Lecture Notes in Statistics

Edited by S. Fienberg, J. Gani, J. Kiefer, and K. Krickeberg

2

Mathematical Statistics and Probability Theory

Proceedings, Sixth International Conference, Wisła (Poland), 1978

Edited by W. Klonecki, A. Kozek, and J. Rosiński

Springer-Verlag
New York Heidelberg Berlin

Lecture Notes in Statistics

Vol. 1: R. A. Fisher: An Appreciation. Edited by S. E. Fienberg and D. V. Hinkley. xi, 208 pages, 1980.

Vol. 2: Mathematical Statistics and Probability Theory. Proceedings 1978. Edited by W. Klonecki, A. Kozek, and J. Rosiński. xxiv, 373, 1980.

Springer Series in Statistics

L. A. Goodman and W. H. Kruskal, Measures of Association for Cross Classifications. x, 146 pages, 1979.

J. O. Berger, Statistical Decision Theory: Foundations, Concepts, and Methods. xiv, 420 pages, 1980.

Lecture Notes in Statistics

Edited by S. Fienberg, J. Gani, J. Kiefer, and K. Krickeberg

2

Mathematical Statistics and Probability Theory

Proceedings, Sixth International Conference, Wisła (Poland), 1978

Edited by W. Klonecki, A. Kozek, and J. Rosiński

Springer-Verlag
New York Heidelberg Berlin

Editors
Dr. Witold Klonecki, Dr. Andrzej Kozek, Dr. Jan Rosiński
Mathematical Institute of the Polish Academy of Science
Kopernika 18, 51-617 Wrocław, Poland

AMS Subject Classifications: 62-XX

Library of Congress Cataloging in Publication Data

Main entry under title:

Mathematical statistics and probability theory.

(Lecture notes in statistics; 2)
Includes bibliographies
1. Mathematical statistics—Congresses.
2. Probabilities—Congresses. I. Klonecki, Witold.
II. Kozek, A. III. Rosiński, Jan. IV. Series.
QA276.A1M3 519.5 80-13322

ISBN 978-0-387-90493-1 ISBN 978-1-4615-7397-5 (eBook)
DOI 10.1007/978-1-4615-7397-5

9 8 7 6 5 4 3 2 1

ON LIMIT THEOREMS FOR SUMS OF DEPENDENT HILBERT SPACE VALUED RANDOM VARIABLES

by

Adam Jakubowski

Nicolaus Copernicus University, Toruń

1. Introduction

Let $\{X_{nk}\}$, $k = 1,2,\ldots,k_n$; $n = 1,2,\ldots,$ be an array of random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. If $\{X_{nk}\}$ are row-wise independent, then there exists a quite satisfactory theory of the weak convergence of sums $S_n' = \sum_{k=1}^{k_n} X_{nk}$. One of the most reasonable trends in the analogous theory for dependent random variables is initiated by papers of Brown [2] and Dvoretzky [4], [5].

This new successful approach (see [3], [6] for generalizations of [2], [5]) can be described very briefly: To obtain limit theorems for dependent random variables one has to replace usual expectations in classical theorems for independent random variables by conditional expectations with respect to a suitably chosen family of σ-subfields of $\mathcal{F}$ and the convergence of numbers by the convergence in probability. This procedure can be observed most explicitly in Theorem C (Section 3) - the Hilbert space version of the Brown's Theorem.

The present paper contains generalizations of theorems of such form for the case when X_{nk} are random variables taking values in real separable Hilbert space. Their proofs are new even in the finite dimensional case and are based on the technics of the regular condi-

adjakubo@mat.umk.pl

tional distributions. Such an approach gives possibility for the use of the Varadhan's theory for weak convergence of convolutions in Hilbert space (see [8], also [7], Chapter VI).

Basic Theorems A and B, which can be treated as modified accompanying laws, are contained in Section 2. In particular, Theorem A is a sufficient tool for quick proofs in the finite dimensional case. In Section 3 it is shown, how to obtain from Theorem B the required results: the Brown's Theorem (Theorem C) and the Hilbert space generalization of theorem of Kłopotowski (Theorem D). More detailed proofs of Theorems B and D will be published elsewhere.

