KYBERNETIKA ČÍSLO 4, ROČNÍK 6/1970

On the Uniform Almost Sure Convergence of Losses for the Repetiton of a Statistical Decision Problem

STANISLAV JÍLOVEC

The uniform almost sure convergence and the convergence rate of losses of two types of decision processes for independent non-Bayesian repetition of a weight-finite statistical decision problem are estıblished.

INTRODUCTION

We shall consider a sequence of statistical decision problems having identical structure known to the statistician. Only the observations, parameter values and statistician's decisions may be changed at each step. Our approach will be non-Bayesian i.e. no probability assumptions are made about the parameter values, they may alter quite arbitrarily. As for the statistician it is assumed that, in every step, he has at his disposal, besides the observable value, an estimate of the one-dimensional distribution (i.e. the relative frequency vector) of the parameter values in the preceding steps.

This problem was first treated by Hannan [3] and developed e.g. in [10], [14], [4] and [5].

In all these papers, it is shown that the difference between average risk of a certain decision process over the first n steps and $\varrho(\bar{a}_n)$, the risk of a decision function which is optimal against the relative frequency vector $\bar{a}_n$ of the parameter values in the first n steps, approaches zero (or has an upper bound approaching zero) if n tends to infinity; the convergence rate is established as well.

As far as it is known to the author, few results are concerned with the convergence of losses, especially with the almost sure convergence. Under the supplementary assumption that the Bayes envelope functional is differentiable, a slowly non-uniform almost sure convergence is proved in [11] and a uniform almost sure convergence for a more complicated decision process based not on Hannan's idea but on Blackwell's one is proved in [12]. Let us note that Van Ryzin [15] and Šubert [13] indicating difficulties at proving a uniform almost sure convergence deal only with a uniform convergence in probability.

In the present paper we shall construct a decision process B_n, $n = 1, 2, \ldots$, where B_n is, roughly speaking, a Bayes decision function against an estimate of $\bar{a}_{n-1}$ or $\bar{a}_n$ so that not only

$$\mathsf{E}\,\frac{1}{n}\sum_{i=1}^{n} w(a_i, B_i) - \varrho(\bar{a}_n) \to 0$$

but also

$$\frac{1}{n}\sum_{i=1}^{n} w(a_i, B_i) - \varrho(\bar{a}_n) \to 0 \quad \text{a.s.}$$

and the convergence is uniform in all sequences $(a_1, a_2, \ldots)$ of parameter values. We shall simultaneously determine the convergence rate in the sense that, for every monotone sequence $\{b_n\}_{n=1}^{\infty}$ of positive numbers satisfying $\sum_{n=1}^{\infty} b_n^2 < \infty$,

$$b_n \sum_{n=1}^{\infty} (w(a_i, B_i) - \varrho(\bar{a}_n)) \to 0 \quad \text{a.s.}$$

uniformly in all sequences of parameter values.

PREREQUISITES

If E and A are non-empty sets, then the symbol E^A denotes the Cartesian product

$$E^A = \mathop{\times}_{a\in A} E_a \quad \text{where} \quad E_a = E\,, \quad a \in A\,.$$

If $x \in E^A$ and $a \in A$, then $(x)_a$ denotes the a-th coordinate of the point x. Therefore $x = \{(x)_a\}_{a\in A}$. Further, if $\mathscr{F}$ is a σ-field of subsets of a set and A is finite non-empty set, then $\mathscr{F}^A$ denotes the minimal σ-field over the class

$$\{\mathop{\times}_{a\in A} E_a : E_a \in \mathscr{F} \quad \text{for all} \quad a \in A\}\,.$$

The symbol $\overline{\overline{E}}$ denotes the cardinality of the set E, and if E is a subset of a fixed set, then χ_E denotes the characteristic function (indicator) of the set E. The letter N will denote the set of all positive integers, the letter R the set of all real numbers and $\mathscr{R}$ the σ-field of all Borel subsets of R.

If A is a finite non-empty set, then R^A is isometrically isomorfic to the $\overline{\overline{A}}$-dimensional Euclidean space if we define addition, real scalar multiplication, inner product and norm by

$$x + y = \{(x)_a + (y)_a\}_{a\in A}\,,$$
$$cx = \{c(x)_a\}_{a\in A}\,,$$
$$xy = \sum_{a\in A} (x)_a\,(y)_a\,,$$

$$\|x\| = (xx)^{1/2},$$

respectively. If x_i, $i = 1, 2, \dots n$, is a sequence of points in R^A, the symbol $\bar{x}_n$ will denote the arithmetic mean

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i , \quad n \in N ,$$

and

$$\bar{x}_0 = 0 \in R^A .$$

The symbol $(\Omega, \mathscr{S}, P)$ will denote a basic probability space. If $X = \{X_t : t \in T\}$ is a family of random variables or random vectors, then $\mathscr{B}(X)$ will denote the σ-field induced by X, i.e. the minimal σ-field with respect to which the X_t's are all measurable. The operation of expectation is denoted by E. If X is a random vector on $(\Omega, \mathscr{S}, P)$ to R^A and if $\mathsf{E}(X)_a$, $a \in A$, are finite, then $\mathsf{E}X$ is defined by $\mathsf{E}X = \{\mathsf{E}(X)_a\}_{a \in A}$. The conditional expectation of a random variable Y given a σ-field $\mathscr{F}$ will be denoted $\mathsf{E}(Y \mid \mathscr{F})$ and the value of $\mathsf{E}(Y \mid \mathscr{F})$ at $\omega \in \Omega$ by $\mathsf{E}(\omega, Y \mid \mathscr{F})$. If $\mathscr{F} = \mathscr{B}(X)$ we shall write $\mathsf{E}(Y \mid X)$ instead of $\mathsf{E}(Y \mid \mathscr{B}(X))$. Similarly, for conditional probability, we shall use the symbols $P(E \mid \mathscr{F})$, $P(\omega, E \mid \mathscr{F})$, $P(E \mid X)$. Other symbols and terminology will be used in accordance with [8] or [2].

Now we shall consider the statistical decision problem $\mathscr{D} = (A, D, w, (X, \mathscr{X}), \{v_a\}_{a \in A})$ where A and D are two finite non-empty sets, $(X, \mathscr{X})$ is a measurable space, $\{v_a\}_{a \in A}$ is a family of probability measures on $\mathscr{X}$ and w is a real function on $A \times D$. The set A will be referred to as the parameter space, D as the decision space, $(X, \mathscr{X})$ as the sample space and w as the weight or loss function. The statistician observes a random variable whose distribution is v_a if the parameter is a. Of course, the parameter is unknown to the statistician. On the basis of the observed value the statistician has to make a decision incurring loss $w(a, d)$ if he decides d and Nature chooses a. Since A and D are finite, the function w is bounded. We shall denote

$$v = \sum_{a \in A} v_a ,$$

$$f_a = \frac{\mathrm{d}v_a}{\mathrm{d}v} , \quad a \in A .$$

The function ϱ defined on R^A by

$$\varrho(\xi) = \inf_{\delta \in \Delta} \sum_{a \in A} (\xi)_a \int w(a, \delta(x)) f_a(x) \,\mathrm{d}v\,(x)$$

where Δ denotes the set of all mappings δ on X to D such that, for every $d \in D$, $\delta^{-1}(\{d\}) \in \mathscr{X}$ will be called the Bayes envelope functional.

A mapping β on $X \times R^A$ to D is called an optimal procedure if there is an ordered set $(D_0, \prec)$ such that

(i) $D_0 \subset D$,

(ii) if $d, d' \in D_0$, $d \neq d'$, then $\max_{a \in A} (w(a, d) - w(a, d')) > 0$ and $\min_{a \in A} (w(a, d) - w(a, d')) < 0$,

(iii) if $d \in D - D_0$, then there is a $d' \in D_0$ such that $\min (w(a, d) - w(a, d') \geqq 0$,

(iv) $\beta(x, \xi) \in D_0$,

(v) $\beta(x, \xi) = d$ if $\sum_{a \in A} (\xi)_a (w(a, d') - w(a, d)) f_a(x) \geqq 0$ for all $d' \succ d$ and $\sum (\xi)_a (w(a, d') - w(a, d)) f_a(x) > 0$ for all $d' \prec d$, $d' \neq d$.

