MATH. SCAND. 13 (1963), 99–104

AN ELEMENTARY INEQUALITY BETWEEN THE PROBABILITIES OF EVENTS

P. ERDÖS, J. NEVEU, A. RENYI

Introduction and summary.

For every positive integer n and any real number α in the open interval $(0,1)$, let $\varepsilon_n(\alpha)$ be the least real number ε with the following property.

PROPERTY (P_1): *For any sequence* $\{A_i, 1 \leqq i \leqq n\}$ *of* n *events in an arbitrary probability space* $(\Omega, \mathscr{A}, P)$ *such that* $P(A_iA_j) \leqq \alpha^2$ *for all* i, j $(i \neq j)$, *the following inequality holds*:

$$\sum_{i=1}^{n} [P(A_i) - \alpha] \leqq \varepsilon \,. \tag{1}$$

The exact value of the constant $\varepsilon_n(\alpha)$ is given in Section 2; however, the essential fact is that it tends to a finite limit as n goes to infinity. More precisely, it is found that $\varepsilon_n(\alpha) = \frac{1}{2}(1-\alpha)$ provided $n\alpha$ or $(n-1)\alpha$ is an integer (hence $\varepsilon_n(\frac{1}{2}) = \frac{1}{4}$ for every n) and that $\varepsilon_n(\alpha) = \frac{1}{2}(1-\alpha) + O(n^{-1})$ for fixed α. It is also shown that it is possible to find, for every n and α, a sequence $\{A_i^*, 1 \leqq i \leqq n\}$ such that

$$P(A_i^*) = \alpha + \varepsilon_n(\alpha)/n, \qquad P(A_i^* A_j^*) = \alpha^2 \quad (i \neq j);$$

for this sequence (1) is an equality. Such an extremal sequence is constructed in $[0,1]$, for the case $\alpha = \frac{1}{2}$, $n \equiv 3 \pmod 4$ by use of the method of quadratic residues; this example originated the present study.

After having completed this work, we became aware of a paper by S. Zubrzycki [2] which deals with the same problem in the special case when $P(A_i) = \omega_1$ for $i = 1, \ldots, n$ and $P(A_iA_j) = \omega_2$ for $i \neq j$; the inequality which he obtains between ω_1 and ω_2 is essentially equivalent with our Corollary 1. Among the extremal cases where the inequality becomes an equality are those in which the $P(A_i)$ and the $P(A_iA_j)$ have the constant values mentioned above and already found by Zubrzycki. Despite the overlapping of the results and the similarity of the methods we have considered the publication of the present paper worth while not

Received March 22, 1963.

only because of the somewhat greater generality of our results but also because of the difference in point of view.

1.

In order to be able to calculate $\varepsilon_n(\alpha)$ easily, we shall first establish the following results.

THEOREM 1. *For every n and α, the constant $\varepsilon_n(\alpha)$ is also the least real number ε having the following property*:

PROPERTY (P_2): *For any random variable N defined on an arbitrary probability space $(\Omega,\mathscr{A},P)$ and with values in the interval $\{0,1,\ldots,n\}$ of integers, such that $E[N(N-1)]\leqq n(n-1)\alpha^2$, the following inequality holds*:

$$E(N) \leqq n\alpha+\varepsilon\,.$$

PROOF. Suppose first that ε is a real number having property (P_2). Given a sequence $\{A_i,\ 1\leqq i\leqq n\}$ of n events, let $N=\sum_{i=1}^n 1_{A_i}$, where 1_A denotes the indicator of A; then N is a random variable taking its values in $\{0,1,\ldots,n\}$ and satisfying the relations

$$E[N] = \sum_i P(A_i), \qquad E[N(N-1)] = \sum_{i\neq j}\sum P(A_iA_j)\,.$$

Hence if the sequence $\{A_i\}$ is such that

$$\sum_{i\neq j}\sum P(A_iA_j) \leqq n(n-1)\alpha^2\,,$$

and a fortiori, if $P(A_iA_j)\leqq\alpha^2$ for every i,j $(i\neq j)$, then by property (P_2)

$$\sum_1^n [P(A_i)-\alpha] = E[N]-n\alpha \leqq \varepsilon\,.$$

This shows that ε has property (P_1).