2. Main Theorems

Let H be a real separable Hilbert space with the inner product $(\cdot,\cdot)$ and let $\mathcal{B}_H$ be the σ-field of Borel subsets of H. All H-valued random variables considered in this paper are defined on fixed probability space $(\Omega, \mathcal{F}, P)$. Let $\{X_{nk}\}$, $k = 1,2,\ldots,k_n; n = 1,2,\ldots$, be an array of random variables and $\{\mathcal{F}_{nk}\}$, $k = 0,1,\ldots k_n$; $n = 1,2,\ldots$, be an array of row-wise increasing σ-subfields of $\mathcal{F}$ (i.e. $\mathcal{F}_{nk} \subset \mathcal{F}_{n,k+1}$ for n fixed and $k = 0,1,\ldots,k_n-1$). The array $\{X_{nk}\}$ is said to be adapted to $\{\mathcal{F}_{nk}\}$ if every X_{nk} is $\mathcal{F}_{nk}$ - measurable.

For $\{X_{nk}\}$ adapted to $\{\mathcal{F}_{nk}\}$ we can define an array $\{\mu_{nk}\}$ of regular random measures by choosing for every (n,k) a regular version of the conditional distribution of X_{nk} given $\mathcal{F}_{n,k-1}$. In other words, for every n, k

$$\mu_{nk} : \mathcal{B}_H \times \Omega \longrightarrow [0,1]$$

is a function such that for every $\omega \in \Omega$ $\mu_{nk}(\cdot,\omega)$ is a probability measure on $\mathcal{B}_H$ and for every $A \in \mathcal{B}_H$ $\mu_{nk}(A,\cdot)$ is a version of $P(X_{nk} \in A \mid \mathcal{F}_{n,k-1})$ (hence $\mathcal{F}_{n,k-1}$-measurable). For some properties of the regular conditional distribution and for the proof of its existence see [1], Chapter 4.

In the sequel we will deal with the arrays $\{X_{nk}\}$ adapted to $\{\mathcal{F}_{nk}\}$ and the regular random measures μ_{nk} defined above and the definitions will not be repeated in the theorems.

Now we can formulate Theorem A, which is sufficient for applications in the finite dimensional case.

Theorem A. Let μ be a distribution on $\mathcal{B}_H$ with the non-vanishing characteristic functional:

$$\forall y \in H \qquad \hat{\mu}(y) := \int e^{i(y,x)}\mu(dx) \neq 0.$$

If for almost all $\omega \in \Omega$, the convolutions

$$\mu_n(\cdot,\omega) := \mu_{n1}(\cdot,\omega) * \mu_{n2}(\cdot,\omega) * \dots \mu_{nk_n}(\cdot,\omega)$$

are weakly convergent to μ ($\mu_n \Rightarrow \mu$ a.s.) then the characteristic functionals of $S_n = \sum_{k=1}^{k_n} X_{nk}$ are pointwise convergent to $\hat{\mu}$:

$$\forall y \in H \qquad E\, e^{i(y,S_n)} \longrightarrow \hat{\mu}(y).$$

P r o o f: Denoting $\hat{\mu}_{nk}(y,\omega) := \widehat{\mu_{nk}(\cdot,\omega)}(y)$ we have

$$\forall y \in H \qquad \prod_{k=1}^{k_n} \hat{\mu}_{nk}(y,\omega) \longrightarrow \hat{\mu}(y) \text{ a.s.}$$

For fixed $y \in H$ the set

$$A_{nk} := \left\{\omega ;\ \prod_{j=1}^{k} \left|\hat{\mu}_{nj}(y,\omega)\right| \geq \tfrac{1}{2}\left|\hat{\mu}(y)\right|\right\}$$

is $\mathcal{F}_{n,k-1}$-measurable $\left(\text{since } \hat{\mu}_{nj}(y,\cdot) = E(e^{i(y,X_{nj})} \mid \mathcal{F}_{n,j-1})\right)$. Moreover $A_{n,k+1} \subset A_{nk}$ and $P\left(\bigcup_{m=1}^{\infty} \bigcap_{n \geq m} A_{nk_n}\right) = 1$.

Putting

$$X^*_{nk} := I_{A_{nk}} X_{nk}, \qquad S^*_n := \sum_{k=1}^{k_n} X^*_{nk},$$

we obtain $S_n - S^*_n \xrightarrow{P} 0$; hence $E\, e^{i(y,S_n)} \to \hat{\mu}(y)$ if and only if $E\, e^{i(y,S^*_n)} \to \hat{\mu}(y)$. But as a regular version of the conditional distribution of X^*_{nk} given $\mathcal{F}_{n,k-1}$ we can choose

$$I_{A^c_{nk}}\, \delta_0(\cdot) + I_{A_{nk}}\, \mu_{nk}(\cdot)$$

so that

$$\prod_{k=1}^{k_n} \left| E(e^{i(y,X^*_{nk})} \mid \mathcal{F}_{n,k-1}) \right| \geqslant \frac{1}{2} \left| \hat{\mu}(y) \right| \quad \text{a.s.}$$

Hence the following computation

$$H_n := E\, e^{i(y,S^*_n)} \left[\prod_{k=1}^{k_n} E(e^{i(y,X^*_{nk})} \mid \mathcal{F}_{n,k-1}) \right]^{-1} = 1$$

is true.

