

UNIVERSITY OF CYPRUS
LIBRARY
UNIVERSITY OF CYPRUS
LIBRARY

ASYMPTOTIC RESULTS FOR NON I.I.D.
MULTIDimensionally INDEXED
RANDOM VARIABLES

**ASYMPTOTIC RESULTS FOR NON I.I.D.
MULTIDimensionally INDEXED
RANDOM VARIABLES**

A thesis submitted in fulfillment of the requirements for the
Degree of Doctor of Philosophy
by

Petroula M. Mavrikiou
Petroula M. Mavrikiou

December 1998

Department of Mathematics and Statistics
University of Cyprus

Abstract

**ASYMPTOTIC RESULTS FOR NON I.I.D.
MULTIDimensionALLY INDEXED
RANDOM VARIABLES**

**A thesis submitted for the degree of
Doctor of Philosophy**

by

Petroula M. Mavrikiou

December 1998

**Department of Mathematics and Statistics
University of Cyprus**

Abstract

A fast expanding field in Probability and Statistics is the field of multidimensionally indexed random variables. In this thesis we introduce asymptotic results for this type of random variables which are not necessarily independent and identically distributed. More specifically a new kind of dependence is introduced, the ρ -radius dependence, which is an extension of the notion of the m -dependence. See for example Berk (1973) and Shergin (1983). For multidimensionally indexed ρ -radius dependent random variables, classical asymptotic results are established.

First, well known asymptotic results related to multidimensionally indexed random variables are stated without proofs. Then, a general technique is given which is subsequently used for the proofs of most of the asymptotic results. The first classical result presented, is the proof of the central limit theorem. Next, the Berry-Esséen theorem for multidimensionally indexed ρ -radius dependent random variables is given. In addition, the strong law of large numbers for multidimensionally indexed ρ -radius dependent random variables using classical techniques is proved.

All the above results are proved for the case of two-dimensionally indexed random variables. The extension to higher dimensions can be easily done even though the notation might become quite complicated.

Finally, various probability inequalities for non identically distributed random variables are established. These inequalities can easily be extended to multidimensionally indexed random variables.

Περίληψη

Ο τομέας των τυχαίων μεταβλητών με πολυδιάστατους δείκτες είναι ένας πολύ γρήγορα αναπτυσσόμενος τομέας στις Πιθανότητες και στη Στατιστική. Η διατριβή αυτή πραγματεύεται ασυμπτωτικά αποτελέσματα για τις πιο πάνω τυχαίες μεταβλητές χωρίς κατ' ανάγκη, οι τυχαίες μεταβλητές να είναι ανεξάρτητες και ισόνομες. Για παράδειγμα δες Berk (1973) και Shergin (1983). Ειδικότερα, παρουσιάζουμε ένα νέο είδος εξάρτησης, την ακτινική εξάρτηση, η οποία είναι μία επέκταση της m -εξάρτησης. Για ακτινικά εξαρτημένες τυχαίες μεταβλητές με πολυδιάστατους δείκτες δίνονται κλασικά ασυμπτωτικά αποτελέσματα.

Αρχικά διατυπώνονται γνωστά ασυμπτωτικά αποτελέσματα σχετικά με τυχαίες μεταβλητές με πολυδιάστατους δείκτες, χωρίς όμως να δίνονται αποδείξεις ή λεπτομέρειες. Στη συνέχεια, δίνεται η γενική τεχνική η οποία χρησιμοποιείται στην απόδειξη των πλείστων εκ των αποτελεσμάτων που παρουσιάζονται. Το πρώτο κλασικό αποτέλεσμα που πραγματεύεται η διατριβή αυτή είναι η απόδειξη του Κεντρικού Οριακού Θεωρήματος. Ακολούθως, παρουσιάζεται το θεώρημα Berry-Esséen για ακτινικά εξαρτημένες τυχαίες μεταβλητές με πολυδιάστατους δείκτες. Επίσης, αποδεικνύεται ο Ισχυρός Νόμος των Μεγάλων Αριθμών για ακτινικά εξαρτημένες τυχαίες μεταβλητές με πολυδιάστατους δείκτες χρησιμοποιώντας κλασικές τεχνικές.

Όλα τα πιο πάνω αποτελέσματα αποδεικνύονται για την περίπτωση των τυχαίων μεταβλητών με διδιάστατους δείκτες. Η επέκταση των αποτελεσμάτων σε διάσταση μεγαλύτερου βαθμού μπορεί να επιτευχθεί εύκολα αν και ο συμβολισμός μπορεί να γίνει αρκετά πολύπλοκος.

Τέλος, αποδεικνύονται διάφορες ανισότητες πιθανότητας για μη ισόνομες τυχαίες μεταβλητές. Οι ανισότητες αυτές μπορούν εύκολα να γενικευθούν και στην περίπτωση των τυχαίων μεταβλητών με πολυδιάστατους δείχτες.

Acknowledgments

I would like to thank my supervisor, Professor Tasos C. Christofides for his continuous support and guidance during my studies. Special thanks should also go to the members of the Department of Mathematics and Statistics for their support.

Finally, I am very grateful to the members of my family for their constant encouragement.

Chapter 1	1
Chapter 2	4
Chapter 3	12
Chapter 4	33
Chapter 5	45
Chapter 6	54
Chapter 7	63
Chapter 8	72
Chapter 9	81
Chapter 10	90
Chapter 11	99
Chapter 12	108
Chapter 13	117
Chapter 14	126
Chapter 15	135
Chapter 16	144
Chapter 17	153
Chapter 18	162
Chapter 19	171
Chapter 20	180
Chapter 21	189
Chapter 22	198
Chapter 23	207
Chapter 24	216
Chapter 25	225
Chapter 26	234
Chapter 27	243
Chapter 28	252
Chapter 29	261
Chapter 30	270
Chapter 31	279
Chapter 32	288
Chapter 33	297
Chapter 34	306
Chapter 35	315
Chapter 36	324
Chapter 37	333
Chapter 38	342
Chapter 39	351
Chapter 40	360
Chapter 41	369
Chapter 42	378
Chapter 43	387
Chapter 44	396
Chapter 45	405
Chapter 46	414
Chapter 47	423
Chapter 48	432
Chapter 49	441
Chapter 50	450
Chapter 51	459
Chapter 52	468
Chapter 53	477
Chapter 54	486
Chapter 55	495
Chapter 56	504
Chapter 57	513
Chapter 58	522
Chapter 59	531
Chapter 60	540
Chapter 61	549
Chapter 62	558
Chapter 63	567
Chapter 64	576
Chapter 65	585
Chapter 66	594
Chapter 67	603
Chapter 68	612
Chapter 69	621
Chapter 70	630
Chapter 71	639
Chapter 72	648
Chapter 73	657
Chapter 74	666
Chapter 75	675
Chapter 76	684
Chapter 77	693
Chapter 78	702
Chapter 79	711
Chapter 80	720
Chapter 81	729
Chapter 82	738
Chapter 83	747
Chapter 84	756
Chapter 85	765
Chapter 86	774
Chapter 87	783
Chapter 88	792
Chapter 89	801
Chapter 90	810
Chapter 91	819
Chapter 92	828
Chapter 93	837
Chapter 94	846
Chapter 95	855
Chapter 96	864
Chapter 97	873
Chapter 98	882
Chapter 99	891
Chapter 100	900

4	The strong law of large numbers	30
5.1	Introduction	30
5.2	Main results	31

Contents

6.1	Introduction	61
6.2	Deterministic inequalities and other results	62
	Abstract	i
	Περίληψη	ii
1	Introduction	1
2	Survey on multidimensionally indexed random variables	4
2.1	Introduction and notation	4
2.2	Literature review on multidimensionally indexed random variables	6
3	The central limit theorem	12
3.1	Introduction	12
3.2	Lattice decomposition and related results	16
3.3	A Liapounov type theorem	23
4	The Berry-Esséen theorem	33
4.1	Introduction	33
4.2	Preliminaries	34
4.3	Berry-Esséen theorem	45

5	The strong law of large numbers	50
5.1	Introduction	50
5.2	Main results	51
6	Probability inequalities	61
6.1	Introduction	61
6.2	Deterministic inequalities and other results	62
6.3	Kolmogorov type inequalities	66
	Applications and future work	71
	Curriculum Vitae	73
	References	74

Chapter 1

Introduction

Arrays of independent multidimensionally indexed random variables constitute a relatively new field in Probability and Statistics. Since 1951 when Dunford made his first approach, many researchers followed and many useful results have been obtained. See for example Wichura (1969), Cairoli (1970), Smythe (1973, 1974), Cairoli and Walsh (1975), Gut (1976, 1978), Shorack and Smythe (1976), Etemadi (1981), Klesov (1981), Christofides and Serfling (1990) and Christofides (1992).

For this special topic in probability called "arrays of independent multidimensionally indexed random variables" or otherwise "independent random fields" there are results concerning weak convergence, almost sure behavior, rates of convergence and asymptotic behavior of partial sums in general.

It is reasonable to ask whether there is a need to study multidimensionally indexed random variables. Couldn't we just easily extend all the well known results for one-dimensionally indexed random variables? The answer is no. The reason for not

being able to generalize the classical results is the lack of total ordering. In the r -dimensional case total ordering is out of the question. We therefore have to deal with partial ordering (see Section 1.2). Cairoli (1970) showed by a counterexample that well known classical maximal inequalities are not valid in the r -dimensional case and therefore a different approach has to be considered.

In Chapter 2, relevant asymptotic results on multidimensionally indexed random variables are stated without proofs or any further details.

In Chapter 3, a general technique is given which is subsequently used for the proofs of most of the asymptotic results. The technique is based on that of Bernstein's, the so called "big blocks technique" first introduced in 1927. The main objective of Chapter 3, is the proof of the central limit theorem. The assumptions made for the proof of the central limit theorem are general and not very restrictive. In fact, it is the kind of assumptions one expects for the central limit theorem to hold in the case of non i.i.d. random variables.

In Chapter 4, the Berry-Esséen theorem for multidimensionally indexed ρ -radius dependent random variables is given. This is a theorem which gives information about the central limit theorem by providing an upper bound for the absolute difference between the distribution of a specific statistic and the standard normal distribution. As it was expected the Berry-Esséen rate achieved is not optimal. For a sequence of independent and identically distributed random variables X_1, X_2, \dots, X_n the optimal rate of convergence is $O(n^{-\frac{1}{2}})$. For two dimensionally indexed ρ -radius dependent random variables the rate of convergence depends on the value of the parameters.

As in the case of the central limit theorem, the conditions imposed for the Berry-Esséen theorem are reasonable for cases of non i.i.d. random variables.

In Chapter 5, we prove the strong law of large numbers for multidimensionally indexed ρ -radius dependent random variables using classical techniques, such as truncation.

All the above results are proved for the case of two-dimensionally indexed random variables. The extension to higher dimensions can be easily done even though the notation might become quite complicated.

In Chapter 6, various probability inequalities for non identically distributed random variables are established. It is generally accepted that probability inequalities for partial sums are extremely useful for asymptotic theory and therefore are very frequently used. These inequalities can easily be extended to multidimensionally indexed random variables.

Chapter 2

Survey on multidimensionally indexed random variables

2.1 Introduction and notation

In this chapter key results referring to multidimensionally indexed random variables which are relevant to this thesis are presented.

Even though the field of multidimensionally indexed random variables is a relatively new one, many asymptotic results have been established as mentioned in Chapter 1. These results are primarily related to the strong law of large numbers, strong convergence in general, and the rate of convergence.

Before we proceed with the literature survey, all basic definitions and notation which will be used throughout this manuscript are presented.

Let $\{X_i, i \in \mathbb{N}^r\}$ be an r -dimensionally indexed array of i.i.d. random variables defined on a probability space (Ω, \mathcal{A}, P) where r is a positive integer and \mathbb{N}^r denotes the r -dimensional positive integer lattice. Clearly, for $r = 1$ we have the classical case of a sequence of independent and identically distributed random variables.

From now on partial ordering is assumed. For $\mathbf{i} = (i_1, \dots, i_r)$ and $\mathbf{j} = (j_1, \dots, j_r) \in \mathbb{N}^r$ the notation $\mathbf{i} \leq \mathbf{j}$ means that $i_k \leq j_k, k = 1, \dots, r$. In addition, if $\mathbf{n} = (n_1, \dots, n_r)$, $|\mathbf{n}|$ denotes the product $\prod_{i=1}^r n_i$ while the notation $|\mathbf{n}| \rightarrow \infty$ means that $n_i \rightarrow \infty$ for $i = 1, \dots, r$ or equivalently, $\min_{1 \leq i \leq r} n_i \rightarrow \infty$.

Associated with any probability space (Ω, \mathcal{A}, P) are the L_p spaces of all measurable functions X and therefore of random variables, for which $E|X|^p < \infty, p > 0$. Specifically, $X \in L_1$ means that $E|X| < \infty$.

In addition, if X is a measurable function, then its positive and negative parts are defined by $X^+ = \max(0, X)$ and $X^- = \max(0, -X)$ respectively.

As a consequence we define, $\log^+ X = \max(0, \log X)$ or $\log^+ X = \log(\max(1, X))$.

The symbols "big oh" and "little oh" are to be used. These symbols compare the magnitude of two functions in the following way. The notation

$$u(n) = O(v(n)), n \rightarrow L, \text{ means that } \lim_{n \rightarrow L} \left| \frac{u(n)}{v(n)} \right| \leq C$$

where L is not necessarily finite and C is a constant. In addition, the notation

$$u(n) = o(v(n)), n \rightarrow L, \text{ means that } \lim_{n \rightarrow L} \frac{u(n)}{v(n)} = 0.$$

2.2 Literature review on multidimensionally indexed random variables

In this section we give briefly, some asymptotic results related to multidimensionally indexed random variables without providing any proofs or details.

Smythe (1973) has studied the strong law of large numbers for r -dimensional arrays of random variables. He approached the problem by stating the following question. Given a probability space (Ω, \mathcal{F}, P) and an r -dimensional array of independent random variables with zero mean defined on (Ω, \mathcal{F}, P) , under what conditions will the sample average converge to zero?

Dunford (1951) proved that for the case of independent and identically distributed r -dimensionally indexed random variables the integrability of $|X_{\mathbf{k}}| \log^+ |X_{\mathbf{k}}|$ is sufficient when $r = 2$ and in general, for $r > 2$ the corresponding condition for almost sure convergence is the integrability of the term $|X_{\mathbf{k}}|(\log^+ |X_{\mathbf{k}}|)^{r-1}$.

Smythe (1973) showed that the above conditions are necessary and sufficient when the random variables are independent and identically distributed. The necessity part of the proof of Smythe, is given by classical arguments, i.e., Fubini's theorem and Borel-Cantelli's lemma while for the sufficiency part of the proof martingale properties and theory are used. The theorem is as follows:

Theorem 2.2.1: *Let $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbb{N}^r\}$ be an array of independent and identically distributed zero mean random variables.*

(1) If $E\{|X_k|(\log^+ |X_k|)^{r-1}\} < \infty$, then $\frac{1}{|k|} \sum_{i \leq k} X_i \rightarrow 0$ a.s.

(2) If $E\{|X_k|(\log^+ |X_k|)^{r-1}\} = \infty$, then $\frac{1}{|k|} \sum_{i \leq k} X_i \not\rightarrow 0$ a.s..

Clearly the first condition is the "sufficiency" part while the second is the "necessity" part. In addition, Smythe (1973) considered the non-identically distributed case and stated that in general, the usual sufficient conditions for convergence in the one-dimensional case are sufficient for convergence in the r -dimensional case as well, provided that they are appropriately stated (or generalized). For example, one of these conditions is the adjusted three-series theorem appropriately generalized to the r -dimensionally indexed random variables.

Etemadi (1981) also presented a proof of the strong law of large numbers for a sequence of pairwise independent random variables, which is elementary and at the same time can be extended to " r -dimensional arrays of random vectors" as it is quoted by his paper. The theorem is stated below.

Theorem 2.2.2: Let $\{X_{mn}, (m, n) \in \mathbb{N}^2\}$ be an array of pairwise independent and identically distributed random variables. Let $S_{mn} = \sum_{i=1}^m \sum_{j=1}^n X_{ij}$. Then,

$$E(|X_{11}| \log^+ |X_{11}|) < \infty$$

implies that

$$\lim_{(m,n) \rightarrow \infty} \frac{S_{mn}}{mn} = EX_{11} \quad \text{a.s. .}$$

The above theorem is the strong law of large numbers for 2-dimensionally indexed random variables. The generalization to r -dimensional arrays is immediate by using the sufficient condition $E\{|X|(\log^+ |X|)^{r-1}\} < \infty$. Notice that the conditions of Smythe (1973) and Etemadi (1981) are more or less the same but the proofs of their main results are different.

