ON THE PROBABILISTIC GENERALIZATION OF THE LARGE SIEVE OF LINNIK

by

ALFRÉD RÉNYI

Let $S = [\Omega, \mathcal{A}, P]$ be a probability space, i. e. Ω an arbitrary abstract space, $\mathcal{A}$ a σ-algebra of subsets of Ω and P a measure on $\mathcal{A}$, for which $P(\Omega) = 1$. The elements of $\mathcal{A}$ will be denoted by $A, B, \ldots$ etc. and called events. Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a finite or infinite sequence of random variables on S. Let us suppose first that the random variables ξ_n are of the discrete type. Let z_{nk} $(k = 1, 2, \ldots)$ denote the possible values of ξ_n; let A_{nk} denote the event $\xi_n = z_{nk}$ $(n, k = 1, 2, \ldots)$. We may suppose $P(A_{nk}) > 0$ for all values of k for which it is defined. Let us define

$$\mathbf{d}(\xi_n, \xi_m) = \sup_{(k,l)} \left| \frac{P(A_{nk} A_{ml})}{P(A_{nk})\, P(A_{ml})} - 1 \right| \qquad \text{for} \quad n \neq m . \tag{1}$$

(We denote by the product of two events the event consisting in the joint occurrence of the two events.) Let us suppose that the quadratic form

$$\sum_{n \neq m}\sum \mathbf{d}(\xi_n, \xi_m)\, x_n x_m$$

is bounded, i. e. there exists a constant $\Delta \geqq 0$ such that for any sequence $\{x_n\}$ for which

$$\sum_n x_n^2 < +\infty ,$$

we have

$$\left| \sum_{n \neq m}\sum \mathbf{d}(\xi_n, \xi_m)\, x_n x_m \right| \leqq \Delta \sum_n x_n^2 . \tag{2}$$

If such a constant Δ exists we call the random variables $\{\xi_n\}$ *(pairwise) almost independent with modulus* Δ. (Clearly $\Delta = 0$ if and only if the variables ξ_n are pairwise independent.) Let us denote by $\mathbf{M}\{\eta\}$ the mean value and by $\mathbf{D}^2\{\eta\}$ the variance of a random variable η, and by $\mathbf{M}\{\eta|\xi\}$ the conditional mean value of η under the condition that the value of ξ is given. In case ξ is a discrete random variable, taking on the values z_k (we denote this event by A_k) with the probability $P(A_k)$ $(k = 1, 2, \ldots)$, $\mathbf{M}\{\eta|\xi\}$ is a random variable which takes on the values $\mathbf{M}\{\eta|A_k\}$ $(K = 1, 2, \ldots)$ with probabilities $P(A_k)$ and thus

(3) $$\mathbf{M}\left\{\mathbf{M}\{\eta|\xi\}\right\} = \mathbf{M}\{\eta\} .$$

Let us put

(4) $$\mathbf{D}^2_\xi(\eta) = \mathbf{M}\left\{(\mathbf{M}\{\eta|\xi\} - \mathbf{M}\{\eta\})^2\right\}$$

i. e. $\mathbf{D}^2_\xi\{\eta\}$ is the variance of the random variable $\mathbf{M}\{\eta|\xi\}$. Let us put further

(5) $$\theta^2_\xi\{\eta\} = \frac{\mathbf{D}^2_\xi\{\eta\}}{\mathbf{D}^2\{\eta\}} .$$

The positive square root $\theta_\xi\{\eta\}$ of the quantity under (5) is the *correlation ratio* introduced by K. PEARSON (see [1], p. 280—281).