By the estimation

$$\left| E\, e^{i(y,S^*_n)} - \hat{\mu}(y) \right| = \left| E\, e^{i(y,S^*_n)} - \hat{\mu}(y)\, H_n \right| \leqslant$$

$$\leqslant E \left| (\hat{\mu}(y))^{-1} - \left[\prod_{k=1}^{k_n} E(e^{i(y,X^*_{nk})} \mid \mathcal{F}_{n,k-1}) \right]^{-1} \right|$$

and the fact, that $\prod_{k=1}^{k_n} E(e^{i(y,X^*_{nk})} \mid \mathcal{F}_{n,k-1}) \to \hat{\mu}(y)$ a.s. the proof is completed.

The next theorem gives the conditions for the weak convergence in the infinite dimensional case.

Theorem B. If μ is an infinitely divisible distribution (for the definition and some properties see [7]), then the following conditions

B1. $\mu_n = \mu_{n1} * \mu_{n2} * \ldots * \mu_{nk_n} \Rightarrow \mu$ a.s.

B2. $\forall \varepsilon > 0 \ \max_{1 \leq k \leq k_n} \mu_{nk}(\|x\| > \varepsilon) \to 0$ a.s.

imply the weak convergence of distributions P_{S_n} of sums S_n to μ: $P_{S_n} \Rightarrow \mu$.

We give only a sketch of the proof.

Since μ is infinitely divisible its characteristic functional is non-vanishing. So B1 together with Theorem A imply $E\, e^{i(y,S_n)} \to \hat{\mu}(y)$ for every $y \in H$. By Lemma 2.10, Chapter VI, [7] it is sufficient to prove that $\{P_{S_n}\}$ form a conditionally compact set of measures.

To accomplish it, let us define:

$$a_{nk} = a_{nk}(\omega) := \int_{[\|x\| \leq 1]} x \, \mu_{nk}(dx, \omega) = E(X_{nk} I(\|X_{nk}\| \leq 1) | \mathcal{F}_{n,k-1}) \text{ a.s.},$$

$$\theta_{nk} := \mu_{nk} * (-a_{nk}).$$

$$\lambda_n := e\left(\sum_{k=1}^{k_n} \theta_{nk}\right) * \left(\sum_{k=1}^{k_n} a_{nk}\right),$$

where $e(F)$ is defined for a finite measure F by the formula

$$e(F) := e^{-F(H)} \sum_{n=0}^{\infty} F^{*n}/n!.$$

Under condition B2 by the accompanying laws (Corollary 6.1. Chapter VI [7]) condition B1 is equivalent to B1', $\lambda_n \Rightarrow \mu$ a.s.

We have introduced λ_n because for such measures we have good criteria of compactness (see [7], paragraph 5, Chapter VI).

Now let us define

$Z_{nk} := X_{nk} - a_{nk}$,

$U_{nk} := Z_{nk} I(\| Z_{nk} \| \leq t) - E(Z_{nk} I(\| Z_{nk} \| \leq t) | \mathcal{F}_{n,k-1})$,

$V_{nk} := Z_{nk} I(\| Z_{nk} \| > t)$,

$W_{nk} := a_{nk} + E(Z_{nk} I(\| Z_{nk} \| \leq t) | \mathcal{F}_{n,k-1})$,

where $t > 0$ is a fixed real number such that $M(\| x \| = t) = 0$ (M is the measure in the Levy's representation of μ, see Section 3 of this paper).

Due to the equality

$$S_n = \sum_{k=1}^{k_n} X_{nk} = \sum_{k=1}^{k_n} U_{nk} + \sum_{k=1}^{k_n} V_{nk} + \sum_{k=1}^{k_n} W_{nk} =$$

$$=: U_n + V_n + W_n$$

the conditional compactness of $\{P_{S_n}\}$ follows from the conditional compactness of the sets $\{P_{U_n}\}$, $\{P_{V_n}\}$, $\{P_{W_n}\}$. For each of the mentioned sets we use the criteria of compactness given by B1: In the proof of the conditional compactness of $\{P_{V_n}\}$ we use the following lemma:

Lemma. Let $\{F_n\}$ be a sequence of finite regular random measures on $\mathcal{B}_X \times \Omega$, where X is a complete separable metric space. If for almost every $\omega \in \Omega$ the family $\{F_n(\cdot, \omega);\ n \in N\}$ is uniformly tight, then for every $\delta > 0$ there exists a set A_δ with the properties

(a) $P(A_\delta) > 1 - \delta$,

(b) the set of measures $\{F_n(\cdot, \omega);\ n \in N,\ \omega \in A_\delta\}$ is uniformly tight.