Theorem 2.2.2 is quite applicable since it only requires the random variables to be pairwise independent and not mutually independent as theorem 2.2.1 requires.

Klesov (1981) gave the strong law of large numbers for independent multidimensionally indexed random variables as a special case of the strong law of large numbers for multidimensionally indexed martingales. His result is stated below.

Theorem 2.2.3: Let $\{X_i, i \leq n \in \mathbf{N}^r\}$ be an r -dimensionally indexed array of independent random variables with zero mean and let $S_n = \sum_{i \leq n} X_i$. If $q \geq 1$ and

$$\sum_i |i|^{-q-1} E|X_i|^{2q} < \infty,$$

then,

$$\frac{1}{|n|} S_n \rightarrow 0 \quad a.s., \quad \text{as } |n| \rightarrow \infty.$$

Using some standard arguments such as symmetrization and desymmetrization, Christofides and Serfling (1990) generalized the classical Kolmogorov strong law of large numbers by the following theorem.

Theorem 2.2.4: Let $S_n = \sum_{k \leq n} X_k$ where the X_k 's are independent random variables with $EX_k = 0$ and $E(X_k^2) < \infty$ for each k . Assume that $\sum_k \frac{1}{|k|^2} EX_k^2 < \infty$. Then,

$$\frac{1}{|n|} S_n \rightarrow 0 \quad \text{a.s., as } |n| \rightarrow \infty.$$

A very interesting class of statistics is the class of U-statistics which was initially introduced and studied by Hoeffding (1948) as a generalization of the notion of the sample mean. The class of U-statistics based on arrays of independent r -dimensionally indexed random variables has been first introduced by Christofides (1987).

Definition 2.2.5: Let $\{X_i, i \leq n \in \mathbb{N}^r\}$ be independent random variables from a distribution F . Let $\theta = \theta(F)$ be a parametric function for which there is an unbiased estimator. The function $h = h(X_{i_1}, \dots, X_{i_m})$ is called kernel and it is assumed that it is symmetric without any loss of generality.

For the estimation of the parametric function θ the following U-statistic is used:

$$U_n = U(X_i, i \leq n) = \binom{|n|}{m}^{-1} \sum_c h(X_{i_1}, \dots, X_{i_m})$$

where \sum_c denotes the summation over the $\binom{|n|}{m}$ combinations of m distinct elements $\{i_1, \dots, i_m\}$ from the set $\{(1, \dots, 1), \dots, (n_1, \dots, n_r)\}$.

Clearly,

$$\theta = \theta(F) = Eh(X_{i_1}, \dots, X_{i_m}) = \int \dots \int h(x_{i_1}, \dots, x_{i_m}) dF(x_{i_1}) \dots dF(x_{i_m})$$

where the set $\{i_1, \dots, i_m\}$ consists of m distinct elements taken from the set

$$\{(1, \dots, 1), \dots, (n_1, \dots, n_r)\}.$$

Apparently, $EU_n = \theta$. What is also clear from the definition is that this class can indeed be considered as a generalization of the sample mean.

Christofides (1992) presented the strong law of large numbers for the class of U-statistics defined above, under the following necessary and sufficient condition

$$E\{|h|(\log^+ |h|)^{r-1}\} < \infty$$

thus generalizing the result of Smythe (1973) and Etemadi (1981).

Theorem 2.2.6: *Let $\{X_i, i \leq n \in \mathbb{N}^r\}$ be a random sample from a distribution F . Let U_n be a U-statistic based on this sample and the kernel h for estimation of the parameter $\theta(F) = E\{h(X_{i_1}, \dots, X_{i_m})\}$. If $E\{|h|(\log^+ |h|)^{r-1}\} < \infty$ then $U_n \rightarrow \theta$ a.s. as $|n| \rightarrow \infty$.*

The proof of the above theorem is based on truncation and martingale theory.

In addition to the strong law of large numbers, Christofides (1997) presented the following theorem which gives the rate of convergence for the strong law of large numbers. The proof is obtained using martingale inequalities.

Theorem 2.2.7: Assume that $E|h(X_{i_1}, \dots, X_{i_m})|^\nu < \infty$ for $\nu \geq 2$ and let U_n be a U -statistic based on a multidimensionally indexed array of random variables and on the kernel $h(X_{i_1}, \dots, X_{i_m})$. Then for $\epsilon > 0$

$$P\left\{ \sup_{k \geq n} |U_k - \theta| > \epsilon \right\} = o(|n|^{1-\nu}), \quad |n| \rightarrow \infty.$$

Furthermore, in Christofides (1998) a Berry-Esséen theorem for U -statistics is presented which gives the rate of convergence for $r = 2$. The theorem is as follows:

Theorem 2.2.8: Let

$$U_n = \binom{|n|}{2}^{-1} \sum_c h(X_{i_1}, X_{i_2})$$

be a U -statistic based on a multidimensionally indexed array of random variables for estimation of the parameter θ . Let σ_n^2 be the variance of U_n and assume that $E|h(X_{i_1}, X_{i_2})|^3 < \infty$. Then, if $\Phi(x)$ denotes the distribution function of the standard normal,

$$\sup_{-\infty < x < \infty} |P\{ \sigma_n^{-1}(U_n - \theta) \leq x \} - \Phi(x)| = O(|n|^{-\frac{1}{2}}), \quad \text{as } |n| \rightarrow \infty.$$

conditional mean or median. Fréchet (1948) also proposed the use of the Lévy metric (distance). A very important kind of dependence considering distance as a measure of dependence, is the m -dependence case. See for example Beck (1973).

Chapter 3

The central limit theorem

3.1 Introduction

One of the most widely studied subjects in probability theory is the concept of dependence. The nature of dependence varies and unless specific assumptions are made about the dependence between random variables, no meaningful statistical model can be assumed.

A measure of dependence indicates how closely two random variables X and Y are, with extremes at mutual independence and complete mutual dependence. Measures of dependence could be conditions based on order or time between random variables, or could be conditions expressed in terms of covariance or correlation coefficient.

Distance is also usually considered as a measure of dependence. For example, Fréchet (1946) proposed as a measure of dependence the use of an average of the distances of the distribution of Y conditional on X from some typical value such as the

conditional mean or median. Fréchet (1948) also proposed the use of the Lévy metric (distance). A very important kind of dependence considering distance as a measure of dependence, is the m -dependence case. See for example Berk (1973).

Definition 3.1.1: A sequence of random variables $\{X_n\}$ is said to be m -dependent if there exists a positive integer r such that any subsequence $\{X_{n_j}, j \geq 1\}$ of $\{X_n\}$ with $\{n_j + m < n_{j+1}\}$ for every $j \geq 1$ and $n_1 \geq r$ is a sequence of independent random variables.

On the other hand, limit theorems in probability theory are of great importance. They are generally regarded as theorems giving convergence of sequences of probability distributions or random variables. Typically, three are the most important classical limit theorems: The classical forms of the strong and weak law of large numbers and the central limit theorem. These theorems generally deal with the asymptotic behavior of the sum of n random variables $S_n = X_1 + \dots + X_n$ taken from a sequence of independent and identically distributed random variables X_1, X_2, \dots .

These classical results have been generalized in various ways during the last fifty years. Such generalizations may refer to the conditions imposed on the random variables such as dependence, extension to higher dimensions than the first, or general "smallness" conditions. See for example the classical books by Loève (1977) and Petrov (1974).

The classical central limit theorem has been extended to the case of dependent random variables by several authors. See for example Bernstein (1927), Heinrich

(1987), and Utev (1989). An important special type of dependent random variables mentioned before and applied to the central limit theorem is the m -dependent case. According to the definition of the m -dependence, two sets of random variables (X_1, X_2, \dots, X_i) and $(X_j, X_{j+1}, \dots, X_n)$ are independent whenever $j - i > m$. On the basis of this case many techniques and variations have appeared concerning the central limit theorem. See for example Shergin (1983).

The first who studied the central limit theorem based on m -dependent random variables were Hoeffding and Robbins (1948). As it is mentioned in their work, the central limit theorem for m -dependent random variables can also be extended to what is called $f(n)$ -dependence. A sequence is said to be $f(n)$ -dependent if for $j - i > f(n)$ the two sets of random variables (X_1, X_2, \dots, X_i) and $(X_j, X_{j+1}, \dots, X_n)$ are independent. The magnitude of $f(n)$ should be of sufficiently lower order than the magnitude of n .

Notice that in the case where $f(n)$ is a constant m we have the m -dependent case.

We now extend the notion of m -dependence to the case of multidimensionally indexed random variables:

Definition 3.1.2: For a positive integer r let \mathbb{N}^r denote the r -dimensional positive integer lattice and $\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^r\}$ be an array of random variables defined on a common probability space (Ω, \mathcal{A}, P) . Let $\rho \geq 0$. The random variables $\{X_{\mathbf{i}}, \mathbf{i} \in \mathbb{N}^r\}$ are said to be ρ -radius dependent if $X_{\mathbf{i}_1}$ and $X_{\mathbf{i}_2}$ are independent whenever $d(\mathbf{i}_1, \mathbf{i}_2) > \rho$, where $d(\mathbf{i}_1, \mathbf{i}_2)$ is the Euclidean distance between \mathbf{i}_1 and \mathbf{i}_2 .

Notice that, if $\rho = 0$ the random variables $\{X_i, i \in \mathbf{N}^r\}$ are independent.

Following Hoeffding and Robbins (1948) we can choose ρ to be a function of $|\mathbf{n}|$, where $|\mathbf{n}| = n_1 n_2 \dots n_r$, and thus we have the extension of $f(n)$ -dependence to multidimensionally indexed random variables. As stated before, for $f(n)$ -dependence the magnitude of ρ should be of sufficiently lower order than that of $|\mathbf{n}|$.

The idea of ρ -radius dependent multidimensionally indexed random variables is a natural extension of the notion of m -dependent random variables. However, the extension of m -dependence to random fields can be defined in a different way. According to Heinrich (1987), a family of random variables $X_z, z \in Z^d = \{\pm 0, \pm 1, \pm 2, \dots\}^d$, defined on a probability space (Ω, \mathcal{A}, P) is said to be a random field (m_1, \dots, m_d) -dependent if for any finite $U, V \subset Z^d$ the random vectors $(X_u)_{u \in U}$ and $(X_v)_{v \in V}$ are independent whenever $\min_{u \in U, v \in V} |u_j - v_j| > m_j$ for at least one $j \in \{1, \dots, d\}$. In the case where $m_1 = \dots = m_d = m (\geq 0)$ the random field is usually called m -dependent.

Although the concept of ρ -radius dependence is geometrically more difficult to deal with than the idea of an m -dependent random field, it appears to be more natural.

There are several applications concerning ρ -radius multidimensionally indexed random variables. One might find applications in media technology, meteorology and in general in situations where spatial statistics should be considered.

The main objective of this chapter is to establish a central limit theorem for an array of multidimensionally indexed ρ -radius dependent random variables.

In the following we are dealing with the case $r = 2$. The case $r > 2$ can be treated similarly although the notation becomes more complicated.

3.2 Lattice decomposition and related results

For a real number x , $[x]$ denotes its integer part. That is, $[x]$ denotes the greatest integer less than or equal to x .

We assume the usual partial ordering for the elements of \mathbf{N}^2 , i.e., the notation $(i_1, j_1) \leq (i_2, j_2)$ means $i_1 \leq i_2$ and $j_1 \leq j_2$.

In addition, the notation $(n_1, n_2) \rightarrow \infty$ means $n_i \rightarrow \infty$ for $i = 1, 2$ or equivalently, $\min_{1 \leq i \leq 2} n_i \rightarrow \infty$.

To avoid any trivialities assume that $\rho \geq 1$. Let $(n_1, n_2) \in \mathbf{N}^2$ and assume that $\{X_{i,j}, (i,j) \leq (n_1, n_2)\}$ is an array of ρ -radius dependent random variables.

Let

$$\rho^* = \begin{cases} \rho & \text{if } \rho \text{ is an integer} \\ [\rho + 1] & \text{if } \rho \text{ is not an integer} \end{cases}$$

i.e., ρ^* is the smallest integer greater than or equal to ρ .

Let D_{i_1, i_2}^ν denote the disk with center at the point (i_1, i_2) and radius ν where ν is a positive integer greater than ρ^* , that is,

$$D_{i_1, i_2}^\nu = \{(j_1, j_2) : d((j_1, j_2), (i_1, i_2)) \leq \nu\}.$$

Let $S_{i_1, i_2}^{(D)} = \sum_{(j_1, j_2) \in D_{i_1, i_2}^\nu} X_{j_1, j_2}$ where the summation is taken over all

lattice points (j_1, j_2) which are in the disk D_{i_1, i_2}^ν . Let

$$T_{i_1, i_2} = \{(j_1, j_2) : i_1 - \nu \leq j_1 \leq i_1 + \nu, i_2 - \nu \leq j_2 \leq i_2 + \nu\}$$

i.e., T_{i_1, i_2} is the square circumscribing the disk D_{i_1, i_2}^ν .

$$\text{Let } S_{i_1, i_2}^{(T)} = \sum_{(j_1, j_2) \in T_{i_1, i_2}} X_{j_1, j_2}.$$

To keep the notation simple, let $k = \nu + 1$, $\lambda = 2\nu + \rho^* + 1$, $d_1 = [n_1/\lambda]$ and $d_2 = [n_2/\lambda]$.

By the definition of ρ -radius dependent random variables the following random variables are mutually independent:

$$S_{k, k}^{(T)}, S_{k, (k+\lambda)}^{(T)}, \dots, S_{k, (k+\lambda d_2)}^{(T)},$$

$$S_{(k+\lambda), k}^{(T)}, S_{(k+\lambda), (k+\lambda)}^{(T)}, \dots, S_{(k+\lambda), (k+\lambda d_2)}^{(T)},$$

.....

$$S_{(k+\lambda(d_1-1)), k}^{(T)}, S_{(k+\lambda(d_1-1)), (k+\lambda)}^{(T)}, \dots, S_{(k+\lambda(d_1-1)), (k+\lambda d_2)}^{(T)},$$

$$S_{(k+\lambda d_1), k}^{(T)}, S_{(k+\lambda d_1), (k+\lambda)}^{(T)}, \dots, S_{(k+\lambda d_1), (k+\lambda d_2)}^{(T)}.$$

Clearly,

$$S_{n_1, n_2} = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}$$

is a sum of independent random variables.

To avoid further complications we assume that

$$n_1 = d_1(2\nu + 1) + (d_1 - 1)\rho^* \quad \text{and} \quad n_2 = d_2(2\nu + 1) + (d_2 - 1)\rho^*.$$

Let

$$A_{n_1, n_2} = \{(i, j) : (1, 1) \leq (i, j) \leq (n_1, n_2)\}$$

and

$$B_{n_1, n_2} = \left\{ \bigcup_{i_1=0}^{d_1-1} \bigcup_{i_2=0}^{d_2-1} T_{(k+\lambda i_1), (k+\lambda i_2)} \right\}^c.$$

In order to make things more clarified we shall give a brief explanation of the elements of the set B_{n_1, n_2} .

First, we divide the two-dimensional positive integer lattice points into independent disks, i.e., we require that the distance between any two points, belonging to two different disks is greater than ρ^* . In other words, the distance between the centers of two different disks is greater than or equal to $\lambda = 2\nu + \rho^* + 1$.

To each of the independent disks we circumscribe a square. Therefore, starting from the lowest left corner, the first row of squares consists of the random variables

$S_{k,k}^{(T)}, S_{(k+\lambda),k}^{(T)}, \dots, S_{(k+\lambda d_1),k}^{(T)}$. Similarly, the first column of disks includes the random variables $S_{k,k}^{(T)}, S_{k,(k+\lambda)}^{(T)}, \dots, S_{k,(k+\lambda d_2)}^{(T)}$. Figure 3.1 shows the special case where $\rho^* = 1$

and $\nu = 2$.

