Stanislav Jílovec; Bruno Šubert

Repetitive play of a game against nature

Aplikace matematiky, Vol. 12 (1967), No. 5, 383–397

Persistent URL: http://dml.cz/dmlcz/103115

REPETITIVE PLAY OF A GAME AGAINST NATURE

STANISLAV JÍLOVEC, BRUNO ŠUBERT

(Received August 29, 1966.)

I. INTRODUCTION

Let us consider a repetitive play of a two-person game, that is a sequence of the component games G identical in structure, with only the moves of players changing at each step. The player I is supposed to possess the character of "Nature", that is his motivations are completely unspecified and, thus, his choices may constitute quite an arbitrary sequence of strategies. The II$^{\text{nd}}$ player, on the other hand, may, at each step, choose his strategy on the basis of the I$^{\text{st}}$ player's past moves and his goal is to minimize the average loss for fixed but arbitrary number of plays n previously unknown to him.

If $\bar{\vartheta}_n$, the relative frequency vector of the I$^{\text{st}}$ player's choices during the n plays, were previously known to the player II, it would be the best for him to choose at each step the strategy optimal with respect to $\bar{\vartheta}_n$, i.e. that minimizing his loss in the component game G where his oponent uses the mixed strategy $\bar{\vartheta}_n$. Hence, if $\Phi(\bar{\vartheta}_n)$ denotes the minimum loss thus obtained in G, the problem in repeated play is to determine the sequence of II$^{\text{nd}}$ player's strategies $s_1, s_2, \ldots, s_n$, where s_k depends only on the first $k - 1$ moves $\vartheta_1, \vartheta_2, \ldots, \vartheta_{k-1}$ of the player I, and such that the average loss approaches the minimum $\Phi(\bar{\vartheta}_n)$ whatever be the I$^{\text{st}}$ player's choices.

This problem was first treated by Hannan [1], who has shown that there does exist a sequence of strategies $s_1, s_2, \ldots$ such that the difference of the (expected) average loss from $\Phi(\bar{\vartheta}_n)$ does not exceed $c : \sqrt{n}$, where c is a constant. The strategies s_k were independent on the number of plays and were defined as optimal strategies in G against "artificially randomized" relative frequency vectors $\bar{\vartheta}_{k-1}$. However, unless the artificial randomization was of a special type, the results were proved for finite games only.

The Hannan's idea was applied by Samuel [4] to the sequential statistical decision problem with two parameter values. Here, the main difference is that the II$^{\text{nd}}$ player (the statistician) cannot observe the values ϑ_k (parameter values) directly but can dispose only with their estimates obtained from observation of the sample variable.

Again, it was proved that if the decision function at each step is Bayes against the artificially randomized estimate of $\bar{\vartheta}_k$ the average risk approaches the Bayes risk $\Phi(\bar{\vartheta}_n)$.

Finally, using different method of the proof, Van Ryzin [3] has shown that the artificial randomization is unnecessary for the sequential statistical decision problem and extended the above result to loss matrices of arbitrary but finite size. Nevertheless, the finiteness of the decision space was essential for the proof and his theorems do not yield Hannan's result as special case.

In this paper, general theorems are proved for the repeated play problem of two-person game with two strategies of the I[st] player but infinite strategy set of the II[nd] player. The theorems yield directly the main results of Hannan and Samuel and it is shown (Theorem 3) that they yield also the main result of Van Ryzin. This is because the decision problem can be imbedded in a game theoretical model and the theorems applied, which could not be made with Hannan's theorems for the sake of assumed finiteness. The restriction in number of I[st] player strategies does not seem to be essential for the proof and we believe it can be removed in the future.

II. PREREQUISITES

Throughout this paper the letter N will denote the set of all positive integers, I the set of all integers and R the set of all real numbers.

If f is a real valued function defined on R then its *variation* $V(f)$ will be defined by

$$V(f) = \sup \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})|$$

where the supremum is taken over the class of all finite increasing sequences $x_0 < < x_1 < \ldots < x_n$, $n \in N$ of real numbers.

The mathematical expectation of a random variable z will be denoted by Ez.

Let A be any nonempty set and let $w(i, a)$ be a real valued function defined on the cartesian product $\{0, 1\} \times A$. Further on, this function will be assumed *bounded by a constant* $K < \infty$. The set A represents the set of II[nd] player's choices and w the payoff function in the component game $G = (\{0, 1\}, A, w)$.

The symbol Θ will denote the set of all mappings ϑ from N into $\{0, 1\}$. The value of the mapping ϑ in $k \in N$ will be denoted by ϑ_k. Thus, every $\vartheta \in \Theta$ is a sequence $\{\vartheta_k\}_{k=1}^{\infty}$ of zeros and ones and represents a possible sequence of I[st] player's choices in the repetitive play. The relative frequency of ones among the first n members of a sequence $\vartheta \in \Theta$ will be denoted $\bar{\vartheta}_n$,

$$\bar{\vartheta}_n = \frac{1}{n} \sum_{k=1}^{n} \vartheta_k\,, \quad n \in N \quad \text{and} \quad \bar{\vartheta}_0 = 0\,.$$

For further purposes we shall consider the linear extension of the function w onto $R \times A$, which will be denoted by the same letter w and defined by the equation

$$w(x, a) = (1 - x)\, w(0, a) + x\, w(1, a)\,, \quad x \in R\,, \quad a \in A\,. \tag{1}$$

The minimum or Bayes loss function Φ is defined by

$$\Phi(x) = \inf_{a \in A} w(x, a)\,, \quad x \in R\,.$$

It is well known that Φ is concave and continuous in R.

In the language of game theory the elements $a \in A$ are called pure strategies of the II$^{\text{nd}}$ player. However, we shall use this term in a broader sense. By the *strategy* we shall understand a mapping from R into A.

Let $\varepsilon \geqq 0$. A strategy b_ε will be called *ε-optimal* if

$$w(x, b_\varepsilon(x)) \leq \Phi(x) + \varepsilon \quad \text{for every} \quad x \in R\,. \tag{P1}$$

It is easily seen that, for every $\varepsilon > 0$, an ε-optimal strategy always exists.

