ON LIMITING DISTRIBUTIONS FOR SUMS OF A RANDOM NUMBER OF INDEPENDENT RANDOM VARIABLES

by

J. MOGYORÓDI[1]

§ 1. Introduction

In paper [1] F. J. ANSCOMBE proved the following

Theorem 1. *Let* $Y_n (n = 1, 2, \ldots)$ *be an infinite sequence of random variables. We suppose that there exists a sequence* $B_n (n = 1, 2, \ldots)$ *of positive numbers and a proper distribution function* $F(x)$, *such that the following conditions hold :*

a) *For any continuity point of* $F(x)$

$$\lim_{n \to +\infty} \mathbf{P}\{Y_n < x B_n\} = F(x) ,$$

where $\mathbf{P}\{A\}$ *denotes the probability of the event* A.

b) *Given any positive* ε *and* η *there is a positive integer* n_0 *and a positive number* c, *such that for any* $n \geqq n_0$

$\mathbf{P}\{|Y_n - Y_{n'}| < \varepsilon B_n$ *simultaneously for all integers* n' *such that* $|n - n'| < cn\} > 1 - \eta$.

Let further $\nu_n (n = 1, 2, \ldots)$ *be an infinite sequence of positive integer-valued random variables and* $k_n (n = 1, 2, \ldots)$ *a sequence of positive integers tending to infinity. We suppose that* ν_n / k_n *converges to* 1 *in probability as* $n \to + \infty$. *Then, if the sequence of random variables* Y_n *satisfies conditions* (a) *and* (b),

$$\lim_{n \to +\infty} \mathbf{P}\{Y_{\nu_n} < x B_{k_n}\} = F(x)$$

at all continuity points of $F(x)$.

We mention that nothing is supposed about the dependence of ν_n on the random variables Y_k.

The importance of the investigation of the behaviour of the sum of a random number of random variables is well known in sequential analysis, in random walk problems and in connection with Monte Carlo methods.

The conditions of the above mentionned Theorem 1 are sufficient. F. J. ANSCOMBE conjectured that condition (b) is also necessary if no further condition than (a) is supposed. The present paper deals with the proof of the necessity in the case when

$$Y_n = \xi_1 + \xi_2 + \ldots + \xi_n$$

where $\xi_1, \xi_2, \ldots, \xi_k, \ldots$ is a sequence of independent random variables, and gives in this case a simpler condition equivalent to (b).

[1] Eötvös Lóránd University, Mathematical Institute, Budapest.

We will prove the following

Theorem 2. *Let us suppose that* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *is a sequence of independent random variables and there exists a sequence* B_n $(B_n \to +\infty$ *as* $n \to +\infty)$ *of positive numbers such that the random variables* ξ_k/B_n $(k = 1, 2, \ldots, n)$ *are infinitesimal* [2] *and*

$$\text{(a)} \qquad \mathbf{P}\left\{\frac{\sum_{k=1}^{n} \xi_k}{B_n} < x\right\} \to F(x) \qquad (n \to +\infty)$$

at all continuity points x *of the proper distribution function* $F(x)$. *Let us suppose further that* ν_n $(n = 1, 2, \ldots)$ *is a sequence of positive integer-valued random variables and* k_n $(n = 1, 2, \ldots)$ *is a sequence of positive integers tending to infinity such that* ν_n/k_n *converges to* 1 *in probability as* $n \to +\infty$. *In order that the distribution function of the sums*

$$\frac{\xi_1 + \ldots + \xi_{\nu_n}}{B_{k_n}}$$

converge to the law $F(x)$, *it is necessary and sufficient that the condition*

$$\text{(b')} \qquad \lim_{\delta \to 0} \limsup_{n \to +\infty} \frac{B_{k_n}}{B_{[k_n(1+\delta)]}} = 1$$

be satisfied.

Here δ is a real number $(\delta > -1)$ and $[x]$ denotes the integral part of the real number x.

We will prove this theorem in the case when $k_n = n$. The general case can be proved similarly.

We mention here without proof two lemmas needed in the proof of Theorem 2.

Lemma 1. *Let us suppose that* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *is a sequence of independent random variables and* $\{B_n\}$ *is a sequence of positive numbers tending to infinity, such that* ξ_k/B_n $(k = 1, 2, \ldots, n)$ *are infinitesimal* [2] *and*

$$\lim_{n \to +\infty} \mathbf{P}\left\{\frac{\sum_{k=1}^{n} \xi_k}{B_n} < x\right\} = F(x),$$

where $F(x)$ *is a proper distribution function. Then there exists also a monotonically increasing sequence* $\{C_n\}$ *of positive numbers tending to infinity such that* $B_n/C_n \to 1$ *and*

$$\lim_{n \to +\infty} \mathbf{P}\left\{\frac{\sum_{k=1}^{n} \xi_k}{C_n} < x\right\} = F(x).$$

Lemma 1 is an easy consequence of Theorem 2, p. 155 of the book [2]. By virtue of Lemma 1 we will suppose in the sequel that B_n $(n = 1, 2, \ldots)$ is a monotonically increasing sequence tending to infinity.