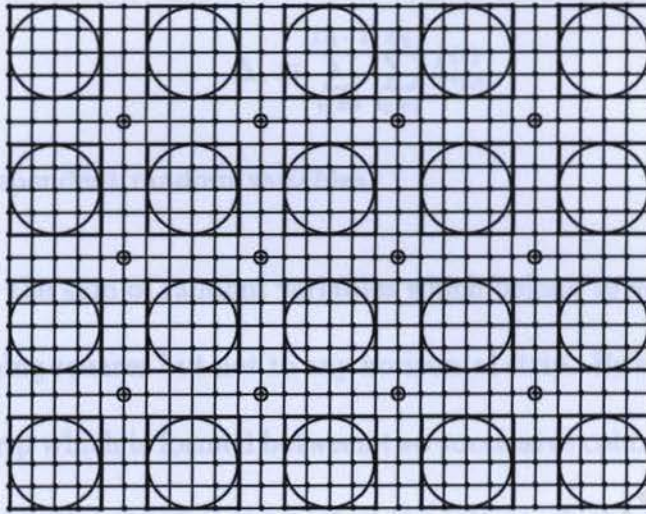


Figure 3.1

The number D_1 of the circumscribing squares in the horizontal dimension is

$$D_1 = \frac{n_1 + \rho^*}{2\nu + 1 + \rho^*} = \frac{n_1 + \rho^*}{\lambda}$$

while the number D_2 of the circumscribing squares in the vertical dimension is

$$D_2 = \frac{n_2 + \rho^*}{2\nu + 1 + \rho^*} = \frac{n_2 + \rho^*}{\lambda}.$$

Now, the set B_{n_1, n_2} includes all lattice points which are not included in any circumscribing square.

Let $\Lambda_{(i_1, j_1 ; i_2, j_2)} = \sum_{s=i_1}^{i_2} \sum_{t=j_1}^{j_2} X_{s,t}.$

Then the sequence $\Lambda_{i_1, i_2}^{(1)} \equiv \Lambda_{(\lambda i_1 - \rho^* + 1, \lambda(i_2 - 1) + 1 ; \lambda i_1, \lambda i_2 - \rho^*)}$ for

$$i_1 = 1, \dots, D_1 - 1 \text{ and } i_2 = 1, \dots, D_2$$

is a sequence of independent random variables and consequently

$$\Delta_c \equiv \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} \Lambda_{i_1, i_2}^{(1)}$$

is a sum of independent random variables.

Note that Δ_c is the sum of random variables which belong to vertical strips between the circumscribing squares but not to any horizontal strip. By saying "vertical strip" we mean the strip which is formed between two successive columns of circumscribing squares. A "horizontal strip" is the part of the lattice which is formed between two successive rows of circumscribing squares.

Similarly, let $\Lambda_{i_1, i_2}^{(2)} \equiv \Lambda_{(\lambda(i_1-1)+1, \lambda i_2 - \rho^* + 1; \lambda i_1 - \rho^*, \lambda i_2)}$ for

$$i_1 = 1, \dots, D_1 \quad \text{and} \quad i_2 = 1, \dots, D_2 - 1$$

and

$$\Delta_r \equiv \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(2)}$$

Then, Δ_r is also a sum of independent random variables.

Δ_r denotes the sum of random variables which belong to horizontal strips between the circumscribing squares but not to any vertical strip.

Let $\Lambda_{i_1, i_2}^{(3)} \equiv \Lambda_{(\lambda i_1 - \rho^* + 1, \lambda i_2 - \rho^* + 1; \lambda i_1, \lambda i_2)}$ for

$$i_1 = 1, \dots, D_1 - 1 \quad \text{and} \quad i_2 = 1, \dots, D_2 - 1.$$

In Figure 3.1 for the special case $\rho^* = 1$, $\nu = 2$ $\Lambda_{i_1, i_2}^{(3)}$ is denoted by \odot . Clearly, $\Lambda_{i_1, i_2}^{(3)}$ is an array of independent random variables and consequently

$$\Delta_b \equiv \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(3)}$$

is a sum of independent random variables.

If Δ_{n_1, n_2} denotes the sum of random variables which do not belong to any circumscribing square then

$$\Delta_{n_1, n_2} = \sum_{(j_1, j_2) \in B_{n_1, n_2}} X_{j_1, j_2} = \Delta_c + \Delta_r + \Delta_b.$$

The following results will be needed for the proof of the central limit theorem in Section 3.3.

Lemma 3.2.1: Let $\{X_{i_1, i_2}, i_1 = 1, \dots, n_1, i_2 = 1, \dots, n_2\}$ be an array of independent random variables with zero mean. Let σ_{i_1, i_2}^2 be the variance of X_{i_1, i_2} while $R_{i_1, i_2}^3 < \infty$ denotes the third absolute moment of X_{i_1, i_2} .

Let

$$r_{n_1, n_2}^3 = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} R_{i_1, i_2}^3 \quad \text{and} \quad s_{n_1, n_2}^2 = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \sigma_{i_1, i_2}^2.$$

If

$$\lim_{(n_1, n_2) \rightarrow \infty} \frac{r_{n_1, n_2}^3}{s_{n_1, n_2}^3} = 0$$

then

(i) $R_{i_1, i_2}^3 < r_{n_1, n_2}$ and $\sigma_{i_1, i_2}^2 < s_{n_1, n_2}$ for $(i_1, i_2) \leq (n_1, n_2)$ and

(ii) $R_{i_1, i_2} < s_{n_1, n_2}$ for all $(n_1, n_2) \geq \mathbf{n}_0$ for some $\mathbf{n}_0 \in \mathbb{N}^2$ and $(i_1, i_2) \leq (n_1, n_2)$.

Proof: The proof of (i) is trivial. For (ii) observe that since

$$\lim_{(n_1, n_2) \rightarrow \infty} \frac{r_{n_1, n_2}}{s_{n_1, n_2}} = 0$$

then $\exists \mathbf{n}_0 \in \mathbb{N}^2$ such that for $(n_1, n_2) \geq \mathbf{n}_0$, $r_{n_1, n_2} < s_{n_1, n_2}$.

Thus, by (i) $R_{i_1, i_2} < r_{n_1, n_2} < s_{n_1, n_2}$ and the proof is complete.

3.3 A Liapounov type theorem

Lemma 3.2.2: Let $\{X_{i_1, i_2}, i_1 = 1, \dots, n_1, i_2 = 1, \dots, n_2\}$ be an array of independent random variables with zero mean. If $EX_{i_1, i_2}^2 = \sigma_{i_1, i_2}^2$ and $E|X_{i_1, i_2}|^3 = R_{i_1, i_2}^3 < \infty$ then,

$$\sigma_{i_1, i_2} \leq R_{i_1, i_2} \text{ for } (i_1, i_2) \leq (n_1, n_2).$$

Proof: By Liapounov's inequality $(E|X|^r)^{\frac{1}{r}}$ is increasing in r , for $r > 0$.

Then, $(E|X_{i_1, i_2}|^2)^{\frac{1}{2}} \leq (E|X_{i_1, i_2}|^3)^{\frac{1}{3}}$ i.e., $(\sigma_{i_1, i_2}^2)^{\frac{1}{2}} \leq (R_{i_1, i_2}^3)^{\frac{1}{3}}$ which implies that $\sigma_{i_1, i_2} \leq R_{i_1, i_2}$ and the proof is complete.

The inequality in the following lemma is the so called Rosenthal inequality the proof of which can be found in Petrov (1994).

Lemma 3.2.3: Let X_1, \dots, X_n be a sequence of independent random variables with $E|X_i|^p < \infty$ for $i = 1, \dots, n$ and for some $p \geq 2$. Without any loss of generality assume that $EX_i = 0$, $i = 1, \dots, n$.

Let

$$S_n = \sum_{i=1}^n X_i, \quad M_{p,n} = \sum_{i=1}^n E|X_i|^p \quad \text{and} \quad B_n = \sum_{i=1}^n EX_i^2.$$

Then,

$$E|S_n|^p \leq c(p)(M_{p,n} + B_n^{p/2})$$

where $c(p)$ is a positive constant depending only on p .

3.3 A Liapounov type theorem

The main result of this chapter is the following Liapounov type theorem.

Theorem 3.3.1: Let \mathbb{N}^2 be the two dimensional positive integer lattice. For $n \in \mathbb{N}^2$ let $\{X_i, i \leq n\}$ be an array of ρ -radius dependent random variables.

Assume that $EX_{i,j} = 0$ for $i = 1, \dots, n_1, j = 1, \dots, n_2$ without any loss of generality.

Let

$$\{S_{(k+\lambda_{i_1}), (k+\lambda_{i_2})}^{(T)}, \quad i_1 = 0, \dots, d_1 - 1, \quad i_2 = 0, \dots, d_2 - 1\}$$

be the array of independent random variables defined in Section 3.2.

Let σ_{i_1, i_2}^2 be the variance of $S_{(k+\lambda_{i_1}), (k+\lambda_{i_2})}^{(T)}$ and let also

$$r_{i_1, i_2}^3 = E|S_{(k+\lambda_{i_1}), (k+\lambda_{i_2})}^{(T)}|^3 \quad \text{and} \quad (\gamma_{i_1, i_2}^{(\zeta)})^3 = E|\Lambda_{i_1, i_2}^{(\zeta)}|^3 \quad \text{for } \zeta = 1, 2, 3.$$

Assume the finiteness of the above quantities and define

$$r_{n_1, n_2}^3 = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} r_{i_1, i_2}^3, \quad \sigma_{n_1, n_2}^2 = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \sigma_{i_1, i_2}^2$$

$$\text{and } (\gamma_{n_1, n_2}^{(\zeta)})^3 = \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} (\gamma_{i_1, i_2}^{(\zeta)})^3, \quad \zeta = 1, 2, 3$$

where

$$\gamma_{D_1, i_2}^{(1)} = \gamma_{i_1, D_2}^{(2)} = \gamma_{D_1, D_2}^{(3)} \equiv 0 \quad \text{for } i_1 = 1, \dots, D_1, \quad i_2 = 1, \dots, D_2.$$

Assume that the following conditions are satisfied.

$$\lim_{(n_1, n_2) \rightarrow \infty} \frac{r_{n_1, n_2}}{\sigma_{n_1, n_2}} = 0, \quad \lim_{(n_1, n_2) \rightarrow \infty} \frac{\gamma_{n_1, n_2}^{(\zeta)}}{\sigma_{n_1, n_2}} = 0,$$

$$\sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(\zeta)})^2 \rightarrow 0 \quad \text{as } (n_1, n_2) \rightarrow \infty \quad \text{for } \zeta = 1, 2, 3.$$

Then,

$$\frac{1}{n_1 n_2} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2} \quad \text{is } AN\left(0, \frac{\sigma_{n_1, n_2}^2}{n_1^2 n_2^2}\right).$$

Proof: The proof is based on the technique first introduced by S. Bernstein (see Hoeffding and Robbins (1948)).

We can express the sum $\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2}$ as

we have

$$\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2} = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)} + \Delta_{n_1, n_2}$$

where Δ_{n_1, n_2} is defined in Section 3.2.

The theorem will be proved if we show that as $(n_1, n_2) \rightarrow \infty$

$$(i) \quad \frac{1}{\sigma_{n_1, n_2}} \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)} \xrightarrow{d} N(0, 1)$$

and

$$(ii) \quad \frac{1}{\sigma_{n_1, n_2}} \sum_{(j_1, j_2) \in B_{n_1, n_2}} X_{j_1, j_2} \xrightarrow{p} 0$$

where B_{n_1, n_2} is defined in Section 3.2.

We proceed to prove (i).

Following Cramèr (1945), we denote by $\Phi_{i_1, i_2}(t)$ the characteristic function of the random variable $S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}$ and by $\Phi_{n_1, n_2}(t)$ the characteristic function of the random variable $\frac{1}{\sigma_{n_1, n_2}} S_{n_1, n_2}$ where

$$S_{n_1, n_2} = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}.$$

Then, by the independence of the random variables

$$\{S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}, i_1 = 0, \dots, d_1 - 1, i_2 = 0, \dots, d_2 - 1\}$$

we have

$$\Phi_{n_1, n_2}(t) = E \exp \left\{ it \frac{1}{\sigma_{n_1, n_2}} \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)} \right\} = \prod_{i_1=0}^{d_1-1} \prod_{i_2=0}^{d_2-1} \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right).$$

Using a standard argument of complex analysis (see Cartan (1963) for an extended treatment of the topic) we take the logarithms of $\prod_{i_1=0}^{d_1-1} \prod_{i_2=0}^{d_2-1} \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right)$ and $\Phi_{n_1, n_2}(t)$ to get

$$\ln \Phi_{n_1, n_2}(t) = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right). \quad (3.3.1)$$

Then using the MacLaurin expansion we can write

$$\begin{aligned} \Phi_{i_1, i_2}(t) &= 1 + E \{ it S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)} \} + \frac{1}{2} E \{ it S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)} \}^2 + \frac{\theta_1}{3!} E |t S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}|^3 \\ &= 1 - \frac{t^2}{2} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{\theta_1}{6} |t|^3 \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3}, \end{aligned}$$

where θ_1 is a quantity with modulus not exceeding unity.

Then

$$\ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) = \ln \left\{ 1 - \frac{t^2}{2} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{\theta_1}{6} |t|^3 \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \right\} \quad (3.3.2)$$

$$= \ln(1 + z) \quad (3.3.2)$$

where

$$z = \frac{-t^2}{2} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{\theta_1}{6} |t|^3 \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3}. \quad (3.3.3)$$

For t fixed, since $\frac{r_{n_1, n_2}}{\sigma_{n_1, n_2}} \rightarrow 0$, then $z \rightarrow 0$ as $(n_1, n_2) \rightarrow \infty$, i.e., $|z| < 1/2$ for (n_1, n_2) sufficiently large. Now,

$$\begin{aligned} \ln(1+z) &= z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \dots \\ &= z + z^2\left(-\frac{1}{2} + \frac{z}{3} - \frac{z^2}{4} + \frac{z^3}{5} - \dots\right) \\ &= z + \theta_2 z^2 \end{aligned}$$

where $\theta_2 = -\frac{1}{2} + \frac{z}{3} - \frac{z^2}{4} + \frac{z^3}{5} - \dots$ and $|\theta_2| < 1$.

This is true since

$$\begin{aligned} |\theta_2| &\leq \frac{1}{2} + \frac{1}{3}|z| + \frac{1}{4}|z^2| + \frac{1}{5}|z^3| + \dots \\ &< \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\left(\frac{1}{2}\right)^2 + \frac{1}{5}\left(\frac{1}{2}\right)^3 + \dots \\ &< \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \dots \\ &= 1. \end{aligned}$$

Equation (3.3.3) can be written

$$z = \theta_3 \frac{t^2}{2} \frac{r_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{\theta_1}{6} \theta_4 |t|^3 \frac{r_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} \quad (3.3.4)$$

where $\theta_3 = \frac{-\sigma_{i_1, i_2}^2}{r_{i_1, i_2}^2}$ and $\theta_4 = \frac{r_{i_1, i_2}}{\sigma_{n_1, n_2}}$.

Notice that by Lemma 3.2.1 for (n_1, n_2) sufficiently large, we have that $|\theta_4| < 1$ and by Lemma 3.2.2 $|\theta_3| \leq 1$.

