THEORY OF PROBABILITY AND ITS APPLICATIONS

Volume XXVIII Number 1

SHORT COMMUNICATIONS

ON RANDOM SEARCH FOR A GLOBAL EXTREMUM

S. M. ERMAKOV AND A. A. ZHIGLYAVSKII

(Translated by K. Durr)

The problem of searching for a global extremum of a function under sufficiently broad assumptions has been investigated relatively little. The existing meaningful results pertain to the simplest methods of random search.

This paper considers more complex probabilistic models by means of which it is possible to analyze several methods [1], [2] suggested because of heuristic considerations that are highly recommended in the solution of practical problems. These models are based on the construction of a sequence of probability distributions that converge for a wide class of functions to a limit distribution concentrated at the global extremum point. The modeling of these distributions is the content of the corresponding algorithms.

Suppose that X is a compact separable metric space, ρ is a metric on X, $\mathcal{B}$ is a σ-algebra of Borel subsets of X, f is a non-negative bounded function on X continuous at all but a finite number of points, and $\mu(dx)$ is a probability measure on B. Consider the problem of seeking the point x^* at which $f(x^*)=\max_{x\in X} f(x)$ from the results of measuring the r.v. $\eta(x_i)=f(x_i)+\xi(x_i)$ at points $x_i\in X$, $i=1,2,\cdots$, defined by the optimization algorithm.

We assume that

a) for any $x\in X$, $\xi(x)$ is a r.v. having distribution $F(x, d\xi)$ with zero mean and concentrated on a finite interval $[-d, d]$, where for any $x_1, x_2, \cdots$ in X the r.v.'s $\xi(x_1), \xi(x_2), \cdots$ are mutually independent;

b) $\eta(x)\geqq c_1>0$ for all $x\in X$ with probability 1. The following algorithm can be used to solve this fairly general problem.

Algorithm 1.

1. *Choose a distribution* $Q(0, N_{-1}; dx)=P_0(dx)$ *on* $\mathcal{B}$, *and set* $s=0$.

2. *Model the distribution* $Q(s, N_{s-1}; dx)$ *a specified number of times* N_s, *and obtain* $x_1^{(s)}, \cdots, x_{N_s}^{(s)}$.

3. *Set* $Q(s+1, N_s; dx)=\sum_{i=1}^{N_s} p_i^{(s)} Q_s(x_i^{(s)}, dx)$, *where* $p_i^{(s)}=\eta(x_i^{(s)})/\sum_{j=1}^{N_s}\eta(x_j^{(s)})$, *and* $Q_s(y, dx)$ *is a given transition probability (i.e. a probability measure on* $\mathcal{B}$ *in the second argument for any* $y\in X$ *and a* $\mathcal{B}$*-measurable function in the first argument).*

4. *Set* $s=s+1$, *and go to step* 2 *if* $s\leqq s_k$.

The quantity x^* is estimated by one of the standard methods from the points of the last iteration. The behavior of Algorithm 1 was investigated in certain special cases in [1], [3].

The distributions $Q(s+1, N_s; dx)$ are the conditional distributions of the random elements $x_i^{(s+1)}$ $(i=1,\cdots, N_{s+1}, s=0,1,\cdots)$ for fixed values of $x_1^{(s)},\cdots, x_{N_s}^{(s)}$, $\xi(x_1^{(s)}),\cdots,\xi(x_{N_s}^{(s)})$. The paper also studies the asymptotic behavior of $P(s+1, N_s; dx)$ (as $N_s\to\infty$, $s\to\infty$), the unconditional distributions of the random elements $x_i^{(s+1)}$ $(i=1,\cdots, N_{s+1}, s=0,1,\cdots)$. We shall prove that the distributions $P(s+1, N_s; dx)$ converge weakly to $\varepsilon_{x^*}(dx)$—a distribution concentrated at x^*. This convergence follows from the fact that for large N_s, the distribution $P(s+1, N_s; dx)$ is an approximation to the distribution $f^s(x)\mu(dx)/\int_X f^s(y)\mu(dy)$. Optimization algorithms of this type were first considered in [4].