Some years ago I have proved (see [3], and for a previous weaker version [2]) the following

Theorem 1. *Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be a sequence of discrete random variables, which are (pairwise) almost independent with modulus Δ. Let η be an arbitrary random variable with finite mean and variance. Then we have*

(6) $$\sum_{n=1}^{\infty} \theta^2_{\xi_n}\{\eta\} \leqq (1 + \Delta) .$$

I obtained Theorem 1. as the probabilistic generalization of the large sieve of Y. V. LINNIK (see [4]). Theorem 1. has important applications, e. g. in number theory.[1]

In [3], Theorem 1. is stated in a somewhat different form ; first of all it is stated not only for discrete random variables ξ_n ; this makes no essential difference as the general case can easily be deduced from the particular case of discrete variables, as will be pointed out below. Besides this, the theorem is not expressed in terms of correlation ratios ; as a matter of fact, when writing the paper [3] it escaped my attention that this is the most simple way of expression.

In the present paper we give an improvement of theorem 1. The improvement consists in that we prove the inequality (6) by supposing instead of the condition that the random variables ξ_n should be „almost-independent" with modulus Δ, only that they should be „weakly-dependent", with modulus B.

The „weak dependence" is defined as follows: Let $\varphi(\xi_n, \xi_m)$ denote the *mean square contingency* [introduced by K. PEARSON (see [1], p. 282)] of the random variables ξ_n and ξ_m, i. e. let $\varphi(\xi_n, \xi_m)$ denote the positive square root of the quantity

(7) $$\varphi^2(\xi_n, \xi_m) = \sum_k \sum_l \frac{(P(A_{nk} A_{ml}) - P(A_{nk}) P(A_{ml}))^2}{P(A_{nk}) P(A_{ml})} .$$

[1] It contains as a special case the theorem which I used in proving (see [5]) that there exists a positive integer K such that every positive integer n can be represented in the form $n = p + P$ where p is a prime number and the number of prime factors of P does not exceed K. An other application of the large sieve has been found by P. T. BATEMAN, S. CHOWLA and P. ERDŐS [6].

We call the sequence $\{\xi_n\}$ *(pairwise) weakly dependent with modulus* B, if the quadratic form

$$\sum_{n \neq m}\sum \varphi(\xi_n, \xi_m)\, x_n\, x_m$$

is bounded with bound B, i. e. if for any sequence $\{x_n\}$ such that

$$\sum_n x_n^2 < +\infty$$

we have

$$\left| \sum_{n \neq m}\sum \varphi(\xi_n, \xi_m)\, x_n\, x_m \right| \leqq B \cdot \sum_n x_n^2 . \tag{8}$$

We prove (Theorem 2.) that under condition (8) we have

$$\sum_{n=1}^{\infty} \theta^2_{\xi_n}\{\eta\} \leqq (1 + B) . \tag{9}$$

As clearly

$$\varphi^2(\xi_n, \xi_m) \leqq \mathbf{d}^2(\xi_n, \xi_m) . \tag{10}$$

we have

$$B \leqq \Delta \tag{11}$$

and thus (9) is stronger than (6). Evidently $B = 0$ if and only if the variables ξ_n are pairwise independent.

It should be mentioned, that if the sequence of random variables ξ_n is finite, then the bound Δ resp. B always exist, and thus the inequalities (6) resp. (9) have a sense for *any* finite sequence of discrete random variables $\{\xi_n\}$.

For the proof we need the following lemma (see [7]), which is a direct generalization of Bessel's inequality for quasi-orthogonal functions (resp. random variables) and which has been already used in [2] and [3].

Lemma. *Let ζ_n be a quasi-orthogonal sequence of random variables, with bound C, i. e. suppose that*

$$\left| \sum_n \sum_m \mathbf{M}\{\zeta_n \zeta_m\}\, x_n\, x_m \right| \leqq C \sum_n x_n^2 \tag{12}$$

for every sequence x_n of real numbers for which

$$\sum_n x_n^2 < +\infty .$$

Then we have for any random variable η, for which $\mathbf{M}\{\eta^2\}$ exists,

$$\sum_n \mathbf{M}^2\{\eta\zeta_n\} \leqq C\, \mathbf{M}\{\eta^2\} . \tag{13}$$

For the sake of completeness we reproduce here the very easy proof of this Lemma.

