DEMONSTRATIO MATHEMATICA
Vol. XX No 3-4 1987

Jacek Wesołowski

CONDITIONING IN LIMIT PROBLEM FOR ORDER STATISTICS

1. Introduction

The Lévy's book [9] contains some results which suggest the following idea of obtaining limit theorems for sums of dependent random variables: to replace all probabilities and expectations in known limit theorems for row-wise idependent double sequences into the analogous conditional quantities, where the conditioning is with respect to some specially chosen σ-fields, and change the convergence of numbers into the convergence in probability. Brown and Eagleson first applied the idea in [1]. They considered the conditioning in respect to the σ-fields from a double row-wise increasing array. Later, the theory with such a conditioning was developed by Eagleson ([2]), Kłopotowski ([3], [4]), Jakubowski ([5]) and others - still for sums only.

Applying this idea and making use of the conditioning proposed in [1] we prove in the present paper two limit theorems for order statistics in the case of dependent random variables. They are generalizations of some results for row-wise independent double sequence which belong to Loéve (see [6]).

We consider a double sequence $\{\{X_{nk}\}\}$, $k = 1,\ldots,k_n$, $n = 1,2,\ldots$, of random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. The double sequence of σ-fields $\{\{\mathcal{F}_{nk}\}\}$ is row-wise increasing and adapted to $\{\{X_{nk}\}\}$. For $\omega \in \Omega$ we order the realizations $X_{n1}(\omega),\ldots,X_{nk_n}(\omega)$, $n \geq 1$, in the increasing sequence $\{z_{nk}\}$, $k = 1,\dots,k_n$. Let us define a random variable Z_{nk} by: $Z_{nk}(\omega) = z_{nk}$, $k = 1,\dots,k_n$, $n \geq 1$. The random variable Z_{n,k_n-k+1} is the k-th order statistic. Obviously

$$Z_{n1} = \min\{X_{n1},\dots,X_{nk_n}\},$$

$$Z_{nk_n} = \max\{X_{n1},\dots,X_{nk_n}\}.$$

In this paper all the equations between random variables are in the almost sure sense; the symbols $\sum_k$, $\prod_k$ denote $\sum_{k=1}^{k_n}$, $\prod_{k=1}^{k_n}$, respectively.

The following lemma taken from [5] is the basic tool in the theory of limit theorems for sums of dependent random variables with the conditioning described above:

L e m m a . Let $\{\{X_{nk}\}\}$, $\{\{\mathcal{F}_{nk}\}\}$ be defined as above. If

$$\prod_k E(\exp(itX_{nk}) \mid \mathcal{F}_{n,k-1}) \xrightarrow{P} z_t \neq 0,$$

then

$$E(\exp(it\sum_k X_{nk})) \to z_t,$$

where t is any real number and z_t is a certain complex number.

We apply the above lemma in the proof of our results for order statistics.

2. The case of r-th order statistics

For the double sequences $\{\{X_{nk}\}\}$, $\{\{\mathcal{F}_{nk}\}\}$ defined above we denote by $F_{nk}(\cdot \mid \mathcal{F}_{n,k-1})$ the conditional distribution function of X_{nk}, $k = 1,\dots,k_n$, and we have

T h e o r e m 1. If the random variables X_{nk}, $k = 1,\dots,k_n$, $n \geq 1$, are row-wise conditionally uniformly assymptotically identically distributed, i.e.

(1) $$\max_k |F_{nk}(x \mid \mathcal{F}_{n,k-1}) - F_n(x)| \xrightarrow{P} 0$$

for certain sequence of distribution functions F_n, $n \geq 1$, and

(2) $$\sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})) \xrightarrow{P} L(x),$$

then

(3) $$P(Z_{n,k_n-r+1} < x) \to \sum_{k=0}^{r-1} \frac{L^k(x)}{k!} e^{-L(x)},$$

where x is a real number.