Conversely, suppose that ε is a real number having property (P_1) and let $\{p_i,\ 0\leqq i\leqq n\}$ be a probability law on the set $\{0,1,\ldots,n\}$. Then it is possible to construct a sequence $\{A_i,\ 1\leqq i\leqq n\}$ of n events in a suitable probability space $(\Omega,\mathscr{A},P)$ such that a) the random variable $N=\sum_{i=1}^n 1_{A_i}$ has $\{p_i,\ 0\leqq i\leqq n\}$ for probability law; b) the sequence $\{A_i\}$ is symmetrically dependent, that is, $P(A_{i_1}\ldots A_{i_e})$ depends only on the number e of different indices $i_1,\ldots,i_e$. In order to construct such a sequence, let $\{A_i,\ 1\leqq i\leqq n\}$ be a sequence of independent equiprobable events in a suitably choosen probability space $(\Omega,\mathscr{A},Q)$ and define a new probability P on $(\Omega,\mathscr{A})$ by

$$P(A) = \sum_{k=0}^{n} p_k Q[A \mid N=k];$$

it is then obvious that $P[N=k]=p_k$; moreover $P[A_{i_1}\dots A_{i_e} \mid N=k]$ depends only on the number of different indices $i_1,\dots,i_e$, because $Q[A_{i_1}\dots A_{i_e} \mid N=k]$ has this property for every k.

If the given law $\{p_i\}$ is such that $\sum_k k(k-1)p_k \leqq n(n-1)\alpha^2$, then the sequence $\{A_i\}$ constructed above is such that for every i_0, j_0 $(i_0 \neq j_0)$

$$P(A_{i_0}A_{j_0}) = \frac{1}{n(n-1)} \sum\sum_{i \neq j} P(A_iA_j)$$
$$= \frac{1}{n(n-1)} \sum_{k=0}^{n} k(k-1)p_k \leqq \alpha^2 .$$

Hence, by the property (P_1) of ε,

$$\sum_k kp_k = \sum_i P(A_i) \leqq n\alpha + \varepsilon_n(\alpha) .$$

This shows that ε has property (P_2) and finishes the proof of the theorem.

COROLLARY 1. *For every n and α, the constant $\varepsilon_n(\alpha)$ is the least real number ε having the following property*:

PROPERTY (P_1'): *For any sequence $\{A_i, 1 \leqq i \leqq n\}$ of n events in an arbitrary probability space $(\Omega, \mathscr{A}, P)$ such that*

$$\frac{1}{n(n-1)} \sum\sum_{i \neq j} P(A_iA_j) \leqq \alpha^2 ,$$

the following inequality holds:

$$\sum_{i=1}^{n} [P(A_i) - \alpha] \leqq \varepsilon . \tag{1}$$

PROOF. If a real number ε satisfies property (P_1'), it clearly satisfies (P_1). The first part of the proof of the preceding theorem also shows that if an ε satisfies (P_2), it satisfies (P_1'). The corollary then follows from the equivalence of (P_1) and (P_2).

2.

In this section, we compute the supremum S of $E(N)$ when N varies in the class of random variables taking their values in the set $\{0, 1, \dots, n\}$ and such that $E[N(N-1)] \leqq C$ for a given constant C. In virtue of Theorem 1 this will give us the constant $\varepsilon_n(\alpha)$ by letting $C = n(n-1)\alpha^2$ and $S = n + \varepsilon_n(\alpha)$.

Let ν be the unique positive integer and θ the unique real number in the semi-closed interval $[-1,1[$ such that $C=\nu(\nu+\theta)$. Then $S=\nu+\frac{1}{2}(1+\theta)$. Moreover the random variable N^* with probability law given by

$$p_\nu^* = \tfrac{1}{2}(1-\theta), \quad p_{\nu+1}^* = \tfrac{1}{2}(1+\theta), \quad p_m^* = 0 \text{ otherwise}$$

is such that $E(N^*)=S$, $E[N^*(N^*-1)]=C$.