Now, using 3.3.2, 3.3.3 and 3.3.4

$$\begin{aligned}
\ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) &= \ln(1+z) \\
&= z + \theta_2 z^2 \\
&= \frac{-t^2}{2} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{\theta_1}{6} |t|^3 \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} + \theta_2 \left(\theta_3 \frac{t^2}{2} \frac{r_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \theta_1 \theta_4 \frac{|t|^3}{6} \frac{r_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} \right)^2 \\
&= \frac{-t^2}{2} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{\theta_1}{6} |t|^3 \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} + \theta_2 \frac{r_{i_1, i_2}^4}{\sigma_{n_1, n_2}^4} \left(\theta_3 \frac{t^2}{2} + \theta_1 \theta_4 \frac{|t|^3}{6} \right)^2 \\
&= \frac{-t^2}{2} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{\theta_1}{6} |t|^3 \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} + \theta_2 \theta_4 \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left(\theta_3 \frac{t^2}{2} + \theta_1 \theta_4 \frac{|t|^3}{6} \right)^2 \\
&= \frac{-t^2}{2} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left\{ \frac{\theta_1}{6} |t|^3 + \theta_2 \theta_4 \left(\theta_3 \frac{t^2}{2} + \theta_1 \theta_4 \frac{|t|^3}{6} \right)^2 \right\}.
\end{aligned}$$

Summing over (i_1, i_2) we have by (3.3.1) that

$$\begin{aligned}
\ln \Phi_{n_1, n_2}(t) &= \frac{-t^2}{2} \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \frac{\sigma_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left\{ \frac{\theta_1}{6} |t|^3 + \theta_2 \theta_4 \left(\theta_3 \frac{t^2}{2} + \theta_1 \theta_4 \frac{|t|^3}{6} \right)^2 \right\} \\
&= \frac{-t^2}{2} + \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left\{ \frac{\theta_1}{6} |t|^3 + \theta_2 \theta_4 \left(\theta_3 \frac{t^2}{2} + \theta_1 \theta_4 \frac{|t|^3}{6} \right)^2 \right\}.
\end{aligned}$$

Therefore,

$$\left| \ln \Phi_{n_1, n_2}(t) + \frac{t^2}{2} \right| \leq \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left| \frac{\theta_1}{6} |t|^3 + \theta_2 \theta_4 \left(\theta_3 \frac{t^2}{2} + \theta_1 \theta_4 \frac{|t|^3}{6} \right)^2 \right|.$$

Observe that,

Using the continuity of characteristic functions of Lévy and Gnani the proof of (i)

is complete.

$$\left| \frac{\theta_1}{6} |t|^3 + \theta_2 \theta_4 \left(\theta_3 \frac{t^2}{2} + \theta_1 \theta_4 \frac{|t|^3}{6} \right)^2 \right|$$

We proceed to prove (ii). Using Markov's inequality we have that,

$$= \left| \frac{\theta_1}{6} |t|^3 + \theta_2 \theta_4 \theta_3^2 \frac{t^4}{4} + \theta_1^2 \theta_2 \theta_4^3 \frac{|t|^6}{36} + \theta_1 \theta_2 \theta_3 \theta_4^2 \frac{|t|^5}{6} \right|$$

$$\leq \frac{|\theta_1|}{6} |t|^3 + |\theta_2 \theta_4 \theta_3^2| \frac{t^4}{4} + |\theta_1^2 \theta_2 \theta_4^3| \frac{|t|^6}{36} + |\theta_1 \theta_2 \theta_3 \theta_4^2| \frac{|t|^5}{6}$$

$$\leq \frac{|t|^3}{6} + \frac{t^4}{4} + \frac{|t|^6}{36} + \frac{|t|^5}{6}$$

since all θ 's are quantities with modulus less than or equal to 1. Thus,

$$\begin{aligned} |\ln \Phi_{n_1, n_2}(t) + \frac{t^2}{2}| &\leq \left(\frac{|t|^3}{6} + \frac{t^4}{4} + \frac{|t|^6}{36} + \frac{|t|^5}{6} \right) \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \frac{r_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \\ &= \left(\frac{|t|^3}{6} + \frac{t^4}{4} + \frac{|t|^6}{36} + \frac{|t|^5}{6} \right) \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3}. \end{aligned}$$

Since

$$\lim_{(n_1, n_2) \rightarrow \infty} \frac{r_{n_1, n_2}}{\sigma_{n_1, n_2}} = 0 \tag{3.3.5}$$

then for every fixed t

$$\lim_{(n_1, n_2) \rightarrow \infty} \ln \Phi_{n_1, n_2}(t) = -\frac{t^2}{2}$$

and thus

$$\lim_{(n_1, n_2) \rightarrow \infty} \Phi_{n_1, n_2}(t) = \exp\left(\frac{-t^2}{2}\right).$$

where inequality (3.3.5) follows from Lemma 3.2 and C denotes a positive constant which is allowed to change from line to line.

Using the continuity of characteristic functions of Lévy and Cramèr the proof of (i) is complete.

We proceed to prove (ii). Using Markov's inequality we have that,

$$P\left\{\left|\frac{1}{\sigma_{n_1, n_2}} \sum_{(j_1, j_2) \in B_{n_1, n_2}} X_{j_1, j_2}\right| > \epsilon\right\} = P\left\{\left|\frac{\Delta_c + \Delta_r + \Delta_b}{\sigma_{n_1, n_2}}\right| > \epsilon\right\}$$

$$\leq P\left\{\left|\frac{\Delta_c}{\sigma_{n_1, n_2}}\right| > \frac{\epsilon}{3}\right\} + P\left\{\left|\frac{\Delta_r}{\sigma_{n_1, n_2}}\right| > \frac{\epsilon}{3}\right\} + P\left\{\left|\frac{\Delta_b}{\sigma_{n_1, n_2}}\right| > \frac{\epsilon}{3}\right\}$$

$$\leq \frac{27}{\sigma_{n_1, n_2}^3 \epsilon^3} (E|\Delta_c|^3 + E|\Delta_r|^3 + E|\Delta_b|^3)$$

$$\begin{aligned} &\leq \frac{C}{\sigma_{n_1, n_2}^3 \epsilon^3} \left\{ \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} E|\Lambda_{i_1, i_2}^{(1)}|^3 + \left\{ \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(1)})^2 \right\}^{\frac{3}{2}} + \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} E|\Lambda_{i_1, i_2}^{(2)}|^3 \right. \\ &\quad \left. + \left\{ \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} E(\Lambda_{i_1, i_2}^{(2)})^2 \right\}^{\frac{3}{2}} + \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} E|\Lambda_{i_1, i_2}^{(3)}|^3 + \left\{ \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} E(\Lambda_{i_1, i_2}^{(3)})^2 \right\}^{\frac{3}{2}} \right\} \quad (3.3.5) \end{aligned}$$

$$\begin{aligned} &= C \left\{ \frac{(\gamma_{n_1, n_2}^{(1)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + \frac{(\gamma_{n_1, n_2}^{(2)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + \frac{(\gamma_{n_1, n_2}^{(3)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + \frac{1}{\epsilon^3} \left\{ \sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(1)})^2 \right\}^{\frac{3}{2}} \right. \\ &\quad \left. + \frac{1}{\epsilon^3} \left\{ \sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} E(\Lambda_{i_1, i_2}^{(2)})^2 \right\}^{\frac{3}{2}} + \frac{1}{\epsilon^3} \left\{ \sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} E(\Lambda_{i_1, i_2}^{(3)})^2 \right\}^{\frac{3}{2}} \right\} \quad (3.3.6) \end{aligned}$$

where inequality (3.3.5) follows from Lemma 3.2.3 and C denotes a positive constant which is allowed to change from line to line.

By the assumptions of the theorem and using the fact that

$$\sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(\zeta)})^2 \rightarrow 0 \text{ as } (n_1, n_2) \rightarrow 0 \text{ for } \zeta = 1, 2, 3$$

we have that

$$P\left\{ \left| \frac{1}{\sigma_{n_1, n_2}} \sum_{(j_1, j_2) \in B_{n_1, n_2}} X_{j_1, j_2} \right| > \epsilon \right\} \rightarrow 0,$$

$$\frac{1}{\sigma_{n_1, n_2}} \sum_{(j_1, j_2) \in B_{n_1, n_2}} X_{j_1, j_2} \xrightarrow{P} 0$$

and therefore the proof of (ii) is complete.

The theorem follows from Slutsky's Theorem.

Remarks

1) Theorem 3.3.1 can be compared to that of Bernstein's (1927) in the sense that the method of "big blocks" used here was first introduced by him. The major difference is that Bernstein imposed assumptions on conditional expectations while in this we don't.

2) In the case of independent and identically distributed multidimensionally indexed random variables, under the second moment assumption, the central limit theorem is proved in Christofides and Serfling (1998). In contrast to the i.i.d. case, in the case of ρ -radius dependent random variables the finiteness of the third moment is needed, in order to prove the result. On the other hand, the finiteness of the third absolute moment of X_{i_1, i_2} which is needed to prove the central limit theorem for

ρ -radius dependent random variables, is also needed for proving the result in the case of independent but not necessarily identically distributed random variables.

3) The conditions under which Theorem 3.3.1 is proved are not very restrictive and are fulfilled in the case $\rho = 0$, i.e., in the case of independent random variables. For example, the assumptions

$$\lim_{(n_1, n_2) \rightarrow \infty} \frac{r_{n_1, n_2}}{\sigma_{n_1, n_2}} = 0 \quad \text{and} \quad \lim_{(n_1, n_2) \rightarrow \infty} \frac{\gamma_{n_1, n_2}^{(\zeta)}}{\sigma_{n_1, n_2}} = 0 \quad \text{for } \zeta = 1, 2, 3$$

are the classical assumptions for the proof of the central limit theorem for independent but not necessarily identically distributed random variables.

In addition, the assumptions

$$\sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(\zeta)})^2 \rightarrow 0 \quad \text{for } \zeta = 1, 2, 3$$

are mild.

4) Theorem 3.3.1 can be extended to an analogous theorem for $f(|\mathbf{n}|)$ -dependence. Of course, $\rho = f(|\mathbf{n}|)$ has to have sufficiently lower order than $|\mathbf{n}|$ and therefore lower order than ν .

5) In this chapter the central limit theorem has been shown for $r = 2$. The result can be extended under the same assumptions to higher dimensions although notation becomes cumbersome and complicated.

Chapter 4

The Berry-Esséen theorem

4.1 Introduction

The exact distribution of a statistic is usually complicated to define and work with. Hence, we usually turn to the approximation of the exact distribution by a simpler distribution with known properties. This is done in Chapter 3. That is, we have approximated the distribution of a sum of multidimensionally indexed ρ -radius dependent random variables by the standard normal distribution under general conditions.

In this chapter we shall show how the values of the parameters affect the speed of convergence to the limit and how large (n_1, n_2) has to be in order the limit distribution (standard normal) to serve as a satisfactory approximation. These two aims are fulfilled to a certain extent by the Berry-Esséen theorem and they have considerable practical and theoretical significance, since one needs to know how

large (n_1, n_2) should be in employing the limit theory. However, estimates of the rate of convergence in the central limit theorem were obtained for the first time by Liapounov (1901), even though the classical result of the Berry-Esséen theorem for a sequence of independent and identically distributed random variables were separately obtained by Berry (1941) and Esséen (1942). We shall therefore give here an explicit upper bound for the difference between the distribution function of a sum of multidimensionally indexed ρ -radius dependent random variables and the standard normal distribution function.

4.2 Preliminaries

The proof of the main theorem involves many estimates. We present these estimates in three different lemmata and then complete the proof of the theorem afterwards.

We will use the same notation as in Chapter 3. First, we present the following lemmata.

Lemma 4.2.1: *Assume that X and Y are arbitrary random variables and $F(x) = P(X < x)$ and $G(x) = P(X + Y < x)$. For any $\epsilon > 0$, $x \in R$ and any distribution function H , we have that*

$$|G(x) - H(x)| \leq \max\{|F(x + \epsilon) - H(x + \epsilon)|, |F(x - \epsilon) - H(x - \epsilon)|\}$$

$$+ \max\{|H(x - \epsilon) - H(x)|, |H(x + \epsilon) - H(x)|\} + P(|Y| \geq \epsilon).$$

Proof: For every real x and $\epsilon > 0$,

$$\begin{aligned} P(X+Y < x) &= P(X+Y < x, Y > -\epsilon) + P(X+Y < x, Y \leq -\epsilon) \\ &\geq -\max\{|F(x+\epsilon) - H(x+\epsilon)|, |F(x-\epsilon) - H(x-\epsilon)|\} \\ &\leq P(X < x + \epsilon) + P(Y \leq -\epsilon) \\ &= \max\{|H(x-\epsilon) - H(x)|, |H(x+\epsilon) - H(x)|\} + P(|Y| \geq \epsilon) \\ &\leq P(X < x + \epsilon) + P(|Y| \geq \epsilon) \end{aligned}$$

which implies that

$$\begin{aligned} G(x) - H(x) &\leq F(x + \epsilon) - H(x) + P(|Y| \geq \epsilon) \\ &= F(x + \epsilon) - H(x) + H(x + \epsilon) - H(x + \epsilon) + P(|Y| \geq \epsilon) \\ &\leq |F(x + \epsilon) - H(x + \epsilon)| + |H(x + \epsilon) - H(x)| + P(|Y| \geq \epsilon) \\ &\leq \max\{|F(x + \epsilon) - H(x + \epsilon)|, |F(x - \epsilon) - H(x - \epsilon)|\} \\ &\quad + \max\{|H(x - \epsilon) - H(x)|, |H(x + \epsilon) - H(x)|\} + P(|Y| \geq \epsilon). \end{aligned}$$

Respectively we have that,

$$\begin{aligned} G(x) &= P(X+Y < x) = 1 - P(X+Y \geq x) \\ &= 1 - P(X+Y \geq x, Y < \epsilon) - P(X+Y \geq x, Y \geq \epsilon) \\ &\geq 1 - P(X \geq x - \epsilon) - P(Y \geq \epsilon) \geq 1 - P(X \geq x - \epsilon) - P(|Y| \geq \epsilon) \\ &= F(x - \epsilon) - P(|Y| \geq \epsilon). \end{aligned}$$

Thus,

$$G(x) - H(x) \geq F(x - \epsilon) - H(x) - P(|Y| \geq \epsilon)$$

$$\begin{aligned}
&= F(x - \epsilon) - H(x - \epsilon) + H(x - \epsilon) - H(x) - P(|Y| \geq \epsilon) \\
&\geq -|F(x - \epsilon) - H(x - \epsilon)| - |H(x - \epsilon) - H(x)| - P(|Y| \geq \epsilon) \\
&\geq -\max\{|F(x + \epsilon) - H(x + \epsilon)|, |F(x - \epsilon) - H(x - \epsilon)|\} \\
&\quad - \max\{|H(x - \epsilon) - H(x)|, |H(x + \epsilon) - H(x)|\} - P(|Y| \geq \epsilon).
\end{aligned}$$

Therefore,

$$\begin{aligned}
|G(x) - H(x)| &\leq \max\{|F(x + \epsilon) - H(x + \epsilon)|, |F(x - \epsilon) - H(x - \epsilon)|\} \\
&\quad + \max\{|H(x - \epsilon) - H(x)|, |H(x + \epsilon) - H(x)|\} + P(|Y| \geq \epsilon).
\end{aligned}$$

Notice that the previous inequality lead us to another very important inequality, namely

$$\begin{aligned}
|G(x) - H(x)| &\leq \sup_x |F(x) - H(x)| \\
&\quad + \max\{|H(x - \epsilon) - H(x)|, |H(x + \epsilon) - H(x)|\} + P(|Y| \geq \epsilon)
\end{aligned} \tag{4.2.1}$$

for every real x and $\epsilon > 0$.

Lemma 4.2.2: Let F be a distribution function and G a real differentiable function with $G(x) \rightarrow 0$ as $x \rightarrow -\infty$ or $G(x) \rightarrow 1$ as $x \rightarrow \infty$.

Let $\sup_x |G'(x)| \leq M$ where M is a positive constant.

If $F - G \in L_1$ and G is of bounded variation on $(-\infty, \infty)$ then for every $T > 0$

$$\sup_x |F(x) - G(x)| \leq \frac{2}{\pi} \int_0^T \frac{|\phi_F(t) - \phi_G(t)|}{t} dt + \frac{24M}{\pi T}$$

where $\phi_F(t)$ and $\phi_G(t)$ are Fourier-Stieltjes transforms of F, G .

Proof: The proof can be found in Chow and Teicher (1988) p. 302.

Lemma 4.2.3: Assume that X and Y are arbitrary random variables and let $F(x) = P(X < x)$ and $G(x) = P(X + Y < x)$.

Let $\epsilon > 0$, $x \in R$ and $\Phi(x)$ be the standard normal distribution function.

If $\sup_x |F(x) - \Phi(x)| \leq M$ then,

$$\sup_x |G(x) - \Phi(x)| \leq M + \max\{|\Phi(x - \epsilon) - \Phi(x)|, |\Phi(x + \epsilon) - \Phi(x)|\} + P(|Y| \geq \epsilon).$$

Proof: The proof is straightforward from Lemma 4.2.1 by setting $H(x) = \Phi(x)$.

Lemma 4.2.4: Let $\{X_{i_1, i_2}, (i_1, i_2) \in \{(1, 1), \dots, (n_1, n_2)\}\}$ be an array of independent random variables with

$$EX_{i_1, i_2} = 0, \quad EX_{i_1, i_2}^2 = s_{i_1, i_2}^2 \quad \text{and} \quad E|X_{i_1, i_2}|^3 = R_{i_1, i_2}^3.$$

Define

$$S_{n_1, n_2} = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2}, \quad \sigma_{n_1, n_2}^2 = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} s_{i_1, i_2}^2,$$

$$r_{n_1, n_2}^3 = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} R_{i_1, i_2}^3 \quad \text{and} \quad \Phi_{n_1, n_2}(t) = E \exp\left(\frac{itS_{n_1, n_2}}{\sigma_{n_1, n_2}}\right).$$

Then, \exists a positive constant C such that

$$|\Phi_{n_1, n_2}(t) - e^{-\frac{t^2}{2}}| \leq C \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} |t|^3 e^{-\frac{t^2}{6}} \quad \text{for} \quad |t| < \frac{\sigma_{n_1, n_2}^3}{2r_{n_1, n_2}^3}.$$

Proof: Following Cramèr (1945), we denote by $\Phi_{i_1, i_2}(t)$ the characteristic function of the random variable X_{i_1, i_2} and by $\Phi_{S_{n_1, n_2}}(t)$ the characteristic function of the random variable $\frac{1}{\sigma_{n_1, n_2}} S_{n_1, n_2}$.

