ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
VIII (1965)

S. TRYBUŁA (Wrocław)

ON THE PARADOX OF n RANDOM VARIABLES

Dealing with random variables we are often interested in the probability $P(X < Y)$ rather than in exact distributions of X and Y. Sometimes these distributions are difficult to obtain while $P(X < Y)$ can be estimated after some trials.

Suppose, for example, that the strength of play of a player is a random variable X and that X defines the results of matches of the player against his opponents in the sense that player I loses against player II if and only if $X_{\mathrm{I}} < X_{\mathrm{II}}$. The number $P(X_{\mathrm{I}} < X_{\mathrm{II}})$ has its precise probabilistic meaning and can be estimated by the results of matches even if the random variables X_{I} and X_{II} are not described in more detail.

Let X, Y, Z be random variables and suppose that $P(X < Y) > 1/2$ and $P(Y < Z) > 1/2$. This does not imply $P(X < Z) > 1/2$ — the title of the paper refers to this phenomenon. But the assumption made about the first two probabilities limits the third probability and it would be interesting to investigate this sort of restrictions.

Thus we have the following problem: For which systems of numbers $\xi_1, \xi_2, \ldots, \xi_n$ there exist independent random variables $X_1, X_2, \ldots, X_n$ such that

$$(1)\qquad \begin{aligned} \xi_i &= P(X_i < X_{i+1}), \quad (i = 1, 2, \ldots, n-1), \\ \xi_n &= P(X_n < X_1). \end{aligned}$$

The case $n = 3$ was solved in [1] and [2]. The general case was considered by Chang Li-chien [3], who obtained the following result: Let us denote

$$\pi_n = \sup \min(\xi_1, \xi_2, \ldots, \xi_n),$$

where the supremum is taken over the set of all random variables $X = (X_1, X_2, \ldots, X_n)$ with independent components satisfying (1). Then

$$\frac{3}{4} - \frac{2n+1}{4n^2} < \pi_n < \frac{3}{4}.$$

In order to avoid trivial complications we shall hold throughout the paper that $P(X_i = X_j) = 0$ for $i \neq j$ $(i, j = 1, 2, \ldots, n)$. Let us define the class $\mathscr{F}$ of random variables $Y = (Y_1, Y_2, \ldots, Y_n)$ as follows: $Y \epsilon \mathscr{F}$ if and only if

a) $Y_1, Y_2, \ldots, Y_n$ are independent;

and

b) $P(Y_i = i) + P(Y_i = i+n) = 1 \quad (i = 1, 2, \ldots, n-1)$,

or

$P(Y_n = n) = 1$,

b') assumption b) is satisfied for $(Y_{T_r(1)}, Y_{T_r(2)}, \ldots, Y_{T_r(n)})$ where

$$T_r(i) = (i+r)(\bmod n).$$

THEOREM 1. *For any set of n independent random variables $X_1, X_2, \ldots$ $\ldots, X_n$ there exist random variables $Y_1, Y_2, \ldots, Y_n$ such that*

$$Y = (Y_1, Y_2, \ldots, Y_n) \epsilon \mathscr{F},$$
$$P(X_i < X_{i+1}) \leqslant P(Y_i < Y_{i+1}) \quad (i = 1, 2, \ldots, n-1),$$
$$P(X_n < X_1) \leqslant P(Y_n < Y_1).$$

To prove the theorem we need the following

LEMMA 1.1. *For any system of numbers $p_1, p_2, \ldots, p_{n-1}$ satisfying the conditions*

$$(2) \qquad \begin{gathered} 0 \leqslant p_i \leqslant 1 \quad (i = 1, 2, \ldots, n-1), \\ p_2 p_3 \ldots p_{n-1} \leqslant (1-p_2)(1-p_3)\ldots(1-p_{n-1}) \end{gathered}$$

there exists a system of numbers $P_2, P_3, \ldots, P_{n-1}$ such that $0 \leqslant P_i \leqslant 1$ $(i = 2, \ldots, n-1)$ and

$$\begin{gathered} p_2(1-p_1) \geqslant P_2, \\ p_3(1-p_2) \geqslant P_3(1-P_2), \\ \cdots\cdots\cdots\cdots\cdots \\ p_{n-1}(1-p_{n-2}) \geqslant P_{n-1}(1-P_{n-2}), \\ 1-p_{n-1} \geqslant (1-p_1)(1-P_{n-1}). \end{gathered}$$

Proof. Suppose, to begin with, that $0 < p_i < 1$ $(i = 2, \ldots, n-1)$ and that

$$(3) \qquad \begin{gathered} P_2 = p_2 \xi, \\ P_{i+1}(1-P_i) = p_{i+1}(1-p_i) \quad (i = 2, \ldots, n-2). \end{gathered}$$

Then $P_3 = p_3(1-p_2)/(1-p_2\xi)$ is a continuous and differentiable function of ξ in the closed interval $\langle 0, 1 \rangle$ and in this interval $0 \leqslant P_3 \leqslant p_3$, $dP_3/d\xi \neq 0$. By inductive argument P_i $(i = 2, \ldots, n-1)$ are continuous and differentiable functions of ξ and $0 \leqslant P_i \leqslant p_i$, $dP_i/d\xi \neq 0$ in the interval $\langle 0, 1\rangle$. We are going to prove that $\xi(1-P_{n-1}) \leqslant 1-p_{n-1}$. Since for $\xi = 1$, $\xi(1-P_{n-1}) = 1-p_{n-1}$ it is sufficient to prove that $\frac{d}{d\xi}\big(\xi(1-P_{n-1})\big) \geqslant 0$. We have

$$(4) \qquad \begin{gathered} P_2P_3\ldots P_{n-1} \leqslant p_2p_3\ldots p_{n-1}, \\ (1-P_2)(1-P_3)\ldots(1-P_{n-1}) \geqslant (1-p_2)(1-p_3)\ldots(1-p_{n-1}). \end{gathered}$$

