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

By

A. PRÉKOPA (Budapest) and A. RÉNYI (Budapest), member of the Academy

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 $ν_k (k = \pm 1, \pm 2, \ldots)$ 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

$$ξ_t = \sum_{k=1}^{∞} k ξ_t^{(k)}$$

where $ξ_t^{(k)}$ is an ordinary homogeneous Poisson process, and the processes $ξ_t^{(k)}$ 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.

§ 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. The independence theorem

We start from a double sequence of random variables

$$ξ_{n1}, ξ_{n2}, \ldots, ξ_{nk_n} \quad (n = 1, 2, \ldots)$$

\(^1\) In [3] the above theorem is formulated more generally.

and suppose always that \( \tilde{\xi}_{n_1}, \tilde{\xi}_{n_2}, \ldots, \tilde{\xi}_{n_k} \) 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 \( f_i(x) = 0 \) are disjoint, or expressed in another way, for which the following relations hold:

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

Let us denote by \( q^{(n)}_{\tilde{\xi}}(u) \) the characteristic function of the random variable \( f_l(\tilde{\xi}_{nk}) \), further let us put

\[
\tilde{\xi}^{(n)}_l = \sum_{k=1}^{k_n} f_l(\tilde{\xi}_{nk}) \quad (l = 1, 2, \ldots, r; n = 1, 2, \ldots).
\]

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

**Theorem 1a.** 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

\[
f_l(\tilde{\xi}_{n_1}), f_l(\tilde{\xi}_{n_2}), \ldots, f_l(\tilde{\xi}_{n_k})
\]

are infinitesimal, i.e.

\[
\lim_{n \to \infty} \sup_{1 \leq k \leq k_n} P(f_l(\tilde{\xi}_{nk}) = 0) = 0.
\]

c) For every \( l \) \( (1 \leq l \leq r) \) the limiting distribution of the random variables \( \tilde{\xi}^{(n)}_l \) exists:

\[
(2) \quad F_l(x) = \lim_{n \to \infty} P(\tilde{\xi}^{(n)}_l < x) \quad (1 \leq l \leq r),
\]

at every point of continuity of \( F_l(x) \).

Under these conditions the random variables \( \tilde{\xi}^{(n)}_1, \tilde{\xi}^{(n)}_2, \ldots, \tilde{\xi}^{(n)}_r \) are asymptotically independent, i.e.

\[
(3) \quad \lim_{n \to \infty} P(\tilde{\xi}^{(n)}_1 < x_1, \tilde{\xi}^{(n)}_2 < x_2, \ldots, \tilde{\xi}^{(n)}_r < x_r) = F_1(x_1) F_2(x_2) \ldots F_r(x_r)
\]

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 \( \tilde{\xi}^{(n)}_l \) \( (l = 1, 2, \ldots, r) \). 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 \( \tilde{\xi}^{(n)}_1, \ldots, \tilde{\xi}^{(n)}_r \) is the following:

\[
(4) \quad \sum_{j_1, j_2, \ldots, j_r} P(\tilde{\xi}^{(n)}_1 = j_1, \ldots, \tilde{\xi}^{(n)}_r = j_r) e^{i(u_1 j_1 + \ldots + u_r j_r)} =
\]

\[
= \prod_{k=1}^{k_n} \left[ 1 + \sum_s P(f_1(\tilde{\xi}_{nk}) = s)(e^{is} - 1) + \cdots + \sum_s P(f_r(\tilde{\xi}_{nk}) = s)(e^{is} - 1) \right].
\]
It follows from (4) that (denoting by $M(\chi)$ the expectation of $\chi$)

$$\mathbf{M}(e^{\sum_{i=1}^n \varepsilon_i u_i}) = \prod_{k=1}^{b_n} \{1 + (q_{ik}^n(u_i) - 1) + \cdots + (q_{ikk}^n(u_i) - 1)\}.\tag{5}$$

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

$$\Phi_l(u_l) = \lim_{n \to \infty} \sum_{l=1}^{b_n} (q_{ik}^n(u_l) - 1) \quad (l = 1, 2, \ldots, r)\tag{6}$$

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

$$\lim_{n \to \infty} \sum_{l=1}^{b_n} \left|1 - q_{ik}^n(u_l)\right|^2 = 0 \quad (l = 1, 2, \ldots, r).\tag{7}$$

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

The suppositions of Theorem 1a 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 $\xi_{nk}$ and the choice of the functions $f_l(x)$, as there are two distinct groups of suppositions, further the convergence of the distribution of $\varepsilon_i^{(n)}$ is not postulated, but is a consequence of the suppositions. This weaker form of Theorem 1a is expressed by the following

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

$$\lim_{n \to \infty} \max_{1 \leq k \leq b_n} \mathbf{P}(\xi_{nk} \neq 0) = 0.$$