Lemma 2. Let $\{B_n\}$ *be a sequence of positive numbers tending monotonically to infinity and suppose that the sequence* $\{B_n\}$ *satisfies condition* (b′) *with* $k_n = n$. *Then*

$$\lim_{\delta\to\infty} \liminf_{n\to+\infty} \frac{B_n}{B_{[n(1+\delta)]}} = 1 .$$

Lemma 2 is almost obvious and therefore we omit here the proof. It follows from Lemma 2 that for any $0 < \varepsilon < 1$ there exist a positive integer n_0 and a positive number δ such that for $n \geqq n_0$

$$1 - \varepsilon < \frac{B_n}{B_{[n(1\pm\delta)]}} < 1 + \varepsilon .$$

This fact will be used in the sufficiency part of the proof of Theorem 2.

§ 2. Proof of Theorem 2

Necessity. Let us suppose that condition (a) holds and that ν_n ($n = 1, 2, \ldots$) is a sequence of positive integer-valued random variables for which ν_n/n converges to 1 in probability. Let us suppose further that

$$\mathbf{P}\left\{\frac{Y_{\nu_n}}{B_n} < x\right\} \to F(x) \qquad \text{as } n \to +\infty$$

at all continuity points of the proper distribution function $F(x)$. Then (b′) necessarily holds.

Let us suppose the contrary. Then

$$\lim_{\delta\to 0} \limsup_{n\to+\infty} \frac{B_n}{B_{[n(1+\delta)]}} = 1$$

does not hold. This means that there exist an $\varepsilon > 0$ and a sequence δ_l ($l = 1, 2, \ldots$) such that $\delta_l \to 0$ and

$$\limsup_{n\to+\infty} \frac{B_n}{B_{[n(1+\delta_l)]}} < 1 - \varepsilon . \tag{1}$$

(The case, when

$$\limsup_{n\to+\infty} \frac{B_n}{B_{[n(1+\delta_l)]}} > 1 + \varepsilon$$

hold, can be treated similarly.)

It follows from (1) that for fixed l there exists an infinite sequence of indices $n_{1l}, n_{2l}, \ldots, n_{kl}, \ldots$ ($l = 1, 2, \ldots$) for which

$$\frac{B_{n_{kl}}}{B_{[n_{kl}(1+\delta_l)]}} < 1 - \frac{\varepsilon}{2} . \tag{2}$$

We have thus for any $l = 1, 2, \ldots$ a monotonically increasing sequence of positive integer-valued indices $\{n_{kl}\}$ ($k = 1, 2, \ldots$) tending to infinity. It is easy to see that one can pick out from these sequences a new sequence m_1, $m_2, \ldots, m_l, \ldots$ for which $m_l < m_{l+1}$ and $m_l \to +\infty$ as $l \to +\infty$.

We define now the random variable $\nu_n (n = 1, 2, \ldots)$ as follows: if $m_l \leqq n < m_{l+1}$ we put

$$\nu_n = \begin{cases} n & \text{with probability } 1/n \\ [n(1+\delta_l)] & \text{with probability } 1 - 1/n\,. \end{cases}$$

It is easy to see that if $n \to +\infty$ then $l \to +\infty$ and thus ν_n/n converges in probability to 1.

Denoting by $F_k(x)$ the distribution function of the random variable Y_k/B_k, the relation

$$\mathbf{P}\left\{\frac{Y_{\nu_n}}{B_n} < x\right\} = \mathbf{P}\left\{\frac{Y_n}{B_n} < x, \nu_n = n\right\} + \mathbf{P}\left\{\frac{Y_{[n(1+\delta_l)]}}{B_n} < x, \nu_n = [n(1+\delta_l)]\right\}$$

gives

$$\mathbf{P}\left\{\frac{Y_{\nu_{m_l}}}{B_n} < x\right\} \leqq F_{[n(1+\delta_l)]}\left(x\frac{B_n}{B_{[n(1+\delta_l)]}}\right) + \frac{1}{n}\,. \tag{3}$$

Now if we take especially $n = m_l$ then for $l = 1, 2, \ldots$ by (2) and (3) we have

$$\mathbf{P}\left\{\frac{Y_{\nu_m}}{B_{m_l}} < x\right\} \leqq F_{[m_l(1+\delta_l)]}\left(x\frac{B_{m_l}}{B_{[m_l(1+\delta_l)]}}\right) + \frac{1}{m_l}\,.$$

Let now l tend to infinity. Then from the preceding inequality we obtain

$$F(x) \leqq F\left(x\left(1 - \frac{\varepsilon}{2}\right)\right)$$

for all continuity points of $F(x)$. This is a contradiction since $F(x)$ is a proper distribution function.