Moreover, the class of all ε-optimal strategies contains a nonempty subclass B of *regular* ε-optimal strategies with the following properties:

1) B contains an ε-optimal strategy for every $\varepsilon > 0$.

2) There is a constant $c_0 < \infty$ such that

$$\max_{i \in \{0,1\}} V(w(i, b_\varepsilon(.))) \leqq c_0 + \varepsilon \tag{P2}$$

for every $b_\varepsilon \in B$.

3) For every two strategies $b_{\varepsilon_1}, b_{\varepsilon_2} \in B$ and every $x \in R$

$$\max_{i \in \{0,1\}} \left| w(i, b_{\varepsilon_1}(x)) - w(i, b_{\varepsilon_2}(x)) \right| \leqq \varepsilon_1 + \varepsilon_2\,. \tag{P3}$$

For the proof, fix $\varepsilon > 0$ and let $\{b_n\}_{n \in N}$ be a sequence of $1/n$-optimal strategies such that the limits

$$w_0(x) = \lim_{n \to \infty} w(0, b_n(x))$$

and

$$w_1(x) = \lim_{n \to \infty} w(1, b_n(x))$$

exist.

Obviously,

$$\Phi(x) = (1 - x)\, w_0(x) + x\, w_1(x)\,. \tag{2}$$

The functions w_i, $i = 0, 1$, are nonincreasing in $(-\infty, i)$ and nondecreasing in (i, ∞).

This follows easily from the inequalities

$$(1 - x) w_0(y) + x w_1(y) = \lim_{n \to \infty} w(x, b_n(y)) \geqq \Phi(x),$$

$$(1 - y) w_0(x) + y w_1(x) = \lim_{n \to \infty} w(y, b_n(x)) \geqq \Phi(y),$$

$x, y \in R$, which, together with (2), yield

$$x \,(w_1(x) - w_1(y)) \leqq (1 - x)(w_0(y) - w_0(x)),$$

$$(1 - y)(w_0(y) - w_0(x)) \leqq y \,(w_1(x) - w_1(y)).$$

Hence, for $x < y < 1$, we have

$$\left(\frac{x}{1 - x} - \frac{y}{1 - y}\right)(w_1(x) - w_1(y)) \leqq 0$$

which implies $w_1(x) \geqq w_1(y)$, and the same with reversed inequality sign for $1 < < x < y$. Similarly for w_0.

The just proved property, together with the assumption that w is bounded, implies that R may be covered by a sequence of nonempty, disjoint intervals $\{E_m\}_{m \in I}$ of at most unit length with the property

(3) $$(m \in I, \quad x, y \in E_m) \Rightarrow \max_{i \in \{0,1\}} |w_i(x) - w_i(y)| < \frac{\varepsilon}{4(1 + d_m)}$$

where $d_m = \sup\{|x| : x \in E_m\}$. Let $\{x_m\}_{m \in I}$ be a sequence of numbers such that $x_m \in \in E_m$, $m \in I$ and $\{n_m\}_{m \in I}$ a sequence of positive integers such that

(4) $$\max_{i \in \{0,1\}} |w_i(x_m) - w(i, b_{n_m}(x_m))| < \frac{\varepsilon}{6 \, . \, 2^{|m|}(1 + d_m)}.$$

We may assume without loss of generality that the sequence $\{x_m\}_{m \in I}$ is increasing. We shall define the strategy b_ε by

$$b_\varepsilon(x) = b_{n_m}(x_m) \quad \text{whenever} \quad x \in E_m.$$

Since, for every $x \in R$ there is $m \in I$ such that $x \in E_m$, and since by (3) and (4)

$$w(x, b_\varepsilon(x)) - \Phi(x) \leqq |1 - x| \, |w(0, b_{n_m}(x_m)) - w_0(x_m)| +$$
$$+ |x| \, |w(1, b_{n_m}(x_m)) - w_1(x_m)| + |1 - x| \, |w_0(x_m) - w_0(x)| +$$
$$+ |x| \, |w_1(x_m) - w_1(x)| < \varepsilon$$

the strategy b_ε is ε-optimal. Further, for any $y_0 < y_1 < \ldots < y_n$, $n \in N$, $i \in \{0, 1\}$,

(3) and (4) imply

$$\sum_{j=1}^{n} \left|w(i, b_\varepsilon(y_j)) - w(i, b_\varepsilon(y_{j-1}))\right| \leqq$$

$$\leqq \sum_{m\in I} \left|w(i, b_\varepsilon(x_m)) - w(i, b_\varepsilon(x_{m-1}))\right| \leqq$$

$$\leqq \sum_{m\in I} \left|w(i, b_\varepsilon(x_m)) - w_i(x_m)\right| + \sum_{m\in I} \left|w_i(x_m) - w_i(x_{m-1})\right| +$$

$$+ \sum_{m\in I} \left|w_i(x_{m-1}) - w(i, b_\varepsilon(x_{m-1}))\right| < \varepsilon + V(w_i)$$

so that b_ε satisfies (P2). However, from (3) and (4) follows easily that (P3) is satisfied as well. Hence b_ε is regular, which was to be proved.

III. GENERAL THEOREMS

Before proving the two general theorems we shall prove the following fundamental

Lemma. *Let F_1 and F_2 be distribution functions, let f be a real valued function defined on R. If f is of bounded variation on R, then*

$$\left|\int f \,\mathrm{d}F_1 - \int f \,\mathrm{d}F_2\right| \leqq V(f) \sup_{x\in R} \left|F_1(x) - F_2(x)\right|$$

where the integrals are Lebesgue-Stieltjes.

Proof: Since f is of bounded variation there are two nondecreasing bounded functions f_1 and f_2 such that $f = f_1 - f_2$ and $V(f) = V(f_1) + V(f_2)$. Hence, it is sufficient to prove the lemma for f nondecreasing and bounded.