Let $E_1, E_2, \ldots, E_r$ denote disjoint subsets of the set of positive integers and let us suppose that $f_l(k)$ ($l = 1, 2, \ldots, r; k = 0, 1, \ldots$) are non-negative in-
teger-valued functions such that \( f_i(0) = 0 \) and \( f_i(k) = 0 \) if \( k \notin E_i \). Let us put \( p_{nks} = P(\xi_{nk} = s) \), \( C_{ns} = \sum_{k=0}^{K} p_{nks} \) 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

\[
\gamma_l^{(n)} = \sum_{k=1}^{K} f_i(\xi_{nk}) \quad (l = 1, 2, \ldots, r; n = 1, 2, \ldots)
\]

we have

\[
\lim_{n \to \infty} \mathbf{P}(\gamma_1^{(n)} < x_1, \gamma_2^{(n)} < x_2, \ldots, \gamma_r^{(n)} < x_r) = F_1(x_1)F_2(x_2)\ldots F_r(x_r)
\]

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

\[
\exp \sum_{s=1}^{\infty} C_s (\gamma_k^{(s)} - 1).
\]

To prove that this Corollary really follows from Theorem 1a, 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 1a is dropped. Our statement is expressed by

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

a) The Borel measurable real functions \( f_i(x) \) \((1 \leq l \leq r)\)

b) For every \( l \) \((1 \leq l \leq r)\)

\[
\lim_{n \to \infty} \sum_{k=1}^{K} |q_{lk}^{(n)}(u_l) - 1|^2 = 0.\]

c) For every \( l \) \((1 \leq l \leq r)\) the random variables

\[
\tilde{f}_i(\xi_{n1}), \tilde{f}_i(\xi_{n2}), \ldots, \tilde{f}_i(\xi_{nk_r})
\]

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

\[
\lim_{n \to \infty} \sup_{1 \leq k \leq k_r} \mathbf{P}(|f_i(\tilde{\xi}_{nk})| > \varepsilon) = 0.
\]

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

\[
\lim_{n \to \infty} \sum_{k=1}^{K} |a_{lk}^{(n)}|^2 = 0,
\]

where

\[
a_{lk}^{(n)} = \int_{[x]} x dF_{lk}^{(n)}(x), \quad F_{lk}^{(n)}(x) = \mathbf{P}(f_i(\tilde{\xi}_{nk}) < x).
\]
d) For every \( l \) (\( 1 \leq l \leq r \)) the limiting distribution of the random variables \( z_{il}^{(n)} \) exists.

Under these conditions the random variables \( z_{i1}^{(n)}, z_{i2}^{(n)}, \ldots, z_{ir}^{(n)} \) are asymptotically independent, i.e. relation (2) holds.

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

\[
M\left( e^{i \sum_{l=1}^{r} n_l f_l(\xi_{nk})} \right) = \prod_{k=1}^{k_n} M\left( e^{i \sum_{l=1}^{r} n_l f_l(\xi_{nk})} \right). 
\]

Let \( A_{pk}^{(n)} \) denote the event consisting in that \( f_i(\xi_{nk}) = 0 \). Then we have

\[
M\left( e^{i \sum_{l=1}^{r} n_l f_l(\xi_{nk})} \right) = \sum_{r=1}^{r_n} \left( M\left( e^{i \sum_{l=1}^{r} n_l f_l(\xi_{nk})} \bigg| A_{pk}^{(n)} \right) - 1 \right) P(A_{pk}^{(n)}) + 1. 
\]

As the event \( A_{pk}^{(n)} \) implies \( f_i(\xi_{nk}) = 0 \) for \( l = r \), we have

\[
M\left( e^{i \sum_{l=1}^{r} n_l f_l(\xi_{nk})} \bigg| A_{pk}^{(n)} \right) = M(e^{i \sum_{l=1}^{r} n_l f_l(\xi_{nk})} \bigg| A_{pk}^{(n)}). 
\]

On the other hand,

\[
[M(e^{i \sum_{l=1}^{r} n_l f_l(\xi_{nk})} \bigg| A_{pk}^{(n)}) - 1] P(A_{pk}^{(n)}) = \varphi_{pk}^{(n)}(u_r) - 1. 
\]

Thus (5) follows from (8)–(11).