Let $P(s, N_{s-1}; dx_1, \cdots, dx_{N_s})$ be the joint distribution of the random elements $x_1^{(s)}, \cdots, x_{N_s}^{(s)}$. Then

$$P(0, N_{-1}; dx_1, \cdots, dx_{N_0}) = P_0(dx_1) \cdots P_0(dx_{N_0}),$$

$$\begin{aligned} &P(s+1, N_s; dx_1, \cdots, dx_{N_{s+1}}) \\ &\quad = \int_{X^{N_s}} \int_{-d}^{d} \cdots \int_{-d}^{d} P(s, N_{s-1}; dy_1, \cdots, dy_{N_s}) \\ &\quad \cdot F(y_1, d\xi_1) \cdots F(y_{N_s}, d\xi_{N_s}) \prod_{k=1}^{N_{s+1}} \sum_{i=1}^{N_s} [f(y_i)+\xi_i]Q(y_i, dx_k) \Big/ \sum_{j=1}^{N_s} [f(y_j)+\xi_j], \end{aligned} \tag{1}$$

$$P(s+1, N_s; dx) = P(s+1, N_s; dx, X, X, \cdots, X).$$

We shall prove two lemmas.

Lemma 1. *Let conditions* a) *and* b) *hold and let*

c) $Q_s(y, dx) = q_s(y, x)\mu(dx), \quad \sup_{x,y\in X} q_s(y, x) \leqq M_s < \infty \quad$ *for all* $s = 0, 1, \cdots$.

Then for any $s = 0, 1, \cdots$ *and* $N_s = 1, 2, \cdots$,

$$\mathbf{P}(s+1, N_s; dx) = \left[\int_X P(s, N_{s-1}; dz)f(z)\right]^{-1} \int_X P(s, N_{s-1}; dy)f(y)R(s, N_s, y; dx), \tag{2}$$

where $R(s, N_s, y; dx) = Q_s(y, dx) + \Delta(s, N_s; dx)$; *for any* $s = 0, 1, \cdots$ *the generalized* (*alternating*) *measures* $\Delta(s, N_s; dx)$ *converge in variation to* 0 *as* $N_s \to \infty$ *at a rate of order* $N_s^{-1/2}$: $\|\Delta(s, N_s; \cdot)\| = O(N_s^{-1/2})(N_s \to \infty)$.

PROOF. Using (1), we can prove by induction that for any $s = 0, 1, \cdots$ the r.v.'s $x_1^{(s)}, \cdots, x_{N_s}^{(s)}$ are symmetrically dependent. We choose the marginal distribution $P(s+1, N_s; dx)$ as follows (for brevity we write $N = N_s$ and $M = N_{s-1}$):

$$\begin{aligned} P(s+1, N; dx) &= \int_{X^N} \int_{-d}^{d} \cdots \int_{-d}^{d} P(s, M; dy_1, \cdots, dy_N)F(y_1, d\xi_1) \cdots \\ &\quad \cdots F(y_N, d\xi_N) \sum_{i=1}^{N} [f(y_i)+\xi_i]Q_s(y_i, dx) \Big/ \sum_{j=1}^{N} [f(y_i)+\xi_j] \\ &= \int_{X^N} \int_{-d}^{d} \cdots \int_{-d}^{d} P(s, M; dy_1, \cdots, dy_N)F(y_1, d\xi_1) \cdots F(y_N, d\xi_N) \\ &\quad \cdot \left[N^{-1} \sum_{j=1}^{N} [f(y_j)+\xi_j]\right]^{-1} [f(y_1)+\xi_1]Q_s(y_1, dx). \end{aligned}$$

