Stability and law of large numbers for sums of a random number of random variables

By D. SZÁSZ in Budapest*)

In [1] BIKELIS and MOGYORÓDI gave sufficient conditions for the stability and the law of large numbers in case of sums of a random number of random variables. They dealt with the case when the random indices have finite mathematical expectations and asserted that the given sufficient conditions are also necessary. However, we shall give a counterexample which shows that the mentioned conditions are not necessary. Moreover, without the condition of the existence of expectations, we can give necessary and sufficient conditions both for the stability and the law of large numbers for sums of a random number of random variables. These conditions follow from the results of the author on convergence of distributions of sums of a random number of random variables [2].

1. The results. For every n, let $\xi_{n1}, \dots, \xi_{nk}, \dots$ be a sequence of independent random variables and ν_n be a random index which is independent of the sequence $\{\xi_{nk}\}_k$. Suppose that

$$\lim_{n\to\infty} \sup_k P(|\xi_{nk}| > \varepsilon) = 0$$

for every $\varepsilon > 0$ and that

$$\text{P-}\lim_{n\to\infty} \nu_n = \infty. \quad ^{1})$$

Let

$$S_k^{(n)} = \xi_{n1} + \cdots \xi_{nk}.$$

Definitions. We say that the sequence of sums $S_{\nu_n}^{(n)}$ is *stable* if there exists a double sequence of constants $A_k^{(n)}$ $(n, k = 1, 2, \dots)$ such that

$$\text{P-}\lim_{n\to\infty} (S_{\nu_n}^{(n)} - A_{\nu_n}^{(n)}) = 0.$$

We say that the sequence of sums $S_{\nu_n}^{(n)}$ satisfies the *law of large numbers* if there exists a sequence of constants C_n such that

$$\text{P-}\lim_{n\to\infty} (S_{\nu_n}^{(n)} - C_n) = 0.$$

---

*) This work was done at the Department of Probability, Lomonosov University, Moscow.

[1]) $\text{P-}\lim_{n\to\infty}$ denotes stochastic convergence.

In [1] it is supposed that $E\nu_n<\infty$ for every n and is proved that the sequence $\{S_{\nu_n}^{(n)}\}_n$ is stable, if

(i)$^+$
$$\lim_{n\to\infty} E\sum_{k=1}^{\nu_n}\int_{|x|>1} dF_{nk}(x+m_{nk})=0,$$

(ii)$^{++}$
$$\lim_{n\to\infty} E\sum_{k=1}^{\nu_n}\int_{|x|\leq 1} x^2\,dF_{nk}(x+m_{nk})=0,$$

where $F_{nk}(x)=P(\xi_{nk}<x)$ and m_{nk} denotes the median of ξ_{nk}.

If in addition to (i)$^+$ and (i)$^{++}$ there exists a sequence of constants C_n such that

(iii)
$$\text{P-}\lim_{n\to\infty}\left(\sum_{k=1}^{\nu_n}\left\{\int_{|x|\leqq 1} x\,dF_{nk}(x+m_{nk})+m_{nk}\right\}-C_n\right)=0,$$

then it is proved that the sequence $\{S_{\nu_n}^{(n)}\}_n$ satisfies the law of large numbers.

Our assertion that the above conditions are not necessary is based on the following example: for every $n>4$ let

$$P(\nu_n=2^k)=\frac{1}{n}\qquad (1\leqq k\leqq n).$$

Clearly $E\nu_n<\infty$ for every n. Further let

$$\xi_{nk}=0$$

if $1\leqq k\leqq 2^{n-1}$ or $k>2^n$ and

$$P(\xi_{nk}=0)=1-\frac{2}{n},$$

$$P(\xi_{nk}=1)=P(\xi_{nk}=a_n)=\frac{1}{n},$$

if $2^{n-1}<k\leqq 2^n$ where $|a_n|>1$.

It is obvious that

$$\text{P-}\lim_{n\to\infty} S_{\nu_n}^{(n)}=0,$$

but

$$E\sum_{k=1}^{\nu_n}\int_{|x|>1} dF_{nk}(x+m_{nk})=E\sum_{k=1}^{\nu_n}\int_{|x|\leqq 1} x^2\,dF_{nk}(x+m_{nk})=\frac{2^n-2^{n-1}}{n^2}=\frac{2^{n-1}}{n^2}\to\infty$$

as $n\to\infty$. So the sequence $\{S_{\nu_n}^{(n)}\}_n$ satisfies the law of large numbers (and is stable), but not conditions (i)$^+$ and (i)$^{++}$.