Proof of the Lemma. We have clearly for any N and for any real sequence a_n

$$\mathbf{M}\left\{\left(\eta-\frac{1}{C}\sum_{n=1}^{N}a_n\zeta_n\right)\right\}=$$

(14)

$$=\mathbf{M}\{\eta^2\}-\frac{2}{C}\sum_{n=1}^{N}a_n\mathbf{M}\{\eta\zeta_n\}+\frac{1}{C^2}\sum_{n=1}^{N}\sum_{m=1}^{N}a_n a_m\mathbf{M}\{\zeta_n\zeta_m\}\ .$$

Putting $a_n=\mathbf{M}\{\eta\zeta_n\}$ and applying (12) to the last term on the right-hand side of (14), we obtain, the left-hand side of (14) being evidently nonnegative, the assertion of the Lemma.

Now we introduce the random variables ξ_{nk} defined as follows:

(15) $$\xi_{nk}=\begin{cases}1 & \text{if}\quad \xi_n=Z_{nk}\\ 0 & \text{otherwise}\end{cases}$$

and put

(16) $$\xi^*_{nk}=\frac{\xi_{nk}-P(A_{nk})}{\sqrt{P(A_{nk})}}\ .$$

Let us put further

(17) $$c_{nmkl}=\mathbf{M}\{\xi^*_{nk}\xi^*_{ml}\}$$

Then we have evidently

(18) $$c_{nnkk}=1-P(A_{nk})$$

(19) $$c_{nnkl}=-\sqrt{P(A_{nk})P(A_{nl})}\quad\text{for}\quad k\neq l$$

(20) $$c_{nmkl}=\frac{P(A_{nk}A_{ml})-P(A_{nk})P(A_{ml})}{\sqrt{P(A_{nk})P(A_{ml})}}\qquad\text{for}\quad n\neq m$$

Let us consider now the quadratic form

(21) $$Q=\sum_n\sum_m\sum_k\sum_l c_{nmkl}x_{nk}x_{ml}$$

where x_{nk} is a double sequence such that

$$\sum_n\sum_k x^2_{nk}<+\infty.$$

Let us suppose that (8) is satisfied. We have

(22) $$Q=\sum_n\sum_k x^2_{nk}-\sum_n\left(\sum_k\sqrt{P(A_{nk})}\,x_{nk}\right)^2+\sum_{n\neq m}\sum_k\sum_l c_{nmkl}x_{nk}x_{ml}$$

Let us put

(23) $$y_n=\left(\sum_k x^2_{nk}\right)^{1/2}.$$

As by the inequality of Schwarz

(24) $$\sum_n\left(\sum_k\sqrt{P(A_{nk})}\,x_{nk}\right)^2\leqq\sum_n\sum_k x^2_{nk}$$

we have (again by the inequality of Schwarz)

$$|Q| \leqq \sum_n \sum_k x_{nk}^2 + \sum_{n \neq m} \sum \varphi(\xi_n, \xi_m) y_n y_m \tag{25}$$

and thus by (8)

$$|Q| < \sum_n \sum_k x_{nk}^2 + B \sum_n y_n^2 . \tag{26}$$

Thus we obtain

$$|Q| \leqq (1 + B) \sum_n \sum_k x_{nk}^2 . \tag{27}$$

Thus the system $\{\xi_{nk}^*\}$ is quasi-orthogonal with bound $C = 1 + B$. As, however, clearly

$$\sum_k \mathbf{M}^2 \{\eta \, \xi_{uk}^*\} = \mathbf{D}_{\xi_n}^2 \{\eta\} \tag{28}$$

it follows by the above Lemma that

$$\sum_n \mathbf{D}_{\xi_n}^2 \{\eta\} \leqq (1 + B) \, \mathbf{M} \{\eta^2\} . \tag{29}$$

We may evidently suppose $\mathbf{M}(\eta) = 0$, because putting $\eta' = \eta - \mathbf{M}(\eta)$ we have $\mathbf{M}(\eta') = 0$ and $\mathbf{D}_{\xi_n}^2(\eta') = \mathbf{D}_{\xi_n}^2(\eta)$; as $\mathbf{M}(\eta) = 0$ implies $\mathbf{M}(\eta^2) = \mathbf{D}^2(\eta)$, dividing both sides of (29) by $\mathbf{D}^2\{\eta\}$ we obtain (9).