Remark 1. If $\{F_n\}$ is a.s. conditionally compact, then the set A_δ can be chosen in such a way, that $\{F_n(\cdot, \omega);\ n \in N, \omega \in A_\delta\}$ is conditionally compact.

The above lemma can be proved analogously as the well known Egorov's Theorem.

3. Consequences

In this section we will give two applications of Theorem B.

First let us remind the Levy's representation of infinitely divisible laws. As in the real case, an infinitely divisible law μ has a unique representation $\mu = l(a, S, M)$ given by the formula

$$\hat{\mu}(y) = \exp\left[i(a,y) - \frac{1}{2}(Sy,y) + \int \left(e^{i(y,x)} - 1 - \frac{i(y,x)}{1+\|x\|^2}\right) M(dx)\right],$$

where $a \in H$, S is an S-operator (i.e. positive and hermitian with the finite trace $\operatorname{tr} S = \sum_{i=1}^{\infty} (Se_i, e_i)$) and M is a σ- finite measure on $\mathcal{B}_H$, which is finite outside every neighbourhood of 0 and has the following properties $M(\{0\}) = 0$, $\int_{[\|x\| \leq 1]} \|x\|^2 M(dx) < +\infty$. IF $M \equiv 0$, then $\mu = G(a, S)$ is called the Gaussian distribution with mean a and covariance operator S.

We need also the notion of martingale difference array (MDA). An array $\{X_{nk}\}$ is called MDA with respect to $\{\mathcal{F}_{nk}\}$ if $\{X_{nk}\}$ is adapted to $\{\mathcal{F}_{nk}\}$, $E\|X_{nk}\|^2 < +\infty$ and $E(X_{nk} \mid \mathcal{F}_{n,k-1}) = 0$.

<u>Theorem C.</u> Let $\{X_{nk}\}$ be MDA with respect to $\{\mathcal{F}_{nk}\}$ and $G(0,S)$ be the Gaussian distribution.

If the following conditions hold:

C1. $\sum_k E(\|X_{nk}\|^2 \mid \mathcal{F}_{n,k-1}) \xrightarrow{P} \operatorname{tr} S$,

C2. $\sum_k E(\|X_{nk}\|^2 I(\|X_{nk}\| > \varepsilon) \mid \mathcal{F}_{n,k-1}) \xrightarrow{P} 0$ for every $\varepsilon > 0$,

C3. $\sum_k E((X_{nk}, e_i)(X_{nk}, e_j) \mid \mathcal{F}_{n,k-1}) \xrightarrow{P} (Se_i, e_j)$ for some orthonormal basis $\{e_l\}$ in H and $i, j \in N$, then $P_{S_n} \Longrightarrow G(0, S)$.

(Here and in the sequel we use the convention $\sum_k \equiv \sum_{k=1}^{k_n}$).

Remark 2. For the stronger conditions see [9]. Theorem 2.

Remark 3. For discussion of the equivalence of the Condition C2 to the Lindeberg Condition see [2]:

$$\text{(LC)} \quad \sum_k E\|X_{nk}\|^2 \, I(\|X_{nk}\| > \varepsilon) \to 0, \quad \varepsilon > 0.$$

P r o o f: Let us choose and fix regular version μ_{nk} of the conditional distributions of X_{nk} given $\mathcal{F}_{n,k-1}$.

Now we can rewrite conditions C1 - C3 in the equivalent form:

C1'. $\sum_k \int \|x\|^2 \ \mu_{nk}(dx) \longrightarrow \operatorname{tr} S$,

C2'. $\sum_k \int_{[\|x\|>0]} \|x\|^2 \ \mu_{nk}(dx) \longrightarrow 0$ for positive rationals q,

C3'. $\sum_k \int (x,e_i)(x,e_j) \ \mu_{nk}(dx) \longrightarrow (Se_i,e_j)$ for $i,j \in N$.