Sufficiency. We will prove the equivalence of conditions (a), (b) and of conditions (a), (b′). Thus we prove that condition (b′) is sufficient, if no further condition than (a) is supposed.

First we shall show that (a) and (b) follow from (a) and (b′). For the sake of brevity and simplicity we restrict our treatment to the case when the variance of random variables ξ_k exists and $Var\,\xi_k = D_k$. We can suppose without loss of generality that the mean value of random variable ξ_k is zero. Now $Y_n = \xi_1 + \xi_2 + \ldots + \xi_n$ and we can put $B_n = \sqrt{D_1^2 + \ldots + D_n^2}$ ([2], p. 153, Theorem 1.). Clearly the probability in condition (b) is larger than

$$\mathbf{P}\{|Y_n - Y_{[n(1-c)]}| + \max_{|n-n'|<cn} |Y_{[n(1-c)]} - Y_{n'}| < \varepsilon B_n\}\,. \tag{4}$$

Thus it is sufficient to show that for arbitrary $\varepsilon > 0$ and $\eta > 0$ there exist a positive integer n_0 and a positive number c such that (4) is larger than $1 - \eta$ if $n \geqq n_0$. For this purpose we will prove that the following inequalities

$$\mathbf{P}\left\{\left|\sum_{k=[n(1-c)]+1}^{n} \xi_k\right| \geqq \frac{\varepsilon}{2} B_n\right\} \leqq \frac{\eta}{2} \tag{5}$$

and

$$\mathbf{P}\left\{\max_{|n-n'|<cn}\left|\sum_{k=[n(1-c)]+1}^{n'}\xi_k\right|\geqq\frac{\varepsilon}{2}B_n\right\}\leqq\frac{\eta}{2} \tag{6}$$

hold for suitably chosen $c > 0$ and $n_0(n_0 \leqq n)$.

By TCHEBYCHEFF's resp. KOLMOGOROFF's inequality we have

$$\mathbf{P}\left\{\left|\sum_{k=[n(1-c)]+1}^{n}\xi_k\right|\geqq\frac{\varepsilon}{2}B_n\right\}\leqq\frac{4}{\varepsilon^2}\left(1-\left(\frac{B_{[n(1-c)]}}{B_n}\right)^2\right) \tag{7}$$

and

$$\mathbf{P}\left\{\max_{|n-n'|<cn}\left|\sum_{k=[n(1-c)]+1}^{n'}\xi_k\right|\geqq\frac{\varepsilon}{2}B_n\right\}\leqq\frac{4}{\varepsilon^2}\,\frac{B^2_{[n(1+c)]}-B^2_{[n(1-c)]}}{B_n^2}. \tag{8}$$

It follows from Lemma 2 that there exist a positive integer n_0 and a positive number c such that for $n \geqq n_0$

$$\left|1-\left(\frac{B_{[n(1\pm c)]}}{B_n}\right)^2\right|\leqq\frac{\varepsilon^2\eta}{16}$$

and thus the right-hand sides of (7) and (8) are smaller than $\eta/2$. We have from (7) and (8)

$$\mathbf{P}\{|Y_n-Y_{[n(1-c)]}|+\max_{|n-n'|<cn}|Y_{[n(1-c)]}-Y_{n'}|\geqq\varepsilon B_n\}\leqq\eta$$

and thus we conclude that (4) is larger than $1-\eta$ if n_0 and $c > 0$ are chosen suitably.

The general case, when the variance of the random variables does not exist, can be treated similarly using the so called "truncation method". We omit here the proof because it is similar to that of the above simple case.