Differentiating both sides of equalities (3) we obtain

$$(5) \qquad \frac{dP_2}{d\xi} = p_2,$$

$$(6) \qquad \frac{dP_i}{d\xi}(1-P_{i-1}) = \frac{dP_{i-1}}{d\xi}P_i \quad (i = 3, \ldots, n-1).$$

Multiplying side by side equality (5) and the equalities (6) for $i = 3, \ldots, n-1$, we get

$$\frac{dP_{n-1}}{d\xi} = \frac{p_2P_3\ldots P_{n-1}}{(1-P_2)(1-P_3)\ldots(1-P_{n-2})}.$$

Then

$$(7) \qquad \begin{aligned} \frac{d}{d\xi}[\xi(1-P_{n-1})] &= 1-P_{n-1}-\xi\frac{dP_{n-1}}{d\xi} \\ &= \frac{(1-P_2)(1-P_3)\ldots(1-P_{n-1})-P_2P_3\ldots P_{n-1}}{(1-P_2)\ldots(1-P_{n-2})}. \end{aligned}$$

A comparison of (2), (4) and (7) brings us to the conclusion that $\frac{d}{d\xi}\big(\xi(1-P_{n-1})\big) \geqslant 0$ and this ends the proof of lemma 1.1 in the case where $0 < p_i < 1$ $(i = 2, \ldots, n-1)$.

Suppose now that there exists an index i $(i = 2, \ldots, n-1)$ such that $p_i = 0$ or $p_i = 1$. Let i_0 be the smallest index for which this supposition holds. If $p_{i_0} = 0$ we put $P_{i_0} = 0$ and $P_i = p_i$ for $i > i_0$. If $p_{i_0} = 1$ then from (2) we have $p_{i_0}p_{i_0+1}\ldots p_{n-1} = 0$ and there exists an index i_1, $i_1 > i_0$, such that $p_{i_1} = 0$. Then it is enough to put $P_{i_0} = P_{i_0+1} = \ldots = P_{i_1} = 0$ and $P_i = p_i$ for $i > i_1$, which completes the proof.

COROLLARY 1.1. *For any system of numbers $p_2, \ldots, p_n$ satisfying the conditions*

$$0 \leqslant p_i \leqslant 1 \quad (i = 2, 3, \ldots, n),$$
$$p_2p_3\ldots p_{n-1} \geqslant (1-p_2)(1-p_3)\ldots(1-p_{n-1})$$

there exists a system of numbers $Q_2, \ldots, Q_{n-1}$ *such that* $0 \leqslant Q_i \leqslant 1$ $(i = 2, \ldots, n-1)$ *and*

$$\begin{aligned} p_2 &\geqslant Q_2 p_n, \\ p_3(1-p_2) &\geqslant Q_3(1-Q_2), \\ &\cdots\cdots\cdots \\ p_{n-1}(1-p_{n-2}) &\geqslant Q_{n-1}(1-Q_{n-2}), \\ p_n(1-p_{n-1}) &\geqslant 1-Q_{n-1}. \end{aligned}$$

Proof. It suffices to substitute $\overline{p}_i$ for $1-p_{n-i+1}$ and $\overline{P}_i$ for $1-Q_{n-i+1}$ and to apply lemma 1.1 to $\overline{p}_i$ and $\overline{P}_i$.

LEMMA 1.2. *For any set of n independent random variables $X_1, X_2, \ldots, X_n$, for which $P(X_i = X_j) = 0$ if $i \neq j$, and for any $\varepsilon > 0$ there exists a set of independent random variables $Y_1, Y_2, \ldots, Y_n$ such that for $i = 1, 2, \ldots, n$*

a) *the value of Y_i with probability one belongs to a finite set of numbers*;

b) $P(Y_{i_1} = Y_{i_2}) = 0$ *if* $i_1 \neq i_2$;

c) $|P(Y_i < Y_{i+1}) - P(X_i < X_{i+1})| < \varepsilon$ $(i = 1, 2, \ldots, n-1)$,
$|P(Y_n < Y_1) - P(X_n < X_1)| < \varepsilon$.

The easy proof of this lemma is omitted.

Proof of theorem 1. By virtue of lemma 1.2 it suffices to consider only discrete random variables with finite sets of values. Let x_{ij} $(j = 1, 2, \ldots, m_i)$ be the values of the random variable X_i $(i = 1, 2, \ldots, n)$. If we assume some of the probabilities $P(X_i = x_{ij})$ to be equal to 0, then, without loss of generality, we may assume that $m_i = m$ and that $x_{ij} < x_{i+1,j} < x_{i,j+1}$. Let us choose the value $x_{i_0 j_0}$ such that

$$P(X_{i_0} < x_{i_0 j_0}) = 0, \quad P(X_{i_0} = x_{i_0 j_0}) > 0$$

and

$$P(X_i < x_{i_0 j_0}) > 0 \quad \text{for} \quad i \neq i_0.$$

New random variables $X_1', X_2', \ldots, X_n'$ may be defined as follows: X_i' are independent and their distributions are identical with those of X_i, respectively, in the interval $(x_{i_0 j_0}, \infty)$. Moreover, we put

$$P(X_i' = x_{i_0 j_0} - r) = P(X_i \leqslant x_{i_0 j_0}) \quad (i = 1, 2, \ldots, n),$$

where $r = (i - i_0) (\operatorname{mod} n)$.