Since we have only countable number of conditions, by the diagonalization procedure from every subsequence $\{S_{n_k}\}$ of $\{S_n\}$ we can choose a further subsequence $\{S_{n_{k_l}}\}$, for which the convergence in probability in conditions C1 - C3' is replaced by the a.s. convergence. Hence there exists a set of probability one Ω' such that for $\omega \in \Omega'$ the convolutions

$$\mu_{n_{k_l}}(\cdot,\omega) := \mu_{n_{k_l}1}(\cdot,\omega) * \mu_{n_{k_l}2}(\cdot,\omega) * \dots$$

are weakly convergent to G(0,S) by the Central Limit Theorem. Therefore by Theorem β (the condition B2 is implied by C2') the distributions of $S_{n_{k_l}}$ weakly converge to G(0,S).

Suppose that $P_{S_n} \not\Rightarrow G(0,S)$. Then we can choose a subsequence $\{P_{S_{n_k}}\}$, any subsequence of which is not weakly convergent to $G(0,S)$. But we have already shown that this is impossible.

Using the same method the following generalization of Theorem 4.2 [12] of A. Kłopotowski can be proved:

Theorem D. If $\mu = l(a,S,M)$ is an infinitely divisible distribution, then the following conditions imply the weak convergence of distributions P_{S_n} to μ:

D1. $\sum_k \mu_{nk}(A) \xrightarrow[n\to\infty]{P} M(A)$

for every $A \in \mathcal{B}_H$ such that $\bar{A} \not\ni 0$ and $M(\partial A) = 0$.

D2. there exists a.s. finite real random variable $C(\omega)$ such that

$P\,(\mathrm{tr}\, T_n^1 > C) \xrightarrow[n\to\infty]{} 0$

D3. $\sup_n \sum_{i=N}^{\infty} (T_n^1 e_i, e_i) \xrightarrow[N\to\infty]{P} 0$

for some orthonormal basis $\{e_i\}$ in H,

D4. $(T_n^{\varepsilon} e_i, e_j) \xrightarrow[n\to\infty]{P} (Se_i, e_j) + \int_{[\|x\|\leq\varepsilon]} (x,e_i)(x,e_j)M(dx)$

for the mentioned basis $\{e_i\}$, $i,j \in N$ and every $\varepsilon > 0$ with $M(\|x\| = \varepsilon) = 0$

D5. $\sum_k \left(a_{nk} + \int \frac{x - a_{nk}}{1 + \|x - a_{nk}\|^2}\, \mu_{nk}(dx)\right) \xrightarrow[n\to\infty]{P} a$

D6. $\max_{1\leq k\leq k_n} \mu_{nk}(\|x\| > \varepsilon) \xrightarrow[n\to\infty]{P} 0$

for every $\varepsilon > 0$.

where a_{nk} is defined by $a_{nk} := \int_{[\|x\| \leq 1]} x \ \mu_{nk}(dx)$ and for every $t>0$ $\left\{T_n^t = T_n^t(\omega)\right\}$ is a set of random S-operators defined by the formulas

$$(T_n^t y, y) := \sum_k \int_{[\|x - a_{nk}\| \leq t]} (y, x - a_{nk})^2 \ \mu_{nk}(dx)$$

Remark 4. Conditions D1 - D6 can be translated into the language of conditional expectations.

References

[1] B r e i m a n, L., Probability, Addison - Wesely, London 1968.

[2] B r o w n, B. M., Martingale central limit theorems, Ann. Math. Statist., 42, 59-66 (1971).

[3] B r o w n, B. M., E a g l e s o n, G. K., Martingale convergence to infinitely divisible laws with finite variances, Trans. Amer. Math. Soc., 162, 449-453 (1971).

[4] D v o r e t z k y, A., The central limit theorems for dependent random variables, Proc. of the Int. Congress of Math. Nice 1970.

[5] D v o r e t z k y, A., Asymptotic normality for sums of dependent random variables, Proc. 6th Berkeley Sympos. Math. Statist. Probab. Univ. Calif., 513-535 (1971).

[6] K ł o p o t o w s k i, A., Limit theorems for sums of dependent, random vectors in R^d, Dissert. Math. CLI, 1-55 (1977).

[7] P a r t h a s a r a t h y, K. R., Probability measures on metric spaces, Academic Press, New York - London 1967.

[8] V a r a d h a n, S. R. S., Limit theorems for sums of independent random variables with values in a Hilbert space, Sankhya 2.4, 213-238 (1962).

[9] W a l k, H., An invariance principle for the Robbins-Monro process in a Hilbert space, Z. Wahrscheinlichkeitstheorie verw. Geb., 39, 135-150 (1977).