Let us note that the existence of an optimal procedure is guaranteed by the finiteness of the set D. Evidently,

$$\{(x, \xi) : \beta(x, \xi) = d\} \in \mathscr{X} \times \mathscr{R}^A , \quad d \in D$$

$$\beta(x, c\xi) = \beta(x, \xi) , \quad c > 0 .$$

If β is an optimal procedure, then r_β will denote the function on $R^A \times R^A$ defined by

$$r_\beta(\xi, \eta) = \sum_{a \in A} (\xi)_a \int w(a, \beta(x, \eta)) f_a(x) \, \mathrm{d}\nu(x) .$$

It is easy to prove the following properties of r_β:

$$r_\beta\Big(\sum_{i=1}^{n} c_i \xi_i, \eta\Big) = \sum_{i=1}^{n} c_i \, r_\beta(\xi_i, \eta) ,$$

$$r_\beta(\xi, \eta) \geqq r_\beta(\xi, \xi) = \varrho(\xi) ,$$

$$r_\beta(\xi, c\eta) = r_\beta(\xi, \eta) , \quad c > 0 ,$$

$$r_\beta(\xi, \eta) - r_\beta(\xi, \zeta) \leqq r_\beta(\eta - \xi, \zeta) - r_\beta(\eta - \xi, \eta) , \tag{1}$$

$$r_\beta(\xi, \eta) - r_\beta(\xi, \zeta) \geqq r_\beta(\zeta - \xi, \zeta) - r_\beta(\zeta - \xi, \eta) . \tag{2}$$

Lemma 1. *If β is an optimal procedure, $a_i \in R^A$, $\eta_i \in R^A$, $i = 1, 2, \ldots, n$, and $\eta_0 = 0 \in R^A$, then*

$$\sum_{i=1}^{n} (r_\beta(a_i, \eta_i) - \varrho(\bar{a}_n)) \leqq$$

$$\leqq r_\beta\Big(\eta_n - \sum_{i=1}^{n} a_i, \bar{a}_n\Big) - \sum_{i=1}^{n} r_\beta(\eta_i - \eta_{i-1} - a_i, \eta_i) .$$

Proof. Using the above properties of r_β we obtain the following chain of equalities and inequalities

$$\sum_{i=1}^{n} (r_\beta(a_i, \eta_i) - \varrho(\bar{a}_n)) =$$

$$= \sum_{i=1}^{n} \left(r_\beta\left(\sum_{j=1}^{i} a_j, \eta_i\right) - r_\beta\left(\sum_{j=1}^{i-1} a_j, \eta_i\right)\right) - r_\beta\left(\sum_{j=1}^{n} a_j, \sum_{j=1}^{n} a_j\right) =$$

$$= r_\beta\left(\sum_{j=1}^{n} a_j, \eta_n\right) - r_\beta\left(\sum_{j=1}^{n} a_j, \sum_{j=1}^{n} a_j\right) +$$

$$+ \sum_{i=1}^{n-1} \left(r_\beta\left(\sum_{j=1}^{i} a_j, \eta_i\right) - r_\beta\left(\sum_{j=1}^{i} a_j, \eta_{i+1}\right)\right) \leqq$$

$$\leqq r_\beta\left(\eta_n - \sum_{j=1}^{n} a_j, \sum_{j=1}^{n} a_j\right) - r_\beta\left(\eta_n - \sum_{j=1}^{n} a_j, \eta_n\right) +$$

$$+ \sum_{i=1}^{n-1} \left(r_\beta\left(\eta_i - \sum_{j=1}^{i} a_j, \eta_{i+1}\right) - r_\beta\left(\eta_i - \sum_{j=1}^{i} a_j, \eta_i\right)\right) =$$

$$= r_\beta\left(\eta_n - \sum_{j=1}^{n} a_j, \sum_{j=1}^{n} a_j\right) - \sum_{i=1}^{n} \left(r_\beta\left(\eta_i - \sum_{j=1}^{i} a_j, \eta_i\right) - r_\beta\left(\eta_{i-1} - \sum_{j=1}^{i-1} a_j, \eta_i\right)\right) =$$

$$= r_\beta\left(\eta_n - \sum_{j=1}^{n} a_j, \bar{a}_n\right) - \sum_{i=1}^{n} r_\beta\left(\eta_i - \eta_{i-1} - a_i, \eta_i\right).$$

This proves the lemma.

Using (2) instead of (1) we obtain

Lemma 2. *If β is an optimal procedure and $a_i \in R^A, \eta_i \in R^A$, $i = 1, 2, \ldots, n$, then*

$$\sum_{i=1}^{n} \left(r_\beta(a_i, \eta_i) - \varrho(\bar{a}_n)\right) \geqq -r_\beta(\eta_1, \eta_1) +$$

$$+ r_\beta\left(\eta_n - \sum_{j=1}^{n-1} a_j, \eta_n\right) + \sum_{i=1}^{n-1} r_\beta\left(\eta_i - \eta_{i+1} + a_i, \eta_i\right).$$

We shall consider that the decision problem $\mathscr{D} = (A, D, w, (X, \mathscr{X}), \{v_a\}_{a \in A})$ is repeated in such a way that A, D, w, $(X, \mathscr{X})$ and $\{v_a\}_{a \in A}$ remain the same, only the observations, parameter values and statistician's decisions may be changed at each step. Nature first selects a sequence $a = \{(a)_i\}_{i \in N}$ of parameters. To every $a \in A^N$ there corresponds a sequence $\{g_{a,i}\}_{i \in N}$ of observations (i.e. measurable mappings on $(\Omega, \mathscr{S})$ to $(X, \mathscr{X})$). We shall assume that the repetition of $\mathscr{D}$ is independent in the sense that

(G1) for every $a \in A^N$, $g_{a,i}$, $i \in N$ are stochastically independent.

Further we shall assume that

$$\text{(G2)} \qquad P g_{a,i}^{-1} = v_{(a)_i}, \quad a \in A^N, \ i \in N$$

which corresponds to our assumption that $\{v_a\}_{a \in A}$ is in all steps the same. As mentioned in introduction, in every step, the statistician knows, besides the observed value, an estimate of the relative frequency vector of parameter values in the preceding steps. Under these assumptions, approximation to the Bayes envelope functional is a good measure of the efficacity of decision processes. Various arguments why this is so are given e.g. in [3], [12] and [15].

SOME AUXILIARY RESULTS

First of all, we shall introduce the almost sure uniform convergence. It is well known that the random variables Z_n, $n \in N$, converge to zero almost surely if, and only if, for every $\varepsilon > 0$

$$\lim_{n\to\infty} P\Big(\bigcup_{k=n}^{\infty} \{|Z_k| \geqq \varepsilon\}\Big) = 0 .$$

Therefore, it is natural to define the uniform almost sure convergence in the following way due to Parzen [9].

Let T be a non-empty set and let $\{\{Z_{t,n}\}_{n\in N} : t \in T\}$ be a family of sequences of random variables on probability space $(\Omega, \mathcal{S}, P)$. We shall say that $\{Z_{t,n}\}_{n\in N}$ converges to zero almost surely uniformly in $t \in T$, and write $Z_{t,n} \to 0$ a. s. uniformly in $t \in T$, if, for every $\varepsilon > 0$,

$$\lim_{k\to\infty} P\Big(\bigcup_{k=n} \{|Z_{t,n}| \geqq \varepsilon\}\Big) = 0$$

uniformly in $t \in T$.

Using the method of the proof of Theorem 16. A in [9] and results of Section 29.1 in [8], we obtain the following form of the strong law of large numbers.