Then, by the independence of the random variables

$\{X_{i_1, i_2}, (i_1, i_2) \in \{(1, 1), \dots, (n_1, n_2)\}\}$ we have that

$$\Phi_{S_{n_1, n_2}}(t) = E \exp\left(\frac{it S_{n_1, n_2}}{\sigma_{n_1, n_2}}\right) = \prod_{i_1=1}^{n_1} \prod_{i_2=1}^{n_2} \Phi_{i_1, i_2}\left(\frac{t}{\sigma_{n_1, n_2}}\right).$$

Using the same argument as that of Chapter 3 we get

$$\ln \Phi_{S_{n_1, n_2}}(t) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \ln \Phi_{i_1, i_2}\left(\frac{t}{\sigma_{n_1, n_2}}\right). \quad (4.2.2)$$

Using the MacLaurin expansion we write

$$\begin{aligned} \Phi_{i_1, i_2}(t) &= 1 + E(itX_{i_1, i_2}) + \frac{1}{2} E(itX_{i_1, i_2})^2 + \frac{\theta}{3!} E|tX_{i_1, i_2}|^3 \\ &= 1 - \frac{t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \frac{\theta}{6} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \end{aligned}$$

where θ is a quantity with modulus not exceeding unity.

Then

$$\begin{aligned} \ln \Phi_{i_1, i_2}\left(\frac{t}{\sigma_{n_1, n_2}}\right) &= \ln\left\{1 - \frac{t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \frac{\theta}{6} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3}\right\} \\ &= \ln(1 + z) \end{aligned} \quad (4.2.3)$$

where $\theta = \frac{s_{i_1, i_2}^2}{R_{i_1, i_2}^2}$ and assume that $|\theta| < 1$.

$$z = \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \frac{\theta}{6} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3}. \quad (4.2.4)$$

Let $\theta_2 = -\frac{s_{i_1, i_2}^2}{R_{i_1, i_2}^2}$. Then, by Lemma 3.2.2, $|\theta_2| \leq 1$.

Therefore, z becomes

$$z = \frac{\theta_2 t^2 R_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \frac{\theta}{6} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3}. \quad (4.2.5)$$

By equation (4.2.3)

$$\ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) = \ln(1 + z) = z + \theta_3 z^2 \quad (4.2.6)$$

with $|\theta_3| < 1$ and where the last equality of (4.2.6) follows using the same arguments as in the proof of the central limit theorem in Chapter 3.

Replacing the two forms of z from (4.2.4) and (4.2.5) into (4.2.6) we have,

$$\begin{aligned} \ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) &= \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \frac{\theta}{6} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} + \theta_3 \left\{ \frac{\theta_2 t^2 R_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \frac{\theta}{6} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \right\}^2 \\ &= \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \frac{\theta}{6} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} + \theta_3 t^4 \frac{R_{i_1, i_2}^4}{\sigma_{n_1, n_2}^4} \left\{ \frac{\theta_2}{2} + \frac{\theta}{6} |t| \frac{R_{i_1, i_2}}{\sigma_{n_1, n_2}} \right\}^2 \\ &= \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left\{ \frac{\theta}{6} + \theta_3 |t| \frac{R_{i_1, i_2}}{\sigma_{n_1, n_2}} \left(\frac{\theta_2}{2} + \frac{\theta}{6} |t| \frac{R_{i_1, i_2}}{\sigma_{n_1, n_2}} \right)^2 \right\}. \end{aligned}$$

Let $\theta_4 = 2|t| \frac{r_{n_1, n_2}}{\sigma_{n_1, n_2}}$ and assume that $\theta_4 < 1$.

Then,

$$\begin{aligned} \ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) &= \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left\{ \frac{\theta}{6} + \frac{\theta_3 \theta_4 R_{i_1, i_2}}{2 r_{n_1, n_2}} \left(\frac{\theta_2}{2} + \frac{\theta}{12} \theta_4 \frac{R_{i_1, i_2}}{r_{n_1, n_2}} \right)^2 \right\} \\ &= \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \left\{ \frac{\theta}{6} + \frac{\theta_3}{2} \theta_4 \theta_5 \left(\frac{\theta_2}{2} + \frac{\theta}{12} \theta_4 \theta_5 \right)^2 \right\} \end{aligned}$$

where $\theta_5 = \frac{R_{i_1, i_2}}{r_{n_1, n_2}}$ and by Lemma 3.2.1 θ_5 is less than 1.

Now put

$$\theta_6 = \frac{\theta}{6} + \frac{\theta_3}{2} \theta_4 \theta_5 \left(\frac{\theta_2}{2} + \frac{\theta}{12} \theta_4 \theta_5 \right)^2.$$

Then,

$$\ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) = \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \theta_6 |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3}.$$

Clearly, $|\theta_6| < \frac{97}{288}$. Thus, put $\theta_7 = \theta_6 \frac{288}{97}$ and $|\theta_7| < 1$ so that we can finally have that

$$\ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) = \frac{-t^2 s_{i_1, i_2}^2}{2 \sigma_{n_1, n_2}^2} + \theta_7 |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \frac{97}{288}.$$

By equation (4.2.2),

$$\ln \Phi_{S_{n_1, n_2}}(t) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right)$$

since $\exp\left(\frac{1}{2}\right) < 2$ and thus,

$$= \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \left\{ \frac{-t^2}{2} \frac{s_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \theta_7 |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \frac{97}{288} \right\}$$

$$= \frac{-t^2}{2} + \theta_7 |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} \frac{97}{288}.$$

Therefore,

$$|\Phi_{S_{n_1, n_2}}(t) - e^{-\frac{t^2}{2}}| = \left| \exp\left(\frac{-t^2}{2} + \theta_7 |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} \frac{97}{288}\right) - e^{-\frac{t^2}{2}} \right|$$

$$= e^{-\frac{t^2}{2}} \left| \exp\left(\theta_7 |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} \frac{97}{288}\right) - 1 \right|$$

$$\leq e^{-\frac{t^2}{2}} |\theta_7| |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} \frac{97}{288} \exp\left(|\theta_7| |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} \frac{97}{288}\right)$$

where the last inequality follows from the elementary inequality $|e^x - 1| \leq |x|e^{|x|}$.

Since $|t| < \frac{\sigma_{n_1, n_2}}{2r_{n_1, n_2}}$ then $|t|^3 < \frac{\sigma_{n_1, n_2}^3}{8r_{n_1, n_2}^3}$. Also, $\frac{97}{288} < \frac{1}{2}$.

Therefore,

$$|\Phi_{S_{n_1, n_2}}(t) - e^{-\frac{t^2}{2}}| \leq e^{-\frac{t^2}{2}} \frac{1}{2} |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} e^{\frac{1}{8} \frac{97}{288}}$$

$$\leq e^{-\frac{t^2}{2}} \frac{1}{2} |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} 2$$

since $\exp(\frac{1}{8} \frac{97}{288}) < 2$ and thus,

$$|\Phi_{S_{n_1, n_2}}(t) - e^{-\frac{t^2}{2}}| \leq e^{-\frac{t^2}{2}} |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3}. \quad (4.2.7)$$

Equation (4.2.7) is valid for $|t| < \frac{\sigma_{n_1, n_2}}{2r_{n_1, n_2}}$.

We shall extend the range of $|t|$ as it is in the statement of Lemma 4.2.4.

Let X_1, X_2 be two independent random variables with corresponding characteristic functions ϕ_1 and ϕ_2 . Clearly, the characteristic function of the sum $X_1 + X_2$ is the product $\phi_1 \phi_2$.

Suppose now that X_1 and X_2 are two independent and identically distributed random variables with $Var(X_1) = \sigma^2$, $E|X_1|^3 = r^3$ and $E(e^{itX_1}) = \phi(t)$. Clearly, the corresponding characteristic function of $X_1 + X_2$ is ϕ^2 . Now, let $Y = X_1 - X_2$. Then, $Var(Y) = Var(X_1 - X_2) = 2\sigma^2$ and $E|X_1 - X_2|^3 \leq 8r^3$. Next, instead of writing that the characteristic function of $X_1 - X_2$, is equal to the square of $|\phi(t)|$, we expand the characteristic function of the random variable Y , with variance $2\sigma^2$ and absolute third moment less than or equal to $8r^3$. In other words we symmetrize the random variable X .

So,

$$\begin{aligned}
 \left| \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) \right|^2 &= 1 - \frac{t^2 (2s_{i_1, i_2}^2)}{2\sigma_{n_1, n_2}^2} + \theta |t|^3 \frac{8R_{i_1, i_2}^3}{6\sigma_{n_1, n_2}^3} \\
 &\leq 1 - \frac{t^2 s_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{4}{3} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3}
 \end{aligned}$$

So, using the triangle inequality

$$\leq \exp \left(-\frac{t^2 s_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{4}{3} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \right)$$

where the last inequality follows from the elementary inequality $1 + x \leq e^x$.

Thus,

$$\ln \left| \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right) \right| \leq \frac{1}{2} \left(-\frac{t^2 s_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{4}{3} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \right). \quad (4.2.8)$$

By equation (4.2.2)

$$\ln \Phi_{S_{n_1, n_2}}(t) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \ln \Phi_{i_1, i_2} \left(\frac{t}{\sigma_{n_1, n_2}} \right)$$

and equation (4.2.8) we have that

$$\left| \Phi_{S_{n_1, n_2}}(t) \right|^2 \leq \exp \left\{ \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \left(-\frac{t^2 s_{i_1, i_2}^2}{\sigma_{n_1, n_2}^2} + \frac{4}{3} |t|^3 \frac{R_{i_1, i_2}^3}{\sigma_{n_1, n_2}^3} \right) \right\} \quad (4.2.9)$$

Combining (4.2.7) and (4.2.9) we have that

$$= \exp \left(-t^2 + \frac{4}{3} |t|^3 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} \right).$$

By the assumption of Lemma 4.2.4, $|t| < \frac{\sigma_{n_1, n_2}^3}{2r_{n_1, n_2}^3}$ and therefore we have that

4.3 Berry-Essén theorem

$$|\Phi_{S_{n_1, n_2}}(t)|^2 \leq \exp(-t^2 + \frac{4}{3} \frac{t^2}{2} \frac{\sigma_{n_1, n_2}^3}{r_{n_1, n_2}^3} \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3})$$

Theorem 4.3.1: Let $\{X_{k, i}, (i, k) \in \{(1, 1), \dots, (n_1, n_2)\}\}$ be an array of r -valued independent random variables. Assume that $EX_{k, i} = 0$ with no loss of generality.

$$= \exp(-t^2 + \frac{2}{3} t^2) = \exp(-\frac{t^2}{3}).$$

Let

So, using the triangle inequality

$$|\Phi_{S_{n_1, n_2}}(t) - e^{-\frac{t^2}{2}}| \leq |\Phi_{S_{n_1, n_2}}(t)| + e^{-\frac{t^2}{2}}$$

is the array of independent

Let $\sigma_{i, j}^2$ be the variance of $S_{(i, j)}$ and let

$$\leq e^{-\frac{t^2}{6}} + e^{-\frac{t^2}{2}} \leq 2e^{-\frac{t^2}{6}}.$$

Let $\frac{\sigma_{n_1, n_2}}{2r_{n_1, n_2}} < |t| < \frac{\sigma_{n_1, n_2}^3}{2r_{n_1, n_2}^3}$.

Then, $8 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} |t|^3 > 1$ and

Assume that the above quantities

$$|\Phi_{S_{n_1, n_2}}(t) - e^{-\frac{t^2}{2}}| \leq 2e^{-\frac{t^2}{6}}$$

$$\leq 8 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} |t|^3 2e^{-\frac{t^2}{6}}$$

$$= 16 \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} |t|^3 e^{-\frac{t^2}{6}}. \quad (4.2.9)$$

Combining (4.2.7) and (4.2.9) we have that,

$$|\Phi_{S_{n_1, n_2}}(t) - e^{-\frac{t^2}{2}}| \leq C \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} |t|^3 e^{-\frac{t^2}{6}} \quad \text{for} \quad |t| < \frac{\sigma_{n_1, n_2}^3}{2r_{n_1, n_2}^3}$$

and the proof of the lemma is complete.

4.3 Berry-Esséen theorem

Theorem 4.3.1: Let $\{X_{i_1, i_2}, (i_1, i_2) \in \{(1, 1), \dots, (n_1, n_2)\}\}$ be an array of ρ -radius dependent random variables. Assume that $EX_{i_1, i_2} = 0$ without any loss of generality.

Let

$$\{S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}, \quad i_1 = 0, \dots, d_1 - 1, \quad i_2 = 0, \dots, d_2 - 1\}$$

be the array of independent random variables defined in Section 3.2.

Let σ_{i_1, i_2}^2 be the variance of $S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}$ and let also

$$r_{i_1, i_2}^3 = E|S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)}|^3 \quad \text{and} \quad (\gamma_{i_1, i_2}^{(\zeta)})^3 = E|\Lambda_{i_1, i_2}^{(\zeta)}|^3 \quad \text{for } \zeta = 1, 2, 3.$$

Assume that the above quantities are finite and define

$$r_{n_1, n_2}^3 = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} r_{i_1, i_2}^3, \quad \sigma_{n_1, n_2}^2 = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \sigma_{i_1, i_2}^2,$$

$$\text{and} \quad (\gamma_{n_1, n_2}^{(\zeta)})^3 = \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} (\gamma_{i_1, i_2}^{(\zeta)})^3, \quad \zeta = 1, 2, 3.$$

Assume that the following condition is satisfied.

$$\sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(\zeta)})^2 = O(\nu^{-1}) \quad \text{for } \zeta = 1, 2, 3$$

with $\nu = O((n_1 n_2)^a)$, $0 < a < 1$.

Then for, $\epsilon_n = \epsilon(\mathbf{n}) > 0$ and $x \in R$ we have that

$$\sup_x \left| P\left(\frac{1}{\sigma_{n_1, n_2}} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2} < x\right) - \Phi(x) \right| \leq C \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} + C \frac{(\gamma_{n_1, n_2}^{(1)})^3}{\sigma_{n_1, n_2}^3 \epsilon_n^3} + C \frac{(\gamma_{n_1, n_2}^{(2)})^3}{\sigma_{n_1, n_2}^3 \epsilon_n^3} + C \frac{(\gamma_{n_1, n_2}^{(3)})^3}{\sigma_{n_1, n_2}^3 \epsilon_n^3} + \frac{C}{\epsilon_n^3} \nu^{-\frac{3}{2}} + \frac{\epsilon_n}{\sqrt{2\pi}}$$

Therefore for some constant C .

Proof: We can express the sum $\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2}$ as

$$\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2} = S_{n_1, n_2} + \Delta_{n_1, n_2}$$

where S_{n_1, n_2} and Δ_{n_1, n_2} are given in Section 3.2.

Put

$$X = \frac{S_{n_1, n_2}}{\sigma_{n_1, n_2}} \quad \text{and} \quad Y = \frac{\Delta_{n_1, n_2}}{\sigma_{n_1, n_2}}$$

Clearly,

In addition, let $F(x) = P(X < x)$ and $G(x) = P(X + Y < x)$.

Using equation (4.2.1) and taking $H(x)$ to be the standard normal distribution function we have that

$$\begin{aligned} |G(x) - \Phi(x)| &\leq \sup_x |F(x) - \Phi(x)| \\ &+ \max\{|\Phi(x - \epsilon) - \Phi(x)|, |\Phi(x + \epsilon) - \Phi(x)|\} + P(|Y| \geq \epsilon). \end{aligned}$$

Observe that $\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ and $\sup_x |\Phi'(x)| \leq \frac{1}{\sqrt{2\pi}}$.

In addition, the random variables with the following corresponding distribution functions $F(x)$ and $\Phi(x)$ both have mean zero and variance one. By Chebychev's inequality $F(x) < \frac{1}{x^2}$ for $x < 0$ and $1 - F(x) \leq \frac{1}{x^2}$ for $x > 0$. Similarly, $\Phi(x) < \frac{1}{x^2}$ for $x < 0$ and $1 - \Phi(x) \leq \frac{1}{x^2}$ for $x > 0$. Then $F - \Phi \in L_1$ and the assumptions of Lemma 4.2.2 are satisfied.

Therefore

$$\sup_x |F(x) - \Phi(x)| \leq \frac{2}{\pi} \int_0^T \left| \frac{\phi_F(t) - \phi(t)}{t} \right| dt + \frac{24}{\pi T \sqrt{2\pi}}$$

where $\phi_F(t)$ and $\phi(t)$ denote the characteristic functions of F and Φ .

