Acta Math. Acad. Sci. Hung. **7** (1956), pp. 319–326.

ON THE INDEPENDENCE IN THE LIMIT OF SUMS DEPENDING ON THE SAME SEQUENCE OF INDEPENDENT RANDOM VARIABLES

András Prékopa (Budapest) and A. Rényi (Budapest), member of the Academy

Mathematical Institute of the Hungarian Academy of Sciences

Received: September 17, 1956

Introduction

Let ξ_t be a stochastic process with independent increments. Suppose that ξ_t is integer-valued and its sample functions are continuous to the left and have a finite number of discontinuities with probability 1. It can be proved (see [3], Theorem 6) that if ν_k is the number of discontinuities of ξ_t of magnitude k in the time interval $I = [a, b]$, then the random variables ν_k $(k = \pm 1, \pm 2, \dots)$ are independent.[1]

This assertion implies, for example, that a homogeneous composed Poisson process ξ_t may be considered as a superposition of independent ordinary Poisson processes, i.e. can be represented in the form

$$\xi_t = \sum_{k=1}^{\infty} k \xi_t^{(k)},$$

where $\xi_t^{(k)}$ is an ordinary homogeneous Poisson process, and the processes $\xi_t^{(k)}$ are independent (see [4]). For a more general form of this statement see [3].

In § 1 of the present paper we prove a general theorem on the asymptotic independence of certain sums of random variables.

In § 2 deals with the application of our independence theorem leading to a theorem somewhat stronger than that formulated above. Further applications will be given in a forthcoming paper[2] of the first named author.

[1] In [3] the above theorem is formulated more generally.

[2] On stochastic set functions. II, *Acta Math. Acad. Sci. Hung.*, **8** (1957), 337–374.

§ 1. The independence theorem

We start from a double sequence of random variables

$$\xi_{n1}, \xi_{n2}, \ldots, \xi_{nk_n} \qquad (n = 1, 2, \ldots)$$

and suppose always that $\xi_{n1}, \xi_{n2}, \ldots, \xi_{n_k}$ are independent for every $n = 1, 2, \ldots$. Let us consider r Borel measurable real functions $f_1(x), f_2(x), \ldots, f_r(x)$ for which the sets defined by $f_k(x) \neq 0$ are disjoint, or expressed in another way, for which the following relations hold:

$$f_j(x) f_k(x) = 0 \quad \text{for} \quad j \neq k \qquad (j, k = 1, 2, \ldots, r). \tag{1}$$

Let us denote by $\varphi_{lk}^{(n)}(u)$ the characteristic function of the random variable $f_l(\xi_{nk})$, further let us put

$$\zeta_l^{(n)} = \sum_{k=1}^{k_n} f_l(\xi_{nk}) \qquad (l = 1, 2, \ldots, r;\ n = 1, 2, \ldots).$$

In order to simplify the understanding of the phenomenon which is described by our theorem, we formulate it first for a special case.

Theorem 1.a *Let us suppose that the following conditions hold:*

a) *The real, Borel measurable functions $f_l(x)$ $(l = 1, 2, \ldots, r)$ are integer-valued and satisfy* (1).

b) *For every l $(1 \leq l \leq r)$ the random variables*

$$f_l(\xi_{n1}), f_l(\xi_{n2}), \ldots, f_l(\xi_{nk_n})$$

are infinitesimal, i.e.

$$\lim_{n \to \infty} \sup_{1 \leq k \leq k_n} \mathbb{P}(f_l(\xi_{nk}) \neq 0) = 0.$$

c.) *For every l $(1 \leq l \leq r)$ the limiting distribution of the random variables $\zeta_l^{(n)}$ exists:*

$$F_i(x_i) = \lim_{n \to \infty} \mathbb{P}(\zeta_i^{(n)} < x_i) \qquad (1 \leq i \leq r), \tag{2}$$

at every point of continuity of $F_i(x)$.

Under these conditions the random variables $\zeta_1^{(n)}, \zeta_2^{(n)}, \ldots, \zeta_r^{(n)}$ are asymptotically independent, i.e.

$$\lim_{n \to \infty} \mathbb{P}(\zeta_1^{(n)} < x_1, \zeta_2^{(n)} < x_2, \ldots, \zeta_r^{(n)} < x_r) = F_1(x_1) F_2(x_2) \ldots F_r(x_r) \tag{3}$$

if x_i is a continuity point of $F_i(x)$ $(i = 1, 2, \ldots, r)$.