Then $V(f) = f(+\infty) - f(-\infty)$, where $f(+\infty) = \lim_{x\to+\infty} f(x)$ and $f(-\infty) = \lim_{x\to-\infty} f(x)$. Let

$$E_{m,j} = \left\{x : f(-\infty) + \frac{j-1}{m} V(f) < f(x) \leqq f(-\infty) + \frac{j}{m} V(f)\right\},$$

$$m = 1, 2, \dots, \quad j = 0, 1, \dots, m$$

and let $\{f_m\}_{m\in N}$ be a sequence of simple functions defined by

$$f_m(x) = \sum_{j=0}^{m} \left(f(-\infty) + \frac{j}{m} V(f)\right) \chi_{E_{m,j}}(x), \tag{5}$$

where $\chi_{E_{m,j}}$ denotes the characteristic function of the set $E_{m,j}$. Obviously, f_m are bounded, Borel measurable functions such that

$$\sup_{x\in R} \left|f_m(x) - f(x)\right| \leqq \frac{V(f)}{m}, \quad m \in N .$$

It follows

$$(6)\qquad \left|\int f\,\mathrm{d}F_1 - \int f\,\mathrm{d}F_2\right| \leqq \int |f - f_m|\,\mathrm{d}F_1 + \int |f - f_m|\,\mathrm{d}F_2 + $$
$$+ \left|\int f_m\,\mathrm{d}F_1 - \int f_m\,\mathrm{d}F_2\right| \leqq \left|\int f_m\,\mathrm{d}F_1 - \int f_m\,\mathrm{d}F_2\right| + \frac{2}{m}\,\mathsf{V}(f)\,.$$

Further, let μ_i, $i = 1, 2$ denote the probability measure induced by the distribution function F_i. From (5) we have

$$\int f_m\,\mathrm{d}F_1 - \int f_m\,\mathrm{d}F_2 = \frac{\mathsf{V}(f)}{m}\sum_{j=1}^{m} j(\mu_1(E_{m,j}) - \mu_2(E_{m,j})) =$$
$$= \frac{\mathsf{V}(f)}{m}\sum_{j=j}^{m}(\mu_1(\bigcup_{k=j}^{m} E_{m,k}) - \mu_2(\bigcup_{k=j}^{m} E_{m,k}))$$

since the sets $\{E_{m,j}\}_{j=0}^{m}$ are disjoint. Obviously,

$$\mu_i(\bigcup_{k=j}^{m} E_{m,k}) = 1 - \mu_i(\bigcup_{k=0}^{j-1} E_{m,k})\,, \quad i = 1, 2\,, \quad j = 1, 2, \ldots, m\,,$$

so that

$$(7)\qquad \int f_m\,\mathrm{d}F_1 - \int f_m\,\mathrm{d}F_2 = \frac{\mathsf{V}(f)}{m}\sum_{j=1}^{m}(\mu_2(\bigcup_{k=0}^{j-1} E_{m,k}) - \mu(\bigcup_{k=0}^{j-1} E_{m,k}))\,.$$

Since, by assumption, f is nondecreasing, the sets $\bigcup_{k=0}^{j-1} E_{m,k}$, $j = 1, \ldots, m$, are either empty sets or intervals of the type $(-\infty, a\rangle$ or $(-\infty, a)$. It follows

$$\left|\mu_2(\bigcup_{k=0}^{j-1} E_{m,k}) - \mu_1(\bigcup_{k=0}^{j-1} E_{m,k})\right| \leqq \sup_{x\in R}\left|F_1(x) - F_2(x)\right|\,,$$

which together with (6) and (7) yields

$$\left|\int f\,\mathrm{d}F_1 - \int f\,\mathrm{d}F_2\right| \leqq \mathsf{V}(f)\sup_{x\in R}\left|F_1(x) - F_2(x)\right| + \frac{2}{m}\,\mathsf{V}(f)\,, \quad m \in N\,.$$

The statement is now obtained by letting $m \to \infty$.

Theorem 1. *Let $\{\gamma_k\}_{k=0}^{\infty}$ be a sequence of real numbers, let $\{\alpha_k\}_{k=0}^{\infty}$ be a sequence of positive real numbers such that*

(C1) $\qquad k\alpha_k < (k + 1)\,\alpha_{k+1}$ *for every* $k = 0, 1, \ldots$

Let $\{z_k\}_{k=0}^{\infty}$ be a sequence of random variables satisfying the conditions

(C2) $$\sup_{k\in N\cup\{0\}} \mathsf{E}|z_k| = c_1 < \infty\,,$$

(C3) $$\mathsf{E}(z_k \mid z_{k+1}) = (1 - \gamma_k)\,z_{k+1} \quad \text{a.s.} \quad k = 0, 1, \ldots$$

let b_{ε_k}, $k = 0, 1, \ldots$, be regular ε_k-optimal strategies.

Then for every $\vartheta \in \Theta$ and $n \in N$,

$$\mathsf{E}\frac{1}{n}\sum_{k=1}^{n} w(\vartheta_k, b_{\varepsilon_{k-1}}(\bar{\vartheta}_{k-1} + \alpha_{k-1}z_{k-1})) - \Phi(\bar{\vartheta}_n) \leqq 2Kc_1\left(2\alpha_n + \frac{1}{n}\sum_{k=1}^{n-1} k\alpha_k|\gamma_k|\right) +$$

$$+ \frac{1}{n}\sum_{k=1}^{n}((k+1)\,\varepsilon_k + \varepsilon_{k-1}) + \frac{1}{n}\sum_{k=1}^{n}(c_0 + \varepsilon_k)\sup_{x\in R}\left|F_k\left(\frac{x - \bar{\vartheta}_k}{\alpha_k}\right) - F_{k-1}\left(\frac{x - \bar{\vartheta}_{k-1}}{\alpha_{k-1}}\right)\right|$$

where F_k, $k = 0, 1, \ldots$ is the distribution function of z_k.

Proof: Notice first that (P2) implies that, for $b_{\varepsilon_{k-1}}$ regular, $w(\vartheta_k, b_{\varepsilon_{k-1}}(\bar{\vartheta}_{k-1} + \alpha_{k-1}x))$ are bounded Borel functions of x and, hence, $w(\vartheta_k, b_{\varepsilon_{k-1}}(\bar{\vartheta}_{k-1} + \alpha_{k-1}z_{k-1}))$ are random variables with finite expectations.