To prove this result, first let ν be the largest integer inferior or equal to S and let $\sigma=S-\nu$. Then notice that the supremum S of $E(N)$ when $E[N(N-1)]\leqq C$ may only be attained by random variables N with law concentrated on the two point set $\{\nu,\nu+1\}$; for suppose that N gives this extremum and let λ (resp. μ) be the smallest (resp. largest) integral value taken by N with positive probability, then, if λ were strictly less than ν it would be possible to modify the law of N making p_λ and p_μ decrease and p_ν increase keeping $E[N(N-1)]$ constant and thus, by the convexity of the function $x(x-1)$, causing $E(N)$ to increase; since this is impossible, one has $\lambda=\nu$ and for a similar reason $\mu=\nu+1$ (except if $\sigma=0$, in which case $\mu=\nu$). Obviously a variable N attaining the supremum S is such that $E[N(N-1)]=C$ (this could be proved, if necessary, by slightly modifying the preceding argument).

By what precedes, the supremum S is attained by the variable N^* concentrated on $\{\nu,\nu+1\}$ and with expectation $E(N^*)=\nu+\sigma$; this variable is unique and its law is given by

$$p_\nu^* = 1-\sigma, \quad p_{\nu+1}^* = \sigma, \quad p_m^* = 0 \text{ otherwise};$$

moreover $C=E[N^*(N^*-1)]=\nu(\nu+2\sigma-1)$.

Finally, let $\theta=2\sigma-1$ to obtain the result indicated above. The representation of the constant C as $\nu(\nu+\theta)$ is unique for it follows from $-1\leqq\theta<1$ that ν is the largest integer such that $\nu(\nu-1)\leqq C$; equivalently ν is the largest integer such that $\nu\leqq\frac{1}{2}+(C+\frac{1}{4})^{\frac{1}{2}}$. It may also be remarked that

$$S = \tfrac{1}{2}+(\nu+\tfrac{1}{2}\theta) = \tfrac{1}{2}+(C+\tfrac{1}{4}\theta^2)^{\frac{1}{2}}$$

so that for large C, one has

$$S = C^{\frac{1}{2}}+\tfrac{1}{2}+O(C^{-\frac{1}{2}})\,.$$

The following final result will be obtained by letting $C=n(n-1)\alpha^2$ and $\varepsilon_n(\alpha)=S-n\alpha$.

Let ν be the largest integer such that $\nu(\nu-1)\leqq n(n-1)\alpha^2$. Then

$$\varepsilon_n(\alpha) = \tfrac{1}{2}(1-\alpha) + \frac{[n\alpha-\nu][(n-1)\alpha-\nu]}{2\nu}\,.$$

The second term vanishes when $n\alpha$ or $(n-1)\alpha$ is an integer (for then ν equals $n\alpha$ or $(n-1)\alpha$, respectively); for $n \to \infty$, it is of the order of $1/n$.

The value of $\varepsilon_n(\alpha)$ given above is obtained by elementary computations from the equations

$$\varepsilon_n(\alpha) = S - n\alpha; \quad S = \nu + \tfrac{1}{2}(1+\theta); \quad \theta = (C-\nu^2)/\nu; \quad C = n(n-1)\alpha^2 .$$

When $n\alpha$ is an integer, it follows from $C = n\alpha(n\alpha - \alpha)$ that $\nu = n\alpha$ and that $\theta = -\alpha$; similarly if $(n-1)\alpha$ is an integer, then $\nu = (n-1)\alpha$ and $\theta = \alpha$; in both cases $\varepsilon_n(\alpha)$ reduces to $\frac{1}{2}(1-\alpha)$. When n increases to infinity, $[n(n-1)\alpha^2]^{\frac{1}{2}} = n\alpha - \frac{1}{2}\alpha + O(1/n)$; hence

$$\varepsilon_n(\alpha) = [n(n-1)\alpha^2]^{\frac{1}{2}} + \tfrac{1}{2} - n\alpha + O(1/n) = \tfrac{1}{2}(1-\alpha) + O(1/n) .$$

3. An example.

Using well known properties of quadratic residues we can construct explicit examples which are extremal in the sense of the preceding paragraph, for $\alpha = \frac{1}{2}$. We shall need the following

THEOREM 2. *Let p be a prime number such that $p \equiv 3 \pmod 4$. Put $p = 4k-1$. Let $r_1, r_2, \ldots, r_{2k-1}$ be all the quadratic residues* mod p. *Then if d is any of the numbers $1, 2, \ldots, p-1$, there are among the numbers $r_j + d$, $j = 1, 2, \ldots, 2k-1$, exactly $k-1$ which are congruent to some r_h, $1 \leqq h \leqq 2k-1$ (and k which are not).*