By Lemma 4.2.3 we get

$$\begin{aligned} \sup_x |G(x) - \Phi(x)| &\leq \frac{2}{\pi} \int_0^T \left| \frac{\phi_F(t) - \phi(t)}{t} \right| dt + \frac{24}{\pi T \sqrt{2\pi}} \\ &+ \max\{|\Phi(x - \epsilon) - \Phi(x)|, |\Phi(x + \epsilon) - \Phi(x)|\} + P(|Y| \geq \epsilon). \end{aligned} \quad (4.3.1)$$

Clearly,

$$\max\{|\Phi(x - \epsilon) - \Phi(x)|, |\Phi(x + \epsilon) - \Phi(x)|\} \leq \frac{\epsilon}{\sqrt{2\pi}}.$$

Using Lemma 4.2.4 and since X represents a sum of independent random variables we have

$$|\phi_F(t) - \phi(t)| = |\phi_F(t) - e^{-\frac{t^2}{2}}| \leq C \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} |t|^3 e^{-\frac{t^2}{6}}, \quad |t| < \frac{\sigma_{n_1, n_2}^3}{2r_{n_1, n_2}^3}.$$

Therefore for $T = \frac{\sigma_{n_1, n_2}^3}{r_{n_1, n_2}^3}$ the right hand side of inequality (4.3.1) is less than or equal to

Since by the assumption of the theorem,

$$\begin{aligned} & \frac{2}{\pi} C \int_0^T \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} t^2 e^{-\frac{t^2}{\epsilon}} dt + \frac{24}{\pi \sqrt{2\pi}} \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} + \frac{\epsilon}{\sqrt{2\pi}} + P(|Y| \geq \epsilon) \\ & \leq \frac{2}{\pi} C \int_0^\infty \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} t^2 e^{-\frac{t^2}{\epsilon}} dt + \frac{24}{\pi \sqrt{2\pi}} \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} + \frac{\epsilon}{\sqrt{2\pi}} + P(|Y| \geq \epsilon), \quad |t| < \frac{\sigma_{n_1, n_2}^3}{2r_{n_1, n_2}^3}. \end{aligned}$$

Then,

$$\sup_x |G(x) - \Phi(x)| \leq$$

$$\frac{2}{\pi} C \int_0^\infty \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} t^2 e^{-\frac{t^2}{\epsilon}} dt + \frac{24}{\sqrt{2\pi}} \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} + \frac{\epsilon}{\sqrt{2\pi}} + P(|Y| \geq \epsilon)$$

Remark

1) For the Δ_{n_1, n_2} identically distributed, the rate of convergence of $G(x)$ to $\Phi(x)$ is equal to

$$\begin{aligned} & = \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} \left\{ \frac{2}{\pi} C \int_0^\infty t^2 e^{-\frac{t^2}{\epsilon}} dt + \frac{24}{\sqrt{2\pi}} \right\} + \frac{\epsilon}{\sqrt{2\pi}} + P(|Y| \geq \epsilon) \\ & = C \frac{r_{n_1, n_2}^3}{\sigma_{n_1, n_2}^3} + \frac{\epsilon}{\sqrt{2\pi}} + P(|Y| \geq \epsilon). \end{aligned}$$

Using Markov's inequality and inequality (3.3.6) we have that

Notice that as it was mentioned at the beginning of this chapter, the rate of con-

$$\begin{aligned} P(|Y| \geq \epsilon) & = P\left(\left|\frac{\Delta_{n_1, n_2}}{\sigma_{n_1, n_2}}\right| \geq \epsilon\right) \leq \\ & C \left\{ \frac{(\gamma_{n_1, n_2}^{(1)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + \frac{(\gamma_{n_1, n_2}^{(2)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + \frac{(\gamma_{n_1, n_2}^{(3)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + \frac{1}{\epsilon^3} \left\{ \sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(1)})^2 \right\}^{\frac{3}{2}} \right. \\ & \left. + \frac{1}{\epsilon^3} \left\{ \sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} E(\Lambda_{i_1, i_2}^{(2)})^2 \right\}^{\frac{3}{2}} + \frac{1}{\epsilon^3} \left\{ \sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} E(\Lambda_{i_1, i_2}^{(3)})^2 \right\}^{\frac{3}{2}} \right\} \end{aligned}$$

where C is a constant.

Since by the assumption of the theorem,

$$\sigma_{n_1, n_2}^{-2} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2} E(\Lambda_{i_1, i_2}^{(\zeta)})^2 = O(\nu^{-1}) \text{ for } \zeta = 1, 2, 3$$

$$\text{with } \nu = O((n_1 n_2)^a), \quad 0 < a < 1$$

$$\sup_x |G(x) - \Phi(x)| \leq C \frac{n_{1, n_2}^3}{\sigma_{n_1, n_2}^3} + C \frac{(\gamma_{n_1, n_2}^{(1)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + C \frac{(\gamma_{n_1, n_2}^{(2)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + C \frac{(\gamma_{n_1, n_2}^{(3)})^3}{\sigma_{n_1, n_2}^3 \epsilon^3} + \frac{C}{\epsilon^3} \nu^{-\frac{3}{2}} + \frac{\epsilon}{\sqrt{2\pi}}$$

and the proof of the Theorem 4.3.1 is complete.

Remarks

1) For the case where the random variables are independent and identically distributed, the rate of convergence of $G(x)$ to $\Phi(x)$ is equal to

$$\min\{|\mathbf{n}|^{-\frac{1}{2}}, |\mathbf{n}|^{-\frac{1}{2}} \epsilon_{\mathbf{n}}^{-3}, |\mathbf{n}|^{-\frac{3}{2}\alpha} \epsilon_{\mathbf{n}}^{-3}, \epsilon_{\mathbf{n}}\}$$

where $0 < \alpha < 1$ and $\epsilon_{\mathbf{n}} = \epsilon(\mathbf{n})$.

Notice that, as it was mentioned at the beginning of this chapter, the rate of convergence is based on the behavior of α and $\epsilon(\mathbf{n})$.

2) The "standard" tricks which are used in the above estimates have their clear origin in Liapounov's proof of his theorem in 1901 as the important problem of error estimation is considered there for the first time.

Chapter 5

The strong law of large numbers

5.1 Introduction

The laws of large numbers have a long history. The standard approach to the strong laws of large numbers for independent random variables is based on truncation and the use of the Kolmogorov's criterion. We distinguish here two fundamental references. The first one is due to Kolmogorov (1933) which includes the well known three-series theorem and the second is a classical paper of Chung (1947). Surveys on the strong law of large numbers for sequences of independent random variables can be found in Stout (1974).

For this classical limit theorem there are various extensions. Such extensions include the strong law of large numbers for multidimensionally indexed random variables and the strong law of large numbers for dependent random variables. For the former extension, several results and references are presented in the first chapter of this

thesis while for the latter see Blum and Brennan (1980), Kuchkarov (1990) and Sharakhmetov (1995). In this chapter we give a strong law of large numbers for multidimensionally indexed ρ -radius dependent random variables using the method of truncation, followed by the adjusted Kronecker's lemma for multidimensionally indexed random variables.

5.2 Main results

For the proof of the strong law of large numbers the following lemma is to be used.

The lemma can be found in Chung (1974) p. 124.

Lemma 5.2.1: *Let ϕ be a positive and even function on R^1 such that as $|x|$ increases $\frac{\phi(x)}{|x|}$ increases and $\frac{\phi(x)}{x^2}$ decreases.*

Then, for $|x| \leq a$

$$\frac{\phi(a)}{a^2} \leq \frac{\phi(x)}{x^2} \quad \text{and therefore} \quad \frac{x^2}{a^2} \leq \frac{\phi(x)}{\phi(a)}. \quad (5.2.1)$$

In addition, for $|x| > a > 0$ we have that

$$\frac{\phi(x)}{|x|} \geq \frac{\phi(a)}{a} \quad \text{and thus} \quad \frac{|x|}{a} \leq \frac{\phi(x)}{\phi(a)}. \quad (5.2.2)$$

Theorem 5.2.2: *Let $\{X_{n_1, n_2}\}$ be an array of independent random variables with $EX_{n_1, n_2} = 0$ for every (n_1, n_2) and assume that $\{a_{n_1, n_2}\}$ is an array such that $0 < a_{n_1, n_2} \uparrow \infty$ as $(n_1, n_2) \rightarrow \infty$. The notation $(n_1, n_2) \rightarrow \infty$ means $n_i \rightarrow \infty$*

for $i = 1, 2$ or equivalently, $\min_{1 \leq i \leq 2} n_i \rightarrow \infty$. If ϕ is a function satisfying the conditions of Lemma 5.2.1 and

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{1}{\phi(a_{n_1, n_2})} E(\phi(X_{n_1, n_2})) < \infty$$

then,

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{X_{n_1, n_2}}{a_{n_1, n_2}} \text{ converges a.s..}$$

Proof: Following Chung (1974) we put

$$Y_{n_1, n_2}(\omega) = \begin{cases} X_{n_1, n_2}(\omega) & \text{if } |X_{n_1, n_2}(\omega)| \leq a_{n_1, n_2} \\ 0 & \text{if } |X_{n_1, n_2}(\omega)| > a_{n_1, n_2}. \end{cases}$$

By the truncation of the $\{X_{n_1, n_2}\}$ we have that

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \text{Var}\left(\frac{Y_{n_1, n_2}}{a_{n_1, n_2}}\right) \leq \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} E\left(\frac{Y_{n_1, n_2}^2}{a_{n_1, n_2}^2}\right)$$

$$= \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} E\left(\frac{X_{n_1, n_2}^2}{a_{n_1, n_2}^2} I(|X_{n_1, n_2}| \leq a_{n_1, n_2})\right)$$

$$\leq \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} E\left(\frac{\phi(X_{n_1, n_2})}{\phi(a_{n_1, n_2})} I(|X_{n_1, n_2}| \leq a_{n_1, n_2})\right)$$

$$\begin{aligned}
&= \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{1}{\phi(a_{n_1, n_2})} E(\phi(X_{n_1, n_2}) I(|X_{n_1, n_2}| \leq a_{n_1, n_2})) \\
&\leq \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{1}{\phi(a_{n_1, n_2})} E(\phi(X_{n_1, n_2})) < \infty
\end{aligned}$$

where the second inequality follows from (5.2.1) and the last from the assumption of the theorem.

Therefore,

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \text{Var}\left(\frac{Y_{n_1, n_2}}{a_{n_1, n_2}}\right) < \infty. \quad (5.2.3)$$

In addition,

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{|EY_{n_1, n_2}|}{a_{n_1, n_2}} = \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \left| E\left(\frac{X_{n_1, n_2}}{a_{n_1, n_2}} I(|X_{n_1, n_2}| \leq a_{n_1, n_2})\right) \right|. \quad (5.2.4)$$

Since $EX_{n_1, n_2} = 0$ we can easily conclude that

$$E\left(\frac{X_{n_1, n_2}}{a_{n_1, n_2}} I(|X_{n_1, n_2}| \leq a_{n_1, n_2})\right) = -E\left(\frac{X_{n_1, n_2}}{a_{n_1, n_2}} I(|X_{n_1, n_2}| > a_{n_1, n_2})\right). \quad (5.2.5)$$

Therefore the right hand side of equality (5.2.4) is equal to

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \left| E\left(\frac{X_{n_1, n_2}}{a_{n_1, n_2}} I(|X_{n_1, n_2}| > a_{n_1, n_2})\right) \right|$$

Inequality (5.2.6) follows from (5.2.2). Therefore we have that,

$$\leq \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} E\left(\frac{|X_{n_1, n_2}|}{a_{n_1, n_2}} I(|X_{n_1, n_2}| > a_{n_1, n_2})\right) \quad (5.2.4)$$

It is known that for a sequence of independent random variables, the three-series theorem is sufficient for convergence of $\sum_{n_1, n_2} X_{n_1, n_2}$. As it is stated in Gnani (1973), the three-series theorem is also sufficient for independent multi-dimensional

indexed random variables. The three-series theorem (5.2.3), (6.2.5) and (6.2.6) are sufficient and thus the sum $\sum_{n_1, n_2} X_{n_1, n_2}$ converges almost surely.

where the second inequality follows from (5.2.2) and thus from the assumption

Corollary 5.2.3: Let (X_{n_1, n_2}) be an array of independent random variables with

$$EX_{n_1, n_2} = 0 \text{ for every } (n_1, n_2) \text{ and } \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{|EY_{n_1, n_2}|}{a_{n_1, n_2}} < \infty. \quad (5.2.5)$$

where $0 < p < 1$. Assuming that

Now,

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} P\{X_{n_1, n_2} \neq Y_{n_1, n_2}\} < \infty$$

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} P\{X_{n_1, n_2} \neq Y_{n_1, n_2}\} = \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} P\{|X_{n_1, n_2}| > a_{n_1, n_2}\}$$

we have that

$$= \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} E[I\{|X_{n_1, n_2}| > a_{n_1, n_2}\}]$$

$$\leq \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} E\left(\frac{\phi(X_{n_1, n_2})}{\phi(a_{n_1, n_2})} I(|X_{n_1, n_2}| > a_{n_1, n_2})\right) \quad (5.2.6)$$

Proof: First we show that the assumptions of Lemma 5.2.1 and Theorem 5.2.2

concerning the function $\phi(x)$ and the array (a_{n_1, n_2}) are satisfied.

$$\leq \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} E\frac{\phi(X_{n_1, n_2})}{\phi(a_{n_1, n_2})} < \infty.$$

For

Inequality (5.2.6) follows from (5.2.2). Therefore we have that,

while for

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} P\{X_{n_1, n_2} \neq Y_{n_1, n_2}\} < \infty. \quad (5.2.7)$$

It is known that for a sequence of independent random variables, the three-series theorem is sufficient for the a.s. convergence of $\sum_n X_n$. As it is stated in Smythe (1973), the three-series theorem is also sufficient for independent multidimensionally indexed random variables. Therefore, inequalities (5.2.3), (5.2.5) and (5.2.7) are sufficient and thus the sum $\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{X_{n_1, n_2}}{a_{n_1, n_2}}$ converges almost surely.

Corollary 5.2.3: *Let $\{X_{n_1, n_2}\}$ be an array of independent random variables with $EX_{n_1, n_2} = 0$ for every (n_1, n_2) and assume that $a_{n_1, n_2} = n_1 n_2$. Let $\phi(x) = |x|^{1+p}$ where $0 < p < 1$. Assuming that*

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{E|X_{n_1, n_2}|^{1+p}}{(n_1 n_2)^{1+p}} < \infty$$

we have that,

$$\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} X_{i, j} \rightarrow 0 \text{ a.s.}$$

Proof: First we show that the assumptions of Lemma 5.2.1 and Theorem 5.2.2 concerning the function $\phi(x) = |x|^{1+p}$ and the array $\{a_{n_1, n_2}\}$ are satisfied.

For

$$0 < p < 1 \text{ and } |x| \leq n_1 n_2 \quad \frac{x^2}{(n_1 n_2)^2} \leq \frac{|x|^{1+p}}{(n_1 n_2)^{1+p}}$$

while for

$$|x| \geq n_1 n_2 \quad \frac{|x|}{(n_1 n_2)} \leq \frac{|x|^{1+p}}{(n_1 n_2)^{1+p}}.$$

and
Since,

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{E|X_{n_1, n_2}|^{1+p}}{(n_1 n_2)^{1+p}} < \infty$$

for
we have by Theorem 5.2.2 that

Then,

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{X_{n_1, n_2}}{n_1 n_2} \text{ converges a.s..}$$

Using the analog of Kronecker's Lemma for multidimensionally indexed random variables, which can be found in Móricz (1981),

Proof: We shall only give the case $r = 2$. The extension to a higher dimension is similar with more complicated. By Corollary

$$\frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} X_{i,j} \rightarrow 0 \text{ a.s.}$$

and the proof is complete.

The extension of Corollary 5.2.3 to the case of r -dimensionally indexed ρ -radius dependent random variables can be easily obtained.

Corollary 5.2.4: Let $\mathbf{n} \in \mathbb{N}^2$ and let $\{X_i, i \leq \mathbf{n}\}$ be an array of ρ -radius dependent random variables such that $EX_{n_1, n_2} = 0$ and $E|X_{n_1, n_2}|^{1+p} < \infty$ for every $n_1 = 1, 2, \dots$ and $n_2 = 1, 2, \dots$.