Let $\vartheta \in \Theta$ and let us denote

$$s_k = b_{\varepsilon_k}(\bar{\vartheta}_k + \alpha_k z_k)\,, \quad k = 0, 1, \ldots,$$

$$Q_1 = \mathsf{E}\frac{1}{n}\sum_{k=1}^{n} w(\vartheta_k, s_k) - \Phi(\bar{\vartheta}_n)\,,$$

$$Q_2 = \mathsf{E}\frac{1}{n}\sum_{k=1}^{n}(w(\vartheta_k, s_{k-1}) - w(\vartheta_k, s_k))\,,$$

so that the left-hand side of the inequality to be proved is $Q_1 + Q_2$. Using (1), the expression Q_1 can be written as follows:

$$Q_1 = \mathsf{E}\frac{1}{n}\sum_{k=1}^{n}(kw(\bar{\vartheta}_k, s_k) - (k-1)\,w(\bar{\vartheta}_{k-1}, s_k)) - \Phi(\bar{\vartheta}_n) = \tag{8}$$

$$= \mathsf{E}\frac{1}{n}\left(\sum_{k=1}^{n} kw(\bar{\vartheta}_k, s_k) - \sum_{k=1}^{n-1} kw(\bar{\vartheta}_k, s_{k+1}) - n\,\Phi(\bar{\vartheta}_n)\right) =$$

$$= \mathsf{E}\frac{1}{n}\left(\sum_{k=1}^{n-1} k(w(\bar{\vartheta}_k, s_k) - w(\bar{\vartheta}_k, s_{k+1})) + n(w(\bar{\vartheta}_n, s_n) - \Phi(\bar{\vartheta}_n))\right).$$

Further, since b_{ε_k} is ε_k-optimal,

$$w(\bar{\vartheta}_k + \alpha_k z_k, s_k) \leqq w(\bar{\vartheta}_k + \alpha_k z_k, s_{k+1}) + \varepsilon_k\,, \quad k = 0, 1, \ldots$$

which, by using (1), yields

$$\begin{aligned} w(\bar{\vartheta}_k, s_k) - w(\bar{\vartheta}_k, s_{k+1}) &\leqq \alpha_k z_k(w(0, s_k) - w(0, s_{k+1})) - \\ &- \alpha_k z_k(w(1, s_k) - w(1, s_{k+1})) + \varepsilon_k\,, \quad k = 1, 2, \ldots \end{aligned} \tag{9}$$

Let b_ε be regular ε-optimal strategy. For the same reason

$$w(\bar{\vartheta}_n + \alpha_n z_n, s_n) \leqq w(\bar{\vartheta}_n + \alpha_n z_n, b_\varepsilon(\bar{\vartheta}_n)) + \varepsilon_n\,,$$

which implies

$$w(\vartheta_n, s_n) - \Phi(\vartheta_n) \leqq \alpha_n z_n(w(0, s_n) - w(0, b_\varepsilon(\vartheta_n))) - \\ - \alpha_n z_n(w(1, s_n) - w(1, b_\varepsilon(\vartheta_n))) + \varepsilon_n + w(\vartheta_n, b_\varepsilon(\vartheta_n)) - \Phi(\vartheta_n) .$$

Letting $\varepsilon \to 0$ yields

$$(10) \qquad w(\vartheta_n, s_n) - \Phi(\vartheta_n) \leqq \alpha_n z_n(w(0, s_n) - w_0(\vartheta_n)) - \\ - \alpha_n z_n(w(1, s_n) - w_1(\vartheta_n)) + \varepsilon_n .$$

Substituing from (9) and (10) into (8) we obtain

$$(11) \qquad Q_1 \leqq \mathsf{E} \frac{1}{n} \sum_{i=0}^{1} (-1)^i \left(\sum_{k=1}^{n-1} k\alpha_k z_k(w(i, s_k) - w(i, s_{k+1})) + \right. \\ \left. + n\alpha_n z_n(w(i, s_n) - w_i(\vartheta_n)) \right) + \frac{1}{n} \sum_{k=1}^{n} k\varepsilon_k \leqq \\ \leqq \frac{1}{n} \sum_{i=0}^{1} \left(\sum_{k=1}^{n} \left|\mathsf{E}(k\alpha_k z_k - (k-1)\, \alpha_{k-1} z_{k-1})\, w(i, s_k)\right| + \left|\mathsf{E} n\alpha_n z_n w_i(\vartheta_n)\right| \right) + \frac{1}{n} \sum_{k=1}^{n} k\varepsilon .$$

Since $w(i, s_k)$, $i = 0, 1$, $k = 1, 2, \ldots$, is a Borel function of the random variable z_k we may apply well-known theorems on conditional expectations. We obtain

$$\mathsf{E} z_{k-1}\, w(i, s_k) = \mathsf{E}\mathsf{E}(z_{k-1}\, w(i, s_k) \mid z_k) = \\ = \mathsf{E}\, w(i, s_k)\, \mathsf{E}(z_{k-1} \mid z_k) = (1 - \gamma_{k-1})\, \mathsf{E} z_k\, w(i, s_k)$$

where the last equality follows from the condition (C3).