It is evident that

$$\begin{aligned} P(X_i' < X_{i+1}') &\geqslant P(X_i < X_{i+1}) \quad (i = 1, 2, \ldots, n-1), \\ P(X_n' < X_1') &\geqslant P(X_n < X_1). \end{aligned}$$

Furthermore, there exists a sequence of numbers

$$(8)\quad \begin{aligned} &x'_{11} < x'_{21} < \ldots < x'_{n1} < x'_{12} < x'_{22} < \ldots < x'_{n2} < \ldots < x'_{1m} \\ &\qquad < x'_{2m} < \ldots < x'_{nm}, \\ \text{or } &x'_{21} < x'_{31} < \ldots < x'_{11} < x'_{22} < x'_{32} < \ldots < x'_{12} < \ldots < x'_{2m} \\ &\qquad < x'_{3m} < \ldots < x'_{1m}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \text{or } &x'_{n1} < x'_{n2} < \ldots < x'_{n-1,1} < x'_{n2} < x'_{12} < \ldots < x'_{n-1,2} < \ldots < x'_{nm} \\ &\qquad < x'_{1m} < \ldots < x'_{n-1,m} \end{aligned}$$

such that

$$\sum_{j=1}^{m} P(X'_i = x'_{ij}) = 1 \quad \text{and} \quad P(X'_i = x'_{i1}) > 0 \quad \text{for} \quad i = 1, 2, \ldots, n.$$

Without loss of generality we may assume that the first case occurs in (8). Write

$$p_{ij} = P(X'_i = x'_{ij}).$$

Then

$$\begin{aligned} P(X'_1 < X'_2) &= p_{11}p_{21} + (p_{11}+p_{12})p_{22} + R_1, \\ P(X'_2 < X'_3) &= p_{21}p_{31} + (p_{21}+p_{22})p_{32} + R_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ P(X'_{n-1} < X'_n) &= p_{n-1,1}p_{n1} + (p_{n-1,1}+p_{n-1,2})p_{n2} + R_{n-1}, \\ P(X'_n < X'_1) &= p_{n1}p_{12} + R_n, \end{aligned}$$

where $R_1, R_2, \ldots, R_n$ do not change their values if $p_{i1}+p_{i2}$ $(i = 1, 2, \ldots, n)$ and p_{ij} $(i = 1, 2, \ldots, n;\ j = 3, \ldots, m)$ do not change theirs.

Write

$$p_i = \frac{p_{i1}}{p_{i1}+p_{i2}} \quad (i = 1, 2, \ldots, n)$$

and suppose that inequality (2) holds. Define distributions of independent random variables $Y'_1, Y'_2, \ldots, Y'_n$ as identical with those of $X'_1, X'_2, \ldots, X'_n$ on the interval $\langle x'_{31}, \infty)$ and suppose that

$$\begin{aligned} P(Y'_i = i+c) &= P_i(p_{i1}+p_{i2}), \\ P(Y'_i = n+i+c) &= (1-P_i)(p_{i1}+p_{i2}) \quad (i = 2, 3, \ldots, n), \\ P(Y'_1 = n+1+c) &= p_{11}+p_{12} \end{aligned}$$

where $2n+1+c < x_{31}$. Then

$$(9)\quad \begin{aligned} P(Y'_1 < Y'_2) - P(X'_1 < X'_2) &= (p_{11}+p_{12})(p_{21}+p_{22})\big(p_2(1-p_1)-P_2\big), \\ P(Y'_i < Y'_{i+1}) - P(X'_i < X'_{i+1}) &= (p_{i1}+p_{i2})(p_{i+1,1}+p_{i+1,2}) \times \\ &\quad \times \big(p_{i+1}(1-p_i) - P_{i+1}(1-P_i)\big) \quad (i = 2, \ldots, n-1), \\ P(Y'_n < Y'_1) - P(X'_n < X'_1) &= (p_{n1}+p_{n2})(p_{11}+p_{12})\big(P_n - p_n(1-p_1)\big). \end{aligned}$$

Assume $P_n = p_n(1-p_1)$. By lemma 1.1 there exist such numbers $P_2, \dots, P_{n-1}$ that all the above differences are non-negative.

Suppose now that

$$p_2 p_3 \dots p_{n-1} > (1-p_2)(1-p_3)\dots(1-p_{n-1}).$$

In this case we also define the distributions of $Y_1', \dots, Y_n'$ as identical with those of $X_1, \dots, X_n$ in the interval $\langle x_{31}', \infty)$ but we suppose that

$$\begin{aligned} P(Y_i' = i+c) &= P_i(p_{i1}+p_{i2}), \\ P(Y_i' = n+i+c) &= (1-P_i)(p_{i1}+p_{i2}) \qquad (i = 1, 2, \dots, n-1), \\ P(Y_n' = n+c) &= p_{n1}+p_{n2}. \end{aligned}$$

Then

$$\begin{aligned} &P(Y_i' < Y_{i+1}') - P(X_i' < X_{i+1}') \\ &= (p_{i1}+p_{i2})(p_{i+1,1}+p_{i+1,2})\big(p_{i+1}(1-p_i)-P_{i+1}(1-P_i)\big) \qquad (i = 1, \dots, n-2), \\ &P(Y_{n-1}' < Y_n') - P(X_{n-1}' < X_n') \\ &\qquad\qquad = (p_{n-1,1}+p_{n-1,2})(p_{n1}+p_{n2})\big(p_n(1-p_{n-1})-P_n\big), \\ &P(Y_n' < Y_1') - P(X_n' < X_1') = (p_{n1}+p_{n2})(p_{11}+p_{12})\big(p_n(1-p_1)+P_1-1\big). \end{aligned}$$

Now putting $1-P_1 = p_n(1-p_1)$ and applying corollary 1.1 we come to the conclusion that also in this case the differences are non-negative.