Proof. Let us consider the characteristic function of the joint distribution of the random variables $\zeta_l^{(n)}$ $(l = 1, 2, \ldots, r)$. Taking the relation (1) into account it can easily be seen by comparing the coefficients on both sides that the r-dimensional characteristic function of $\zeta_1^{(n)}, \ldots, \zeta_r^{(n)}$ is the following:

$$\sum_{j_1, j_2, \ldots, j_r} \mathbb{P}(\zeta_1^{(n)} = j_1, \ldots, \zeta_r^{(n)} = j_r) e^{i(u_1 j_1 + \cdots + u_r j_r)}$$

$$= \prod_{k=1}^{k_n} \left\{ 1 + \sum_s \mathbb{P}(f_1(\xi_{nk}) = s)(e^{iu_1 s} - 1) + \cdots + \sum_s \mathbb{P}(f_r(\xi_{nk}) = s)(e^{iu_r s} - 1) \right\}. \tag{4}$$

It follows from (4) that (denoting by $\mathbb{M}(\chi)$ the expectation of χ)

$$\mathbb{M}(e^{i(\zeta_1^{(n)} u_1 + \cdots + \zeta_r^{(n)} u_r)}) = \prod_{k=1}^{k_n} \left\{ 1 + \left(\varphi_{1k}^{(n)}(u_1) - 1 \right) + \cdots + \left(\varphi_{rk}^{(n)}(u_r) - 1 \right) \right\}. \tag{5}$$

Conditions a), b) and c) imply that the limits

$$\Phi_l(u_l) = \lim_{n \to \infty} \sum_{k=1}^{k_n} (\varphi_{lk}^{(n)}(u_l) - 1) \qquad (l = 1, 2, \ldots, r) \tag{6}$$

exist (see [1], § 24, Theorem 1) and $e^{\Phi_l(u_l)}$ is the characteristic function of the limiting distribution $F_l(x_l)$ $(l = 1, 2, \ldots, r)$. Moreover, by Condition b) we have

$$\lim_{n \to \infty} \sum_{k=1}^{k_n} |1 - \varphi_{lk}^{(n)}(u_l)|^2 = 0 \qquad (l = 1, 2, \ldots, r). \tag{7}$$

According to (6) and (7) the sequence (5) converges to the r-dimensional characteristic function

$$e^{\Phi_1(u_1)} \cdots e^{\Phi_r(u_r)}$$

and thus relation (3) holds.

A heuristic argument in favour of Theorem 1.a can be given as follows: Our suppositions a), b) and c) imply that in general only a small number of terms of the sum $\zeta_l^{(n)} = \sum_{k=1}^{k_n} f_l(\xi_{nk})$ are different from 0 for each l. Supposition (1) ensures that the sums $\zeta_l^{(n)}$ $(l = 1, 2, \ldots, r)$ will always depend on disjoint subsets of the independent random variables $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nk_n}$; of course, these sets are random, and therefore the sums $\zeta_l^{(n)}$ are not independent, only almost independent. Nevertheless in the limit their dependence disappears.

The suppositions of Theorem 1.a may be replaced by a set of more special suppositions which, however, have the advantage that no supposition restricts at the same time the choice of the random variables ξ_{nk} and the choice of the functions $f_l(x)$, as there are two distinct groups of suppositions, further the convergence of the distribution of $\zeta_l^{(n)}$ is not postulated, but is a consequence of the suppositions. This weaker form of Theorem 1.a is expressed by the following

Corollary *Let $\xi_{n1}, \xi_{n2}, \dots, \xi_{nk_n}$ denote a double sequence of independent non-negative integer-valued random variables which are infinitesimal, i.e.*

$$\lim_{n\to\infty} \max_{1\le k\le k_n} \mathbb{P}(\xi_{nk} \ne 0) = 0.$$

Let $E_1, E_2, \dots, E_r$ denote disjoint subsets of the set of positive integers and let us suppose that $f_l(k)$ $(l = 1, 2, \dots, r;\ k = 0, 1, \dots)$ are non-negative integer-valued functions such that $f_l(0) = 0$ and $f_l(k) = 0$ if $k \notin E_l$. Let us put $p_{nks} = \mathbb{P}(\xi_{nk} = s)$, $C_{ns} = \sum_{k=0}^{k_n} p_{nks}$ and suppose that there exists a convergent series of non-negative numbers $\sum_{s=1}^{\infty} C_s$ such that

$$\lim_{n\to\infty} \sum_{s=1}^{\infty} |C_{ns} - C_s| = 0.$$

It follows that putting

$$\zeta_l^{(n)} = \sum_{k=1}^{k_n} f_l(\xi_{nk}) \qquad (l = 1, 2, \dots, r;\ n = 1, 2, \dots)$$

we have

$$\lim_{n\to\infty} \mathbb{P}(\zeta_1^{(n)} < x_1,\ \zeta_2^{(n)} < x_2, \dots, \zeta_r^{(n)} < x_r) = F_1(x_1)F_2(x_2)\dots F_r(x_r),$$

where the distribution function $F_k(x)$ has the generating function

$$\exp \sum_{s=1}^{\infty} C_s(z^{f_k(s)} - 1).$$

To prove that this Corollary really follows from Theorem 1.a, we have to apply Theorem 3 of the paper [2].