Let us consider now two arbitrary random variables ζ_1 and ζ_2. Let us define $\varphi(\zeta_1, \zeta_2)$ as the least upper bound of $\varphi(f_1(\zeta_1), f_2(\zeta_2))$ where f_1 and f_2 are arbitrary (Borel measurable) step functions and thus $f_1(\zeta_1)$ and $f_2(\zeta_2)$ are discrete random variables. In view of the fact, that denoting by I and J arbitrary real intervals, the two-dimensional interval function

$$\nu(I \times J) = \frac{(\mathbf{P}\{\zeta_1 \in I, \zeta_2 \in J\} - \mathbf{P}\{\zeta_1 \in I, \zeta_2 \in J\})^2}{\mathbf{P}\{\zeta_1 \in I\} \; \mathbf{P}\{\zeta_2 \in J\}} \tag{30}$$

is clearly subadditive, it follows (see the Lemma on p. 139. of [3] which is evidently valid for interval functions on the plane too) that

$$\varphi^2(\zeta_1, \zeta_2) = \iint \nu(I \times J) \tag{31}$$

where the integral of the interval-function on the right-hand side of (31) is to be understood in the sense of BURKILL, and extended over the whole plane. We shall call the positive square root $\varphi(\zeta_1, \zeta_2)$ of the quantity (31) the *mean square contingency* of the random variables ζ_1 and ζ_2.

Taking into account, that if ξ is an arbitrary random variable, $\theta_\xi(\eta)$ is equal to the least upper bound of $\theta_{f(\xi)}(\eta)$ where f is an arbitrary Borel-measurable step-function, it follows, that (9) holds also if the variables ξ_n are not supposed to be of the discrete type.

Thus we have proved the following

Theorem 2. *Let $\{\xi_n\}$ be a sequence of random variables. Let $\varphi(\xi_n, \xi_m)$ denote the mean square contingency of ξ_n and ξ_m. Let us suppose that*

$$\left| \sum_{n \neq m} \sum \varphi(\xi_n, \xi_m) \, x_n \, x_m \right| \leqq B \sum_n x_n^2$$

provided that

$$\sum_n x_n^2 < +\infty.$$

Let η *be an arbitrary random variable with finite second moment. Then we have, denoting by* $\theta_\xi\{\eta\}$ *the correlation ratio of* η *on* ξ,

$$\sum_n \theta_{\xi_n}^2\{\eta\} \leqq (1+B)\ . \tag{32}$$

(Received July 23, 1958.)

REFERENCES

[1] CRAMÉR, H. : *Mathematical methods of statistics*. Princeton University Press, 1946.
[2] RÉNYI, A. : „Un nouveau théorème concernant les fonctions indépendantes et ses applications a la théorie des nombres". *Journal de Mathématiques Pures et Appliquées* **28** (1949) 137—149.
[3] RÉNYI, A. : „Sur un théorème général de probabilité". *Annales de l'Institut Fourier* **1** (1949) 43—52.
[4] LINNIK, Y. V. : „The large sieve". *Comptes Rendus de l'Académie des Sciences de l'URSS* **30** (1941) 292—294.
[5] RÉNYI, A. : "О представлении четных чисел в виде суммы простого и почти простого числа". *Известия Академии Наук СССР, сер. мат.* **12** (1948) 57—58.
[6] BATEMAN, P. T.—CHOWLA, S.—ERDŐS, P. : „Remarks on the size of $L(1, \chi)$" *Publicationes Mathematicae (Debrecen)* **1** (1950) 165—182.
[7] BOAS, R. P. JR. : „A general moment problem". *American Journal of Mathematics* **63** (1941) 361—370.