It should be emphasized that the conditions (1) and (2) are the conditional versions of those from the theorem for row-wise independent double sequence. As in [6] the proof is divided into three parts, and also similarly as in [6] the result obtained in the first part (a) follows from the known limit results - see the papers [7], [8] or the book [10]. We present it to make our paper selfcontained and to show how useful Jakubowski's lemma is in proving limit results.

P r o o f . Since $(P(Z_{n,k_n-r+1} < x))$ is a compact sequence, then for its convergence to s it suffices to prove that from every subsequence we can choose a subsequence convergent to s. The sequence $\{F_n(x)\}$ is also compact and consequently we can assume that

(4) $$\max_k |F_{nk}(x \mid \mathcal{F}_{n,k-1}) - F(x)| \xrightarrow{P} 0$$

for certain distribution function $F(x)$. Applying the routine technique of subsequences we can change the convergence in probability in (2) and (4) into the a.e. convergence.

Let us observe that

(5) $$P(Z_{n,k_n-r+1} < x) = P\Big(\sum_k I(X_{nk} > x) < r\Big)$$

for any $r = 1,\dots,k_n$, $n \geq 1$. Now we consider three possible cases: (a) $F(x) = 1$, $L(x) < \infty$, (b) $0 < F(x) \leq 1$, $L(x) = \infty$, (c) $F(x) = 0$. (Obviously for $0 \leq F(x) < 1$ we have $L(x) = \infty$).

(a) From the expansion of the exponential function we get

$$\log \prod_k E_{n,k-1}(\exp(itI(X_{nk} > x))) =$$

$$= \sum_k \log(1 + (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1}))(e^{it} - 1)) =$$

$$= \sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1}))(e^{it} - 1) +$$

$$+ \frac{\theta(x)}{2}(e^{it} - 1)^2 \sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1}))^2,$$

where $|\theta(x)| \leq 1$ and $E_{n,k-1}(\cdot) = E(\cdot \mid \mathcal{F}_{n,k-1})$ for any $k = 1,\dots,k_n$, $n \geq 1$. On the other hand

$$\sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1}))^2 \leq$$

$$\leq \max_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})) \sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})).$$

Consequently (2) implies

$$\prod_k E_{n,k-1}(\exp(itI(X_{nk} > x))) \xrightarrow{P} \exp(L(x)(e^{it} - 1)) > 0.$$

Now applying the lemma with X_{nk} changed into $I(X_{nk} > x)$ we see that the random variable $\sum_k I(X_{nk} > x)$ is Poissonian with the mean $L(x)$ in limit. Hence (3) follows from (5).

(b) Since

$$P(Z_{n,k_n-r+1} < x) = \sum_{i=0}^{r-1} P\left(\sum_k I(X_{nk} > x) = i\right),$$

then it suffices to show that for any $i = 0,1,\dots,r-1$ we have

$$(6) \qquad P\Big(\sum_k I(X_{nk} > x) = i\Big) \longrightarrow 0.$$

At first let us observe that

$$(7) \qquad P\Big(\sum_k I(X_{nk} > x) = i\Big) =$$

$$= \sum_{1 \le j_1 < \dots < j_i \le k_n} E\Big(\prod_{k \ne j_1,\dots,j_i} I(X_{nk} \le x) \prod_{l=1}^{i} I(X_{nj_l} > x)\Big).$$

Now, we introduce the following definition

$$A = \Big\{ \omega \in \Omega : \max_k \big| F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) - F(x)\big| \longrightarrow 0$$

$$\text{and} \quad \sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega)) \longrightarrow \infty \Big\}.$$