Assume that

$$T_1 = \frac{1}{\beta} \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \frac{E|S_{(k+\lambda_{i_1}), (k+\lambda_{i_2})}^{(T)}|^{1+p}}{(i_1 i_2)^{1+p}} < \infty$$

and

where β denotes the number of lattice points which belong to a square with length

ν i.e., β is equal to $(2\nu+1)^2$, $h(\rho^*, \nu) = (\rho^*)^\nu$, $d_i = d_i(n_i)$ and

$$\sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \frac{E|\Lambda_{i_1, i_2}^{(\zeta)}|^{1+p}}{(i_1 i_2)^{1+p}} < \infty$$

for

Now,

$$\zeta = 1, 2, 3 \text{ and } 0 < p < 1.$$

Then,

$$n_1 n_2 = \beta D_1 D_2 + h(\rho^*, \nu)(D_1 - 1)D_2 +$$

$$h(\rho^*, \nu)D_1(D_2 - 1) + g(\rho^*)(D_1 - 1)(D_2 - 1)$$

whereas

$$\frac{1}{n_1 n_2} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2} \rightarrow 0 \text{ a.s.}$$

Proof: We shall only give the proof for the case $r = 2$. The extension to a higher dimension is similar although notation becomes quite complicated. By Corollary 5.2.3 and since the following four terms are sums of independent random variables, as $\min_{1 \leq i \leq 2} n_i \rightarrow \infty$ $i = 1, 2$ we have that

$$\bar{T}_1 = \frac{1}{\beta D_1 D_2} \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda_{i_1}), (k+\lambda_{i_2})}^{(T)} \rightarrow 0 \text{ a.s}$$

$$\bar{T}_2 = \frac{1}{h(\rho^*, \nu)(D_1 - 1)D_2} \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} \Lambda_{i_1, i_2}^{(1)} \rightarrow 0 \text{ a.s}$$

$$\bar{T}_3 = \frac{1}{h(\rho^*, \nu)D_1(D_2 - 1)} \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(2)} \rightarrow 0 \text{ a.s}$$

Clearly,

$$\bar{T}_4 = \frac{1}{g(\rho^*)(D_1 - 1)(D_2 - 1)} \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(3)} \rightarrow 0 \text{ a.s.}$$

where β denotes the number of lattice points which belong to a square with length ν i.e., β is equal to $(2\nu + 1)^2$, $h(\rho^*, \nu) = (2\nu + 1)\rho^*$, $g(\rho^*) = (\rho^*)^2$, $d_i = d_i(n_i)$ and $D_i = D_i(n_i)$ for $i = 1, 2$.

Now,

$$n_1 n_2 = \beta D_1 D_2 + h(\rho^*, \nu)(D_1 - 1)D_2 + h(\rho^*, \nu)D_1(D_2 - 1) + g(\rho^*)(D_1 - 1)(D_2 - 1)$$

whereas

$$\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2} = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)} + \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} \Lambda_{i_1, i_2}^{(1)} + \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(2)} + \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(3)}$$

Clearly, any linear combination of continuous functions is continuous. We keep ρ^*

and ν fixed and we write $\frac{1}{n_1 n_2} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2}$ as

$$\frac{1}{n_1 n_2} \left\{ \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} S_{(k+\lambda i_1), (k+\lambda i_2)}^{(T)} + \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2} \Lambda_{i_1, i_2}^{(1)} + \sum_{i_1=1}^{D_1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(2)} + \sum_{i_1=1}^{D_1-1} \sum_{i_2=1}^{D_2-1} \Lambda_{i_1, i_2}^{(3)} \right\} =$$

$$\frac{1}{n_1 n_2} \left\{ \beta D_1 D_2 \bar{T}_1 + h(\rho^*, \nu)(D_1 - 1)D_2 \bar{T}_2 + h(\rho^*, \nu)D_1(D_2 - 1) \bar{T}_3 + g(\rho^*)(D_1 - 1)(D_2 - 1) \bar{T}_4 \right\}$$

Clearly,

$$\frac{\beta D_1 D_2}{n_1 n_2} \rightarrow c_1, \quad \bar{T}_1 \rightarrow 0 \text{ a.s.},$$

converges.

2) Let the array $\{X_{i_1, i_2}\}$ be an array of independent multidimensionally indexed random variables. Then for $p \geq 1$, Corollary 5.2.4 coincides with Theorem 5.2.3 for $p = 1$ and $r = 1$ by the central Kolmogorov's strong law of large numbers.

$$\frac{h(\rho^*, \nu)(D_1 - 1)D_2}{n_1 n_2} \rightarrow c_2, \quad \bar{T}_2 \rightarrow 0 \text{ a.s.},$$

$$\frac{h(\rho^*, \nu)D_1(D_2 - 1)}{n_1 n_2} \rightarrow c_3, \quad \bar{T}_3 \rightarrow 0 \text{ a.s.},$$

$$\frac{g(\rho^*)(D_1 - 1)(D_2 - 1)}{n_1 n_2} \rightarrow c_4, \quad \bar{T}_4 \rightarrow 0 \text{ a.s.}$$

and since $c_i, i = 1, 2, 3, 4$ are fixed constants

$$\frac{1}{n_1 n_2} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} X_{i_1, i_2} \rightarrow 0 \text{ a.s.}$$

and the proof is complete.

Remarks

1) For the case of independent and identically distributed random variables, the following assumption in Corollary 5.2.3 and consequently in Corollary 5.2.4

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{E|X_{n_1, n_2}|^{1+p}}{(n_1 n_2)^{1+p}} < \infty$$

is valid since

$$\sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} \frac{E|X_{n_1, n_2}|^{1+p}}{(n_1 n_2)^{1+p}} = E|X_{1,1}|^{1+p} \sum_{n_1=1}^{\infty} \frac{1}{n_1^{1+p}} \sum_{n_2=1}^{\infty} \frac{1}{n_2^{1+p}}$$

converges.

2) Let the array $\{X_i, i \leq n\}$ in Corollary 5.2.4 be an array of independent multidimensionally indexed random variables. Then for $p = 1$ Corollary 5.2.4 coincides with Theorem 2.2.4, where for $p = 1$ and $r = 1$ we have the classical Kolmogorov's strong law of large numbers.

6.1 Introduction

The behavior of sums $(S_n, n \geq 1)$ of independent random variables $(X_i, i = 1, 2, \dots)$ is of great interest in probability theory. In particular, in the case where the random variables $(X_i, i = 1, 2, \dots)$ are independent and identically distributed, there are many interesting results. In this chapter two Kolmogorov inequalities for the sample average of independent (but not necessarily identically distributed) Bernoulli random variables are presented.

For a sequence of independent and identically distributed Bernoulli random variables X_1, X_2, \dots with $E(X_i) = p$, Kolmogorov (1933) provided the following inequality:

$$P\left(\sup_{k \leq n} |\bar{X}_k - p| > \epsilon\right) \leq 2e^{-2n\epsilon^2(1-p)}$$

where

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i, \quad k \geq 1.$$

Improvements, extensions and many related results can be found in Hoeffding (1963).

Chapter 6

Probability inequalities

6.1 Introduction

The behavior of sums $\{S_n, n \geq 1\}$ of independent random variables $\{X_i, i = 1, 2, \dots\}$ is of great interest in probability theory. In particular, in the case where the random variables $\{X_i, i = 1, 2, \dots\}$ are independent and identically distributed, there are many interesting results. In this chapter two Kolmogorov inequalities for the sample average of independent (but not necessarily identically distributed) Bernoulli random variables are presented.

For a sequence of independent and identically distributed Bernoulli random variables X_1, X_2, \dots with $E(X_1) = p$, Kolmogorov (1963) provided the following inequality:

$$P\left\{\sup_{k \geq n} |\bar{X}_k - p| > \epsilon\right\} \leq 2e^{-2n\epsilon^2(1-\epsilon)}$$

where

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i, \quad \epsilon > 0.$$

Improvements, extensions and many related results can be found in Hoeffding (1963), Kambo and Kotz (1966), Young, Seaman and Marco (1987), Turner, Young and Seaman (1992), Young, Turner and Seaman (1988), Christofides (1991), Christofides (1994) and Banjevic (1985). In this final chapter, we provide two Kolmogorov inequalities for the case of independent but not necessarily identically distributed Bernoulli random variables.

6.2 Deterministic inequalities and other results

We will make use of the following results. Let $\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i$ where $p_1, \dots, p_n \in R^+$.

Lemma 6.2.1: For $t > 0$

$$\prod_{i=1}^n (p_i e^t + 1 - p_i) \leq (\bar{p} e^t + 1 - \bar{p})^n.$$

Proof: The proof of the lemma is quite trivial. From the arithmetic-geometric mean inequality, we have that

$$\prod_{i=1}^n (p_i e^t + 1 - p_i) \leq \left\{ \sum_{i=1}^n \frac{1}{n} (p_i e^t + 1 - p_i) \right\}^n$$

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} = (\bar{p}e^t + 1 - \bar{p})^n.$$

The following result is due to Christofides (1994).

Lemma 6.2.2: Let $\epsilon < \frac{1}{2}$ and

$$g(\bar{p}, \epsilon) = (1 - \bar{p} - \epsilon) \ln\left(\frac{1 - \bar{p} - \epsilon}{1 - \bar{p}}\right) + (\bar{p} + \epsilon) \ln\left(\frac{\bar{p} + \epsilon}{\bar{p}}\right).$$

Then, for $\bar{p} + \epsilon < \frac{1}{2}$ or $\frac{1}{2} + \frac{1}{3}\epsilon \leq \bar{p} \leq 1$

$$g(\bar{p}, \epsilon) \geq -\frac{1}{2} \ln(1 - 4\epsilon^2).$$

Lemma 6.2.3: Let $x = 2(\bar{p} + \epsilon) - 1$ and $y = 1 - 2\bar{p}$. Then,

$$\sum_{r=1}^{\infty} \frac{1}{2r(2r-1)} \{x^{2r} + (2r-1)y^{2r} + 2rxy^{2r-1}\} = g(\bar{p}, \epsilon).$$

where $g(\bar{p}, \epsilon)$ is defined in Lemma 6.2.2.

Proof: We have $\bar{p} + \epsilon = \frac{x+1}{2}$ and $\bar{p} = \frac{1-y}{2}$. Then,

$$g(\bar{p}, \epsilon) = \left(\frac{1-x}{2}\right) \ln\left(\frac{1-x}{1+y}\right) + \left(\frac{1+x}{2}\right) \ln\left(\frac{1+x}{1-y}\right).$$

Using the Taylor series expansions

$$\ln(1+x) = \sum_{r=1}^{\infty} (-1)^{r-1} \frac{x^r}{r} \quad \text{and} \quad \ln(1-x) = -\sum_{r=1}^{\infty} \frac{x^r}{r}$$

we have that

Proof: For the derivative of $H(y)$ we have

$$g(\bar{p}, \epsilon) = \frac{(1-x)}{2} \left\{ -\sum_{r=1}^{\infty} \frac{x^r}{r} - \sum_{r=1}^{\infty} (-1)^{r-1} \frac{y^r}{r} \right\} + \frac{(1+x)}{2} \left\{ \sum_{r=1}^{\infty} (-1)^{r-1} \frac{x^r}{r} + \sum_{r=1}^{\infty} \frac{y^r}{r} \right\}.$$

After algebraic manipulations we arrive at the desired result.

Lemma 6.2.4: *Let ν be a positive integer and $x > 1$. Then*

$$2^{2\nu-2}(x^{2\nu-1} + 1) - (x+1)^{2\nu-1} > 0.$$

Proof: Let

$$F(x) = 2^{2\nu-2}(x^{2\nu-1} + 1) - (x+1)^{2\nu-1}.$$

Then,

$$\begin{aligned} F'(x) &= 2^{2\nu-2}(2\nu-1)x^{2\nu-2} - (2\nu-1)(x+1)^{2\nu-2} \\ &= (2\nu-1)\{2^{2\nu-2}x^{2\nu-2} - (x+1)^{2\nu-2}\} \\ &= (2\nu-1)\{(2x)^{2\nu-2} - (x+1)^{2\nu-2}\} > 0 \quad \text{since } x > 1. \end{aligned}$$

Thus, $F(x)$ is an increasing function and $F(x) > F(1) = 0$.

Lemma 6.2.5: *Let $y \geq 1$ and $\nu = 1, 2, \dots$. Then,*

Taking $c = x/y$ we have

$$H(y) = y^{2\nu} + 2\nu - 1 + 2\nu y - \frac{4}{2^{2\nu}}(y+1)^{2\nu} \geq 0.$$

Proof: For the derivative of $H(y)$ we have

$$H'(y) = 2\nu y^{2\nu-1} + 2\nu - \frac{4}{2^{2\nu}} 2\nu (y+1)^{2\nu-1}$$

$$= 2\nu \frac{4}{2^{2\nu}} \left\{ \frac{2^{2\nu}}{4} y^{2\nu-1} + \frac{2^{2\nu}}{4} - (y+1)^{2\nu-1} \right\}$$

$$= 2^{3-2\nu} \nu \{ 2^{2\nu-2} y^{2\nu-1} + 2^{2\nu-2} - (y+1)^{2\nu-1} \}$$

$$= 2^{3-2\nu} \nu \{ 2^{2\nu-2} (y^{2\nu-1} + 1) - (y+1)^{2\nu-1} \}.$$

By Lemma 6.2.4, $H'(y) > 0$ implying that H is increasing and therefore $H(x) \geq H(1) = 0$.

Lemma 6.2.6: Let x and y be as in Lemma 6.2.3 and $r = 1, 2, \dots$. Then

$$x^{2r} + (2r-1)y^{2r} + 2rxy^{2r-1} \geq 4\left(\frac{x+y}{2}\right)^{2r}.$$

Proof: By lemma 6.2.5 for $c \geq 1$

$$c^{2r} + 2r - 1 + 2rc - \frac{4}{2^{2r}}(c+1)^{2r} \geq 0.$$

Taking $c = x/y$ we have

$$\frac{x^{2r}}{y^{2r}} + 2r - 1 + 2r \frac{x}{y} \geq \frac{4}{2^{2r}} \left(\frac{x}{y} + 1\right)^{2r}$$

and therefore

$$x^{2r} + (2r - 1)y^{2r} + 2rxy^{2r-1} \geq 4\left(\frac{x+y}{2}\right)^{2r}.$$

6.3 Kolmogorov type inequalities

The following theorems provide exponential bounds for \bar{Y} , the sample average of independent Bernoulli random variables.

Theorem 6.3.1: Let Y_1, Y_2, \dots, Y_n be a sequence of independent Bernoulli random variables with $E(Y_i) = p_i$, $i = 1, \dots, n$ and $\epsilon < \frac{1}{2}$. Then for $\bar{p} + \epsilon < \frac{1}{2}$ or $\frac{1}{2} + \frac{1}{3}\epsilon \leq \bar{p} \leq 1$,

$$P\{\bar{Y} - \bar{p} > \epsilon\} \leq (1 - 4\epsilon^2)^{\frac{n}{2}}$$

where $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ and $\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i$.

Proof: Let $s > 0$. Then,

$$P\{\bar{Y} - \bar{p} > \epsilon\} = P\{s(\bar{Y} - \bar{p} - \epsilon) > 0\}$$

$$\leq E(e^{s(Y-\bar{p}-\epsilon)})$$

and by Lemma 6.2.2

$$= e^{-s(\bar{p}+\epsilon)} E\{e^{sY}\} = e^{-s(\bar{p}+\epsilon)} E\{e^{s\frac{1}{n}\sum_{i=1}^n Y_i}\}$$

$$= e^{-s(\bar{p}+\epsilon)} \prod_{i=1}^n E(e^{\frac{s}{n}Y_i}) \quad (6.3.1)$$

The following theorem gives an exponential bound under different conditions on \bar{p} and ϵ .

$$= e^{-s(\bar{p}+\epsilon)} \prod_{i=1}^n (p_i e^{\frac{s}{n}} + 1 - p_i)$$

Theorem 6.3.2: Let Y_1, Y_2, \dots, Y_n be a sequence of independent Bernoulli random variables, with $B(Y_i) \leq e^{-s(\bar{p}+\epsilon)} (\bar{p} e^{\frac{s}{n}} + 1 - \bar{p})^n = e^{-f(s)}$ for $\bar{p} < 1$ and $Y_i < 1$,

where $f(s) = s(\bar{p} + \epsilon) - n \ln(\bar{p} e^{\frac{s}{n}} + 1 - \bar{p})$. Observe that the last inequality follows from Lemma 6.2.1.