A LINNIK-FÉLE NAGY SZITA VALÓSZÍNŰSÉGSZÁMÍTÁSI ÁLTALÁNOSÍTÁSÁRÓL

RÉNYI A.

Kivonat

A szerző 1947 és 1949 között több irányban általánosította LINNIK nagy szitának nevezett számelméleti módszerét, és bebizonyított egy új valószínűségszámítási tételt, amely a Linnik-féle nagy szitát speciális esetként tartalmazza. A Linnik-féle nagy szita általánosítása tette lehetővé, hogy a szerző bebizonyíthatta a következő számelméleti tételt :

Létezik egy olyan K állandó, hogy minden n természetes szám előállítható $n = p + P$ alakban, ahol p törzsszám és P törzstényezőinek száma legfeljebb K.

E dolgozatban a szerző a szóbanforgó, a nagy szitát tartalmazó valószínűségszámítási tétel egy élesítését bizonyítja be.

Legyenek $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ diszkrét eloszlású valószínűségi változók. Jelölje z_{nk} $(k = 1, 2, \ldots)$ azokat az értékeket, amelyeket ξ_n pozitív valószínű-

séggel vesz fel ; jelölje A_{nk} azt az eseményt, hogy ξ_{nk} a z_{nk} értéket veszi fel, és jelölje $P(A_{nk})$ az A_{nk} esemény valószínűségét.

Jelölje $\varphi(\xi_n, \xi_m)$ a ξ_n és ξ_m változók közötti függőség PEARSON-féle átlagos négyzetes mérőszámát (mean square contingency), vagyis legyen

$$\varphi(\xi_n, \xi_m) = \left(\sum_k \sum_l \frac{(P(A_{nk} A_{ml}) - P(A_{nk}) P(A_{ml}))^2}{P(A_{nk}) P(A_{ml})} \right)^{\frac{1}{2}} \tag{1}$$

Legyen η egy tetszőleges valószínűségi változó, jelölje $\mathbf{M}\{\eta\}$ az η változó várható értékét, $\mathbf{M}\{\eta^2\}$ η második momentumát és $\mathbf{D}^2\{\eta\}$ η szórásnégyzetét. Jelölje $\mathbf{M}\{\eta|A\}$ az η változó feltételes várható értékét az A feltétel mellett. Jelölje $\mathbf{M}\{\eta|\xi_n\}$ az η változó feltételes várható értékét ξ_n rögzített értéke mellett ; $\mathbf{M}\{\eta|\xi_n\}$ tehát olyan valószínűségi változó, amely az $\mathbf{M}\{\eta|A_{nk}\}$ értéket $P(A_{nk})$ valószínűséggel veszi fel.

Jelölje $\theta_{\xi_n}\{\eta\}$ az η valószínűségi változó ξ_n-re vonatkozó PEARSON-féle korrelációs hányadosát, vagyis legyen

$$\theta^2_{\xi_n}\{\eta\} = \frac{\mathbf{D}^2\{\mathbf{M}\{\eta|\xi_n\}\}}{\mathbf{D}^2(\eta)} \tag{2}$$

ahol $\mathbf{D}^2\{\mathbf{M}\{\eta|\xi_n\}\}$ az $\mathbf{M}\{\eta|\xi_n\}$ valószínűségi változó szórásnégyzete. Akkor fennáll a következő

