A central limit theorem for equivalent random variables

By P. RÉVÉSZ (Budapest)

Introduction

The author of the present paper gave a talk in the Seminar of the Mathematical Institute of the Hungarian Academy of Sciences at November 17, 1964 in the presence of Professor A. N. KOLMOGOROV. In this talk Theorem 2 of this paper was proved under the condition that the equivalent random variables are uniformly bounded. KOLMOGOROV proposed a way to prove Theorem 2 in this stronger form. In this paper we follow the advice of KOLMOGOROV to prove Theorem 2.

First of all let us mention some definitions.

Definition 1. (See [1], [2], [3]). The events $A_1, A_2, \dots$ are called equivalent if the probability of the event $A_{i_1}A_{i_2}\dots A_{i_k}$ ($i_j \neq i_l$ if $j \neq l$) depends only on k and it does not depend on the indices $i_1, i_2, \dots, i_k$. The numbers

$$\alpha_k = P(A_{i_1}A_{i_2}\dots A_{i_k}) \qquad (k=1, 2, \dots)$$

are called the moments of the sequence $A_1, A_2, \dots$.

Definition 2. The random variables $\xi_1, \xi_2, \dots$ are equivalent if the joint distribution function of the random variables $\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_k}$ ($i_j \neq i_l$ if $j \neq l$) depends only on k and it does not depend on the indices $i_1, i_2, \dots, i_k$ i. e.

$$P\{\xi_{i_1} < x_1, \xi_{i_2} < x_2, \dots, \xi_{i_k} < x_k\} = F_k(x_1, x_2, \dots, x_k) \qquad (k=1, 2, \dots)$$

In our paper [4] we studied the properties of equivalent events. Our main result is the following: *if the indicator function of A_k is $a_k(\omega)$, i. e.*[1])

$$a_k(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in A_k \\ 0 & \text{if} \quad \omega \notin A_k \end{cases} \qquad (k=1, 2, \dots)$$

then the sequence a_k converges weakly to a random variable $\lambda(\omega)$ in the Hilbert-space of the square integrable random variables, i. e.

$$M(a_k\eta) \to M(\lambda\eta) \qquad (k \to \infty)$$

[1]) All random variables and events are defined in a probability space $\{\Omega, S, P\}$, i. e. all events are elements of the σ-algebra S, all random variables are measurable functions in the space $\{\Omega, S\}$. So $a_k(\omega)$ is also a random variable.

for any square integrable η, and the events $A_1, A_2, \ldots$ are independent under the condition that λ takes a fixed value, i. e.

$$P(A_{i_1}A_{i_2}\ldots A_{i_k}|\lambda) = P(A_{i_1}|\lambda)\,P(A_{i_2}|\lambda)\ldots P(A_{i_k}|\lambda) = \lambda^k$$

with probability 1.

Using this result the full characterization of the sequences of equivalent events is very easy.

The study of equivalent random variables is not so simple. The aim of § 2 is to generalize the above mentioned theorem of [4]. (Theorem 1). Actually this generalization already was mentioned in [4] in a not very exact form without any proof.

In § 3 we prove our central limit theorem using Theorem 1.

§ 1. Lemmas and notations

In this § we prove some simple lemmas and introduce some notations. These lemmas will be used in § 2 and § 3, but some of them are interesting in themselves.

Our first lemma is known (see [5], [6]).

Lemma 1. *If H is a Hilbert space and $\{f_n\}$ is a sequence of elements of H such that*

$$\lim_{n\to\infty}(f_n, f_k) = \lambda_k \qquad (k = 1, 2, \ldots)$$

and

$$\|f_n\| \leqq C \qquad (n = 1, 2, \ldots)$$

where C is a positive constant and $\{\lambda_k\}$ is a sequence of real numbers. Then f_n converges weakly to an element f of the Hilbert space H, i. e.

$$(f_n, g) \to (f, g) \qquad (n \to \infty)$$

for every element g of H.

The next lemma is a trivial consequence of Lemma 1.

Lemma 2. *Let $\xi_1, \xi_2, \ldots$ be a sequence of equivalent random variables with finite variances, then ξ_n converges weakly to a random variable μ in the Hilbert space of square integrable random variables, i. e.*

$$M(\xi_n\eta) \to M(\mu\eta) \qquad (n \to \infty)$$

for any square integrable random variable η.