The function f is maximized at $s^{max} = n \ln\left\{\frac{(1-\bar{p})(\bar{p}+\epsilon)}{\bar{p}(1-\bar{p}-\epsilon)}\right\}$ and

Proof: From the proof of Theorem 6.3.1

$$f(s^{max}) = n(\bar{p} + \epsilon) \ln\left\{\frac{(\bar{p} + \epsilon)(1 - \bar{p})}{\bar{p}(1 - \bar{p} - \epsilon)}\right\} - n \ln\left\{\frac{1 - \bar{p}}{1 - \bar{p} - \epsilon}\right\} = ng(\bar{p}, \epsilon)$$

where

$$g(\bar{p}, \epsilon) = (\bar{p} + \epsilon) \ln\left\{\frac{(\bar{p} + \epsilon)(1 - \bar{p})}{\bar{p}(1 - \bar{p} - \epsilon)}\right\} - \ln\left\{\frac{1 - \bar{p}}{1 - \bar{p} - \epsilon}\right\}.$$

Thus

By Lemma 6.2.3

$$g(\bar{p}, \epsilon) = \sum_{r=0}^{\infty} \frac{P\{\bar{Y} - \bar{p} > \epsilon\}}{2^r(2^r - 1)} \leq e^{-ng(\bar{p}, \epsilon)}$$

and by Lemma 6.2.2 using Lemma 6.2.6 we have

$$P\{\bar{Y} - \bar{p} > \epsilon\} \leq (1 - 4\epsilon^2)^{\frac{n}{2}}.$$

The following theorem gives an exponential bound under different conditions on \bar{p} and ϵ .

Theorem 6.3.2: Let Y_1, Y_2, \dots, Y_n be a sequence of independent Bernoulli random variables, with $E(Y_i) = p_i$, $i = 1, \dots, n$. Then for $\bar{p} + \epsilon > \frac{1}{2}$ or $\bar{p} < \frac{1}{2}$ and $\forall \epsilon < 1$,

$$P\{\bar{Y} - \bar{p} > \epsilon\} \leq e^{-n(2\epsilon^2 + \frac{1}{3}\epsilon^4 e^{\frac{\epsilon^2}{4}})}$$

where $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ and $\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i$.

Proof: From the proof of Theorem 6.3.1

$$P\{\bar{Y} - \bar{p} > \epsilon\} \leq e^{-ng(\bar{p}, \epsilon)}$$

where

$$g(\bar{p}, \epsilon) = (1 - \bar{p} - \epsilon) \ln\left(\frac{1 - \bar{p} - \epsilon}{1 - \bar{p}}\right) + (\bar{p} + \epsilon) \ln\left(\frac{\bar{p} + \epsilon}{\bar{p}}\right).$$

By Lemma 6.2.3

and the proof of the theorem is complete.

Remark $g(\bar{p}, \epsilon) = \sum_{r=1}^{\infty} \frac{1}{2r(2r-1)} \{x^{2r} + (2r-1)y^{2r} + 2rxy^{2r-1}\}.$

Since $x + y = 2\epsilon$ and using Lemma 6.2.6 we have

$$g(\bar{p}, \epsilon) \geq \sum_{r=1}^{\infty} \frac{1}{2r(2r-1)} 4\left(\frac{2\epsilon}{2}\right)^{2r}$$

$$P\{\sup_{1 \leq k \leq n} (\bar{Y}_k - \bar{y}_k) > \epsilon\} \text{ where } \bar{Y}_k = \sum_{i=1}^k Y_i \text{ and } \bar{y}_k = \frac{1}{k} \sum_{i=1}^k y_i$$

$$= \sum_{r=1}^{\infty} \frac{2\epsilon^{2r}}{r(2r-1)}$$

This is possible because by applying Lemma 1 of Turner, Young and Soutman (1994)

we arrive at (6.3.1) having the required inequality $P\{\sup_{1 \leq k \leq n} (\bar{Y}_k - \bar{y}_k) > \epsilon\}$ as our

left hand side. Then, the exact same as above we have

$$= 2\epsilon^2 + \epsilon^4 \sum_{r=2}^{\infty} \frac{2(\epsilon)^{2r-4}}{r(2r-1)}$$

2) In view of the previous remark both Theorem 6.3.1 and 6.3.2 provide sharper bounds than that of the main result of Turner, Young and Soutman (1994), under of

$$\geq 2\epsilon^2 + 2\epsilon^4 \sum_{r=2}^{\infty} \frac{(\epsilon^2)^{r-2}}{4^{r-2}(r-2)!6}$$

since $r(2r-1) \leq 4^{r-2}(r-2)!6$ for $r \geq 2$. Then,

3) Theorem 6.3.1 is an extension of Corollary 3.3 of Christofides (1994).

4) It is straightforward that both Theorem 6.3.1 and 6.3.2 can easily be generalized

$$g(\bar{p}, \epsilon) \geq 2\epsilon^2 + 2\epsilon^4 \frac{1}{6} \sum_{r=2}^{\infty} \left(\frac{\epsilon^2}{4}\right)^{r-2} \frac{1}{(r-2)!}$$

to the case of multidimensionally indexed random variables.

$$= 2\epsilon^2 + \frac{1}{3}\epsilon^4 e^{\frac{\epsilon^2}{4}}.$$

Thus,

$$P\{\bar{Y} - \bar{p} > \epsilon\} \leq e^{-n(2\epsilon^2 + \frac{1}{3}\epsilon^4 e^{\frac{\epsilon^2}{4}})}$$

and the proof of the theorem is complete.

Remarks

1) The left hand side of Theorem 6.3.1 and that of Theorem 6.3.2 can in fact be replaced by the stronger version

$$P\{\sup_{k \geq n} (\bar{Y}_k - \bar{p}_k) > \epsilon\} \text{ where } \bar{Y}_k = \sum_{i=1}^k Y_i \text{ and } \bar{p}_k = \frac{1}{k} \sum_{i=1}^k p_i.$$

This is possible because by applying Lemma 1 of Turner, Young and Seaman (1994) we arrive at (6.3.1) having the required quantity $P\{\sup_{k \geq n} (\bar{Y}_k - \bar{p}_k) > \epsilon\}$ as our left hand side. Then, the exact same steps can be followed.

2) In view of the previous remark both Theorem 6.3.1 and 6.3.2 provide sharper bounds than that of the main result of Turner, Young and Seaman (1994), under of course restrictions on \bar{p} and ϵ .

3) Theorem 6.3.1 is an extension of Corollary 3.2 of Christofides (1994).

4) It is straightforward that both Theorems 6.3.1 and 6.3.2 can easily be generalized to the case of multidimensionally indexed random variables.

Applications and future work

The importance of multidimensionally indexed ρ -radius dependent random variables is due to the fact that they can be very applicable. In real life, most of the time we have to deal with spatial random variables which very often as expected are not independent.

Let us take as an example data in meteorology. Rainfall is measured at rainfall stations. The location of each station is defined by height, latitude and longitude. Therefore, we can assume that rainfall in each station is a three-dimensionally indexed random variable, i.e., X_{i_1, i_2, i_3} . Apparently, measurements of rainfall are associated to each other according to the location of the stations. This association can be interpreted as ρ -radius dependence.

A television screen can be thought of as a two-dimensional lattice, with n_1, n_2 , sufficiently large. For each lattice point we can associate a two dimensionally indexed random variable which measures the intensity or the brightness of the picture at the specific point. Clearly, lattice points which are close to each other are expected to have similar intensity or brightness whereas these characteristics for distant points should be independent.

Multidimensionally indexed ρ -radius dependent random variables and tools like the central limit theorem and strong laws, play an important role in the statistical inference of the above cases.

For research purposes, multidimensionally indexed ρ -radius dependent random vari-

ables could serve as the basis for other interesting results. For example, we could extend the above results to the case of U-statistics, that is, we can derive asymptotic results for U-statistics based on multidimensionally indexed ρ -radius dependent random variables.

Furthermore, one might consider the possibility of exploring a more general notion of dependence between the random variables. For example, we could investigate the asymptotic behavior of multidimensionally indexed random variables satisfying a condition which is analogous to a mixing condition in the case of one dimensionally indexed random variables.

Curriculum Vitae

Petroula Mavrikiou was born at Katokopia, Cyprus in 1970. After finishing high school in 1987 she entered the Department of Civil Engineering of the National Technical University of Athens. In 1992 she graduated with an overall grade 7.33. In 1993 she was employed by the University of Cyprus as a Graduate Assistant and at the same time she started working on her Ph.D. under the supervision of Professor Tasos C. Christofides, in the Department of Mathematics and Statistics of the University of Cyprus.

Berry, A.C. (1941): The accuracy of the Gaussian approximation to the sum of independent variables, *Trans. Amer. Math. Soc.* 40, 136.

Blum, J.B. and Brennan, M. (1987): On the weak law of large numbers for dependent random variables, *Israel J. Math.* 54, 193.

Carroll, R. (1970): Une loi locale pour martingales à infinis multiples et ses applications, *Seminaire de Probabilités IV - Lecture Notes in Mathematics*, Vol. 124, (Springer, Berlin), pp. 127.

Carroll, R. and Walsh, J. B. (1975): Stochastic integrals in the plane, *Acta Math.* 147, 109-154.

Cartan, H. (1961): *Elementary theory of analytic functions of one or several complex variables*, Addison-Wesley, Reading, Mass.

Chow, Y.S. and Teicher, H. (1988): *Probability theory. Independence, interchangeability, martingales*, Springer-Verlag, New York.

References

- Banjevic, D. (1985): On a Kolmogorov inequality, *Theory Probab. Appl.* **29**, 391-394.
- Berk, Kenneth N. (1973): A central limit theorem for m -dependent random variables with unbounded m , *Ann. Probab.* **1**, 352-354.
- Bernstein, S. (1927): Sur l' extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes, *Mathematische Annalen* **97**, 1-59.
- Berry, A.C. (1941): The accuracy of the Gaussian approximation to the sum of independent variates, *Trans. Amer. Math. Soc.* **49**, 122-136.
- Blum, J.R. and Brennan, M. (1980): On the strong law of large numbers for dependent random variables, *Israel J. Math.* **37**, 241-245.
- Cairoli, R. (1970): Une inégalité pour martingales a indices multiples et ses applications, *Seminaire de Probabilités IV Lecture Notes in Mathematics. Vol 124*, (Springer, Berlin) p.p. 1-27.
- Cairoli, R. and Walsh, J. B. (1975): Stochastic integrals in the plane, *Acta Math.* **143**, 111-183.
- Cartan, H. (1963): *Elementary theory of analytic functions of one or several complex variables*, Addison-Wesley, Reading, Mass.
- Chow, Y.S. and Teicher, H. (1988): *Probability theory. Independence, interchangeability, martingales*, Springer-Verlag, New York.

Christofides, T.C. (1987): Maximal probability inequalities for multidimensionally indexed semimartingales and convergence theory of U-statistics, Ph.D. dissertation, Dept. Math. Sci., Johns Hopkins University.

Christofides, T.C. (1991): Probability inequalities with exponential bounds for U-statistics, *Statist. Probab. Lett.* **12**, 257-261.

Christofides, T.C.(1992): A strong law of large numbers for U-statistics, *J. Statist. Plann. Inference* **31**, 133-145.

Christofides, T.C. (1994): A Kolmogorov inequality for U-statistics based on Bernoulli kernels, *Statist. Probab. Lett.* **21**, 357-362.

Christofides, T.C. (1997): Rate of convergence in the strong law of large numbers for U-statistics and von Mises statistics, *Statist. Probab. Lett.* **31**, 225-231.

Christofides, T.C. (1997): A Berry-Esséen theorem for U-statistics, *Proceedings of the 10th Panhellenic Conference of the Greek Statistical Institute*, 251-264.

Christofides, T.C. and Serfling, R. (1990): Maximal inequalities for multidimensionally indexed submartingale arrays, *Ann. Probab.* **18**, 630-641.

Christofides, T.C. and Serfling, R. (1998): U-statistics on a lattice of i.i.d. random variables, *Statist. Probab. Lett.* **40**, 293-303.

Chung, K.L. (1947): Note on some strong law of large numbers, *Amer. J. Math.* **69**, 189-192.

Chung, K.L. (1974): *A Course in Probability Theory*, 2nd ed., Academic Press, New York.

Cramèr, H. (1945): *Mathematical Methods of Statistics*, Princeton Univ. Press.

Dunford, N. (1951): An individual ergodic theorem for non-commutative transformations, *Acta Sci. Math. (Szeged)* **14**, 1-4.

Esséen, C. G. (1942): On the Liapounov limit of error in the theory of probability, *Arkiv Mat. Astr. och Fysik* **28A(2)**, 1-19.

Etemadi, N. (1981): An elementary proof of the strong law of large numbers, *Z. Wahrsch. Verw. Gebiete* **55**, 119-122.

Fréchet, M. (1946): *Proc. Math. Phys. Soc. Egypt* **3(2)**, 13-20.

Fréchet, M. (1948): *Proc. Math. Phys. Soc. Egypt* **3**, 73-74.

Gut, A. (1976): Convergence of reverse martingales with multidimensional indices, *Duke Math. J.* **43**, 269-275.

Gut, A. (1978): Marcinkiewicz laws and convergence rates in the law of large numbers for random variables with multidimensional indices, *Ann. Probab.* **6**, 469-482.

Heinrich, L. (1987): Asymptotic expansions in the central limit theorem for a special class of m -dependent random fields I, *Math. Nachr.* **134**, 83-106.

Hoeffding, W. (1948): A class of statistics with asymptotically normal distribution, *Ann. Math. Statist.* **19**, 293-325.

Hoeffding, W. (1963): Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58**, 13-30.

Hoeffding, W. and Robbins, H. (1948): The central limit theorem for dependent random variables, *Duke Math. J.* **15**, 773-780.

Kambo, N.S. and Kotz, S. (1966): On exponential bounds for binomial probabilities, *Ann. Inst. Statist. Math.* **18**, 277-287.

Klesov, O.I. (1981): The Hájek-Rényi inequality for random fields and the strong law of large numbers, *Theor. Prob. and Stat.* **22**, 63-71.

Kolmogorov, A. (1933): *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer-Verlag, Berlin.

Kolmogorov, A.N. (1963): On tables of random numbers, *Sankhya*, Ser. A **25**, 369-376.

Kuchkarov, Ya. Kh. (1990): The strong law of large numbers for dependent random variables, *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk* No. 2, 21-26.

Liapounov, A. M. (1901): Nouvelle forme du théorème sur la limite de probabilité, *Mém. Acad. Imp. Sci. St.-Petersbourg* **12**(5), 1-24.

Loève, M. (1977): *Probability Theory I*, 4th ed., Springer-Verlag, New York.

Móricz, F. (1981): The Kronecker lemmas for multiple series and some applications, *Acta Math. Acad. Sci. Hung.* **37**, 39-50.

Petrov, V.V. (1974): *Sums of independent random variables*, Springer-Verlag, New York.

Petrov, V.V. (1994): *Limit theorems of probability theory. Sequences of independent random variables*, Oxford Science Publications.

Sharakhmetov, Sh. (1995): The strong law of large numbers for dependent random variables, *Teor. Mat. Stat.*, No. 53 168-174.

Shergin, V.V. (1983): The global form of the central limit theorem for m -dependent random variables, *Teor. Veroyatnost. i Mat. Statist.* **29**, 122-128.

Shorack, G. R. and Smythe, R.T. (1976): Inequalities for $\max |S_k|/b_k$ where $k \in \mathbb{N}^r$, *Proc. Amer. Math. Soc.* **54**, 331-336.

Smythe, R.T. (1973): Strong laws of large numbers for r -dimensional arrays of random variables, *Ann. Probab.* **1**, 164-170.

Smythe, R.T. (1974): Sums of independent random variables on partially ordered sets, *Ann. Probab.* **2**, 906-917.

Stout, W.F. (1974): *Almost Sure Convergence*, Academic Press, New York.

Turner, D.W., Young, D.M. and Seaman, J.W. (1992): Improved Kolmogorov inequalities for the binomial distribution, *Statist. Probab. Lett.* **13**, 223-227.

Turner, D.W., Young, D.M. and Seaman, J.W. (1994): A Kolmogorov inequality for the sum of independent Bernoulli random variables with unequal means, *Statist. Probab. Lett.* **23**, 243-245.

Utev, S.A. (1989): A central limit theorem for dependent random variables, *Probability theory and mathematical statistics*, Vol. II, 519-528.

Wichura, M.J. (1969): Inequalities with applications to the weak convergence of random processes with multidimensional time parameters, *Ann. Math. Statist.* **40**, 681-687.

Young, D.M., Seaman, J.W. and Marco, V.R. (1987): A note on a Kolmogorov inequality, *Statist. Probab. Lett.* **5**, 217-218.

Young, D.M., Turner, D.W. and Seaman, J.W. (1988): A note on a Kolmogorov inequality for the binomial distribution, *Theory Probab. Appl.* **33**, 747-749.