Now we turn to the general case in which the first part of Condition a) of Theorem 1.a is dropped. Our statement is expressed by

Theorem 1.b *Let us suppose that the following conditions hold:*

a) *The Borel measurable real functions $f_l(x)$ $(l = 1, 2, \dots, r)$ satisfy* (1).

b) *For every l $(1 \le l \le r)$*

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} |\varphi_{lk}^{(n)}(u_l) - 1|^2 = 0.$$ [3]

[3] It can be seen that if Conditions c) and d) hold, then Condition b) holds also if for some $\tau > 0$

$$\lim_{n\to\infty} \sum_{k=1}^{k_n} |\alpha_{lk}^{(n)}|^2 = 0,$$

where

$$\alpha_{lk}^{(n)} = \int_{|x|<\tau} x \, \mathrm{d}F_{lk}^{(n)}(x), \qquad F_{lk}^{(n)}(x) = \mathbb{P}(f_l(\xi_{nk}) < x).$$

c) *For every l $(1 \le l \le r)$ the random variables*

$$f_l(\xi_{n1}), f_l(\xi_{n2}), \dots, f_l(\xi_{nk_n})$$

are infinitesimal, i.e. for every $\varepsilon > 0$

$$\lim_{n\to\infty} \sup_{1\le k\le k_n} \mathbb{P}(|f_l(\xi_{nk})| > \varepsilon) = 0.$$

d) *For every l $(1 \le l \le r)$ the limiting distribution of the random variables $\zeta_l^{(n)}$ exists.*

Under these conditions the random variables $\zeta_1^{(n)}, \zeta_2^{(n)}, \dots, \zeta_r^{(n)}$ are asymptotically independent, i.e. relation (2) holds.

Proof. First we observe that (5) holds without the restriction that the $f_l(x)$ are integer-valued. This can be shown as follows: By virtue of the independence of the variables ξ_{nk} we obtain

$$\mathbb{M}\left(e^{i\sum_{l=1}^r u_l \zeta_l^{(n)}}\right) = \prod_{k=1}^{k_n} \mathbb{M}\left(e^{i\sum_{l=1}^r u_l f_l(\xi_{nk})}\right). \tag{8}$$

Let $A_{lk}^{(n)}$ denote the event consisting in that $f_l(\xi_{nk}) \neq 0$. Then we have[4]

$$\mathbb{M}\left(e^{i\sum_{l=1}^r u_l f_l(\xi_{nk})}\right) = \sum_{\nu=1}^{r}\left(\mathbb{M}\left(e^{i\sum_{l=1}^r u_l f_l(\xi_{nk})} \,\middle|\, A_{\nu k}^{(n)}\right) - 1\right)\mathbb{P}(A_{\nu k}^{(n)}) + 1. \tag{9}$$

As the event $A_{\nu k}^{(n)}$ implies $f_l(\xi_{nk}) = 0$ for $l \neq \nu$, we have

$$\mathbb{M}\left(e^{i\sum_{l=1}^r u_l f_l(\xi_{nk})} \,\middle|\, A_{\nu k}^{(n)}\right) = \mathbb{M}(e^{iu_\nu f_\nu(\xi_{nk})} \mid A_{\nu k}^{(n)}). \tag{10}$$

On the other hand,

$$[\mathbb{M}(e^{iu_\nu f_\nu(\xi_{nk})} \mid A_{\nu k}^{(n)}) - 1]\mathbb{P}(A_{\nu k}) = \varphi_{\nu k}^{(n)}(u_\nu) - 1. \tag{11}$$

Thus (5) follows from (8)–(11).

Condition d) implies the existence of

$$\Psi_l(u_l) = \lim_{n\to\infty} \prod_{k=1}^{k_n} \varphi_{lk}^{(n)}(u_l) \qquad (l = 1, 2, \dots, r). \tag{12}$$

As $\Psi_l(u_l)$ is the characteristic function of an infinitely divisible distribution (see [1], § 24, Theorem 2), we have

$$\Psi_l(u_l) \neq 0 \qquad (l = 1, 2, \dots, r)$$

[4] $\mathbb{M}(\eta \mid A)$ denotes the conditional expectation of η under the condition A.