Theorem 1. *Let $\{b_n\}_{n\in N}$ be a monotone sequence of positive numbers converging to zero, T be a non-empty set, $\{Z_{t,n} : t \in T, n \in N\}$ be a family of random variables on $(\Omega, \mathcal{S}, P)$ and $\{\mathcal{F}_{t,n} : t \in T, n \in N\}$ be a family of sub-σ-fields of $\mathcal{S}$ such that*

$$\mathcal{F}_{t,1} = \{0, \Omega\} ,$$

$$\mathcal{B}(Z_{t,i} : i = 1, 2, \ldots, n) \subset \mathcal{F}_{t,n+1} \subset \mathcal{F}_{t,n+2} , \quad t \in T, n \in N .$$

If the series

$$\sum_{n=1}^{\infty} b_n^2\, \sigma^2(Z_{t,n}) , \quad t \in T ,$$

are convergent and uniformly bounded, then

$$b_n \sum_{i=1}^{n} (Z_{t,n} - \mathrm{E}(Z_{t,n} \mid \mathcal{F}_{t,n})) \to 0$$

a.s. uniformly in $t \in T$.

Theorem 2. *Let $\{b_n\}_{n\in N}$ be a monotone sequence of positive numbers such that*

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

Let T be a non-empty set and $\{Z_{t,n} : t \in T,\ n \in N\}$ be a family of random variables on $(\Omega, \mathcal{S}, P)$. If there is a constant $c < \infty$ such that, for every $t \in T$ and $n \in N$,

$$\mathsf{E}\Big(\sum_{i=1}^{n} |Z_{t,i}|\Big)^2 \leqq cn$$

then

$$b_n \sum_{i=1}^{n} Z_{t,i} \to 0$$

a.s. uniformly in $t \in T$.

Proof. The convergence of the series $\sum_{n=1}^{\infty} b_n^2$ and the monotony of the sequence $\{b_n\}_{n\in N}$ imply

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

Therefore, for every $\varepsilon > 0$, there is a $l_\varepsilon \in N$ such that

$$\frac{2c}{\varepsilon^2} \sum_{k=l_\varepsilon}^{\infty} 2^k b_{2^k}^2 < \varepsilon .$$

If $l \geqq 2^{l_\varepsilon}$, then

$$P\Big(\bigcup_{n\geqq l} \{|b_n \sum_{i=1}^{n} Z_{t,i}| \geqq \varepsilon\}\Big| \leqq$$
$$\leqq \sum_{k=l_\varepsilon}^{\infty} P\Big(\bigcup_{2^k \leqq n < 2^{k+1}} \{|b_n \sum_{i=1}^{n} Z_{t,i}| \geqq \varepsilon\}\Big) \leqq$$
$$\leqq \sum_{k=l_\varepsilon}^{\infty} P\{\max_{2^k \leqq n \leqq 2^{k+1}} |b_n \sum_{i=1}^{n} Z_{t,i}| \geqq \varepsilon\} \leqq$$
$$\leqq \sum_{k=l_\varepsilon}^{\infty} P\{b_{2^k} \sum_{i=1}^{2^{k+1}} |Z_{t,i}| \geqq \varepsilon\} \leqq$$
$$\leqq \sum_{k=l_\varepsilon} \frac{2^{k+1} c b_{2^k}^2}{\varepsilon^2} < \varepsilon$$

which proves the theorem. Let us note that the last but one inequality follows from the Tchebychev inequality.

Treating conditional expectations and conditional probabilities we shall often use the following lemma without further reference.

Lemma 3. *Let h_i, $i = 1, 2$, be measurable mappings on $(\Omega, \mathscr{F}_i)$ to $(X_i, \mathscr{X}_i)$ respectively, and φ be a Borel measurable function on $(X_1 \times X_2, \mathscr{X}_1 \times \mathscr{X}_2)$. If $\mathscr{F}_1$ and $\mathscr{F}_2$ are independent sub-σ-fields of $\mathscr{S}$, then*

$$\mathsf{E}(\omega, \varphi(h_1, h_2) \mid \mathscr{F}_2) = \int \varphi(h_1(\omega'), h_2(\omega)) \, \mathrm{d}P(\omega') = \mathsf{E}\varphi(h_1, h_2(\omega)) .$$

Proof. Using decomposition of a measurable function into its positive and negative parts, monotone convergence theorem and additivity of integrals, it is easy to reduce the general case to the case where φ is a characteristic function. But the class $\mathscr{C}$ of all sets C such that the asserted equality is true for $\varphi = \chi_c$ is a monotone system containing all sets of the form $\bigcup_{i=1}^{n} F_{1,i} \times F_{2,i}$ where $F_{1,i} \in \mathscr{F}_1$ and $F_{2,i} \in \mathscr{F}_2$, $i = 1, 2, \ldots, n$. Therefore, $\mathscr{C} = \mathscr{F}_1 \times \mathscr{F}_2$. This concludes the proof.

Lemma 4. *Let Z_k, $k = 1, 2, \ldots, r$, be integrable random variables on $(\Omega, \mathscr{S}, P)$ such that $Z_k - \mathsf{E}Z_k$, $k = 1, 2, \ldots, r$, are linearly independent*. Let c_k, $k = 0, 1, \ldots, r$, be real numbers. If $\max_{1 \leq k \leq r} |c_k| > 0$, then*

$$P\{\sum_{k=1}^{r} c_k Z_k = c_0\} < 1 .$$

Proof. Let us assume that

$$\sum_{k=1}^{r} c_k Z_k = c_0 \quad \text{a.s.}$$

Then

$$\sum_{k=1}^{r} c_k \mathsf{E}Z_k = c_0$$

and therefore

$$\sum_{k=1}^{r} c_k(Z_k - \mathsf{E}Z_k) = 0 \quad \text{a.s.}$$

which contradicts the linear independence of $Z_k - \mathsf{E}Z_k$, $k = 1, 2, \ldots, r$.

Lemma 5. *Let Z_k, $k = 1, 2, \ldots, r$, be integrable random variables on $(\Omega, \mathscr{S}, P)$ such that $Z_k - \mathsf{E}Z_k$, $k = 1, 2, \ldots, r$, are linearly independent. Then there exist real constants $l > 0$ and $\varepsilon > 0$ such that*

$$\sup_{C} \sup_{x \in R} P\{x \leqq \sum_{k=1}^{r} c_k Z_k \leqq x + l\} \leqq 1 - \varepsilon$$

* As far as we speak about the linear independence of random variables, we identify the equivalent random variables.

where

$$C = \{(c_1, c_2, \ldots, c_r) : c_i \in R, i = 1, \ldots, r, \sum_{i=1}^{r} c_i^2 = 1\}.$$

Proof. Let us assume that the lemma does not hold. Then, for every $n \in N$, there are real numbers $x_n, c_{n,1}, c_{n,2}, \ldots, c_{n,r}$ such that

$$\sum_{k=1}^{r} c_{n,k}^2 = 1,$$

$$P\left\{x_n \leqq \sum_{k=1}^{r} c_{n,k} Z_k \leqq x_n + \frac{1}{n}\right\} > 1 - 2^{-n}.$$

Since Z_n are finite functions and $|c_{n,k}| \leqq 1$ the sequence $\{x_n\}_{n \in N}$ must be bounded. Hence, there is a subsequence $\{n_i\}_{i \in N}$ such that

$$x_{n_i} \to x,$$

$$c_{n_i,k} \to c_k.$$

Let us denote

$$F = \bigcup_{j=1}^{\infty} \bigcap_{i=j}^{\infty} \left\{x_{n_i} \leqq \sum_{k=1}^{r} c_{n_i,k} Z_k \leqq x_{n_i} + \frac{1}{n_i}\right\}.$$

Evidently

$$P(F) = 1,$$

$$F \subset \{\sum_{k=1}^{r} c_k Z_k = 1\}.$$

Therefore

$$P\{\sum_{k=1}^{r} c_k Z_k = x\} = 1.$$

Since this contradicts the Lemma 1, the proof is complete.

Lemma 6. *Let $L \in R$ and let $\{Z_i\}_{i \in N}$ be a sequence of independent and equally distributed random vectors on $(\Omega, \mathscr{S}, P)$ to R^A. If $(Z_1 - \mathsf{E}Z_1)_a$, $a \in A$, are linearly independent, then there exist real constants $l > 0$ and $\varkappa > 0$ such that, for every $n \in N$ and $c \in R^A - \{0\}$,*

$$\sup_{x \in R} P\{x \leqq \sum_{i=1}^{n} cZ_i \leqq x + L\} \leqq \frac{\varkappa \max\left(\frac{|L|}{\|c\|}, l\right)}{\sqrt{n}}.$$

Proof. It follows from Lemma 5 and Kolmogorov-Rogozin inequality (see e.g. [1]).

