates of convergence in the strong law of large numbers. We also derive moment inequalities of order $p$ for $p \in ]1,2[$ from these inequalities.

Chapter 4 is devoted to the classical central limit theorem for partial sums of random variables. We start by considering strictly stationary sequences. In this context, we apply projective criteria derived from Gordin’s martingale approximation theorem (1969) to get the central limit theorem for partial sums of a strongly mixing sequence. We then give a uniform functional central limit theorem in the sense of Donsker for the normalized partial sum process associated to a stationary and strongly mixing sequence. The proof of tightness is based on the maximal inequalities of Chap. 3. Section 4.4*, at the end of the chapter, is devoted to the central limit theorem for triangular arrays.

In Chap. 5, we give coupling results for weakly dependent sequences, under assumptions of strong mixing or $\beta$-mixing. In particular, we recall and we prove Berbee’s coupling Lemma (1979), which characterizes the $\beta$-mixing coefficient between a $\sigma$-field $\mathcal{A}$ and the $\sigma$-field $\sigma(X)$ generated by some random variable $X$ with values in some Polish space. If $(\Omega, \mathcal{F}, \mathbb{P})$ contains an auxiliary atomless random variable independent of $\mathcal{A} \lor \sigma(X)$, then one can construct a random variable $X^*$ with the same law as $X$, independent of $\mathcal{A}$ and such that

$$\mathbb{P}(X = X^*) = 1 - \beta(\mathcal{A}, \sigma(X)). \quad (5)$$

We give a constructive proof of (5) for random variables with values in $[0,1]$. This proof is more technical than the usual proof. Nevertheless, the constructive proof is more informative than the usual proof. In particular, using a comparison theorem between $\alpha$-mixing coefficients and $\beta$-mixing coefficients for purely atomic $\sigma$-fields due to Bradley (1983), one can obtain (see Exercise 1) the following upper bound for the so-constructed random variables:

$$\mathbb{E}(|X - X^*|) \leq 4\alpha(\mathcal{A}, \sigma(X)). \quad (6)$$

In Sect. 5.2, we give a direct proof of (6) with an improved constant. Our method of proof is based on the conditional quantile transformation. Section 5.6* contains an extension of (6) to unbounded random variables.
Chapters 6–8 are devoted to the applications of these coupling results. In Chap. 6, we prove that Inequality (6) yields efficient deviation inequalities for partial sums of real-valued random variables. In particular, we generalize the Fuk–Nagaev deviation inequalities (1971) to partial sums of strongly mixing sequences of real-valued random variables. For example, for sums $S_k = X_1 + \cdots + X_k$ of real-valued and centered random variables $X_i$ satisfying $||X_i||_{\infty} \leq 1$, we prove that, for any $\lambda > 0$ and any $r \geq 1$,

$$
P \left( \sup_{k \in [1,n]} |S_k| \geq 4\lambda \right) \leq 4 \left( 1 + \frac{\lambda^2}{r s_n^2} \right)^{-r/2} + \frac{n\alpha(|\lambda|/r)}{\lambda},
$$

(7)

where

$$
s_n^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} |\text{Cov}(X_i, X_j)|.
$$

This inequality is an extension of the Fuk–Nagaev inequality to weakly dependent sequences. Theorem 6.2 provides an extension in the general case of unbounded random variables. Choosing $r = 2 \log \log n$, we then apply (7) to the bounded law of the iterated logarithm. In Chaps. 7 and 8, we apply (5)–(7) to empirical processes associated to dependent observations. We refer the reader to Dudley (1984) and to Pollard (1990) for more about the theory of functional limit theorems for empirical processes. In Chap. 7, we give uniform functional central limit theorems for the normalized and centered empirical distribution function associated to real-valued random variables or to random variables with values in $\mathbb{R}^d$. We prove that the uniform functional central limit theorem for the normalized and centered empirical distribution function holds true under the strong mixing condition $\alpha_n = O(n^{-1-\epsilon})$ for any $d > 1$. The strong mixing condition does not depend on the dimension, in contrast to the previous results. The proof is based on Inequality (7). This inequality does not provide uniform functional central limit theorems for empirical processes indexed by large classes of sets. For this reason, we give a more general result in Chap. 8, which extends Dudley’s theorem (1978) for empirical processes indexed by classes of sets to $\beta$-mixing sequences. The proof of this result is based on the maximal coupling theorem of Goldstein (1979).

Chapter 9, which concludes these lecture notes, is devoted to the mixing properties of irreducible Markov chains and the links between ergodicity, return times, absolute regularity and strong mixing. We also prove the optimality of some of the results of the previous chapters on various examples of Markov chains. The Annexes are devoted to convex analysis, exponential inequalities for sums of independent random variables, tools for empirical processes, upper bounds for the weighted moments introduced in Chaps. 1 and 2, measurability questions and quantile transformations.