Thus, the summands in (11) become

$$\left|\mathsf{E}(k\alpha_k z_k - (k-1)\, \alpha_{k-1} z_{k-1})\, w(i, s_k)\right| = \\ = \left|(k\alpha_k - (k-1)\, \alpha_{k-1}(1 - \gamma_{k-1}))\, \mathsf{E} z_k\, w(i, s_k)\right| \leqq \\ \leqq (k\alpha_k - (k-1)\, \alpha_{k-1})\, \mathsf{E}(z_k\, w(i, s_k)| + \\ + (k-1)\, \alpha_{k-1} |\gamma_{k-1}|\, \mathsf{E}|z_k\, w(i, s_k)| , \quad k = 1, 2, \ldots ,$$

since $k\alpha_k - (k-1)\, \alpha_{k-1}$ is positive according to (C1). However, (C2) and the assumption that w is bounded by K yield

$$\mathsf{E}|z_k\, w(i, s_k)| \leqq Kc_1 , \quad k = 1, 2, \ldots$$

and

$$\mathsf{E}|z_n\, w_i(\vartheta_n)| \leqq Kc_1 .$$

Substituting back into (11) we obtain at last

$$Q_1 \leqq 2Kc_1 \left(2\alpha_n + \frac{1}{n} \sum_{k=1}^{n-1} k\alpha_k |\gamma|_k \right) + \frac{1}{n} \sum_{k=1}^{n} k\varepsilon_k .$$

It remains to establish a bound for Q_2. Writing

$$\begin{aligned}&\left|E(w(\vartheta_k, s_{k-1}) - w(\vartheta_k, s_k))\right| \leqq \\ &\leqq \left|E(w(\vartheta_k, b_{\varepsilon_{k-1}}(\bar{\vartheta}_{k-1} + \alpha_{k-1} z_{k-1})) - w(\vartheta_k, b_{\varepsilon_k}(\bar{\vartheta}_{k-1} + \alpha_{k-1} z_{k-1})))\right| + \\ &+ \left|E(w(\vartheta_k, b_{\varepsilon_k}(\bar{\vartheta}_{k-1} + \alpha_{k-1} z_{k-1})) - w(\vartheta_k, b_{\varepsilon_k}(\bar{\vartheta}_k + \alpha_k z_k)))\right|\end{aligned}$$

and applying (P3) to the first term and the lemma and (P2) to the second one, we obtain the bounds

$$\varepsilon_k + \varepsilon_{k-1}$$

for the first term, and

$$(c_0 + \varepsilon_k) \sup_{x \in R} \left| F_k\left(\frac{x - \bar{\vartheta}_k}{\alpha_k}\right) - F_{k-1}\left(\frac{x - \bar{\vartheta}_{k-1}}{\alpha_{k-1}}\right)\right|$$

for the second one. Hence we have

$$Q_2 \leqq \frac{1}{n} \sum_{k=1}^{n} (\varepsilon_k + \varepsilon_{k-1}) + \frac{1}{n} \sum_{k=1}^{n} (c_0 + \varepsilon_k) \sup_{x \in R} \left| F_k\left(\frac{x - \bar{\vartheta}_k}{\alpha_k}\right) - F_{k-1}\left(\frac{x - \bar{\vartheta}_{k-1}}{\alpha_{k-1}}\right)\right|$$

and the assertion is proved.

Theorem 2. *Let $\{\delta_k\}_{k \in N}$ be a sequence of real numbers, let $\{\alpha_k\}_{k=0}^{\infty}$ be a sequence of positive real numbers satisfying the condition* (C1) *of Theorem* 1.

Let $\{z_k\}_{k=0}^{\infty}$ be a sequence of random variables satisfying the condition (C2) *of Theorem* 1 *and the condition*

(C4) $$E(z_{k+1} \mid z_k) = (1 - \delta_{k+1})\, z_k \quad a.s. \quad k \in N\,,$$

let b_{ε_k}, $k = 0, 1, \ldots$ be regular ε_k-optimal strategies.

Then for every $\vartheta \in \Theta$ and $n \in N$,

$$\begin{aligned}&E \frac{1}{n} \sum_{k=1}^{n} w(\vartheta_k, b_{\varepsilon_{k-1}}(\bar{\vartheta}_{k-1} + \alpha_{k-1} z_{k-1})) - \Phi(\bar{\vartheta}_n) \geqq \\ &\geqq - 2Kc_1 \left(\frac{2\alpha_1}{n} + 2\alpha_n + \frac{1}{n} \sum_{k=1}^{n-1} (k+1)\, \alpha_{k+1} |\delta_{k+1}|\right) - \frac{1}{n} \sum_{k=1}^{n} k\varepsilon_k\,.\end{aligned}$$

Proof: Let us denote again $s_k = b_{\varepsilon_k}(\bar{\vartheta}_k + \alpha_k z_k)$. Then

(12) $$\begin{aligned}&E \frac{1}{n} \sum_{k=1}^{n} w(\vartheta_k, s_{k-1}) - \Phi(\bar{\vartheta}_n) = \\ &= E \frac{1}{n} \sum_{k=1}^{n} k(w(\bar{\vartheta}_k, s_{k-1}) - w(\bar{\vartheta}_k, s_k)) + E(w(\bar{\vartheta}_n, s_n) - \Phi(\bar{\vartheta}_n))\,.\end{aligned}$$

Since b_{ε_k} is ε_k-optimal

$$w(\bar{\vartheta}_k + \alpha_k z_k, s_k) \leqq w(\bar{\vartheta}_k + \alpha_k z_k, s_{k-1}) + \varepsilon_k\,, \quad k \in N$$

whence, denoting for short $q_k = w(1, s_k) - w(0, s_k)$,

$$w(\vartheta_k, s_{k-1}) - w(\vartheta_k, s_k) \geqq \alpha_k z_k(q_k - q_{k-1}) - \varepsilon_k .$$

It follows

$$(13)\quad \mathsf{E}\frac{1}{n}\sum_{k=1}^{n} k(w(\vartheta_k, s_{k-1}) - w(\vartheta_k, s_k)) \geqq \mathsf{E}\frac{1}{n}\sum_{k=1}^{n} k\alpha_k z_k(q_k - q_{k-1}) - \frac{1}{n}\sum_{k=1}^{n} k\varepsilon_k =$$

$$= \frac{1}{n}\sum_{k=1}^{n-1} \mathsf{E}(k\alpha_k z_k - (k+1)\,\alpha_{k+1}z_{k+1})\, q_k + \alpha_n \mathsf{E} z_n q_n - \frac{\alpha_1}{n}\,\mathsf{E} z_1 q_0 - \frac{1}{n}\sum_{k=1}^{n} k\varepsilon_k .$$