The method described above permits us to reduce one value of one random variable but orders all the values as follows:

$$y_2, y_3, \dots, y_n, \overbrace{\overbrace{y_1, y_2, \dots, y_n}, \dots, \overbrace{y_1, y_2, \dots, y_n}}^{m-1 \text{ times}}$$

in the first case and

$$y_1, y_2, \dots, y_n, \underline{y_1, y_2, \dots, y_{n-1}}, \overbrace{\overbrace{\underline{y_1, y_2, \dots,}\; y_n}, \dots, \overbrace{y_1, y_2, \dots, y_n}}^{m-2 \text{ times}}$$

in the second case (y_i is any value assumed by the random variable Y_i). The second case requires further ordering, which we perform on the underlined values of Y_i. We simply transfer the higher value of the Y_i to the lower one, preserving the probability.

To the distributions thus obtained we apply again the same method and repeat this procedure until we get the random variables belonging to the class $\mathscr{F}$. This ends the proof of theorem 1.

COROLLARY 1.2. *For any set of the independent random variables $X_1, X_2, \dots, X_n$ satisfying the condition $P(X_i = X_j) = 0$ for $i \neq j$ there* *exist random variables* $Y_1, Y_2, \ldots, Y_n$ *such that*

$$Y = (Y_1, \ldots, Y_n) \in \mathscr{F},$$
$$P(X_i < X_{i+1}) \geqslant 1 - P(Y_i < Y_{i+1}) \quad (i = 1, 2, \ldots, n-1),$$
$$P(X_n < X_1) \geqslant 1 - P(Y_n < Y_1).$$

Proof. Put $X_i' = X_i$. Then

$$P(X_i < X_{i+1}) = P(X_i \leqslant X_{i+1}) = 1 - P(X_i' < X_{i+1}')$$
$$P(X_n < X_1) = 1 - P(X_n' < X_1').$$

Now, it is enough to apply theorem 1 to the random variables X_i'.

Let us denote by D_n a subset of the n-dimensional Euclidean space E_n such that $\xi = (\xi_1, \xi_2, \ldots, \xi_n) \in D$ if and only if there exist independent random variables $X_1, X_2, \ldots, X_n$ and $\xi_i = P(X_i < X_{i+1})$ $(i = 1, \ldots, n-1)$ $\xi_n = P(X_n < X_1)$. Obviously, if $\xi \in D_n$, then $0 \leqslant \xi_i \leqslant 1$, but it is not the only restriction imposed on D_n. For $X \in \mathscr{F}$ we have either

$$(10) \qquad \begin{aligned} \xi_i &= 1 - p_{i+1}(1-p_i) \quad (i = 1, \ldots, n-2), \\ \xi_{n-1} &= p_{n-1}, \\ \xi_n &= 1 - p_1, \end{aligned}$$

or equations (10) hold for indices $T_r(i)$ where

$$T_r(i) = (i+r)(\mathrm{mod}\, n) \quad (r = 1, 2, \ldots, n-1).$$

The system of equations (10) defines an $(n-1)$-dimensional surface in E_n. Changing r we obtain n surfaces $S_0, S_1, \ldots, S_{n-1}$ for $r = 1, 2, \ldots, n-1$, respectively. These surfaces limit the set D_n in the sense that if $\xi \in D_n$ then there exist such a surface S_r and such a point $\xi^0 \in S_r$ that $\xi_i \leqslant \xi_i^0$ for $i = 1, 2, \ldots, n$. Taking into account corollary 1.2 we can obtain corresponding restrictions from below. On the other hand, we have

THEOREM 2. *If there exist independent random variables* $X_1, X_2, \ldots, X_n$ *such that* $P(X_i < X_{i+1}) = \xi_i$ $(i = 1, \ldots, n-1)$ *and* $P(X_n < X_1) = \xi_n$ *then for any* $0 \leqslant a \leqslant 1$ *there exist independent random variables* $Y_1, Y_2, \ldots, Y_n$ *such that*

$$P(Y_i < Y_{i+1}) = a\xi_i + (1-a)(1-\xi_i),$$
$$P(Y_n < Y_1) = a\xi_n + (1-a)(1-\xi_n).$$

Proof. Without loss of generality we may suppose that $X_1, X_2, \ldots, X_n$ are bounded. Otherwise we may put, for example, $X_i' = \arctan X_i$ which does not change the probabilities ξ_i. Let $\bar{\mu}_1, \mu_2, \ldots, \mu_n$ denote, respectively, distributions of $X_1, X_2, \ldots, X_n$ and let $\bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_n$ denote, respectively, distributions of $2M - X_1, 2M - X_2, \ldots, 2M - X_n$ where