Our results do not suppose the finiteness of expectations. We prove:

Theorem 1. *The sequence* $\{S_{\nu_n}^{(n)}\}_n$ *is stable iff*

(i)
$$\text{P-}\lim_{n\to\infty} \sum_{k=1}^{\nu_n} \int_{|x|>1} dF_{nk}(x+m_{nk}) = 0$$

and

(ii)
$$\text{P-}\lim_{n\to\infty} \sum_{k=1}^{\nu_n} \int_{|x|\leq 1} x^2\, dF_{nk}(x+m_{nk}) = 0.$$

Theorem 2. *The sequence* $\{S_{\nu_n}^{(n)}\}_n$ *satisfies the law of large numbers iff conditions* (i), (ii), *and for a suitable sequence of constants* C_n *condition* (iii) *are satisfied.*

We remark that conditions (i) and (ii) are respectively equivalent to the following conditions: for every $q\in[0, 1)$,

(1.1)
$$\lim_{n\to\infty} \sum_{k\leq l_n(q)} \int_{|x|>1} dF_{nk}(x+m_{nk}) = 0$$

and

(1.2)
$$\lim_{n\to\infty} \sum_{k\leq l_n(q)} \int_{|x|\leq 1} x^2\, dF_{nk}(x+m_{nk}) = 0.$$

Here $l_n(q)$ denotes the lower bound of q-quantiles of the random variable ν_n.

We remark that an analogous result for the stability of sums $S_{\nu_n}^{(n)}$ was recently reached in a different way by A. V. PETCHINKIN (to be published later).

2. Preliminaries to the proofs. In [2] we have proved

Proposition 1 ([2], Corollary to Theorem 1). *Suppose that for almost every* $q\in[0, 1]$ *there exists a distribution function* $\Phi_q(x)$ *such that*

$$\lim_{n\to\infty} P(S_{l_n(q)}^{(n)} < x) = \Phi_q(x).$$

Then there exists a measurable stochastic process $\{\chi(t): t\in[0, 1)\}$ *with independent increments such that for almost every t*

$$P(\chi(t) < x) = \Phi_t(x)$$

and

$$\lim_{n\to\infty} P(S_{\nu_n}^{(n)} < x) = P(\chi(\pi) < x).$$

Here the random variable π *is uniformly distributed in* [0, 1] *and is independent of the process* $\{\chi(t): t\in[0, 1)\}$.

For an arbitrary random variable X there exists a unique real number $\Delta(X)$ such that

$$E \operatorname{arc\,tg}(X-\Delta(X)) = 0.$$

This number $\Delta(X)$ is called Doob's center or briefly center. A stochastic process $\{\chi(t): t\in T\}$ is centered, if $\Delta(\chi(t))=0$ for every $t\in T$. [3].

Proposition 2 ([2], Corollary 1 of Theorem 2). *Suppose that the random variables ξ_{nk} are symmetric for every n and k, and the sums $S_{\nu_n}^{(n)}$ have a limit distribution as $n\to\infty$. Then there exist a decomposition of indices to a finite or infinite number sequences: $N_1, \ldots, N_j$, and centered, measurable stochastic processes $\{\chi^{(i)}(t): t\in[0, 1)\}$ with independent increments $(1\leqq i\leqq j)$ such that for almost every t*

(2.1) $$\lim_{\substack{n\to\infty\\ n\in N_i}} P(S_{l_n(t)}^{(n)}<x)=P(\chi^{(i)}(t)<x)$$

and also

(2.2) $$\lim_{\substack{n\to\infty\\ n\in N_i}} P(S_{\nu_n}^{(n)}<x)=P(\chi^{(i)}(\pi)<x) \qquad (1\leqq i\leqq j).$$

Lemma ([2], Lemma 4). *If the distributions of the random variables X_n weakly converge to the distribution of the random variable X as $n\to\infty$, then*

$$\lim_{n\to\infty} \Delta(X_n)=\Delta(X).$$

2. The proofs. Let for every n $\{\eta_{nk}\}_k$ be a sequence of independent random variables such that it is independent of the sequence $\{\xi_{nk}\}_k$ and of the random variable ν_n and that

$$P(\xi_{nk}<x)=P(\eta_{nk}<x) \qquad (n, k=1, 2 \ldots).$$

Let

$$\zeta_{nk} = \xi_{nk}-\eta_{nk}, \quad Z_k^{(n)} = \zeta_{n1}+\cdots+\zeta_{nk}.$$

Proof of Theorem 1. *Necessity.* If the assertion is not true then there exist some $q_0<1$ and a subsequence $\{n'\}$ of indices such that either

(3.1) $$\liminf_{n'\to\infty} \sum_{k\leqq l_{n'}(q_0)} \int_{|x|>1} dF_{n'k}(x+m_{n'k}) > 0$$

or

(3.2) $$\liminf_{n'\to\infty} \sum_{k\leqq l_{n'}(q_0)} \int_{|x|\leqq 1} x^2\, dF_{n'k}(x+m_{n'k}) > 0.$$

We can suppose and for the sake of simplicity we do suppose that the subsequence $\{n'\}$ coincides with the entire sequence $\{n\}$ of positive integers. Thus instead of n' we can always write n.

