ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XVII, 2 (1982)

R. ZIELIŃSKI (Warszawa)

A CLASS OF STOPPING RULES FOR FIXED PRECISION SEQUENTIAL ESTIMATES

1. Introduction. Let $(\Omega, \mathscr{F}, \{P_\theta : \theta \in \Theta\})$ be a statistical space and g a mapping from Θ into a metric space $\mathscr{X}$ with the distance between $x, y \in \mathscr{X}$ denoted by $\|x-y\|$. Let $(X_t), t = 1, 2, \ldots$, be a sequence of $\mathscr{X}$-valued random elements on $(\Omega, \mathscr{F})$ which is assumed to converge to $g(\theta)$ whenever the distribution of (X_t) is generated by P_θ. Given $\varepsilon > 0$ and $\gamma \in (0, 1)$, the problem consists in fixed precision estimation of $g(\theta)$, i.e. in finding a stopping variable $\tau : \Omega \to \{1, 2, \ldots\}$ such that

$$P_\theta\{\|X_\tau - g(\theta)\| < \varepsilon\} \geqslant 1-\gamma \quad \text{for all } \theta \in \Theta.$$

As in [3] suppose that we can observe k independent copies $(X_t^{(i)})$, $i = 1, 2, \ldots, k$, of (X_t). Intuitively, one might expect that if k is large, then all the sequences $(X_t^{(i)})$ "meet together" only in a narrow neighborhood of $g(\theta)$. This suggests the stopping rule

$$\tau = \inf\{t \geqslant 1 : \varrho_k(X_t^{(1)}, \ldots, X_t^{(k)}) < \delta\},$$

where δ is a (small) positive number and $\varrho_k\colon \mathscr{X}^k \to R_+^1$ is a measure of concentration of points in $\mathscr{X}$ (e.g., the radius of the smallest ball containing all points $X_t^{(1)}, \ldots, X_t^{(k)}$). We shall show that under some conditions concerning the speed of convergence of (X_t) and the distribution of (X_t) the stopping variable τ leads to the desired result: Theorem 1 in Section 2 establishes the existence of fixed precision estimates, and Theorem 2 suggests a method of the effective construction of such estimates. The method given in Theorem 2 is then applied in the example in which a "usual" fixed precision estimate does not exist.

2. Results. The following assumptions are relevant to the Lemma and Theorem 1:

(A1) $(\forall \eta > 0) \sum_{t=1}^{\infty} P_\theta\{\|X_t - g(\theta)\| \geqslant \eta\} < \infty$ uniformly in $\theta \in \Theta$, i.e.

$$(\forall \eta > 0)(\forall \eta' > 0)(\exists T)(\forall t > T)(\forall \theta \in \Theta) \sum_{j=t}^{\infty} P_\theta\{\|X_j - g(\theta)\| \geqslant \eta\} \leqslant \eta'.$$

(A2) Given $\theta \in \Theta$ and t, the random elements $X_t^{(i)}$ for $i = 1, 2, \ldots$ are i.i.d.

(A3) For each $k = 2, 3, \ldots, \varrho_k$ is a non-negative real-valued function on $\mathscr{X}^k$ such that for any set of points $x^{(1)}, \ldots, x^{(k)} \in \mathscr{X}$

$$c_1 \max_{i,j} \|x^{(i)} - x^{(j)}\| \geqslant \varrho_k(x^{(1)}, \ldots, x^{(k)}) \geqslant c_2 \max_{i,j} \|x^{(i)} - x^{(j)}\|$$

with some positive constants $c_1 = c_1(k)$ and $c_2 = c_2(k)$.

(A4) For each $k = 2, 3, \ldots, \hat{g}_k$ is a mapping from $\mathscr{X}^k$ into $\mathscr{X}$ such that for every set of points $x^{(1)}, \ldots, x^{(k)} \in \mathscr{X}$ and for each $y \in \mathscr{X}$

$$\|\hat{g}_k(x^{(1)}, \ldots, x^{(k)}) - y\| \leqslant \max_{1 \leqslant i \leqslant k} \|x^{(i)} - y\|.$$

For example, if $\mathscr{X}$ is a convex subset of a normed linear space (with norm $\|\cdot\|$), then $\hat{g}_k$ might be defined as a "sample mean" $(x^{(1)} + \ldots + x^{(k)})/k$ (see [3]).

LEMMA. *Under assumptions* (A1)-(A3) *we have* $P_\theta\{\tau < \infty\} = 1$ *for each* $\delta > 0$ *and for all* $\theta \in \Theta$.