Lemma 3. *If the random variable $\psi(\omega)$ is a symmetric Baire function of the equivalent random variables $\xi_1, \xi_2, \ldots$ and $g(x, y)$ is a Borel measurable function defined on the Euclidean plane then $g(\xi_1, \psi), g(\xi_2, \psi), \ldots$ is a sequence of equivalent random variables.*

PROOF. The distribution functions [2] $F_1(x_1)$, $F_2(x_1, x_2)$, ... uniquely determine the distribution function of $\psi(\omega)$ and the joint distribution function of $g(\xi_{i_1}, \psi)$, $g(\xi_{i_2}, \psi)$, ..., $g(\xi_{i_n}, \psi)$ which does not depend on the indices $i_1, i_2, \ldots, i_k$

Lemma 4. *Let $\xi_1, \xi_2, \ldots$ be a sequence of equivalent random variables with finite variances, then*

$$P\left\{\frac{\xi_1+\xi_2+\ldots+\xi_n}{n} \to \mu\right\} = 1 \tag{1}$$

and

$$\frac{(\xi_1-\mu)^2+(\xi_2-\mu)^2+\ldots+(\xi_n-\mu)^2}{n} \tag{2}$$

converges to an integrable random variable $\sigma^2(\omega)$ with probability 1, *where μ is the weak limit of ξ_n.*

PROOF. (1) and (2) are trivial consequences of the Birkhoff's individual ergodic theorem and the Lemma 3. A very different proof of (1) can be found in [7].

Let $\xi_1, \xi_2, \ldots$ be a sequence of equivalent random variables, we use the following notations:

1. $A_n(x)$ is the event that $\xi_n < x$, i. e. $A_n(x)$ is the set of those points $\omega \in \Omega$ for which $\xi_n(\omega) < x$.
2. $a_n^{(x)}(\omega)$ is the indicator function of $A_n(x)$, i. e.

$$a_n^{(x)}(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in A_n(x) \\ 0 & \text{if} \quad \omega \notin A_n(x). \end{cases}$$

Our next lemma characterizes the behaviour of the weak limits of the sequences $a_n^{(x)}(\omega)$.

Lemma 5. *We can define a stochastic process $\lambda_x(\omega)$ $(-\infty < x < \infty)$ such that*

1° *λ_x is the weak limit of $\{a_n^{(x)}\}_{n=1}^{\infty}$ $(-\infty < x < \infty)$*
2° *$\lambda_x(\omega)$ is a distribution function for each $\omega \in \Omega$.*

PROOF. Evidently

$$P\{\lambda_x \leqq \lambda_y\} = 1$$

if $x < y$. Let the sequence of rational numbers be $r_1, r_2, \ldots$ We define the random variable λ_{r_1} as any [3] weak limit of the sequence $\{a_n^{(r_1)}\}_{n=1}^{\infty}$. If λ_{r_1} is already defined for each ω, then we can define the random variable λ_{r_2} as that weak limit of $\{a_n^{(r_2)}\}_{n=1}^{\infty}$ what is not larger (not smaller) than λ_{r_1} everywhere if $r_2 < r_1$ $(r_1 < r_2)$ respectively. If $\lambda_{r_1}, \lambda_{r_2}, \ldots, \lambda_{r_k}$ is already defined then we can define the random variable $\lambda_{r_{k+1}}$ such that

a) $\lambda_{r_{k+1}}$ is the weak limit of $\{a_n^{(r_{k+1})}\}_{n=1}^{\infty}$

b) $\lambda_{r_{k+1}} \leqq \lambda_{r_j}$ if $j \leqq k$ and $r_j > r_{k+1}$

$\lambda_{r_{k+1}} \geqq \lambda_{r_j}$ if $j \leqq k$ and $r_j < r_{k+1}$.

[2] $$F_k(x_1, x_2, \ldots, x_k) = P\{\xi_1 < x_1, \xi_2 < x_2, \ldots, \xi_k < x_k\}$$

[3] The weak limit of a sequence is uniquely determined *except* for a set of measure 0.