We remark that from the stability of sequence $\{S_{\nu_n}^{(n)}\}_n$ it follows that

(3.3) $$\text{P-}\lim_{n\to\infty} Z_{\nu_n}^{(n)}=0.$$

Let us use Proposition 2 for the sums $Z_{\nu_n}^{(n)} = \zeta_{n1} + \dots + \zeta_{n\nu_n}$. In consequence of (3. 3) and (2. 2), every process $\{\chi^{(i)}(t): t \in [0, 1)\}$ appearing in Proposition 2 vanishes, i.e. for all i and t $(1 \leqq i \leqq j,\ t \in [0, 1))$

$$P(\chi^{(i)}(t) = 0) = 1.$$

Because of (2. 1), for a suitable q_1 $(q_0 \leqq q_1 < 1)$ we have

$$\operatorname*{P-lim}_{\substack{n\to\infty \\ n \in N_1}} Z_{l_n(q_1)}^{(n)} = 0.$$

Here we can use the following simple remark: if for every n X_n and Y_n are independent random variables and

$$\operatorname*{P-lim}_{n\to\infty} (X_n + Y_n) = 0,$$

then there exists a sequence of constants a_n such that

$$\operatorname*{P-lim}_{n\to\infty} (X_n - a_n) = 0.$$

On the basis of this remark, there exist constants a_n $(n \in N_1)$ such that

$$\operatorname*{P-lim}_{\substack{n\to\infty \\ n \in N_1}} (S_{l_n(q_1)}^{(n)} - a_n) = 0.$$

On the basis of the theorem of § 22 [4], we get that

$$\lim_{\substack{n\to\infty \\ n \in N_1}} \sum_{k \leqq l_n(q_1)} \int_{|x|>1} dF_{nk}(x + m_{nk}) = 0,$$

$$\lim_{\substack{n\to\infty \\ n \in N_1}} \sum_{k \leqq l_n(q_1)} \int_{|x|\leqq 1} x^2 dF_{nk}(x + m_{nk}) = 0.$$

These relations contradict both (3. 1) and (3. 2).

Sufficiency. From conditions (1. 1) and (1. 2) it follows that for every $q \in [0, 1)$ there exist sequences $\{m_n(q)\}_n$ of constants such that

(3. 4) $$\operatorname*{P-lim}_{n\to\infty} \left(S_{l_n(q)}^{(n)} - m_n(q)\right) = 0.$$

Our lemma asserts that $m_n(q)$ can be chosen as

$$m_n(q) = \Delta(S_{l_n(q)}^{(n)}).$$

So if

$$A_k^{(n)} = m_n(A_n(k)),$$

where $A_n(x)=P(\nu_n<x)$, then Proposition 1 and (3. 4) give that

$$\text{P-}\lim_{n\to\infty}(S_{\nu_n}^{(n)}-A_{\nu_n}^{(n)})=0.$$

We remark that here Proposition 1 is applied instead of the summands ξ_{nk} to the summands $\xi_{nk}-(A_k^{(n)}-A_{k-1}^{(n)})$.

Proof of Theorem 2. Necessity of conditions (i) and (ii) follows from the preceding theorem. If (i) and (ii) are satisfied, then in (3. 4) $m_n(q)$ can be chosen also in the following manner

$$m_n(q)=\sum_{k\le l_n(q)}\left\{\int_{|x|\le 1}x\,dF_{nk}(x+m_{nk})+m_{nk}\right\},$$

and so in consequence of Proposition 1,

$$\text{P-}\lim_{n\to\infty}\left(S_{\nu_n}^{(n)}-\sum_{k=1}^{\nu_n}\left\{\int_{|x|\le 1}x\,dF_{nk}(x+m_{nk})+m_{nk}\right\}\right)=0. \tag{3. 5}$$

However, at the same time

$$\text{P-}\lim_{n\to\infty}(S_{\nu_n}^{(n)}-C_n)=0 \tag{3. 6}$$

and from (3. 5) and (3. 6) condition (iii) follows obviously.

To prove the sufficiency we remark that (3. 6) and (iii) together give (3. 6).

Literature

[1] A. P. Bikelis—J. Mogyoródi, On a problem of B. V. Gnedenko, *Acta Sci. Math.*, **30** (1969), 241—245.

[2] D. Szász, Limit theorems for the distributions of the sums of a random number of random variables, *Annals of Math. Statistics* (to appear).

[3] K. Ito, *Stochastic processes*. Vol. 1 (Moscow, 1960) (in Russian).

[4] B. V. Gnedenko—A. N. Kolmogorov, *Limit distributions for sums of independent random variables* (Reading, 1954).

(Received July 22, 1971)