Let us now define a sequence of subsets of the interval $[0, 1)$ as follows. Let A_0 be the union of the intervals

$$\left[0, \frac{3}{4p}\right] \quad \text{and} \quad \left[\frac{r_j}{p}, \frac{r_{j+1}}{p}\right), \quad j = 1, 2, \ldots, 2k-1 .$$

Let further A_d be obtained from A_0 by shifting it mod 1 by the distance d/p, $d = 1, 2, \ldots, p-1$. For $p = 7$, the sets $A_0, A_1, \ldots, A_6$ are shown in Fig. 1. Denoting by $P(A)$ the Lebesgue measure of the set A we have

$$P(A_j) = \frac{p-1}{2p} + \frac{3}{4p} = \frac{1}{2} + \frac{1}{4p} . \tag{2}$$

On the other hand, according to the theorem on quadratic residues mentioned above we have

$$P(A_j A_{j+d}) = P(A_0 A_d) = \frac{p-3}{4p} + \frac{3}{4p} = \frac{1}{4} \tag{3}$$

for $d = 1, 2, \ldots, p-1-j$; $j = 0, 1, \ldots, p-2$. As a matter of fact the interval $[(r_j + d)/p, (r_j + d + 1)/p)$ belonging to A_d coincides (mod 1) with an

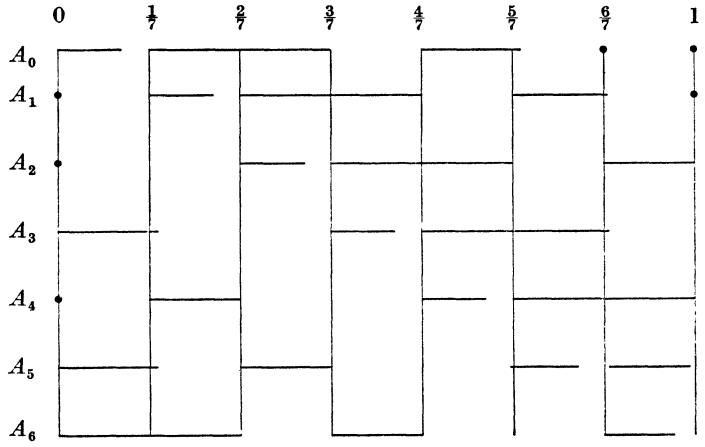

Fig. 1.

interval of A_0 if and only if r_j+d is a quadratic residue, and thus for $k-1=(p-3)/4$ values of j; further, if d is a quadratic residue then the interval $[d/p, (d+\frac{3}{4})/p)$ belonging to A_d coincides with $\frac{3}{4}$ of an interval belonging to A_0, while if d is not a quadratic residue then $p-d$ is a quadratic residue mod p (p being of the form $4k-1$) and thus the interval $(0, 3/(4p))$ belonging to A_0 coincides with $\frac{3}{4}$ of an interval belonging to A_d (namely of that obtained from the interval

$$[(p-d)/p, (p-d+1)/p)$$

by shifting it by d/p). Thus our system of sets $A_0, A_1, \ldots, A_{p-1}$ satisfies the conditions $P(A_j A_h)=\frac{1}{4}$ for $j \neq h$ and

$$P(A_j) = \frac{1}{2} + \frac{1}{4p} \quad \text{for} \quad j = 0, 1, \ldots, p-1 .$$

Hence if we consider the sets $A_0, \ldots, A_{p-1}$ as events in the probability space $(\Omega, \mathscr{A}, P)$ where Ω is the interval $[0, 1)$, $\mathscr{A}$ the set of all measurable subsets of $\mathscr{A}$, and P the Lebesgue measure, then $\{A_0, \ldots, A_{p-1}\}$ is an extremal system of events (for $\alpha=\frac{1}{2}$, $n=p$) of the required type.

REFERENCES

1. O. Perron, *Bemerkung über die Verteilung der quadratischen Reste*, Math. Z. 56 (1952), 122–130.

2. S. Zubrzycki, *Les inégalités entre les moments des variables aléatoires équivalentes*, Studia Math. 14 (1954), 232–242.

MATHEMATICAL INSTITUTE, HUNGARIAN ACADEMY OF SCIENCES
UNIVERSITY OF PARIS, FRANCE