M is a constant large enough to make $P(|X_1| \leqslant M) = P(|X_2| \leqslant M) = \ldots = P(|X_n| \leqslant M) = 1$. Suppose that the distributions of $Y_1, \ldots, Y_n$ are $\alpha\mu_1 + (1-\alpha)\bar{\mu}_1, \ldots, \alpha\mu_n + (1-\alpha)\bar{\mu}_n$ respectively. Then

$$P(Y_i < Y_{i+1}) = \alpha^2 P(X_i < X_{i+1}) + \alpha(1-\alpha)P(X_i < 2M - X_{i+1}) + \\ + \alpha(1-\alpha)P(2M - X_i < X_{i+1}) + (1-\alpha)^2 P(2M - X_i < 2M - X_{i+1}) \\ = \alpha^2 \xi_i + \alpha(1-\alpha) + (1-\alpha)^2(1-\xi_i) = \alpha\xi_i + (1-\alpha)(1-\xi_i).$$

In the same way we obtain

$$P(Y_n < Y_1) = \alpha\xi_n + (1-\alpha)(1-\xi_n),$$

which completes the proof.

Theorem 1 may be used to calculate the maximum of any non-decreasing function $\varphi(\xi_1, \ldots, \xi_n)$ over D_n. We have

$$\sup_{D_n} \varphi(\xi_1, \ldots, \xi_n) = \sup_{S} \varphi(\xi_1, \ldots, \xi_n) = \sup_{r} \sup_{S_r} \varphi(\xi_1, \ldots, \xi_n) \tag{11}$$

where $S = S_0 \cup S_1 \cup \ldots \cup S_{n-1}$. Formula (11) becomes simpler when φ is a symmetrical function of ξ_i. In this case

$$\sup_{D_n} \varphi(\xi_1, \ldots, \xi_n) \\ = \sup_{\substack{0 \leqslant p_i \leqslant 1 \\ i=1,\ldots,n-1}} \varphi\big(1-p_2(1-p_1), \ldots, 1-p_{n-1}(1-p_{n-2}), p_{n-1}, 1-p_1\big).$$

In this way one can obtain the inequalities

$$\xi_1 \xi_2 \ldots \xi_n \leqslant 1/4, \tag{12}$$

$$\xi_1 + \xi_2 + \ldots + \xi_n \leqslant n-1 \tag{13}$$

and there exist random variables such that equality holds in (12) and (13) respectively. The last inequality may be obtained (see [2]) without the assumption of the independence of X_i.

Using also theorem 1 we shall solve the extremal problem mentioned at the beginning of the paper.

THEOREM 3. *Let $X_1, X_2, \ldots, X_n$ be independent random variables and let*

$$\xi_i = P(X_i < X_{i+1}) \quad (i = 1, \ldots, n-1), \qquad \xi_n = P(X_n < X_1),$$
$$\pi_n = \sup \min(\xi_1, \xi_2, \ldots, \xi_n),$$

where supremum is taken over the set of all random variables $X = (X_1, \ldots, X_n)$. Then π_n is the only number satisfying the conditions

$$p_1 = p_2(1-p_1) = \ldots = p_{n-1}(1-p_{n-2}) = 1-p_{n-1} = 1-\pi_n, \tag{14}$$

where

$$1/4 < p_i < 3/4 \quad (i = 1, 2, \ldots, n-1).$$

We shall first prove two lemmas. To state the first one the definition of the p-quantile will be necessary. The number $x^{(p)}$ will be called a *p-quantile* of the random variable X if

$$P(X \leqslant x^{(p)}) \geqslant p \quad \text{and} \quad P(X \geqslant x^{(p)}) \geqslant 1-p.$$

LEMMA 3.1. *Let X and Y be independent random variables. If $P(X < Y) > a$ and $q(1-p) \geqslant 1-a$, $0 \leqslant p \leqslant 1$, $0 \leqslant q \leqslant 1$, then $x^{(p)} < y^{(q)}$.*

Proof. Let $q(1-p) \geqslant 1-a$ and suppose that $x^{(p)} \geqslant y^{(q)}$. Then

$$\begin{aligned} P(X < Y) &= P(X < Y,\, Y > y^{(q)}) + P(X < Y,\, X \geqslant x^{(p)},\, Y \leqslant y^{(q)}) + \\ &\quad + P(X < Y,\, X < x^{(p)},\, Y \leqslant y^{(q)}) \leqslant P(Y > y^{(q)}) + P(X < x^{(p)},\, Y \leqslant y^{(q)}) \\ &= 1 - P(X \geqslant x^{(p)},\, Y \leqslant y^{(q)}) \leqslant 1 - q(1-p) \leqslant a, \end{aligned}$$

against the supposition that $P(X < Y) > a$.

LEMMA 3.2. *The system of equations*

$$p_1 = p_2(1-p_1) = \ldots = p_{n-1}(1-p_{n-2}) = 1-p_{n-1} \tag{15}$$

has one and only one solution satisfying the conditions

$$\tfrac{1}{4} < p_1 < \tfrac{1}{2}, \quad 0 \leqslant p_i \leqslant 1,$$

and in this case $\frac{1}{4} < p_i < \frac{3}{4}$. When $0 \leqslant p_1 \leqslant \frac{1}{4}$ there exists no solution of (15) such that $0 \leqslant p_i \leqslant 1$ $(i = 2, \ldots, n-1)$.