Let λ_t for an irrational t be defined by

$$\lim \lambda_{r_j} = \lambda_t$$

where $\{r_j\}$ is an increasing sequence of rational numbers going to t.
To prove our lemma we have to show that:

I. λ_t is the weak limit of $\{a_n^{(t)}\}_{n=1}^{\infty}$ (t is irrational)

II. $P\{\lim_{h\to 0} \lambda_{t-h} = \lambda_t\} = 1$ (for each t)

III. $P\{\lim_{t\to -\infty} \lambda_t = 0\} = P\{\lim_{t\to\infty} \lambda_t = 1\} = 1$.

The proof of I: Let the weak limit of $a_n^{(t)}$ be β_t. Then evidently

$$P\{\beta_t \geqq \lambda_t \geqq \lambda_r\} = 1$$

if $r < t$. For any $\varepsilon > 0$ we can find a rational r such that $r < t$ and

$$\int_{\Omega} (a_N^{(t)} - a_N^{(r)})\, dP < \varepsilon \quad \text{for each } N \tag{3}$$

but we have

$$\lim_{N\to\infty} \int_{\Omega} (a_N^{(t)} - a_N^{(r)})\, dP = \int_{\Omega} (\beta_t - \lambda_r)\, dP \geqq \int_{\Omega} (\beta_t - \lambda_t)\, dP \tag{4}$$

(3) and (4) together imply I.

The proofs of II and III are so similar that we omit these proofs.

Our next lemma is a generalization of a theorem of Dynkin ([8]) and our Theorem 3 in [4].

Lemma 6.

$$M\left(a_{i_1}^{(x_1)} a_{i_2}^{(x_2)} \dots a_{i_k}^{(x_k)} \lambda_{t_1}^{r_1} \lambda_{t_2}^{r_2} \dots \lambda_{t_k}^{r_k}\right) = M\left(\lambda_{x_1} \lambda_{x_2} \dots \lambda_{x_k} \lambda_{t_1}^{r_1} \lambda_{t_2}^{r_2} \dots \lambda_{t_k}^{r_k}\right)$$

provided that the indices $i_1, i_2, \dots, i_k$ are different.

Proof of this lemma is exactly the same as the proof of our Theorem 1 in [4], therefore we omit it.

§ 2. The conditional independce of equivalent random variables

Let $\xi_1, \xi_2, \dots$ be a sequence of equivalent random variables. The problem of this § is to find a "small and sufficiently concrete" σ-algebra F such that

$$P\{\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n | F\} =$$
$$= P\{\xi_1 < x_1 | F\} P\{\xi_2 < x_2 | F\} \dots P\{\xi_n < x_n | F\}$$

with probability 1. Knowing the mentioned result of [4], it is a very natural conjecture that the σ-algebra generated by the stochastic process $\lambda_x(\omega)$ is good from this purpose. More exactly[4]:

Theorem 1. *Let $\xi_1, \xi_2, \ldots$ be a sequence of equivalent random variables, then*

(A) $$P\{\xi_{i_1} < x_1, \xi_{i_2} < x_2, \ldots, \xi_{i_k} < x_k | F\} = \lambda_{x_1}\lambda_{x_2}\ldots\lambda_{x_k}$$

with probability 1, *where F is the smallest σ-algebra containing the sets*

(5) $$A = \{\omega: a_1 \leqq \lambda_{x_1} < b_1, a_2 \leqq \lambda_{x_2} < b_2, \ldots, a_n \leqq \lambda_{x_n} < b_n\}$$

and $\lambda_x(\omega)$ is the stochastic process defined in Lemma 5.

PROOF. Let us suppose that

(6) $$P\{\xi_{i_1} < x_1, \xi_{i_2} < x_2, \ldots, \xi_{i_k} < x_k | F\} = \lambda_{x_1}\lambda_{x_2}\ldots\lambda_{x_k} + \varepsilon_{i_1 i_2 \ldots i_k}(x_1, x_2, \ldots, x_k).$$

Here the random variable $\varepsilon = \varepsilon_{i_1 i_2 \ldots i_k}(x_1, x_2, \ldots, x_k)$ is evidently measurable with respect to F.