(see [1], § 17, Theorem 1). It follows hence and from (12) that if $|\varphi_{lk}^{(n)}(u_l)-1| \le \dfrac{1}{2}$, then

$$\begin{aligned} &\left|\log \Psi_l(u_l) - \sum_{k=1}^{k_n}(\varphi_{lk}^{(n)}(u_l)-1)\right| \\ (13)\qquad &\le \left|\log \Psi_l(u_l) - \prod_{k=1}^{k_n}\varphi_{lk}^{(n)}(u_l)\right| + \sum_{k=1}^{k_n}|\varphi_{lk}^{(n)}(u_l)-1|^2 \qquad (l=1,2,\dots,r). \end{aligned}$$

The member on the right-hand side of (13) tends to 0, hence

$$\Phi_l(u_l) = \log \Psi_l(u_l) = \lim_{n\to\infty}\sum_{k=1}^{k_n}(\varphi_{lk}^{(n)}(u_l)-1) \qquad (l=1,2,\dots,r). \tag{14}$$

By (5) (14) and Condition b) it follows finally

$$\lim_{n\to\infty}\mathbb{M}(e^{i(\zeta_1^{(n)}u_1+\ldots+\zeta_r^{(n)}u_r)}) = \prod_{l=1}^{r} e^{\Phi_l(u_l)}.$$

Thus Theorem 1.b is proved. □

§ 2. Application to stochastic processes

In this § we consider a stochastic process with independent increments ξ_t. For the sake of simplicity we suppose that ξ_t is defined in the time interval $[0,1]$. We suppose furthermore that the sample functions of ξ_t are continuous to the left for $0 \le t \le 1$, with probability 1. Let $\nu(I)$ denote the random variable giving the number of discontinuities of ξ_t of magnitudes $h \in 1$. We prove the following

Theorem 2 *If the process ξ_t is weakly continuous, i.e. for every $\varepsilon > 0$*

$$\lim_{\Delta t\to 0}\mathbb{P}(|\xi_{t+\Delta t}-\xi_t| \ge \varepsilon) = 0 \tag{15}$$

uniformly in t and $I_1, I_2, \dots, I_r$ are pairwise disjoint intervals with positive distances from the point 0, then the random variables

$$\nu(I_1), \nu(I_2), \dots, \nu(I_r)$$

are independent.

Proof. Let $f_l(x)$ denote the characteristic function (in the sense of set theory) of the interval I_l. We define the random variables

$$\eta_{n,k+1} = \xi_{\frac{k+1}{n}} - \xi_{\frac{k}{n}} \qquad (k=0,1,2,\dots,n-1). \tag{16}$$

Obviously,

$$\mathbb{P}\left(\nu(I_l) = \lim_{n\to\infty} \sum_{k=1}^{n} f_l(\eta_{n,k})\right) = 1, \tag{17}$$

hence Condition c) of Theorem 1.a is satisfied. Since

$$\mathbb{P}(f_l(\eta_{n,k+1}) \neq 0) \leq \mathbb{P}\left(\left|\xi_{\frac{k+1}{n}} - \xi_{\frac{k}{n}}\right| \geq \delta\right),$$

where δ is the minimal distance of the intervals I_l from the point 0, the random variables

$$f_l(\eta_{n,1}), f_l(\eta_{n,2}), \ldots, f_l(\eta_{n,n})$$

are infinitesimal for every l. As Condition a) is obviously satisfied, the relations (2) and (17) imply our assertion. $\square$

If instead of the intervals $I_1, I_2, \ldots, I_r$ we choose pairwise disjoint Borel measurable sets with positive distances from the point 0, then Theorem 2 holds obviously without any change. By choosing for $f_l(x)$ other functions, further results can be obtained this way. For related results see [5].

References

[1] Gnedenko, B. V. and A. N. Kolmogorov (1949). *Predel'nüe raszpredelenija dlja szumm nezaviszimüh szlucsajnüh velicsin*, Moszkva–Leningrad.

[2] Rényi, A. (1951). On Composed Poisson Distributions. II, *Acta Math. Acad. Sci. Hung.*, **2**, 83–98.

[3] Prékopa, A. (1957). On Poisson and Composed Poisson Stochastic Set Functions, *Studia Math.*, **16**, 142–157.

[4] Jánossy, L., J. Aczél and A. Rényi (1950). On Composed Poisson Distributions. I, *Acta Math. Acad. Sci. Hung.*, **1**, 209–224.

[5] Ito, K. (1942). Stochastic Processes. I, *Jap. J. Math.*, **18**, 261–301.