Conditions (C1) and (C4) yield

$$\mathsf{E}(k\alpha_k z_k - (k+1)\,\alpha_{k+1}z_{k+1})\, q_k \geqq (k\alpha_k - (k+1)\,\alpha_{k+1})\, 2Kc_1 - (k+1)\,\alpha_{k+1}|\delta_{k+1}|\, 2Kc_1 ,$$

$$\alpha_n \mathsf{E} z_n q_n \geqq -2\alpha_n K c_1 , \qquad -\frac{\alpha_1}{n}\,\mathsf{E} z_1 q_0 \geqq -\frac{2\alpha_1}{n}\, K c_1 .$$

Substituting into (13) we obtain

$$\mathsf{E}\frac{1}{n}\sum_{k=1}^{n} k(w(\vartheta_k, s_{k-1}) - w(\vartheta_k, s_k)) \geqq$$

$$\geqq -2Kc_1\left(\frac{2\alpha_1}{n} + 2\alpha_n + \frac{1}{n}\sum_{k=1}^{n-1}(k+1)\,\alpha_{k+1}|\delta_{k+1}|\right) - \frac{1}{n}\sum_{k=1}^{n} k\varepsilon_k .$$

This, together with the nonnegativeness of the second term in (12), terminates the proof.

IV. APPLICATION TO THE REPETITIVE PLAY

Theorems 1 and 2 can be applied directly to the repetitive play problem. By a suitable choice of the sequences $\{\alpha_k\}$, $\{\varepsilon_k\}$ and the random variables z_k we can obtain both upper and lower bounds for the difference of the expected average loss from Bayes loss of the type $C/\sqrt{n}$. For example we may choose $\alpha_k = k^{-1/2}$ and $\varepsilon_k = k^{-2}$ for $k = 1, 2, \ldots$ and $\alpha_0 = \varepsilon_0 = 1$. The radnom variables z_k represent what is called "artificial randomization". The simplest choice is $z_k = z_0$, $k \in N$ with z_0 taking values from a finite interval and the distribution function F_0 satisfying the Lipschitz condition

$$|F_0(x_1) - F_0(x_2)| \leqq C_3|x_1 - x_2| , \quad x_1, x_2 \in R , \quad C_3 < +\infty .$$

This is the case treated by Hannan and the main results of [1] immediately follow. Of course, the choices of the player II are based on the relative frequencies ϑ_k.

In case of a sequential statistical decision problem, however, we usually have only random estimates of these relative frequencies at disposal. The estimates are defined as arithmetic means of a sequence of independent random variables y_k. As was mentioned in [3], these random variables may, under quite general conditions, be defined as $y_k = \vartheta_k + h_k$, where h_k satisfy the conditions of the following theorem.

Theorem 3. *Let* $\{h_k\}_{k\in N}$ *be a sequence of independent, identically distributed random variables such that*

$$\mathsf{E}h_k = 0\,,$$
$$\mathsf{E}h_k^2 = \sigma^2 > 0\,,$$
$$\mathsf{E}|h_k|^3 < \infty\,, \quad k \in N\,.$$

Let $\{\varepsilon_k\}_{k=0}^{\infty}$ *be a sequence of nonnegative numbers such that*

$$\frac{1}{n}\sum_{k=1}^{n}((k+1)\,\varepsilon_k + \varepsilon_{k-1}) < \frac{c_2}{\sqrt{n}}$$

for a constant $c_2 < \infty$ *and every* $n \in N$, *let* b_{ε_k}, $k = 0, 1, \ldots,$ *be regular* ε_k*-optimal strategies.*

Then there is a constant $c < \infty$ *such that for every* $\vartheta \in \Theta$ *and* $n \in N$,

$$\left|\mathsf{E}\frac{1}{n}\sum_{k=1}^{n} w\left(\vartheta_k, b_{\varepsilon_{k-1}}\left(\bar{\vartheta}_{k-1} + \frac{1}{k-1}\sum_{j=1}^{k-1} h_j\right)\right) - \Phi(\bar{\vartheta}_n)\right| < \frac{c}{\sqrt{n}}\,^{1})\,.$$

Proof: First, we shall prove that the difference is bounded from above.

As h_k, $k \in N$ are inedpendent, identically distributed random variables, we have

$$\mathsf{E}(h_k \mid \sum_{j=1}^{n} h_j) = \mathsf{E}(h_1 \mid \sum_{j=1}^{n} h_j) \quad \text{a.s. for every}$$

$k = 1, 2, \ldots, n$, $n \in N$. Obviously, for $n \in N$,

$$\sum_{k=1}^{n} h_k = \mathsf{E}(\sum_{k=1}^{n} h_k \mid \sum_{j=1}^{n} h_j) = \sum_{k=1}^{n} \mathsf{E}(h_k \mid \sum_{j=1}^{n} h_j) \quad \text{a.s.}$$

These two relations yield

$$\mathsf{E}(h_k \mid \sum_{j=1}^{n} h_j) = \frac{1}{n}\sum_{j=1}^{n} h_j \quad \text{a.s.}\,,$$

$k = 1, 2, \ldots, n$, $n \in N$, which implies

$$\mathsf{E}\,(\sum_{k=1}^{n-1} h_k \mid \sum_{j=1}^{n} h_j) = \frac{n-1}{n}\sum_{j=1}^{n} h_j \quad \text{a.s.} \tag{14}$$

---

1) We put $(k-1)^{-1}\sum_{j=1}^{k-1} h_j = 0$ for $k = 1$.