The last relation can be written in the form (2), where

$$\begin{aligned} \Delta(s, N; dx) &= \int_{X^N} \int_{-d}^{d} \cdots \int_{-d}^{d} P(s, M; dy_1, \cdots, dy_N)F(y_1, d\xi_1) \cdots F(y_N, d\xi_N) \\ &\quad \cdot [f(y_1)+\xi_1]Q_s(y_1, dx) \\ &\quad \cdot \left\{\left[N^{-1} \sum_{j=1}^{N} [f(y_j)+\xi_j]\right]^{-1} - \left[\int_X P(s, M; dz)f(z)\right]^{-1}\right\}. \end{aligned} \tag{3}$$

Let us show that $\Delta(s, N; \cdot) \to 0$ in variation as $N \to \infty$. In this case this means ([5], p. 118), that $\int_X \delta(N, x)\mu(dx) \to 0$ as $N \to \infty$, where

$$\delta(N, x) = \Bigg| \int_{X^N} \int_{-d}^{d} \cdots \int_{-d}^{d} P(s, M; dy_1, \cdots, dy_N) F(y_1, d\xi_1) \cdots$$
$$\cdots F(y_N, d\xi_N)[f(y_1) + \xi_1] q_s(y_1, x)$$
$$\cdot \left\{ \left[N^{-1} \sum_{j=1}^{N} [f(y_j) + \xi_j] \right]^{-1} - \left[\int_X P(s, M; dz) f(z) \right]^{-1} \right\} \Bigg|.$$

To this end, it is sufficient for any $\varepsilon_1 > 0$ and any $x \in X$ that there exist an $N_* = N_*(\varepsilon_1)$ such that for $N \geqq N_*$,

$$\delta(N, x) < \varepsilon_1. \tag{4}$$

We shall prove this.

Since the random elements $x_1^{(s)}, \cdots, x_N^{(s)}$ for any N are symmetrically dependent, then so are the r.v.'s $\eta(x_i^{(s)}) = f(x_i^{(s)}) + \xi(x_i^{(s)})$, $i = 1, \cdots, N$. Hence the r.v.'s ([6], p. 400) $\varkappa_N = N^{-1} \sum_{j=1}^{N} \eta(x_j^{(s)})$ converge in the mean to some r.v. $\varkappa$ which is independent of all the $\varkappa_i$, $i = 1, 2, \cdots$, with $\mathbf{E}\varkappa = \mathbf{E}\eta(x_1^{(s)}) = \int_X P(s, M; dz) f(z)$. We can formulate this as follows: for any $\varepsilon_2 > 0$ there is an $N^* \geqq 1$ such that $\mathbf{E}|\varkappa_N - \varkappa| < \varepsilon_2$ for $N \geqq N^*$. Write $\zeta = [f(x_1^{(s)}) + \xi(x_1^{(s)})] q_s(x_1^{(s)}, x)$. It is clear that $\operatorname{ess\,sup} \zeta \leqq (\max_{x \in X} f(x) + d) M_s = M_*$. Using the independence of $\varkappa$ from $\varkappa_N$ and ζ and conditions a)–c), we obtain

$$\delta(N, x) = \left| \mathbf{E}\left(\frac{\zeta}{\varkappa_N}\right) - \frac{\mathbf{E}\zeta}{\mathbf{E}\varkappa} \right| = \frac{1}{\mathbf{E}\varkappa} \left| \mathbf{E}\left(\frac{\varkappa\zeta}{\varkappa_N}\right) - \mathbf{E}\zeta \right|$$
$$\leqq [\inf \mathbf{E}\varkappa]^{-1} \left| \mathbf{E}\left(\frac{\varkappa\zeta}{\varkappa_N} - \zeta\right) \right| \leqq c_1^{-1} \mathbf{E}\left[\left| \frac{\zeta}{\varkappa_N} \right| |\varkappa - \varkappa_N| \right]$$
$$\leqq c_1^{-1} \operatorname{ess\,sup} \zeta [\operatorname{ess\,sup} \varkappa_N]^{-1} \mathbf{E}|\varkappa_N - \varkappa| \leqq c_1^{-2} M_* \mathbf{E}|\varkappa_N - \varkappa|.$$