Up to the end of this paper we shall assume that there is given a statistical decision problem $\mathscr{D} = (A, D, w, (X, \mathscr{X}), \{v_a\}_{a \in A})$ with finite A and D. Further, if $a \in A^N$ and

 $i \in N$, then the symbol a_i will denote a point of R^A such that

$$(a_i)_m = 1 \quad \text{for} \quad (a)_i = m\,, \qquad 0 \quad \text{for} \quad (a)_i \neq m\,.$$

Lemma 7. *Let* $\{\{h_{a,i}\}_{i\in N} : a \in A^N\}$ *and* $\{\{h'_{a,i}\}_{i\in N} : a \in A^N\}$ *be two families of sequences of independent integrable random vectors on* $(\Omega, \mathcal{S}, P)$ *to* $(R^A, \mathcal{R}^A)$ *such that*

(i) *for every* $a \in A^N$, *the families* $\{h_{a,i} : i \in N\}$ *and* $\{h'_{a,i} : i \in N\}$ *are independent and equally distributed,*

(ii) $Ph_{a,i}^{-1} = Ph_{a',i'}^{-1}$ *if* $a_i = a'_{i'}$.

Let β *be an optimal procedure and* $d, d' \in D$, $d \neq d'$. *Let us denote*

$$\begin{aligned}
F_{a,i}(x) &= \{\omega : \beta(x, \sum_{k=1}^{i} h_{a,k}(\omega)) = d,\ \beta(x, \sum_{k=1}^{i-1} h_{a,k}(\omega) + h'_{a,i}(\omega)) = d'\}\,,\\
I_{i,j}(a) &= \{k : a_k = a_i, j < k < i\}\,,\\
\mathcal{F}_{i,j}(a) &= \mathcal{B}(h'_{a,i}, h_{a,k}, k \in \{1, 2, \ldots, i\} - I_{i,j}(a))\,,\\
\pi_{i,j}(a) &= (\max\{1, \bar{\bar{I}}_{i,j}(a)\})^{-1/2}\,, \quad x \in X,\ a \in A^N\,,\\
&\qquad i \in N,\ j \in N \cup \{0\},\ i > j\,.
\end{aligned}$$

Then there exist real constants $\varkappa$ *and* l *such that, for every* $a \in A^N$, $x \in X$, $i \in N$, $j \in N \cup \{0\}$, $i > j$, *there holds*

$$P(\omega, F_{a,i}(x) \mid \mathcal{F}_{i,j}(a)) \leqq \varkappa \pi_{i,j}(a) \max(\|h_{a,i}(\omega) - h'_{a,i}(\omega)\|, l) \quad \text{a.s.}$$

Proof. Let us recall that, for $x, y \in R^A$, the symbol $xy = \sum_{m\in A} (x)_m (y)_m$ denotes the inner product in R^A. From the definition of optimal procedures it follows that, for $d \prec d'$, $d \neq d'$,

$$\{\omega : \beta(x, \sum_{k=1}^{i} h_{a,k}(\omega)) = d\} \subset \{\omega : \sum_{k=1}^{i} v\, h_{a,k}(\omega) \leqq 0\}\,,$$

$$\{\omega : \beta(x, \sum_{k=1}^{i-1} h_{a,k}(\omega) + h'_{a,i}(\omega) = d'\} \subset \{\omega : \sum_{k=1}^{i-1} v\, h_{a,k}(\omega) > -v\, h'_{a,i}(\omega)\}$$

where $v = \{(w(a, d) - w(a, d'))\, f_a(x)\}_{a\in A}$.

If in these formulae „$\leqq$" is replaced by „$<$" and conversely, the modified formulae are valid for $d' \prec d$, $d \neq d'$. Hence, denoting by $J(\alpha, \lambda)$ either of the intervals $\langle \alpha, \alpha + \lambda)$, $(\alpha, \alpha + \lambda\rangle$ we have

$$\text{(3)} \qquad F_{a,i}(x) \subset \{\omega : \sum_{k=1}^{i-1} v\, h_{a,k}(\omega) \in J(-v\, h'_{a,i}(\omega), -v(h_{a,i}(\omega) - h'_{a,i}(\omega)\} =$$

$$= \{\omega : \sum_{k \in I_{i,j}(a)} v(h_{a,k}(\omega) - \mathsf{E} h_{a,k}) \in J(y(\omega), v(h'_{a,i}(\omega) - h_{a,i}(\omega)))\}$$

where

$$y(\omega) = -v\, h'_{a,i}(\omega) - \sum_{k \in I_{ij}(a)} v \mathsf{E} h_{a,k} - \sum_{\substack{1 \leq k < i \\ k \notin I_{i,j}(a)}} v\, h_{a,k}(\omega)\,.$$

From (ii) it follows that for every $m \in A$ there are a subset $A_m \subset A$ and real constants $c_{j,k}^{(m)}$, $j \in A$, $k \in A_m$ such that, for every $i \in N$ and $a \in A^N$ satisfying $(a)_i = m$ the functions $(h_{a,i} - \mathsf{E}h_{a,i})_k$, $k \in A_m$, are linearly independent and

$$(4) \qquad (h_{a,i} - \mathsf{E}h_{a,i})_j = \sum_{k \in A_m} c_{j,k}^{(m)} (h_{a,i} - \mathsf{E}h_{a,i})_k \quad \text{a.s.}, j \in A\,.$$

If $A_{(a)_i} = \emptyset$, then $h_{a,i} = h'_{a,i} = \text{const} \in R^A$ a.s., therefore* $F_{a,i}(x) \sim \emptyset$ and the lemma is proved.

Assuming $A_{(a)_i} \neq \emptyset$ and denoting

$$\tilde{h}_{a,i} = \{(h_{a,i} - \mathsf{E}h_{a,i})_m\}_{m \in A_{(a)_i}}\,,$$

$$\tilde{h}'_{a,i} = \{(h'_{a,i} - \mathsf{E}h'_{a,i})_m\}_{m \in A_{(a)_i}}\,,$$

$$\tilde{v}_{a,i} = \{\sum_{j \in A} c_{j,k}^{(a)_i}(v)_j\}_{k \in A_{(a)_i}}$$

we obtain, according to (3) and (4),

$$F_{a,i}(x) \sim \{\omega : \sum_{k \in I_{i,j}(a)} \tilde{v}_{a,k}\, \tilde{h}_{a,k}(\omega) \in J(y(\omega), \tilde{v}_{a,i}(\tilde{h}'_{a,i}(\omega) - \tilde{h}_{a,i}(\omega))\}\,.$$

If $\tilde{v}_{a,i} = 0$ or $I_{i,j}(a) = 0$, the lemma obviously holds. In the opposite case, the desired result follows from the finiteness of D and Lemma 6.

Lemma 8. *Let $\{\{h_{a,i}\}_{i \in N} : a \in A^N\}$ be a family of sequences of independent integrable random vectors on $(\Omega, \mathscr{S}, P)$ to $(R^A, \mathscr{R}^A)$ such that*

(i) *for every $a \in A^N$ and $i \in N$, one of two following conditions holds:*

(i_1) *the functions $(h_{a,i} - \mathsf{E}h_{a,i})_m$, $m \in A$, are linearly independent,*

(i_2) *there is $m_{a,i} \in A$ such that $(h_{a,i} - \mathsf{E}h_{a,i})_m$, $m \in A - \{m_{a,i}\}$, are linearly independent and*

$$\sum_{m \in A} (h_{a,i} - \mathsf{E}h_{a,i})_m = 0 \quad \text{a.s.}$$

and

$$(\mathrm{ii}) \qquad Ph_{a,i}^{-1} = Ph_{a',i'}^{-1} \quad \textit{if} \quad a_i = a'_{i'}\,.$$

* For $F, G \in \mathscr{S}$, we write $F \sim G$ if $P((F - G) \cup (G - F)) = 0$.