By Lemma 6 and our condition (6) we have evidently

$$M\big(\lambda_{t_1}^{r_1}\lambda_{t_2}^{r_2}\ldots\lambda_{t_j}^{r_j}\lambda_{x_1}\lambda_{x_2}\ldots\lambda_{x_k}\big) = M\big(a_{i_1}^{(x_1)}a_{i_2}^{(x_2)}\ldots a_{i_k}^{(x_k)}\lambda_{t_1}^{r_1}\lambda_{t_2}^{r_2}\ldots\lambda_{t_j}^{r_j}\big) =$$
$$= M\big(M(a_{i_1}^{(x_1)}a_{i_2}^{(x_2)}\ldots a_{i_k}^{(x_k)}\lambda_{t_1}^{r_1}\lambda_{t_2}^{r_2}\ldots\lambda_{t_j}^{r_j}| F)\big) = M\big(\lambda_{t_1}^{r_1}\lambda_{t_2}^{r_2}\ldots\lambda_{t_j}^{r_j} M(a_{i_1}^{(x_1)}a_{i_2}^{(x_2)}\ldots a_{i_k}^{(x_k)}| F)\big) =$$
$$= M\big[\lambda_{t_1}^{r_1}\lambda_{t_2}^{r_2}\ldots\lambda_{t_j}^{r_j}\big(\lambda_{x_1}\lambda_{x_2} + \ldots + \lambda_{x_k} + \varepsilon\big)\big].$$

Therefore we get

(7) $$M\big(\lambda_{t_1}^{r_1}\lambda_{t_2}^{r_2}\ldots\lambda_{t_j}^{r_j}\varepsilon\big) = 0$$

for any sequence $\{t_i\}_{i=1}^{j}$ of real numbers and any sequence $\{r_i\}_{i=1}^{j}$ of integers.

To prove our theorem it is enough to see that

(8) $$\int_A \varepsilon dP = \int_\Omega \varepsilon\alpha dP = 0$$

for any A of type (5), where α is the indicator function of A. But (8) follows from (7) using the fact that α can be approximated in mean by a polynomial

$$\Sigma c_{r_1 r_2 \ldots r_j}\lambda_{t_1}^{r_1}\lambda_{t_2}^{r_2}\ldots\lambda_{t_j}^{r_j}.$$

§ 3. A central limit theorem

In this § we prove that the partial sums of a sequence of equivalent random variables are asymptotically normal distributed under the condition that the random variables μ and σ take fixed values.

---

[4] A weaker version of this theorem is proved in [8]. Namely the author proves that the expectation of the right hand side of (A) is equal to the expectation of the left i. e.

(B) $$P(\xi_1 < x_1, \xi_2 < x_2, \ldots \xi_n < x_n) = M(\lambda_{x_1}\lambda_{x_2}\ldots\lambda_{x_n}).$$

In [9] the relation (A) is stated but only (B) is proved.

More exactly:

Theorem 2. *Let $\xi_1, \xi_2, \ldots$ be a sequence of equivalent random variables having finite variances. Then*

$$P\left\{\frac{\xi_1+\xi_2+\ldots+\xi_n-n\mu}{\sqrt{n}\sigma}<x\,\middle|\,\mu,\sigma\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\,dt \qquad (n\to\infty) \tag{9}$$

with probability 1 *and similarly*

$$P\left\{\frac{(\xi_1-\mu)+(\xi_2-\mu)+\ldots+(\xi_n-\mu)}{\sqrt{n}\sigma}<x\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}\,dt \qquad (n\to\infty) \tag{10}$$

provided that $P\{\sigma=0\}=0$ (where the random variables μ and σ were defined in Lemma 4.)

Remark. It is easy to see that if $P\{\sigma=0\}>0$ then

$$P\{\xi_i=\mu|\sigma=0\}=1.$$

Therefore from now it will be assumed that $P\{\sigma=0\}=0$.

Before the proof of Theorem 2 we have to prove two lemmas. First of them is evident:

Lemma 7. $\frac{(\xi_1-\mu)^2}{\sigma^2}, \frac{(\xi_2-\mu)^2}{\sigma^2}, \ldots$ *is a sequence of equivalent random variables with*

$$M\left(\frac{(\xi_i-\mu)^2}{\sigma^2}\right)=1 \qquad (i=1,2,\ldots)$$

The second one gives the connection between the random variables μ, σ and the stochastic process λ_x.