Next we turn to the proof of the second part of the equivalance. We prove that (b′) follows from (a) and (b). First we show that if x is any continuity point of $F(x)$ then

$$F(x)=\lim_{\delta\to0}\liminf_{n\to+\infty}F_{[n(1+\delta)]}\left(x\frac{B_n}{B_{[n(1+\delta)]}}\right)=\lim_{\delta\to0}\limsup_{n\to+\infty}F_{[n(1+\delta)]}\left(x\frac{B_n}{B_{[n(1+\delta)]}}\right). \tag{9}$$

It follows from (b) that for arbitrary chosen $\varepsilon > 0$ and $\eta > 0$ there exist a positive integer n_0 and a positive number c such that if $n \geqq n_0$ and $|\delta| < c$

$$\mathbf{P}\{|Y_{[n(1+\delta)]}-Y_n|<\varepsilon B_n\}>1-\eta.$$

An easy calculation shows that from this inequality

$$\mathbf{P}\left\{\frac{Y_n}{B_n}<x-\varepsilon\right\}-\eta\leqq\mathbf{P}\left\{\frac{Y_{[n(1+\delta)]}}{B_n}<x\right\}\leqq\mathbf{P}\left\{\frac{Y_n}{B_n}<x+\varepsilon\right\}+\eta,$$

or, expressing this by the language of distribution functions,

$$F_n(x-\varepsilon)-\eta \leqq F_{[n(1+\delta)]}\left(x \frac{B_n}{B_{[n(1+\delta)]}}\right) \leqq F_n(x+\varepsilon)+\eta .$$

We obtain from this inequality by virtue of (a) for $n \to +\infty$

$$F(x-\varepsilon)-\eta \leqq \liminf_{n\to+\infty} F_{[n(1+\delta)]}\left(x \frac{B_n}{B_{[n(1+\delta)]}}\right) \leqq \limsup_{n\to+\infty} F_{[n(1+\delta)]}\left(x \frac{B_n}{B_{[n(1+\delta)]}}\right) \leqq$$
$$\leqq F(x+\varepsilon)+\eta .$$

Let now ε and η tend to zero. Then if $\delta \to 0$ we obtain (9). It is easy to see that condition (b′) follows immediately from (9). Q. e. d.

Remarks. Condition (b′) is satisfied if the monotonically increasing sequence $\{B_n\}$ is in the sense of KARAMATA "slowly oscillating", i. e. $B_n = n^a L(n)$, $(\alpha > 0)$, where $L([cn])/L(n) \to 1$ for any positive number c.

Recently A. RÉNYI [3] proved that if $\frac{\nu_n}{n}$ converges in probability to a positive random variable λ, having a discrete distribution and the random variables ξ_k are independent and identically distributed with mean value 0 and variance 1,

$$\lim_{n\to+\infty} \mathbf{P}\left\{\frac{\xi_1+\xi_2+\ldots+\xi_{\nu_n}}{\sqrt{\nu_n}} < x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy .$$

In a following paper we shall generalize this theorem in the case when ν_n/n converges to a positive random variable λ, having arbitrary distribution.

I am indebted to Prof. A. RÉNYI and P. RÉVÉSZ for their valuable remarks and helpful criticism.

(Received May 18, 1961.)

REFERENCES

[1] ANSCOMBE, F. J.: "Large sample theory of sequential estimation". *Proceedings of the Cambridge Philosophical Society* **48** (1952) 600—607.

[2] GNEDENKO, B. V. and KOLMOGOROFF, A. N.: *Limit distributions for sums of independent random variables.* Addison-Wesley Cambridge, 1954.

[3] RÉNYI, A.: "On the central limit theorem for the sum of a random number of independent random variables". *Acta Methematica Academiae Scientiarum Hungaricae* **11** (1960) 97—102.

[4] RÉNYI, A.: "On the asymptotic distribution of the sum of a random number of independent random variables". *Acta Mathematica Academiae Scientiarum Hungaricae* **8** (1957) 193—199.

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ СУММ СЛУЧАЙНОГО ЧИСЛА НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

J. MOGYORÓDI

Резюме

Доказывается следующая

Теорема. *Пусть* $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ *последовательность независимых случайных величин, такая что*

$$\lim_{n\to+\infty} \mathbf{P}\left\{\frac{\xi_1+\ldots+\xi_n}{B_n} < x\right\} = F(x),$$

где $F(x)$ *не вырожденная функция распределения, и случайные величины* $\xi_k/B_n (k=1,2,\ldots,n)$ *бесконечно малы.*

Пусть далее ν_n $(n=1,2,\ldots)$ *последовательность неотрицательных и целочисленных случайных величин, таких что* ν_n/k_n *стремится по вероятности к* 1, *где* k_n *последовательность положительных целых чисел, стремящаяся к бесконечности. Независимость величин* ν_n *от величин* ξ_k *не предполагается. Тогда для того, чтобы функция распределения случайной величины*

$$\frac{\xi_1+\xi_2+\ldots+\xi_{\nu_n}}{B_{k_n}}$$

стремилась к закону распределения $F(x)$, *необходимо и достаточно выполнение условия*

$$\lim_{\delta\to 0}\limsup_{n\to+\infty}\frac{B_{k_n}}{B_{[k_n(1+\delta)]}} = 1 .$$