 Let β be an optimal procedure and $d, d' \in D$, $d \neq d'$. Let us denote

$$G_{a,i}(x) = \{\omega : \beta(x, \sum_{k=1}^{i} h_{a,k}(\omega)) = d, \beta(x, \sum_{k=1}^{i-1} h_{a,k}(\omega)) = d'\},$$
$$I_{i,j}(a) = \{k : a_k = a_i, j < k < i\},$$
$$\mathscr{G}_{i,j}(a) = \mathscr{B}(h_{a,k} : k \in \{1, 2, \ldots, i\} - I_{i,j}(a)),$$
$$\pi_{i,j}(a) = (\max\{1, I_{i,j}(a)\})^{-1/2},$$
$$x \in X, \quad a \in A^N, \quad i \in N, \quad j \in N \cup \{0\}, \quad i > j.$$

Then there exist positive real constants $\varkappa$ and l such that, for every $a \in A^N$, $x \in X$, $i \in N$, $j \in N \cup \{0\}$, $i > j$, it holds

$$P(\omega, G_{a,i}(x) | \mathscr{G}_{i,j}(a)) \leqq \varkappa\, \pi_{i,j}(a) \max(\|h_{a,i}(\omega)\|, l) \quad \text{a.s.}$$

Proof. Using the method of the proof of Lemma 7, we obtain

$$G_{a,i}(x) \subset \{\omega : \sum_{k \in I_{i,j}(a)} v(h_{a,k}(\omega) - \mathsf{E}h_{a,k}) \in J(z(\omega), -v\, h_{a,i}(\omega))\}$$

where

$$z(\omega) = -\sum_{k \in I_{i,j}(a)} v\mathsf{E}h_{a,k} - \sum_{\substack{1 \leq k < i \\ i \notin I_{i,j}(a)}} v\, h_{a,k}(\omega).$$

If $v = 0$ or $I_{i,j}(a) = \emptyset$ then the lemma evidently holds. Therefore, we shall assume $v \neq 0$ and $I_{i,j}(a) \neq \emptyset$.

Under (i_1), the lemma follows directly from Lemma 6.

Under (i_2) we have

$$G_{a,i}(x) \sim \{\omega : \sum_{k \in I_{i,j}(a)} \tilde{v}\, \tilde{h}_{a,k}(\omega) \in J(z(\omega), -v\, h_{a,i}(\omega))\}$$

where

$$\tilde{h}_{a,i} = \{(h_{a,i} - \mathsf{E}h_{a,i})_m\}_{m \in A - \{m_{a,i}\}},$$
$$\tilde{v}_{a,i} = \{(v)_m - (v)_{m_{a,i}}\}_{m \in A - \{m_{a,i}\}}.$$

Since the (ii)-property of optimal procedures guarentees $\tilde{v} \neq 0$ and $\|v\|/\|\tilde{v}_{a,i}\| < \bar{A} + 1$, the lemma again follows from Lemma 6.

Lemma 9. *Let $\{b_n\}_{n \in N}$ be a monotone sequence of positive numbers such that $\sum_{n=1}^{\infty} b_n^2 < \infty$, β be an optimal procedure and $\{\psi_{a,i} : a \in A^N, i \in N\}$ be a family of measurable mappings on $(R^A, \mathscr{R}^A)$ to $(R^A, \mathscr{R}^A)$. Further, let $\{\{h_{a,i}\}_{i \in N} : a \in A^N\}$ be a family of sequences of independent random vectors on $(\Omega, \mathscr{S}, P)$ to $(R^A, \mathscr{R}^A)$ satisfying conditions* (i) *and* (ii) *of Lemma* 8 *and the condition*

(iii) $$\mathsf{E}(h_{a,i})_m^2 < \infty\,, \quad \mathsf{E}(\psi_{a,i}(h_{a,i}))_m^2 < \infty$$

for every $a \in A^N$, $i \in N$ *and* $m \in A$.

Then

$$b_n \sum_{i=1}^{n} \big(r_\beta(\psi_{a,i}(h_{a,i}), \sum_{j=1}^{i} h_{a,j}) - r_\beta(\psi_{a,i}(h_{a,i}), \sum_{j=1}^{i-1} h_{a,j})\big) \to 0$$

a.s. *uniformly in* $a \in A^N$.

Proof. Let us choose $m \in A$, $d, d' \in D$, $d \neq d'$. Let $\mathscr{G}_{i,j}(a)$ and $G_{a,i}(x)$ be the symbols introduced in Lemma 8 and let

$$\psi_{a,i}^{(m)} = (\psi_{a,i}(h_{a,i}))_m\,,$$

$$\zeta_{a,i} = \varkappa \max(\|h_{a,i}\|, l)\,,$$

$\mathscr{S}_{i,j}(a)$ be the minimal σ-field containing $\mathscr{G}_{j,0}(a) \cup \mathscr{B}(h_{a,i})$. Using smoothing properties of conditional expectations, Lemma 8, our independence assumption and Schwarz inequality, we get

$$\begin{aligned}
(5)\qquad & \mathsf{E}(|\psi_{a,i}^{(m)}|\,|\psi_{a,j}^{(m)}|\,\chi_{G_{a,i}(x)}\,\chi_{G_{a,j}(x)}) = \\
&= \mathsf{E}(|\psi_{a,i}^{(m)}|\,|\psi_{a,j}^{(m)}|\,\chi_{G_{a,j}(y)}\,P(G_{a,i}(x)|\,\mathscr{G}_{i,j}(a))) \leqq \\
&\leqq \pi_{i,j}(a)\,\mathsf{E}(|\psi_{a,i}^{(m)}|\,|\psi_{a,j}^{(m)}|\,\chi_{G_{a,j}(y)}\,\zeta_{a,i}) = \\
&= \pi_{i,j}(a)\,\mathsf{E}(|\psi_{a,i}^{(m)}|\,|\psi_{a,j}^{(m)}|\,\zeta_{a,i}\,P(G_{a,j}(y)|\,\mathscr{S}_{i,j}(a))) = \\
&= \pi_{i,j}(a)\,\mathsf{E}(|\psi_{a,i}^{(m)}|\,|\psi_{a,j}^{(m)}|\,\zeta_{a,i}\,P(G_{a,j}(y)\mid\mathscr{G}_{j,0}(a))) \leqq \\
&\leqq \pi_{i,j}(a)\,\pi_{j,0}(a)\,\mathsf{E}(|\psi_{a,i}^{(m)}|\,|\psi_{a,j}^{(m)}|\,\zeta_{a,i}\zeta_{a,j}) \leqq \\
&\leqq \pi_{i,j}(a)\,\pi_{j,0}(a)\,\mathsf{E}(|\psi_{a,i}^{m}|\,\zeta_{a,i})\,\mathsf{E}(\psi_{a,j}^{(m)}\zeta_{a,j}) \leqq \\
&\leqq \pi_{i,j}(a)\,\pi_{j,0}(a)\,(\mathsf{E}|\psi_{a,i}^{m}|^2\,\mathsf{E}\zeta_{a,i}^2\,\mathsf{E}|\psi_{a,j}^{(m)}|^2\,\mathsf{E}\zeta_{a,j}^2)^{1/2}\,.
\end{aligned}$$

Obviously

$$\sum_{k=1}^{x} k^{-1/2} \leqq 2x^{1/2} - 1\,, \quad x \in N\,,$$

$$\sum_{m\in A} x_m^{1/2} \leqq (\bar{A} \sum_{m\in A} x_m)^{1/2}\,, \quad x_m \in R\,.$$

Using these inequalities, we obtain

$$\sum_{i=j+1}^{n} \pi_{i,j}(a) = \sum_{m\in A} \sum_{\substack{i=j+1\\(a)_i=m}}^{n} \pi_{i,j}(a) \leqq \sum_{m\in A} \big(1 + \sum_{k=1}^{x_{n,j}^{(m)}(a)} k^{-1/2}\big) \leqq$$

$$\leqq 2\sum_{m\in A} (x_{n,j}^{(m)}(a))^{1/2} \leqq (\bar{A}(n-j))^{1/2} \leqq (\bar{A}n)^{1/2}$$

where

$$x_{n,j}^{(m)}(a) = \sum_{t=j+1}^{n} \chi_{\{m\}}((a)_t)$$

 and, hence

$$(6) \qquad \sum_{\substack{i,j=1 \\ i>j}}^{n} \pi_{i,j}(a)\, \pi_{j,0}(a) = \sum_{j=1}^{n-1} \pi_{j,0}(a) \sum_{i=j+1}^{n} \pi_{ij}(a) \leqq \bar{A} n\,.$$