Tétel. *Ha a*

$$\sum_{n \neq m} \sum \varphi(\xi_n, \xi_m) x_n x_m$$

kvadratikus alak korlátos és korlátja B, vagyis ha bármely olyan x_n számsorozatra, amelyre

$$\sum x_n^2 < +\infty$$

érvényes a

$$\left| \sum_{n \neq m} \sum \varphi(\xi_n, \xi_m) x_n x_m \right| \leq B \sum_{n=1}^{\infty} x_n^2 \tag{3}$$

egyenlőtlenség, akkor

$$\sum_{n=1}^{\infty} \theta^2_{\xi_n}\{\eta\} \leq (1 + B) \ . \tag{4}$$

О ТЕОРЕТИКО-ВЕРОЯТНОСТНОМ ОБОБЩЕНИИ БОЛЬШОГО РЕШЕТА ЛИННИКА

А. RÉNYI

Резюме

Автор в 1947—49 годах в нескольких направлениях обобщил метод введенный в теорию чисел Ю. В. Линником и называемый им большим решетом, и доказал одно новую теоретико-вероятностную теорему, содержащую большое решето Линника в качестве частного случая. Обобщение большого решета Линника позволило автору доказать следующую теорему теории чисел :

Существует такая постоянная K, что всякое натуральное число n может быть представленно в виде $n = p + P$, где p простое число, а число простых делителей P не превосходит K.

Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ суть случайные величины с дискретным распределением. Обозначим через $z_{nk}\,(k = 1, 2, \ldots)$ значения, принимаемые ξ_n с положительной вероятностью, через A_{nk} событие, заключающееся в том, что ξ_n принимает значение z_{nk}, а через $P(A_{nk})$ вероятность события A_{nk}.

Обозначим через $\varphi(\xi_n, \xi_m)$ средную квадратичную меру (mean square contingency) Pearson-а зависимости между величинами ξ_n и ξ_m, т. е. пусть

$$\varphi(\xi_n, \xi_m) = \left(\sum_k \sum_l \frac{(P(A_{nk} A_{ml}) - P(A_{nk})\, P(A_{ml}))^2}{P(A_{nk}) P(A_{ml})}\right)^{\frac{1}{2}} \tag{1}$$

Пусть $\mathbf{M}\{\eta\}$ обозначает математическое ожидание, $\mathbf{M}\{\eta^2\}$ второй момент, а $\mathbf{D}^2\{\eta\}$ дисперсию случайной величины η. Пусть $\mathbf{M}\{\eta | A\}$ обозначает условное математическое ожидание величины η при условии A. Обозначим через $\mathbf{M}\{\eta | \xi_n\}$ условное математическое ожидание величины η при фиксированном значении ξ_n; таким образом $\mathbf{M}\{\eta | \xi_n\}$ такая случайная величина, которая принимает значение $\mathbf{M}\{\eta | A_{nk}\}$ с вероятностью $P(A_{nk})$.

Пусть $\theta_{\xi_n}\{\eta\}$ обозначает корреляционное отношение Pearson-а случайной величины η относительно ξ_n, т. е. пусть

$$\theta^2_{\xi_n}\{\eta\} = \frac{\mathbf{D}^2\{\mathbf{M}\{\eta | \xi_n\}\}}{\mathbf{D}^2\{\eta\}}, \tag{2}$$

где $\mathbf{D}^2\{\mathbf{M}\{\eta | \xi_n\}\}$ дисперсия случайной величины $\mathbf{M}\{\eta | \xi_n\}$. Тогда имеет место следующая

Теорема. *Если квадратичная форма*

$$\sum_{n \neq m} \sum \varphi(\xi_n, \xi_m)\, x_n x_m$$

ограниченна и B ее граница, т. е. если для всякой последовательности x_n для которой

$$\sum_{n=1}^{\infty} x_n^2 < \infty$$

имеет место неравенство

$$\left| \sum_{n \neq m} \sum \varphi(\xi_n, \xi_m)\, x_n x_m \right| \leq B \sum_{n=1}^{\infty} x_n^2, \tag{3}$$

то

$$\sum_{n=1}^{\infty} \theta^2_{\xi_n}\{\eta\} \leq (1 + B). \tag{4}$$