Proof. Suppose that

$$p_{i+1} = \frac{p_1}{1-p_i} \quad (i = 1, 2, \ldots). \tag{16}$$

Write

$$\Delta p_i = p_{i+1} - p_i.$$

We obtain

$$\Delta p_1 = p_2 - p_1 = \frac{p_1}{1-p_1} - p_1 = \frac{p_1^2}{1-p_1} \geqslant 0,$$

$$\Delta p_i = p_{i+1} - p_i = \frac{p_1}{1-p_i} - \frac{p_1}{1-p_{i-1}} = \frac{p_i}{1-p_i}\Delta p_{i-1}.$$

Then $p_k \geqslant 0$ if $0 \leqslant p_i < 1$ $(i = 1, 2, \ldots, k)$.

Consider the case $0 \leqslant p_1 \leqslant \frac{1}{4}$. We have

$$p_2 = \frac{p_1}{1-p_1} < \frac{1}{2},$$

$$p_{i+1} = \frac{p_1}{1-p_i} < \frac{1}{2} \quad \text{if} \quad p_i < \frac{1}{2}.$$

Then $p_i < \frac{1}{2}$ for $i = 1, 2, \ldots$ and the system of numbers $p_1, p_2, \ldots, p_n$ cannot be a solution of (15).

Suppose now that $\frac{1}{4} < p_1 < \frac{1}{2}$. For $n = 3$ there exists only one solution of (15), $p_1 = \frac{1}{2}(\sqrt{5}-1)$, $p_2 = 1-p_1$. Suppose that there exists a solution $p_1^0, p_2^0, \ldots, p_{n-1}^0$ of (15) for $n = n_0$. Then for $p_i = p_i^0$ formula (16) holds and $p_{n_0} = p_1/(1-p_{n_0}) = 1$. On the other hand, for $p_1 = \frac{1}{4}$, $p_{n_0} < \frac{1}{2}$. The number p_{n_0} defined by (16) can be considered as a function of p_1 increasing and continuous in the interval $(\frac{1}{4}, p_1^0)$. Since for $p_1 = p_1^0$, $p_n > 1-p_1$, in the interval $(\frac{1}{4}, p_1^0)$ there exists a point p_1 such that $p_n = 1-p_1$. The solution of (15) is uniquely determined by this point and it is the only solution because for $0 < p_i < 1$ $(i = 1, 2, \ldots, k)$ p_k is a strictly increasing function of p_1.

Proof of theorem 3. By theorem 1

$$\pi_n = \sup_{Y \in \mathcal{F}} \min(\xi_1, \xi_2, \ldots, \xi_n).$$

Since $\min(\xi_1, \xi_2, \ldots, \xi_n)$ is a symmetrical function of $\xi_1, \ldots, \xi_n$, then without loss of generality we may suppose that

$$P(Y_i = i) + P(Y_i = i+n) = 1, \quad P(Y_n = n) = 1.$$

Let $p_1, \ldots, p_{n-1}$ be the solution of (14). Then $p_{i+1}(1-p_i) = 1-p_{n-1}$ $(i = 1, \ldots, n-2)$. Suppose that $\xi_i > p_{n-1}$ $(i = 1, \ldots, n)$. Applying lemma 3.1 we obtain

$$y_1^{(p_1)} < y_2^{(p_2)} < \ldots < y_{n-1}^{(p_{n-1})} \tag{18}$$

where $y_i^{(p_i)}$ is a p_i-quantile of the random variable Y_i. Moreover, since $P(Y_{n-1} < Y_n) > p_{n-1}$ and $P(Y_n = n) = 1$, we have $y_{n-1}^{(p_{n-1})} < n$. On the other hand, $P(Y_n < Y_1) > p_{n-1} = 1-p_1$; then $y_1^{(p_1)} > n$ and we have $y_1^{(p_1)} > y_{n-1}^{(p_{n-1})}$. But this is impossible in view of (18) and the inequality $\pi_n \leqslant p_{n-1}$ is proved.

We shall prove that $\pi_n \geqslant p_{n-1}$. Suppose that

$$P(Y_i = i) = p_i, \quad P(Y_i = i+n) = 1-p_i \quad (i = 1, \ldots, n-1),$$
$$P(Y_n = n) = 1,$$

where p_i satisfy (14). Then

$$P(Y_i < Y_{i+1}) = 1-p_{i+1}(1-p_i) = p_{n-1},$$
$$P(Y_{n-1} < Y_n) = p_{n-1},$$
$$P(Y_n < Y_1) = 1-p_1 = p_{n-1},$$

and $\min(\xi_1, \xi_2, \ldots, \xi_n) = p_{n-1}$, which completes the proof.

Here are the numerical values of the first 30 π_n's as calculated from equation (15).

n	π_n	n	π_n	n	π_n
3	0,61803	12	0,73870	21	0,74567
4	0,66667	13	0,74011	22	0,74601
5	0,69202	14	0,74126	23	0,74631
6	0,71688	15	0,74228	24	0,74658
7	0,72361	16	0,74304	25	0,74683
8	0,72844	17	0,74373	26	0,74704
9	0,73205	18	0,74432	27	0,74724
10	0,73481	19	0,74483	28	0,74741
11	0,73698	20	0,74528	29	0,74757
				30	0,74772