According to (5), (6), (iii) and (ii), there is a real constant c such that, for every $n \in N$ and $a \in A^N$, it holds

$$\mathsf{E}\Big(\sum_{i=1}^{n} |\psi_{a,i}^{(m)}| \int \chi_{G_{a,i}(x)}\, \mathrm{d}\nu_m(x)\Big)^2 \leqq cn\,.$$

Therefore, according to Theorem 2,

$$(7) \qquad b_n \sum_{i=1}^{n} \psi_{a,i}^{(m)} \int \chi_{G_{a,i}(x)}\, \mathrm{d}\nu_m(x) \to 0 \quad \text{a.s. uniformly in} \quad a \in A^N\,.$$

Since for $\xi, \eta, \vartheta \in R^A$,

$$r_\beta(\xi, \eta) - r_\beta(\xi, \vartheta) = \sum_{m \in A} \sum_{\substack{d,d' \in D_0 \\ d \neq d}} (w(m, d) - w(m, d'))\, (\xi)_m \int \chi_{M^{dd'}}(x)\, \mathrm{d}\nu_m(x)$$

where

$$M^{d,d'} = \{x : \beta(x, \eta) = d,\ \beta(x, \vartheta) = d'\}\,,$$

the lemma follows from (7) and the finiteness of A and D.

If we use Lemma 7 instead of Lemma 8 in the method of proof of Lemma 9, we obtain,

Lemma 10. *Let $\{b_n\}_{n \in N}$ be a monotone sequence of positive numbers such that $\sum_{n=1}^{\infty} b_n^2 < \infty$, β be an optimal procedure and $\{\psi_{a,i} : a \in A^N, i \in N\}$ be a family of measurable mappings on $(R^A, \mathscr{R}^{A'})$ to $(R^A, \mathscr{R}^A)$. Further, let $\{\{h_{a,i}\}_{i \in N} : a \in A^N\}$ and $\{\{h'_{a,i}\}_{i \in N} : a \in A^N\}$ be two families of sequences of independent random vectors on $(\Omega, \mathscr{S}, P)$ to $(R^A, \mathscr{R}^A)$ satisfying conditions* (i) *and* (ii) *of Lemma 7 and condition* (iii) *of Lemma 9. Then*

$$b_n \sum_{i=1}^{n} \Big(r_\beta\big(\psi_{a,i}(h_{a,i}), \sum_{j=1}^{i} h_{a,j}\big) - r_\beta\big(\psi_{a,i}(h_{a,i}), \sum_{j=1}^{i-1} h_{a,j} + h'_{a,i}\big)\Big) \to 0$$

a.s. uniformly in $a \in A^N$.

MAIN THEOREMS

Throughout this section we shall assume that there is given a decision problem $\mathscr{D} = (A, D, w, (X, \mathscr{X}), \{\nu_a\}_{a \in A})$ with finite A and D and a family $\{g_{a,i} : a \in A^N, i \in N\}$

of generalized random variables i.e. measurable mappings on $(\Omega, \mathcal{S}, P)$ to $(X, \mathcal{X})$ satisfying conditions (G1) and (G2).

Theorem 3. *Let* $\{b_n\}_{n\in N}$ *be a monotone sequence of positive numbers satisfying* $\sum_{n=1}^{\infty} b_n^2 < \infty$, *$\beta$ be an optimal procedure and* $\{\{h_{a,i}\}_{i\in N} : a \in A^N\}$ *be a family of sequences of independent random vectors on* $(\Omega, \mathcal{S}, P)$ *to* $(R^A, \mathcal{R}^A)$ *such that, for every* $i \in N$ *and* $a \in A^N$, *the following conditions hold*

(i) $\mathsf{E}h_{a,i} = a_i$,

(ii) $Ph_{a,i}^{-1} = Ph_{a',i'}^{-1}$ *if* $a_i = a'_{i'}$,

(iii) $\mathsf{E}(h_{a,i})_m^2 < \infty$, $m \in A$,

(iv) *the σ-fields* $\mathcal{B}(h_{a,k}, g_{a,k} : k = 1, 2, \ldots, i)$ *and* $\mathcal{B}(g_{a,i+1})$ *are independent,*

(v) *the functions* $(h_{a,i} - \mathsf{E}h_{a,i})_m$, $m \in A$, *are linearly independent,*

or there is $m_{a,i} \in A$ *such that* $(h_{a,i} - \mathsf{E}h_{a,i})_m$, $m \in A - \{m_{a,i}\}$, *are linearly independent and*

$$\sum_{m\in A} (h_{a,i})_m = 1 .$$

Then

$$b_n \sum_{i=1}^{n} (w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i-1})) - \varrho(\bar{a}_n)) \to 0$$

a.s. uniformly in $a \in A^N$.

Note. If Z_k, $k = 1, 2, \ldots, r$, are random variables with finite dispersions, λ_k, $k = 1, 2, \ldots, r$, are real numbers and

$$c_{i,j} = \mathsf{E}(Z_i - \mathsf{E}Z_i)(Z_j - \mathsf{E}Z_j), \quad i, j = 1, 2, \ldots, r,$$

then

$$\max_{1\leq j\leq r} \left| \sum_{i=1}^{r} \lambda_i c_{i,j} \right| = 0$$

if, and only if,

$$\sum_{i=1}^{r} \lambda_i (Z_i - \mathsf{E}Z_i) = 0 \quad \text{a.s.}$$

Hence it follows that the covariance matrix $C = (c_{i,j})_{i,j=1}^r$ has rank m if, and only if, the linear subspace $\mathcal{L}(Z_1, Z_2, \ldots, Z_r)$ spanned by $Z_1, Z_2, \ldots, Z_r$ has rank m. Therefore (v) is implied both by (A_4) and by (A_5) in [14].

Proof. For $a \in A^N$ and $n \in N$, let us denote

$$Q_n(a) = b_n \sum_{i=1}^{n} \left(w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i-1})) - \varrho(\bar{a}_n)\right),$$

$$Q_n^{(1)}(a) = b_n \sum_{i=1}^{n} \left(r_\beta(a_i, \bar{h}_{a,i}) - \varrho(\bar{a}_n)\right),$$

$$Q_n^{(2)}(a) = b_n \sum_{i=1}^{n} \left(w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i-1})) - r_\beta(a_i, \bar{h}_{a,i-1})\right),$$

$$Q_n^{(3)}(a) = b_n \sum_{i=1}^{n} \left(r_\beta(a_i, \bar{h}_{a,i-1}) - r_\beta(a_i, \bar{h}_{a,i})\right),$$

$$Q_n^{(4)}(a) = r_\beta\left(b_n \sum_{i=1}^{n} (h_{a,i} - a_i), \bar{a}_n\right),$$

$$Q_n^{(5)}(a) = b_n \sum_{i=1}^{n} \left(r_\beta(h_{a,i} - a_i, \bar{h}_{i-1}) - r_\beta(h_{a,i} - a_i, \bar{h}_{a,i})\right),$$

$$Q_n^{(6)}(a) = -b_n \sum_{i=1}^{n} r_\beta(h_{a,i} - a_i, \bar{h}_{a,i-1}),$$

$$Q_n^{(7)}(a) = r_\beta\left(b_n \sum_{j=1}^{n} (h_{a,j} - a_j), \bar{h}_{a,n}\right),$$

$$Q_n^{(8)}(a) = b_n \sum_{i=1}^{n} \left(r_\beta(a_i, \bar{h}_{a,i-1}) - \varrho(\bar{a}_n)\right).$$

First we shall prove that, for every $l = 2, 3, \ldots, 7$,

$$Q_n^{(l)}(a) \to 0 \quad \text{a.s. uniformly in} \quad a \in A^N. \tag{8}$$

For $l = 3$ and $l = 5$, (8) follows from Lemma 9 if we set $\psi_{a,i} \equiv a_i$ and $\psi_{a,i}(x) = x - a_i$, respectively.