THEOREM 1. *If* (A1)-(A4) *hold and if for each* t

$$a_{\theta,t}(\delta) = \sup_{y \in \mathscr{X}} P_\theta\{\|X_t - y\| < \delta\}$$

tends to zero as $\delta \to 0$, *uniformly in* $\theta \in \Theta$, *then for every* $\varepsilon > 0$, $\gamma \in (0, 1)$ *and* $k \geqslant 2$ *there exists a* $\delta > 0$ *such that*

$$P_\theta\{\|\hat{g}_k(X_\tau^{(1)}, \ldots, X_\tau^{(k)}) - g(\theta)\| < \varepsilon\} \geqslant 1 - \gamma$$

for all $\theta \in \Theta$.

The proofs are slight modifications of those given in [3] and will be omitted.

THEOREM 2. *Let* (Y_t) *and* (r_t) *be two sequences of* $\mathscr{X}$*-valued and real-valued random elements, respectively, such that the distribution of* $S_t = r_t^{-1}\|Y_t - g(\theta)\|$ *does not depend on* θ. *If there exists a sequence* (a_t) *of real numbers such that*

$$\sum_{t=1}^{\infty} P_\theta\{S_t \geqslant a_t\} < \infty$$

and

$$P_\theta\text{-}\lim_{t \to \infty} a_t r_t = 0 \quad \textit{for all } \theta \in \Theta,$$

then for every $\varepsilon > 0$ *and* $\gamma \in (0, 1)$ *there exists an integer* T *such that* $\tau = \inf\{t > T : a_t r_t \leqslant \varepsilon\}$ *is a (finite) stopping rule and*

$$P_\theta\{\|Y_\tau - g(\theta)\| < \varepsilon\} \geqslant 1 - \gamma \quad \textit{for all } \theta \in \Theta.$$

Proof. $P_\theta\text{-}\lim_{t\to\infty} a_t r_t = 0$ implies $P_\theta\{\tau < \infty\} = 1$. We have

$$P_\theta\{\|Y_\tau - g(\theta)\| \geqslant \varepsilon\} = \sum_{t=T}^{\infty} P_\theta\{\|Y_t - g(\theta)\| \geqslant \varepsilon,\ \tau = t\}$$
$$\leqslant \sum_{t=T}^{\infty} P_\theta\{\|Y_t - g(\theta)\| \geqslant \varepsilon,\ \varepsilon \geqslant a_t r_t\} \leqslant \sum_{t=T}^{\infty} P_\theta\{S_t \geqslant a_t\}$$

and the theorem follows.

3. Examples. 1. Let $\Theta = \{\theta = (\mu, \sigma): -\infty < \mu < \infty,\ \sigma' \leqslant \sigma \leqslant \sigma''\}$ with $0 < \sigma' \leqslant \sigma'' < \infty$ and let (X_t) be a sequence of real-valued random variables. Suppose that P_θ is a probability measure such that

$$P_\theta\{X_t \leqslant x\} = \Phi\left(\frac{x-\mu_t}{\beta_t \sigma}\right),$$

where Φ is the probability distribution function of the standard normal variable $N(0,1)$, $\mu_t = \mu + m_t$, $m_t \to 0$ and $0 < \beta_t \to 0$ fast enough as $t \to \infty$. This is a typical stochastic approximation case.

Suppose that $m_t = O(\beta_t)$. Then, for every positive η,

$$\sum_{t=1}^{\infty} P_\theta\{|X_t - \mu| \geqslant \eta\} = \sum_{t=1}^{\infty}\left[\Phi\left(-\frac{\eta}{\beta_t\sigma} - \frac{m_t}{\beta_t\sigma}\right) + 1 - \Phi\left(\frac{\eta}{\beta_t\sigma} - \frac{m_t}{\beta_t\sigma}\right)\right],$$

which converges uniformly in $\theta \in \Theta$ so that (A1) holds. For $a_{\theta,t}(\delta)$ we have

$$a_{\theta,t}(\delta) = 2\Phi\left(\frac{\delta}{\beta_t\sigma}\right) - 1,$$

which for each t tends to zero as $\delta \to 0$, uniformly in $\theta \in \Theta$.