THEOREM 4. *Let $X_1, X_2, \ldots, X_n$ be independent random variables and let*

$$\xi_i = P(X_i < X_{i+1}) \quad (i = 1, \ldots, n-1), \qquad \xi_n = P(X_n < X_1),$$

$$\pi_n = \sup\min(\xi_1, \xi_2, \ldots, \xi_n),$$

where supremum is taken over the set of all random variables $X = (X_1, \ldots, X_n)$. Then

$$\frac{3}{4} - \frac{3}{n(n+4)} \leqslant \pi_n < \frac{3}{4}.$$

Proof. The inequality $\frac{3}{4} > \pi_n$ follows immediately from lemma 3.2 and theorem 3. We shall prove that $\pi_n \geqslant \frac{3}{4} - \frac{3}{n(n+4)}$. It is enough to find n independent random variables such that $\xi_i > \frac{3}{4} - \frac{3}{n(n+4)}$ for $i = 1, \ldots, n$. Suppose that

$$P(X_i = i) = p_i, \quad P(X_i = n+i) = 1-p_i \quad (i = 1, 2, \ldots, n-1),$$

$$P(X_n = n) = 1,$$

where

$$p_i = \frac{i}{2(i+1)}\left(1 + \frac{4(i+2)}{n(n+4)}\right) \quad \text{for} \quad i \leqslant n/2,$$

$$p_i = 1 - p_{n-i} \quad \text{for} \quad i > n/2 \ (i = 1, 2, \ldots, n-1).$$

Put $\beta_i = 4(i+2)/n(n+4)$ and suppose that $i \leqslant \frac{1}{2}n - 1$. We have

$$p_{i+1}(1-p_i) = \frac{i+1}{2(i+2)}(1+\beta_{i+1})\left(1 - \frac{(1+\beta_i)i}{2(i+1)}\right)$$

$$= \frac{1}{4} + \frac{1}{4}\left(\beta_{i+1} - \frac{i\beta_i}{i+2}(1+\beta_{i+1})\right) < \frac{1}{4} + \frac{1}{4}\left(\beta_{i+1} - \frac{i\beta_i}{i+2}\right)$$

$$= \frac{1}{4} + \frac{3}{n(n+4)}.$$

Then

$$\xi_i = P(X_i < X_{i+1}) = 1-p_{i+1}(1-p_i) > \frac{3}{4} - \frac{3}{n(n+4)} \quad \text{for } i \leqslant \frac{n}{2}-1$$

and

$$\xi_i = 1-p_{i+1}(1-p_i) = 1-p_{n-i}(1-p_{n-i+1}) > \frac{3}{4} - \frac{3}{n(n+4)}$$

for $\frac{1}{2}n \leqslant i < n-1$.

The case $\frac{1}{2}n-1 < i < \frac{1}{2}n$ requires special consideration. It can occur when n is an odd number. Taking into account that $p_{n-i} = 1-p_i$, we obtain

$$\xi_{n/2-1/2} = 1-p_{n/2+1/2}(1-p_{n/2-1/2}) = 1-(1-p_{n/2-1/2})^2 = p_{n/2-1/2}(2-p_{n/2-1/2})$$

$$= \left(\frac{1}{2} - \frac{2n+3}{n(n+1)(n+4)}\right)\left(\frac{3}{2} + \frac{2n+3}{n(n+1)(n+4)}\right) > \frac{3}{4} - \frac{3}{4(n+4)} \quad \text{for} \quad n > 2.$$

Moreover,

$$\xi_{n-1} = P(X_{n-1} < X_n) = p_{n-1} = 1-p_1 = \frac{3}{4} - \frac{3}{n(n+4)},$$

$$\xi_n = 1-p_1 = \frac{3}{4} - \frac{3}{n(n+4)}.$$

Then inequality $\xi_i \geqslant \frac{3}{4} - \frac{3}{n(n+4)}$ holds for $i = 1, 2, \ldots, n$, which completes the proof.

References

[1] H. Steinhaus and S. Trybuła, *On a paradox in applied probabilities* Bull. Acad. Polon. Sci. 7 (1959), p. 67.

[2] S. Trybuła, *On the paradox of three random variables*, Zastosow. Mat. 5 (1961), p. 321.

[3] Chang Li-chien, *On the maximin probability of cyclic random inequalities*, Scientia Sinica 10(1961), pp. 499-504.

WROCŁAW TECHNICAL INSTITUTE, DEPARTMENT OF MATHEMATICS
POLITECHNIKA WROCŁAWSKA, KATEDRA MATEMATYKI

Received on 23. 11. 1964

S. TRYBUŁA (Wrocław)

O PARADOKSIE n ZMIENNYCH LOSOWYCH

STRESZCZENIE

Praca poświęcona jest następującemu problematowi: Dla jakich układów liczb $(\xi_1, \xi_2, \ldots, \xi_n)$ istnieją niezależne zmienne losowe $X_1, X_2, \ldots, X_n$ takie, że

$$
\begin{aligned}
P(X_i < X_{i+1}) &= \xi_i \quad (i = 1, 2, \ldots, n-1), \\
P(X_n < X_1) &= \xi_n.
\end{aligned} \tag{1}
$$

Zdefiniujmy klasę $\mathscr{F}$ zmiennych losowych $Y = (Y_1, \ldots, Y_n)$ jak następuje: $Y \in \mathscr{F}$ wtedy i tylko wtedy, gdy

a) zmienne losowe $Y_1, Y_2, \ldots, Y_n$ są niezależne

i

b) $P(Y_i = i) + P(Y_i = i+n) = 1$ $(i = 1, 2, \ldots, n-1), P(Y_n = n) = 1$

lub

b′) założenie b) jest spełnione dla układu $(Y_{T_r(1)}, Y_{T_r(2)}, \ldots, Y_{T_r(n)})$, gdzie

$$T_r(i) = (i+r)(\mathrm{mod}\, n).$$

W pracy udowodniono, że dla każdego układu niezależnych zmiennych losowych $X_1, X_2, \ldots, X_n$ można dobrać zmienne $Y_1, Y_2, \ldots, Y_n$ takie, że

$$
\begin{gathered}
Y = (Y_1, \ldots, Y_n) \in \mathscr{F}, \\
P(X_i < X_{i+1}) \leqslant P(Y_i < Y_{i+1}) \quad (i = 1, 2, \ldots, n-1), \\
P(X_n < X_1) \leqslant P(Y_n < Y_1).
\end{gathered}
$$