Since the double sequence $\{\{\mathcal{F}_{nk}\}\}$ is row-wise increasing and $P(A) = 1$ (consequently $EX = EXI(A)$ for any integrable random variable X) so this together with (7) implies that

$$(8) \qquad P\Big(\sum_k I(X_{nk} > x) = i\Big) \le$$

$$\le \sum_{1 \le j_1 < \dots < j_i \le k_n} \sup_{\omega \in A} \prod_{k \ne j_1,\dots,j_i} F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) \times$$

$$\times \prod_{l=1}^{i} (1 - F_{nj_l}(x \mid \mathcal{F}_{n,j_l-1})(\omega)).$$

For any $\omega \in A$ and sufficiently large n we have

$$F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) \ge \frac{q}{2},$$

where $q = F(x)$, $k = 1,\dots,k_n$. Consequently

$$\prod_{k\neq j_1,\dots,j_i} F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) =$$

$$= \Big(\prod_{l=1}^{i} F_{nj_l}(x \mid \mathcal{F}_{n,j_l-1})(\omega)\Big)^{-1} \prod_k F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) \leqslant$$

$$\leqslant \left(\frac{2}{q}\right)^{i} \prod_k F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega).$$

On the other hand

$$\log \prod_k F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) = \sum_k \log(1 - (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega)))$$

$$\leqslant - \sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega))$$

and thus for any $\omega \in A$ and sufficiently large n we get

$$\sum_{1\leqslant j_1<\dots<j_i\leqslant k_n} \prod_{k\neq j_1,\dots,j_i} F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) \times$$

$$\times \prod_{l=1}^{i} (1 - F_{nj_l}(x \mid \mathcal{F}_{n,j_l-1})(\omega)) \leqslant$$

$$\leqslant \left(\frac{2}{q}\right)^{i} \exp\Big(-\sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega))\Big) \times$$

$$\times \sum_{1\leqslant j_1<\dots<j_i\leqslant k_n} \prod_{l=1}^{i} (1 - F_{nj_l}(x \mid \mathcal{F}_{n,j_l-1})(\omega)) \leqslant$$

$$\leqslant \left(\frac{2}{q}\right)^{i} \exp\Big(-\sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega))\Big) \times$$

$$\times \Big(\sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega))\Big)^{i}.$$

For any $\omega \in A$ we have $\sum_k (1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega)) \to \infty$, consequently (8) and the above inequality imply (6).

(c) Similarly as in the case (b) we get

$$P(Z_{n,k_n-r+1} > x) = P\left(\sum_k I(X_{nk} > x) \geqslant r\right) \geqslant$$

$$\geqslant E \prod_{i=1}^{r} I(X_{ni} > x) \geqslant \inf_B \prod_{i=1}^{r} (1 - F_{ni}(x \mid \mathcal{F}_{n,i-1})(\omega)),$$

where $B = \left\{\omega \in \Omega : \max_k F_{nk}(x \mid \mathcal{F}_{n,k-1})(\omega) \to 0\right\}$. Since for any $\omega \in B$

$$\prod_{i=1}^{r} (1 - F_{ni}(x \mid \mathcal{F}_{n,i-1})(\omega)) \to 1,$$

then $P(Z_{n,k_n-r+1} < x) \to 0$.

Since $L(x) = \infty$, we obtain (3). The proof is completed.

3. The case of (k_n-r)-th order statistics

For the double sequences $\{\{X_{nk}\}\}$, $\{\{\mathcal{F}_{nk}\}\}$ defined above we have

T h e o r e m 2. If the condition (1) is fulfilled and

$$\sum_k F_{nk}(x \mid \mathcal{F}_{n,k-1}) \xrightarrow{P} M(x), \tag{9}$$

x is a real number, then

$$P(Z_{n,r+1} < x) \to \sum_{k=r+1}^{\infty} \frac{M^k(x)}{k!} e^{-M(x)}. \tag{10}$$

P r o o f . Let us observe that

$$P(Z_{n,r+1} < x) = 1 - P \sum_k I(X_{nk} < x) < r + 1). \tag{11}$$

Repeating the proof of Theorem 1 with $F_{nk}(x \mid \mathcal{F}_{n,k-1})$, F, L and $I(X_{nk} < x)$ changed respectively into $1 - F_{nk}(x \mid \mathcal{F}_{n,k-1})$, $1 - F$, M and $I(X_{nk} > x)$ we get from (1) and (9)

$$P\left(\sum_k I(X_{nk} < x) < r + 1\right) \to \sum_{k=0}^{r} \frac{M^k(x)}{k!} e^{-M(x)}.$$

Consequently (11) implies (10).