Lemma 8. *If $\xi_1, \xi_2, \ldots$ is a sequence of equivalent random variables with finite variances, then*

$$M(\xi_i|F)=\int_{-\infty}^{+\infty}x\,d\lambda_x=\mu \qquad (i=1,2,\ldots) \tag{11}$$

and

$$M((\xi_i-\mu)^2|F)=\int_{-\infty}^{+\infty}(x-\mu)^2\,d\lambda_x=\sigma^2 \tag{12}$$

with probability 1.

Proof. We prove only (11) because the proof of (12) is exactly the same. (11) evidently follows from Theorem 1 and Lemma 4. In fact Theorem 1 and the Kolmogorov's strong law of large numbers imply

$$P\left\{\frac{\xi_1+\ldots+\xi_n}{n}\to\int_{-\infty}^{+\infty}t\,d\lambda_t\,\middle|\,F\right\}=1 \tag{13}$$

and

(14) $$M(\xi_k|F) = \int_{-\infty}^{+\infty} t\, d\lambda_t$$

with probability 1. (13) implies

(15) $$P\left\{\frac{\xi_1 + \dots + \xi_n}{n} \to \int_{-\infty}^{+\infty} t\, d\lambda_t\right\} = 1$$

(14), (15) and (1) together give (11).

Proof of Theorem 2. By our Theorem 1, Lemma 8 and the simplest form of the central limit theorem we have

(16) $$P\left\{\frac{\xi_1 + \xi_2 + \dots + \xi_n - n\mu}{\sqrt{n}\sigma} < x \,\middle|\, F\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt \qquad (n \to \infty)$$

with probability 1. Therefore

$$P\left\{\frac{\xi_1 + \xi_2 + \dots + \xi_n - n\mu}{\sqrt{n}\sigma} < x \,\middle|\, \mu, \sigma\right\} = M\left\{P\left(\frac{\xi_1 + \xi_2 + \dots + \xi_n - n\mu}{\sqrt{n}\sigma} < x \,\middle|\, F\right) \,\middle|\, \mu, \sigma\right\} \to$$

$$\to M\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt \,\middle|\, \mu, \sigma\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$$

so we have (9). (10) evidently follows from (9) and Lemma 8.

Let us mention the following sharper form of the previous result:

Theorem 3. *Let $\xi_1, \xi_2, \dots$ be a sequence of equivalent random variables with finite variances. Then*

$$P\left\{\frac{c_1(\xi_1 - \mu) + c_2(\xi_2 - \mu) + \dots + c_n(\xi_n - \mu)}{S_n \sigma} < x\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt \qquad (n \to \infty)$$

where $\{c_k\}$ is any sequence of real numbers for which

$$\lim_{n \to \infty} \frac{\max\limits_{1 \le k \le n} c_k}{S_n} = 0$$

and

$$S_n = \sqrt{c_1^2 + c_2^2 + \dots + c_n^2} \qquad (n = 1, 2, \dots).$$

Proof of this theorem is exactly the same as the proof of Theorem 2, only we have to apply the Lindeberg's central limit theorem instead of the simplest form of the central limit theorem.

References

[1] B. DE FINETTI, La prevision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7** (1937), 1–68.
[2] А. Я. Хинчин, О классах эквывалентных событий, *Математический Сборник* **39** (1932) 40–43.
[3] I. HAAG, Sur un problème général de probabilités et ses diverses applications, *Proc. of the International Congress of Mathematicians, Toronto*, 1924, *Toronto* 1928, 659–674.
[4] A. RÉNYI–P. RÉVÉSZ, A study of sequences of equivalent events as a special stable sequence, *Publ. Math. Debrecen* **10** (1963), 319–325.
[5] W. SCHMEIDLER, Lineare Operatoren im Hilbertschen Raum, *Stuttgart*, 1954.
[6] A. RÉNYI, On stable sequences of events, *Sankya, Ser. A.* **25** (1963), 293–302.
[7] P. RÉVÉSZ, The solution of a problem of Steinhaus, *in print*.
[8] Е. Б. Дынкин, Классы эквивалентных случайных величин, *Успехи Мат. Наук* **5** 2(54) (1953), 125–130.
[9] M. LOÉVE, Probability theory, *New York*, 1955.

(Received December 11, 1964.)