Powyższy rezultat pozwala obliczyć maksimum funkcji $\varphi(\xi_1, \ldots, \xi_n)$, dla ξ^i określonych wzorem (1), niemalejącej ze względu na każdą ze zmiennych ξ_i. W ten sposób można np. otrzymać nierówności

$$
\begin{gathered}
\xi_1 \xi_2 \cdots \xi_n \leqslant \tfrac{1}{4}, \\
\xi_1 + \xi_2 + \ldots + \xi_n \leqslant n-1,
\end{gathered}
$$

przy czym stałych występujących po prawej stronie nierówności nie można poprawić.

Innym zastosowaniem powyższego rezultatu jest odpowiedź na zagadnienie postawione przez Hugona Steinhausa. Wyznaczyć wartość

$$\pi_n = \sup_X \min(\xi_1, \ldots, \xi_n),$$

gdzie supremum wzięte jest na zbiorze wszystkich zmiennych losowych $X = (X_1, X_2, \ldots, X_n)$ o niezależnych komponentach spełniających warunek (1).

W pracy udowodniono, że dla każdego n, liczba π_n jest jednoznacznie wyznaczona przez warunki

$$p_1 = p_2(1-p_1) = \ldots = p_{n-1}(1-p_{n-2}) = 1-p_{n-1} = 1-\pi_n, \quad \tfrac{1}{4} < p_i < \tfrac{3}{4}.$$

Tabelka na str. 153 podaje wartości π_n obliczone z powyższego układu równań dla $n = 3, 4, \ldots, 30$. Ponadto pokazano, że dla $n > 2$

$$\frac{3}{4} - \frac{3}{n(n+4)} \leqslant \pi_n < \frac{3}{4}.$$

С. ТРЫБУЛА (Вроцлав)

О ПАРАДОКСЕ n СЛУЧАЙНЫХ ВЕЛИЧИН

РЕЗЮМЕ

Работа посвящена следующей задаче: для каких систем чисел $(\xi_1, \xi_2, \ldots, \xi_n)$ существуют независимые случайные величины $X_1, X_2, \ldots, X_n$ такие, чтобы

$$(1)\qquad \begin{aligned} P(X_i < X_{i+1}) &= \xi_i \quad (i = 1, 2, \ldots, n-1),\\ P(X_n < X_1) &= \xi_n. \end{aligned}$$

Определим класс $\mathscr{F}$ случайных величин $Y = (Y_1, Y_2, \ldots, Y_n)$ следующим образом: $y \in \mathscr{F}$ тогда и только тогда, когда

а) случайные величины $Y_1, Y_2, \ldots, Y_n$ являются независимыми,

б) $P(Y_i = i) + P(Y_i = i+n) = 1$ $(i = 1, 2, \ldots, n-1)$, $P(Y_n = n) = 1$,

либо

б′) предположение б) выполняется для системы $(Y_{T_r(1)}, Y_{T_r(2)}, \ldots, Y_{T_r(n)})$, где

$$T_r(i) = (i+r)(\mathrm{mod}\, n).$$

В работе доказано, что для любой системы независимых случайных величин $X_1, X_2, \ldots, X_n$ можно подобрать так переменные $Y_1, Y_2, \ldots, Y_n$, чтобы

$$Y = (Y_1, Y_2, \ldots, Y_n) \in \mathscr{F},$$
$$P(X_i < X_{i+1}) \leqslant P(Y_i < Y_{i+1}) \quad (i = 1, 2, \ldots, n-1),$$
$$P(X_n < X_1) \leqslant P(Y_n < Y_1).$$

Указанный выше результат дает возможность вычисления максимума функции $\varphi(\xi_1, \xi_2, \ldots, \xi_n)$ для ξ_i определенных по формуле (1), неубывающей относительно каждой переменной ξ_i. Таким образом можно например получить неравенства

$$\xi_1 \xi_2 \cdots \xi_n \leqslant \tfrac{1}{4},$$
$$\xi_1 + \xi_2 + \ldots + \xi_n \leqslant n-1.$$

При этом постоянные в правой части неравенств нельзя улучшить. Другим применением этого результата является ответ на вопрос поставленный Хуго Штайнгаусом. Определить величину

$$\pi_n = \sup_X \min(\xi_1, \xi_2, \ldots, \xi_n),$$

где супремум берется в множестве всех случайных величин $X = (X_1, X_2, \ldots, X_n)$ с независимыми составляющими и удовлетворяющими условию (1).

В работе доказано, что для любого n число π_n определяется однозначно условиями

$$p_1 = p_2(1-p_1) = \ldots = p_{n-1}(1-p_{n-2}) = 1-p_{n-1} = 1-\pi_n, \quad \tfrac{1}{4} < p_i < \tfrac{3}{4}.$$

Таблица на стр. 153 дает величины π_n вычисленные по указанной выше системе уравнений при $n = 3, 4, \ldots, 30$. Кроме того показано, что при $n \geqslant 2$

$$\frac{3}{4} - \frac{3}{n(n+4)} \leqslant \pi_n < \frac{3}{4}.$$