Condition d) implies the existence of

\[
\Psi_i(u_l) = \lim_{n \to \infty} \prod_{k=1}^{k_n} \varphi_{lk}^{(n)}(u_l) \quad (l = 1, 2, \ldots, r).
\]

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

\[
\Psi_i(u_l) = 1 \quad (l = 1, 2, \ldots, r)
\]

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

\[
| \log \Psi_i(u_l) - \sum_{k=1}^{k_n} (\varphi_{lk}^{(n)}(u_l) - 1) | \leq
\]

\[
\leq | \log \Psi_i(u_l) - \prod_{k=1}^{k_n} \varphi_{lk}^{(n)}(u_l) | + \sum_{k=1}^{k_n} | \varphi_{lk}^{(n)}(u_l) - 1 |^2 \quad (l = 1, 2, \ldots, r).
\]

\(^1\) \( M(\xi | A) \) denotes the conditional expectation of \( \xi \) under the condition \( A \).
The member on the right-hand side of (13) tends to 0, hence

\begin{equation}
\varphi_i(u_l) = \log \Psi_i(u_l) = \lim_{n \to \infty} \sum_{k=1}^{k_n} (q^{(n)}_{lk}(u_l) - 1) \quad (l = 1, 2, \ldots, r).
\end{equation}

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

\[ \lim_{n \to \infty} M(e^{\varphi_i(u_{1i}, \ldots, u_{ri})}) = \prod_{i=1}^{r} e^{\varphi_i(u_i)}. \]

Thus Theorem 1b is proved.

§ 2. Application to stochastic processes

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

**Theorem 2.** If the process \( \xi_t \) is weakly continuous, i.e., for every \( \varepsilon > 0 \)

\begin{equation}
\lim_{M \to 0} P(\left| \xi_{t+M} - \xi_t \right| > \varepsilon) = 0
\end{equation}

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

\[ r(I_1), r(I_2), \ldots, r(I_r) \]

are independent.

**Proof.** Let \( f_i(x) \) denote the characteristic function (in the sense of set theory) of the interval \( I_i \). We define the random variables

\begin{equation}
\eta_{n, k+1} = \frac{\xi_{k+1} - \xi_k}{n} \quad (k = 0, 1, 2, \ldots, n-1).
\end{equation}

Obviously,

\begin{equation}
P\left( r(I_i) = \lim_{n \to \infty} \sum_{k=1}^{n} f_i(\eta_{n, k}) \right) = 1,
\end{equation}

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

\[ P(f_i(\eta_{n, k+1}) = 0) \leq P\left( \left| \xi_{k+1} - \xi_k \right| \geq \delta \right) \]

where \( \delta \) is the minimal distance of the intervals \( I_i \) from the point 0, the random variables

\[ f_i(\eta_{n, 1}), f_i(\eta_{n, 2}), \ldots, f_i(\eta_{n, n}) \]
are infinitesimal for every $l$. As Condition a) is obviously satisfied, the relations (2) and (17) imply our assertion.

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 $f_i(x)$ other functions, further results can be obtained this way. For related results see [5].

MatheMatiCal Institute
Of The Hungarian Academy Of Sciences

(Received 17 September 1956)

Bibliography


О СОВМЕСТНОМ ПРЕДЕЛЬНОМ РАСПРЕДЕЛЕНИИ СУММ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

А. Прекопа и А. Реньи (Будапешт)

(Резюме)

Пусть $\tilde{z}_{n1}, \tilde{z}_{n2}, \ldots, \tilde{z}_{nk_z}$ — последовательность сумм независимых случайных величин, $a f_1(x), f_2(x), \ldots, f_k(x)$ вещественные и измеримые по Борелю функции таковы, что $f_i(x)f_j(x) = 0$, если $i \neq k$.

Теорема 1а. Предположим, что выполняются следующие условия:

a) Функции $f_i(x)$ ($l = 1, 2, \ldots, r$) принимают лишь целые значения;

b) для всех $l$ ($l = 1, 2, \ldots, r$) случайные величины $f_l(z_{n1}), f_l(z_{n2}), \ldots, f_l(z_{nk_z})$ бесконечно малы (см. [1], § 20);

c) для всех $l$ ($l = 1, 2, \ldots, r$) существует предельное распределение последовательности случайных величин

$$ s^{(n)}_{\frac{1}{k_l}} = \sum_{k=1}^{k_n} f_l(z_{nk}). $$
Из этих условий следует предельная независимость случайных величин $z_{ij}^{(n)} (l = 1, 2, \ldots, r)$ при $n \to \infty$, т. е. выполнение соотношения (3), где функция $F_i(x)$ дается формулой (2).

Если вместо условия а) требовать выполнения условия

$$\lim_{n \to \infty} \frac{1}{k_n} \sum_{k=1}^{k_n} \left| \varphi_{ik}^{(n)}(u_i) - 1 \right|^2 = 0,$$

где $\varphi_{ik}^{(n)}$ характеристическая функция случайной величины $f_i(x_{nk})$, то наше утверждение остается в силе. Это утверждается в теореме 1b.

Теорема 2, доказываемая с помощью упомянутых теорем, утверждает, что если реализации процесса с независимыми приращениями $\xi_i$, удовлетворяющими условию (15), суть непрерывные слева ступенчатые функции, то числа скачков, попадающих в интервалы без общих точек, находящихся от точки 0 на положительном расстоянии, являются независимыми случайными величинами.