Applying Theorem 1 with $\mathscr{F}_{a,n+1} = \mathscr{B}(h_{a,i}, g_{a,i} : i = 1, 2, \ldots, n)$ and $\mathscr{F}_{a,n+1} = \mathscr{B}(h_{a,i} : i = 1, 2, \ldots, n)$, we obtain (8) for $l = 2$ and $l = 6$, respectively. Evidently, there is a real constant c such that

$$r_\beta\left(b_n \sum_{i=1}^{n} (h_{a,i} - a_i), \eta\right) \leqq c \sum_{m \in A} \left|b_n \sum_{i=1}^{n} (h_{a,i} - a_i)_m\right|, \quad \eta \in R^A.$$

Using this inequality and Theorem 1, we find that (8) holds for $l = 4$ and $l = 7$, too. Obviously,

$$Q_n^{(2)}(a) + Q_n^{(8)}(a) = Q_n(a) = \sum_{l=1}^{3} Q_n^{(l)}(a).$$

By Lemma 1 and Lemma 2,

$$Q_n^{(1)}(a) \leqq Q_n^{(4)}(a) + Q_n^{(5)}(a) + Q_n^{(6)}(a),$$

(9)
$$Q_n^{(8)}(a) \geqq \frac{b_n}{b_{n-1}} (Q_{n-1}^{(6)}(a) + Q_{n-1}^{(7)}(a)).$$

Therefore,

$$Q_n^{(2)}(a) + \frac{b_n}{b_{n-1}} (Q_{n-1}^{(6)}(a) + Q_{n-1}^{(7)}(a)) \leqq Q_n(a) \leqq \sum_{l=2}^{6} Q_n^{(l)}(a).$$

Since

$$\liminf_{n\to\infty} \frac{b_n}{b_{n-1}} \leqq 1,$$

the proof is complete.

Let us emphasize that the decision procedure $\beta(g_{a,i}, \bar{h}_{a,i-1})$ depends on $h_{a,1}, h_{a,2}, \ldots, h_{a,i-1}$ by means of their sum only, therefore the statistician needs to know for his decision in the i-th step, besides the observed value $g_{a,i}(\omega)$ only the sum $\sum_{j=1}^{i-1} h_{a,j}(\omega)$ or the arithmetic mean $\bar{h}_{a,i-1}(\omega)$.

If, in every step, the statistician knows the relative frequency vector of the parameter values in the preceding steps, then there exists a sequence $\{\zeta_i\}_{i\in N}$ of independent random vectors (to R^A) such that the random vectors $h_{a,i}$ defined by

$$h_{a,i} = \zeta_i + a_i, \quad a \in A^N, i \in N,$$

satisfy the conditions (i), (ii), (iii), (iv), and (v) of Theorem 3. Let us note that the assumption (v) is not fulfilled if we put $\zeta_i = 0$ and, as we shall see, (v) cannot be omitted.

If the Radon-Nikodym derivatives f_m, $m \in A$, are linearly independent, then there exists a measurable mapping ψ on $(X, \mathscr{X})$ to $(R^A, \mathscr{R}^A)$ such that $\mathsf{E}\,\psi(g_{a,i}) = a_i$ $\mathsf{E}\|\psi(g_{a,i})\|^2 < \infty$, $a \in A^N$, $i \in N$ (see [14]). Therefore, in every step, the statistician knows an unbiased estimate of the parameter value and he can use this knowledge for his decision, as the following theorem shows.

Theorem 4. *Let $\{b_n\}_{n\in N}$ be a monotone sequence of positive numbers such that $\sum_{n=1}^{\infty} b_n^2 < \infty$, β be an optimal procedure and $\{\{h_{a,i}\}_{i\in N} : a \in A^N\}$ be a family of sequences of independent random vectors on $(\Omega, \mathscr{S}, P)$ to $(R^A, \mathscr{R}^A)$ such that, for every $a \in A^N$ and $i \in N$, the conditions (i), (ii), (iii), and (iv) of Theorem 3 are satisfied. Then*

$$(b_n \sum_{i=1}^{n} (w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i})) - \varrho(\bar{a}_n)))^+ \to 0$$

a.s. uniformly in $a \in A^N$. If, moreover, for every $i \in N$ and $a \in A^N$, the condition (v) *of Theorem* 3 *holds, then*

$$b_n \sum_{i=1}^{n} (w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i})) - \varrho(\bar{a}_n)) \to 0$$

a.s. uniformly in $a \in A^N$.

Proof. Let $\{h'_{a,i} : a \in A^N,\ i \in N\}$ be a family of random vectors such that, for every $a \in A^N$, the families $\{h_{a,i} : i \in N\}$ and $\{h'_{a,i} : i \in N\}$ are equally distributed and the families

$$(\{h_{a,i} : i \in N\} \cup \{g_{a,i} : i \in N\}) \quad \text{and} \quad \{h'_{a,i} : i \in N\}$$

are independent. Let us denote

$$S_n(a) = b_n \sum_{i=1}^{n} (w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i})) - \varrho(\bar{a}_n)),$$

$$S_n^{(1)}(a) = b_n \sum_{i=1}^{n} (r_\beta(a_i, \bar{h}_{a,i}) - \varrho(\bar{a}_n)),$$

$$S_n^{(2)}(a) = b_n \sum_{i=1}^{n} (r_\beta(a_i, \sum_{j=1}^{i-1} h_{a,j} + h'_{a,i}) - r_\beta(a_i, \bar{h}_{a,i})),$$

$$S_n^{(3)}(a) = b_n \sum_{i=1}^{n} (w((a)_i, \beta(g_{a,i}, \sum_{j=1}^{i-1} h_{a,j} + h'_{a,i})) - r_\beta(a_i, \sum_{j=1}^{i-1} h'_{a,i} + h'_{a,i})),$$

$$S_n^{(4)}(a) = b_n \sum_{i=1}^{n} (w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i})) - w((a)_i, \beta(g_{a,i}, \sum_{j=1}^{i-1} h_{a,j} + h'_{a,i}))),$$

$$S_n^{(5)}(a) = r_\beta(b_n \sum_{i=1}^{n} (h_{a,i} - a_i), \bar{a}_n),$$

$$S_n^{(6)}(a) = b_n \sum_{i=1}^{n} (r_\beta(h_{a,i} - a_i, \bar{h}_{a,i}) - r_\beta(h_{a,i} - a_i, \sum_{j=1}^{i-1} h_{a,j} + h'_{a,i})),$$

$$S_n^{(7)}(a) = b_n \sum_{i=1}^{n} r_\beta(h_{a,i} - a_i, \sum_{j=1}^{i-1} h_{a,j} + h'_{a,i}),$$

$$a \in A^N, \quad n \in N.$$

First we shall prove that, for $l = 2, 3, \ldots, 7$,

$$S_n^{(l)}(a) \to 0 \quad \text{a. s. uniformly in} \quad a \in A^N. \tag{10}$$

For $l = 2$ and $l = 6$, (10) follows from Lemma 10 if we put $\psi_{a,i} \equiv a_i$ and $\psi_{a,i}(x) = x - a_i$, respectively. Applying Theorem 1 with $\mathscr{F}_{a,n+1} = \mathscr{B}(g_{a,i}, h_{a,i}, h'_{a,i} : i = 1, 2, \ldots, n)$, $\mathscr{F}_{a,n+1} = \mathscr{B}(h_{a,i} : i = 1, 2, \ldots, n)$ and $\mathscr{F}_{a,n+1} = \mathscr{B}(h'_{a,i}, h'_{a,i+1}, h_{a,i} : i = 1, 2, \ldots, n)$ we obtain (10) for $l = 3, 5$, and 7, respectively.