Suppose that k independent realizations $(X_t^{(i)})$, $i = 1, 2, \ldots, k$, of (X_t) are available. Then, by Theorem 1, as a sequential estimate of μ we can take

$$\hat{\mu}_t = \sum_{i=1}^{k} X_t^{(i)}/k$$

and stop the process of estimation using the stopping rule τ with

$$\varrho_k(X_t^{(1)}, \ldots, X_t^{(k)}) = \sum_{i=1}^{k} (X_t^{(i)} - \hat{\mu}_t)^2/k.$$

2. Let ξ_n $(n = \ldots, -1, 0, 1, \ldots)$ be a doubly infinite sequence of i.i.d. random variables $N(0,1)$ and let $X_t = \mu + \xi_t + Z_{t,m}$, where

$$Z_{t,m} = \begin{cases} 0 & \text{if } m = 0, \\ m^{-1/2} \sum_{j=1}^{m} \xi_{t-j} & \text{if } m = 1, 2, \ldots \end{cases}$$

Our aim is to construct a fixed length (ε, say), fixed confidence level ($1-\gamma$, say) confidence interval for the mean μ with m being unknown. It is well known (see [1]) that if ε and γ satisfy some additional requirements, then there exist no functions $\hat{\mu}_t(X_1, \ldots, X_t)$, $t = 1, 2, \ldots$, and no finite stopping rule which could give a solution.

Take k independent copies $(X_t^{(i)})$, $i = 1, 2, \ldots, k$, of (X_t) and put

$$\bar{X}_t^{(i)} = \sum_{n=1}^{t} X_n^{(i)}/t.$$

We apply Theorem 2 with

$$Y_t = \sum_{i=1}^{k} \bar{X}_t^{(i)}/k \quad \text{and} \quad r_t^2 = \sum_{i=1}^{k} (\bar{X}_t^{(i)} - Y_t)^2/(k-1).$$

The random variable $S_t = r_t^{-1}(Y_t - \mu)\sqrt{k}$ is distributed as t_{k-1} — Student's t with $k-1$ degrees of freedom — and depends neither on $\theta = (m, \mu)$ nor on t. Take the following approximation for the cumulative distribution function of t_{k-1} (see [2]):

$$\mathrm{P}\{t_{k-1} \geqslant x\} = cx^{-(k-1)}\big(1 + O(x^{-2})\big) \qquad (c = \text{const}).$$

It follows that for $a_t = t^a$ the series $\sum_{t=1}^{\infty} P_0\{|S_t| \geqslant a_t\}$ converges if $a > 1/(k-1)$.

By the Tchebycheff inequality we have

$$P_0\{a_t^2 r_t^2 > \eta\} \leqslant c' t^{-1} a_t^2 \qquad (c' = \text{const}),$$

so that $a_t r_t \to 0$ in probability if $a < 1/2$. By Theorem 2, a fixed precision estimate for μ exists whenever the number of copies is $k \geqslant 4$.

Acknowledgment. The author is greatly indebted to Dr. Jacek Koronacki for helpful remarks and stimulating discussion.

References

[1] J. R. Blum and J. Rosenblatt, *On fixed precision estimation in time series*, Ann. Math. Statist. 40 (1969), p. 1021-1032.

[2] L. N. Bol'šev and N. V. Smirnov (Л. Н. Большев и Н. В. Смирнов), *Таблицы математической статистики*, Москва 1968.

[3] R. Zieliński, *A stopping rule for sequential estimation processes*, Bull. Acad. Polon. Sci., Sér. Sci. Math., Astronom. Phys., 26 (1978), p. 441-444.

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
00-950 WARSZAWA

Received on 11. 1. 1979

R. ZIELIŃSKI (Warszawa)

PEWNA KLASA REGUŁ ZATRZYMYWANIA PROCESÓW ESTYMACJI SEKWENCYJNEJ

STRESZCZENIE

Niech $(\Omega, \mathscr{F}, \{P_\theta: \theta \in \Theta\})$ będzie przestrzenią statystyczną i niech g będzie odwzorowaniem Θ w przestrzeń metryczną $\mathscr{X}$, gdzie odległość między punktami $x, y \in \mathscr{X}$ oznaczona jest przez $\|x-y\|$. Niech (X_t), $t = 1, 2, \ldots$, będzie ciągiem elementów losowych o wartościach w $\mathscr{X}$, zbieżnym do $g(\theta)$, gdy „prawdziwy" rozkład jest P_θ. (Ciąg (X_t) jest estymatorem sekwencyjnym parametru $g(\theta)$.) Niech $\varepsilon > 0$ oraz $\gamma \in (0, 1)$ będą ustalonymi liczbami. Zakładając, że można obserwować jednocześnie $k \geqslant 2$ niezależnych realizacji $(X_t^{(i)})$, $i = 1, 2, \ldots, k$, ciagu (X_t), skonstruowano taki estymator sekwencyjny $\hat{g}_t = g(X_t^{(1)}, \ldots, X_t^{(k)})$ oraz taki moment zatrzymywania τ, że $P_\theta\{\|\hat{g}_\tau - g\| < \varepsilon\} \geqslant 1-\gamma$ dla wszystkich $\theta \in \Theta$.