Let us define a sequence of random variables $\{z_k\}_{k=0}^{\infty}$ as follows:

$$z_0 = 0\,, \quad z_k = \frac{1}{\sqrt{k}} \sum_{j=1}^{k} h_j\,, \quad k \in N\,.$$

Since

$$\mathsf{E}(z_k \mid z_{k+1}) = z_{k+1} \sqrt{\frac{k}{k+1}} \quad \text{a.s.}\,, \quad k \in N\,,$$

according to (14), the random variables z_k satisfy the conditions (C1)–(C3) of Theorem 1 with

$$\alpha_k = \frac{1}{\sqrt{k}}\,, \quad \gamma_k = 1 - \sqrt{\frac{k}{k+1}}\,, \quad k \in N\,, \quad \alpha_0 = \gamma_0 = 1\,.$$

Hence, the theorem yields

$$(15) \qquad \mathsf{E} \frac{1}{n} \sum_{k=1}^{n} w\left(\vartheta_k, b_{\varepsilon_{k-1}}\left(\bar{\vartheta}_{k-1} + \frac{1}{k-1} \sum_{j=1}^{k-1} h_j\right)\right) - \Phi(\bar{\vartheta}_n) \leqq$$

$$\leqq 2Kc_1 \left(\frac{2}{\sqrt{n}} + \frac{1}{n} \sum_{k=1}^{n-1} \left(1 - \sqrt{\frac{k}{k+1}}\right) \sqrt{k}\right) + \frac{1}{n} \sum_{k=1}^{n} ((k+1)\,\varepsilon_k + \varepsilon_{k-1}) +$$

$$+ \frac{1}{n} \sum_{k=1}^{n} (c_0 + \varepsilon_k) \sup_{x \in R} \left|F_k((x - \bar{\vartheta}_k)\sqrt{k}) - F_{k-1}((x - \bar{\vartheta}_{k-1})\sqrt{(k-1)})\right|$$

for every $\vartheta \in \Theta$ and $n \in N$.

Further, since for every $n \in N$,

$$(16) \qquad \frac{1}{n} \sum_{k=1}^{n-1} \left(1 - \sqrt{\frac{k}{k+1}}\right) \sqrt{k} < \frac{1}{n} \sum_{k=1}^{n-1} (\sqrt{(k+1)} - \sqrt{k}) \leqq \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}$$

it follows that the first term on the right-hand side of (15) is bounded by $6Kc_1/\sqrt{n}$. The second term is bounded by $c_2/\sqrt{n}$ by assumption and so it remains to prove that there is a constant $c_4 < \infty$ such that

$$\frac{1}{n} \sum_{k=1}^{n} \sup_{x \in R} \left|F_k((x - \bar{\vartheta}_k)\sqrt{k}) - F_{k-1}((x - \bar{\vartheta}_{k-1})\sqrt{(k-1)})\right| \leqq \frac{c_4}{\sqrt{n}}$$

for every $\vartheta \in \Theta$, $n \in N$.

Denoting $G(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$ the normal distribution function, writing

$$(17) \qquad \sup_{x \in R} \left|F_k((x - \bar{\vartheta}_k)\sqrt{k}) - F_{k-1}((x - \bar{\vartheta}_{k-1})\sqrt{(k-1)})\right| \leqq$$

$$\leqq \sup_{x \in R} \left|F_k(x) - G\left(\frac{x}{\sigma}\right)\right| + \sup_{x \in R} \left|F_{k-1}(x) - G\left(\frac{x}{\sigma}\right)\right| +$$

$$+ \sup_{x \in R} \left|G\left(\frac{\sqrt{k}}{\sigma}(x - \bar{\vartheta}_k)\right) - G\left(\frac{\sqrt{(k-1)}}{\sigma}(x - \bar{\vartheta}_{k-1})\right)\right|,$$

and applying the Berry-Esseen normal approximation theorem ([2], p. 288) to the sequence $\{h_k\}_{k\in N}$ we obtain

$$\sup_{x\in R}\left|F_k(x) - G\left(\frac{x}{\sigma}\right)\right| \leqq \beta\,\frac{\mathsf{E}|h_1|^3}{\sigma^3\sqrt{k}}, \quad k \in N$$

where β is the Berry-Esseen constant. Since

$$\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\sqrt{k}} \leqq \frac{2}{\sqrt{n}} \tag{18}$$

we obtained the desired bounds for the first two terms of (17). For the last one, let $x \in R$. Since, for every $x_1, x_2 \in R$,

$$\left|\int_{x_1}^{x_2} \mathrm{e}^{-t^2/2}\,\mathrm{d}t\right| \leqq |x_1 - x_2|$$

we have, for every $k \in N$,

$$\left|G\left(\frac{\sqrt{k}}{\sigma}(x - \bar{\vartheta}_k)\right) - G\left(\frac{\sqrt{(k-1)}}{\sigma}(x - \bar{\vartheta}_{k-1})\right)\right| \leqq \tag{19}$$

$$\leqq \left|\frac{\sqrt{k}}{\sigma}(x - \bar{\vartheta}_k) - \frac{\sqrt{(k-1)}}{\sigma}(x - \bar{\vartheta}_{k-1})\right| \leqq \frac{|x|}{\sigma}(\sqrt{k} - \sqrt{(k-1)}) +$$

$$+ \frac{1}{\sigma}\left|\frac{1}{\sqrt{k}}\sum_{j=1}^{k-1}\vartheta_j - \frac{1}{\sqrt{(k-1)}}\sum_{j=1}^{k-1}\vartheta_j\right| + \frac{\vartheta_k}{\sigma\sqrt{k}} \leqq$$

$$\leqq \frac{|x|}{\sigma}(\sqrt{k} - \sqrt{(k-1)}) + \frac{1}{\sigma}[\sqrt{k} - \sqrt{(k-1)}] + \frac{\vartheta_k}{\sigma\sqrt{k}}.$$

Let x_k, $k = 2, 3, \ldots$, be a number for which

$$\sup_{x\in R}\left|G\left(\frac{\sqrt{k}}{\sigma}(x - \bar{\vartheta}_k)\right) - G\left(\frac{\sqrt{(k-1)}}{\sigma}(x - \bar{\vartheta}_{k-1})\right)\right| = \tag{20}$$

$$= \left|G\left(\frac{\sqrt{k}}{\sigma}(x_k - \bar{\vartheta}_k)\right) - G\left(\frac{\sqrt{(k-1)}}{\sigma}(x_k - \bar{\vartheta}_{k-1})\right)\right|.$$

Such a number must satisfy the equation

$$x^2 - 2x\vartheta_k + k\bar{\vartheta}_k^2 - (k-1)\,\bar{\vartheta}_{k-1}^2 - \sigma^2\log\frac{k}{k-1} = 0$$

and hence

$$|x_k| < 2 + \sigma.$$

Applying this inequality to (19) and (20) we obtain

$$\sup_{\in x R} \left| G\left(\frac{\sqrt{k}}{\sigma}(x - \vartheta_k)\right) - G\left(\frac{\sqrt{(k-1)}}{\sigma}(x - \vartheta_{k-1})\right)\right| <$$
$$< \frac{3 + \sigma}{\sigma}[\sqrt{k} - \sqrt{(k-1)}] + \frac{\vartheta_k}{\sigma\sqrt{k}},$$

which, together with (18), yields the bound for the last term of (17). Hence, the upper bound is established.