Obviously,

$$(11) \qquad S_n^{(4)}(a) = \sum_{\substack{d,d' \in D_0 \\ d \neq d}} b_n \sum_{i=1}^{n} (w((a)_i, d) - w((a)_i, d'))\, \chi_{M_{a,i}^{d,d'}}$$

where

$$M_{a,i}^{d,d'} = \{\omega : \beta(g_{a,i}(\omega), \bar{h}_{a,i}(\omega)) = d\,, \quad \beta(g_{a,i}(\omega), \sum_{j=1}^{i-1} h_{a,j}(\omega) + h'_{a,i}(\omega)) = d'\}\,.$$

From our independence assumptions it follows that

$$(12) \quad P(\omega, M_{a,i}^{d,d'} \mid g_{a,i}, h'_{a,i}, h_{a,k} : k \in I_{i,j}(a)) = P(\omega, F_{a,i}(g_{a,i}(\omega)) \mid \mathscr{F}_{i,j}(a)) \quad \text{a.s.}$$

where the symbols $I_{i,j}(a)$, $F_{a,i}(\cdot)$ and $\mathscr{F}_{i,j}(a)$ are introduced in Lemma 7. Using the method of the proof of Lemma 9, (11), (12) and Lemma 7, we find that (10) is valid for $l = 4$, too. According to Lemma 1 we have

$$S_n^{(1)}(a) \leqq \sum_{l=5}^{7} S_n^{(l)}(a)\,.$$

Since

$$(13) \qquad S_n(a) = \sum_{l=1}^{4} S_n^{(l)}(a)\,,$$

it follows that

$$S_n(a) \leqq \sum_{l=2}^{7} S_n^{(l)}(a)\,.$$

This proves the first assertion. As for the second one, according to (13), it suffices to prove

$$S_n^{(1)}(a) \to 0 \quad \text{a.s. uniformly in} \quad a \in A^N\,.$$

Since

$$S_n^{(1)}(a) = Q_n^{(8)}(a) - Q_n^{(3)}(a)$$

where $Q_n^{(8)}(a)$ and $Q_n^{(3)}(a)$ are defined in the proof of Theorem 9, the desired result follows from (9) and (8).

If the assumption (v) is dropped, Theorem 3 and the second part of Theorem 4 need not be true as the following example illustrates.

Example. Let $A = D = \{1, 2\}$, $X = \{x_0\}$,

$$w(1, 1) = w(2, 2) = 0\,, \quad w(1, 2) = w(2, 1) = 1\,,$$

$h_{a,i} \equiv a_i$. Then, for the optimal procedure β defined by

$$\beta(x_0, \xi) = \begin{matrix} 1 & \text{if} & (\xi)_1 > (\xi)_2\,, \\ 2 & \text{if} & (\xi)_1 \leqq (\xi)_2\,, \end{matrix}$$

and the point α such that $(\alpha)_{2i-1} = 1$, $(\alpha)_{2i} = 2$, $i \in N$, it holds

$$w((\alpha)_i, \beta(g_{\alpha,i}, \bar{h}_{\alpha,i})) = 0\,, \quad i = 1, 2, \ldots\,,$$
$$w((\alpha)_i, \beta(g_{\alpha,i}, \bar{h}_{\alpha,i-1})) = 1\,, \quad i = 2, 3, \ldots\,,$$
$$\varrho(\bar{\alpha}_n) \to \tfrac{1}{2}\,.$$

Under the assumptions of Theorem 3, there is a constant $c < \infty$ such that, for every $n \in N$ and $a \in A^N$, it holds

$$\left| \mathsf{E} \frac{1}{n} \sum_{i=1}^{n} w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i-1})) - \varrho(\bar{a}_n) \right| \leqq \frac{c}{\sqrt{n}}\,.$$

Under the assumptions of the first part of Theorem 4, there is a constant $c < \infty$ such that, for every $n \in N$ and $a \in A^N$,

$$\mathsf{E} \frac{1}{n} \sum_{i=1}^{n} w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i})) - \varrho(\bar{a}_n) \leqq \frac{c}{\sqrt{n}}\,.$$

If, moreover, (v) holds, then, for every $n \in N$ and $a \in A^N$,

$$\left| \mathsf{E} \frac{1}{n} \sum_{i=1}^{n} w((a)_i, \beta(g_{a,i}, \bar{h}_{a,i})) - \varrho(\bar{a}_n) \right| \leqq \frac{c}{\sqrt{n}}\,.$$

We shall not prove these assertions. They have been proved in [14] under slightly stronger assumptions. Weakening of assumptions is reached by the use of Kolmogorov-Rogozin inequality instead of Berry-Esseen one.

If we assume that, in every step, the statistician knows the unbiased estimate of the k-dimensional $(k > 1)$ empirical distribution of parameter values in the preceding steps, it would be possible to construct a decision process which is "better" that those constructed in Theorem 3 and 4. We shall deal with this problem in a subsequent paper.

(Received January 26, 1970.)

REFERENCES

[1] C. G. Esseen: On the Kolmogorov-Rogozin Inequality for the Concentration Function. Z. Wahrscheinlichkeitstheorie und Verw. Geb. *5* (1966), 210—216.

[2] P. R. Halmos: Measure Theory. New York 1950.

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

[4] S. Jílovec, B. Šubert: Repetitive Play of a Game against Nature. Aplikace matematiky *12* (1967), 383—397.

[5] S. Jílovec: Последовательное повторение игры при неопределености. Kybernetika *3* (1967), 560—586.

[6] S. Jílovec: Равномерная сходимость потерь в небайесовской модели статистического решения. Research report. ÚTIA-ČSAV, 1967.

[7] S. Jílovec: O vlastnostech jednoho strategického postupu při opakovaném rozhodování. Research repor. ÚTIA-ČSAV, 1969.
[8] M. Loève: Probability Theory. New York 1963.
[9] E. Parzen: On Uniform Convergence of Families of Sequences of Random Variables. Univ. of California Publ. in Statist. *2* (1954), 23—54.
[10] E. Samuel: Asymptotic Solutions of the Sequential Compound Decision Problem. Ann. Math. Statist. *34* (1963), 1079—1094.
[11] E. Samuel: Convergence of the Losses of Certain Decision Rules for the Sequential Compound Decision Problem. Ann. Math. Statist. *35* (1964), 1606—1621.
[12] B. Šubert: An Asymptotically Optimal Decision Procedure. Transaction of the Fourth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes. Prague 1967, 573—587.
[13] B. Šubert: Repetitive Play of an Unknown Game Against Nature. Stanford University Report No. 6151—6153, December 1967.
[14] J. Van Ryzin: The Sequential Compound Decision Problem with $m \times n$ Finite Loss Matrix. Ann. Math. Statist. *37* (1966), 954—975.
[15] J. Van Ryzin: Repetitive Play in Finite Statistical Games with Unknown Distributions. Ann. Math. Statist. *37* (1966), 976—994.

VÝTAH

O skoro jistě stejnoměrné konvergenci ztrát při opakování statistického rozhodovacího problému

Stanislav Jílovec

Článek se zabývá opakováním statistického rozhodovacího problému, jehož struktura je statistikovi známa a zůstává stále stejná. Pouze hodnoty parametrů, pozorování a statistikova rozhodnutí se mohou měnit v každém kroku. Předpokládá se, že statistik má v každém kroku kromě pozorované hodnoty k dispozici odhad relativních četností hodnot parametrů v předcházejících krocích. Uvažovaný přístup je nebayesovský, tj. parametr se nepovažuje za náhodnou proměnnou. Kromě toho se předpokládá, že při daných hodnotách parametrů jsou odpovídající pozorování stochasticky nezávislá.

Autor navazuje na problematiku formulovanou Hannanem a rozvíjenou dále v řadě prací. Na rozdíl od jiných autorů studuje však konvergenci ztrát a nikoliv konvergenci rizik. Konstruuje dva typy takových rozhodovacích procesů, že rozdíl průměrné ztráty za prvních n kroků a rizika optimální rozhodovací funkce vzhledem k relativním četnostem hodnot parametrů v prvních n krocích konverguje k nule skoro jistě. Tato konvergence je stejnoměrná a „dosti" rychlá (Teorém 3 a 4).

Dr. Stanislav Jílovec, CSc., Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.