4. Remarks

Now we present an example of a double sequence of rv's which is row-wise conditionally uniformly assymptotically identically distributed.

We consider two independent double sequences $\{\{Y_{nk}\}\}$, $\{\{Z_{nk}\}\}$ of row-wise iid rv's. The common distribution function on n-th row in $\{\{Y_{nk}\}\}$ is F_n. Let us take a double sequence $\{\{a_{nk}\}\}$ of positive real numbers such that $\sum_k a_{nk} \to 0$ and for any k and n consider such a set A_{nk} that $P(Y_{nk} \in A_{nk}^c) < a_{nk}$ and $P(Z_{nk} \in A_{nk}^c) < a_{nk}$. We define $X_{n1} = Y_{n1}$,

$$X_{nk} = \begin{cases} Y_{nk} & \text{if} \quad X_{n,k-1} \in A_{n,k-1}, \\ Z_{nk} & \text{if} \quad X_{n,k-1} \in A_{n,k-1}^c, \end{cases}$$

$k = 2,\dots,k_n$, $n \geq 1$. Then it is easy to see that $\{\{X_{nk}\}\}$ is row-wise conditionally uniformly assymptotically identically distributed.

Theorems 1 and 2 are generalizations of the results for row-wise independent double sequence. To obtain the ones from our theorems it suffices to take $\mathcal{F}_{nk} = \{\emptyset, \Omega\}$ for every $k = 1,\dots,k_n$, $n \geq 1$; then the conditions (1), (2) and (9) change into the ones from Loéve's theorems (see [6]),

There still remains an open question if Lévy's idea is applicable to the limit theorems for r_n-th order statistics, where r_n and $k_n - r_n$ tend to infinity as $n \to \infty$. Another question, still unsolved also for sums is the following: describe the class of σ-fields for which the above procedure of obtaining limit theorems for dependent random variables from the known results in the independent case works.

Acknowledgement. I wish to express my thanks to Doc. A. Kłopotowski for many materials which direct my interest to this field. I am also greatly indebted to the referee for valuable remarks (on the formulation of Theorem 1, on the example and others).

REFERENCES

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

[2] E a g l e s o n G.K. : Martingale convergence to mixtures of infinitely divisible laws, Ann. Prob. 3 (1977), 557-562.

[3] 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 (1977), 1-58.

[4] K ł o p o t o w s k i A.: Mixtures of infinitely divisible distributions as limit laws for sums of dependent random variables, Z. Wahr. verw. Geb. 51 (1980), 101-113.

[5] J a k u b o w s k i A.: On limit theorems for sums of dependent Hilbert space valued random variables, Lecture Notes in Stat. 2 (1980), 178-187.

[6] L o é v e M.: Ranking limit problem, Proc. Third Berkeley Symp. 2 (1956), 177-194.

[7] F r e e d m a n D.: The Poisson approximation for dependent events, Ann. Prob. 2 (1974), 256-269.

[8] S e r f l i n g R.J.: A general Poisson approximation theorem, Ann. Prob. 3 (1975), 726-731.

[9] L é v y P.; Théorie de l'Addition des Variables Alé-toires, Paris, Gauthier-Villars (1937).

[10] L e a d b e t t e r M.R., L i n d g r e n G.L., R o o t z é n H.: Extremes and related properties of random sequences and processes, Series in Statistics, Springer Verlag, New York, Berlin 1983.

INSTITUTE OF MATHEMATICS, TECHNICAL UNIVERSITY OF WARSAW,
00-661 WARSZAWA

Received October 2, 1985; revised version April 10, 1986.