The proof for the lower bound of the difference is easy since

$$E(z_{k+1} | z_k) = \frac{1}{\sqrt{(k+1)}} E\left(\sum_{j=1}^{k+1} h_j \Big| \sum_{j=1}^{k} h_j\right) = \frac{1}{\sqrt{(k+1)}} \sum_{j=1}^{k} h_j =$$
$$= z_k \sqrt{\frac{k}{k+1}} \quad \text{a.s.}, \quad k \in N.$$

Thus, the random variables z_k satisfy also the condition (C4) of Theorem 2 with $\delta_{k+1} = 1 - \sqrt{[k/(k+1)]} = \gamma_k$ and the lower bound follows immediately from Theorem 2 and (16).

The proof is complete.

References

[1] *J. Hannan:* Approximation to Bayes Risk in Repeated Play. Contributions to the Theory of Games 3 (1957), 97—139.

[2] *M. Loève:* Probability Theory, 3^d ed., Van Nostrand, New York 1963.

[3] *J. Van Ryzin:* The Sequential Compound Decision Problem with $n \times m$ Finite Loss Matrix. Ann. Math. Statist. 37 (1966), 4, 954—975.

[4] *E. Samuel:* Asymptotic Solutions of the Sequential Compound Decision Problem. Ann. Math. Statist. 34 (1963), 1079—1094.

Souhrn

OPAKOVANÉ HRANÍ HRY PROTI PŘÍRODĚ

STANISLAV JÍLOVEC, BRUNO ŠUBERT

Článek se zabývá opakováním hry dvou hráčů. V k-tém kroku (tj. v k-tém opakování hry) druhý hráč zná relativní četnosti strategií, kterých použil I. hráč v předcházejících $k - 1$ krocích. Cílem II. hráče je minimalizovat svoji střední ztrátu pro daný počet opakování, který je mu však neznámý. I. hráč může své strategie volit zcela libovolně; nepředpokládá se ani, že chování I. hráče lze popsat pravděpodobnostním způsobem. Za těchto předpokladů může II. hráč postupovat takovým způsobem, že jeho střední ztráta v n krocích nepřevyšuje $\Phi(\vartheta_n)$ o více než $c : \sqrt{n}$. (Zde $\Phi(.)$ označuje

Bayesovo riziko, $\bar{\vartheta}_n$ vektor relativních četností, kterých použil I. hráč v prvních n hrách, a c konstantu, závislou pouze na výplatní funkci.) To znamená, že II. hráč může volit své strategie takovým způsobem, že jeho střední ztráta je nejvýše o $c : \sqrt{n}$ horší než minimální ztráta, které by mohl dosáhnout, kdyby předem věděl, jaké budou relativní četnosti strategií I. hráče v prvních n hrách.

Tohoto výsledku však II. hráč nemusí dosáhnout, volí-li v každém kroku strategii, která je optimální vzhledem k relativním četnostem I. hráče v předcházejících hrách. Relativní četnosti je třeba vhodným způsobem znáhodnit a volit optimální strategie vzhledem k těmo znáhodnělým relativním četnostem. Věta 1 udává obecný tvar takového znáhodnění a v odstavci 4 je ukázáno, že speciálními případy jsou způsoby znáhodnění použité v [1] a [3].

Резюме

ПОСЛЕДОВАТЕЛЬНЫЕ ПОВТОРЕНИЯ ИГРЫ ПРОТИВ ПРИРОДЫ

СТАНИСЛАВ ЙИЛОВЕЦ, БРУНО ШУБЕРТ (Stanislav Jílovec, Bruno Šubert)

Статья занимается повторением игры двух игроков. В k-м шагу (т.е. в k-м повторении игры) II-й игрок знает частоты стратегий, принятых I-м игроком в предшествующих k-1 играх. Цель II-го игрока – минимализировать свои средние потери для фиксированного числа повторений, которое ему неизвестно. I-й игрок может выбирать свои стратегии совсем произвольно; не предполагается даже, что поведение I-го игрока можно описать вероятностным образом. При этих предположениях II-й игрок может поступать таким образом, что его средняя потеря при n повторениях игры не превосходит $\Phi(\bar{\vartheta}_n)$ на больше чем $c : \sqrt{n}$. (Здесь $\Phi(.)$ – функция Байесовского риска, $\bar{\vartheta}_n$ – вектор частот стратегий, использованных I-м игроком в первых n повторениях игры, c – постоянная, зависящая только от вида платежной функции.) Значит, II-й игрок может выбирать свои стратегии таким образом, что его средняя потеря при n повторениях игры максимально на $c : \sqrt{n}$ выше минимальной потери, достижимой в предположении, что ему заранее известно, какие будут частоты стратегий I-го игрока в первых n шагах. Но этого результата II-й игрок не должен обязательно достигнуть, если в каждом шагу принимает стратегию, которая оптимальна относительно частот стратегий I-го игрока в предыдущих шагах. Частоты стратегий надо подходящим образом рандомизировать и II-му игроку придется выбирать стратегию оптимальную относительно этих рандомизированных частот. Теорема 1 показывает общий вид такой рандомизации и в параграфе 4 показано, что частными случаи являются способы рандомизации применяемые в [1] и [3].

Authors' addresses: RNDr. *Stanislav Jílovec* C.Sc., Ing. *Bruno Šubert* C.Sc., Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.