Thus if we put $\varepsilon_2 = \varepsilon_1 c_1^2 M_*^{-1}$ and $N_* = N^*$, then (4) will hold for $N \geqq N_*$. Moreover, from the last chain of inequalities, it follows that $\|\Delta(s, N; \cdot)\| \leqq c_1^{-2} M_* \mathbf{E}|\varkappa_N - \varkappa|$. From the central limit theorem for symmetrically dependent r.v.'s (see [7]) and the inequality

$$\mathbf{E}|\varkappa_N - \varkappa| \leqq N^{-1/2} + \operatorname{ess\,sup} |\varkappa_N - \varkappa| \mathbf{P}\{|\varkappa_N - \varkappa| \geqq N^{-1/2}\},$$

which is a consequence of an inequality given in [6], p. 157, it follows that $\mathbf{E}|\varkappa_N - \varkappa| = O(N^{-1/2})$ $(N \to \infty)$ and hence $\|\Delta(s, N; \cdot)\| = O(N^{-1/2})$ $(N \to \infty)$. The lemma is proved.

Lemma 2. *Let the following conditions hold:*

d) x^*, *the point at which the function f has a global maximum, is unique and f is continuous in some neighborhood of this point;*

e) $\mu\{B_\varepsilon(x)\} > 0$ *for any* $\varepsilon > 0$ *and* $x \in X$, *where* $B_\varepsilon(x) = \{y \in X | \rho(x, y) \leqq \varepsilon\}$;

f) *there exists an* $\varepsilon_0 > 0$ *such that for any* ε, $0 < \varepsilon \leqq \varepsilon_0$, *the set* $A(\varepsilon) = \{x \in X | f(x^*) - f(x) \leqq \varepsilon\}$ *is simply connected and* $\mu\{a(\varepsilon) > 0\}$.

Then the sequence of distributions $f^m(x)\mu(dx)/\int_X f^m(z)\mu(dz)$ *converges weakly to the distribution* $\varepsilon_{x^*}(dx)$ *as* $m \to \infty$.

PROOF. Let $B_i = B_{\varepsilon_i}(x^*)$,

$$D_i = \overline{X \backslash B_i} = \{x \in X | \rho(x, x^*) \geqq \varepsilon_i\}, \quad i = 1, 2, 4, \quad K_1 = \sup_{x \in D_1} f(x).$$

Choose an arbitrary $\varepsilon_1 > 0$. From d), it follows that for any $\varepsilon_1 > 0$ there exists an ε_2 $(0 < \varepsilon_2 < \varepsilon_1)$ such that $K_2 = \inf_{x \in B_2} f(x) > K_1$. For any $m > 0$, we have

$$\int_{B_1} \left(\frac{f(x)}{K_1}\right)^m \mu(dx) > \int_{B_2} \left(\frac{f(x)}{K_1}\right)^m \mu(dx) \geqq \int_{B_2} \left(\frac{K_2}{K_1}\right)^m \mu(dx).$$

Passing to the limit (as $m \to \infty$) in the inequality

$$\int_{D_1} \mu(dx) \Big/ \left(\frac{K_2}{K_1}\right)^m \int_{B_2} \mu(dx) \geqq \int_{D_1} \left(\frac{f(x)}{K_1}\right)^m \mu(dx) \Big/ \int_{B_1} \left(\frac{f(x)}{K_1}\right)^m \mu(dx) \geqq 0,$$

we obtain $\int_{D_1} f^m(x)\mu(dx)/\int_{B_1} f^m(x)\mu(dx) \to 0$, which implies that ($c_m = [\int_X f^m(x)\mu(dx)]^{-1}$)

$$c_m \int_{D_1} f^m(x)\mu(dx) \to 0, \qquad c_m \int_{B_1} f^m(x)\mu(dx) \to 1 \quad (m \to \infty). \tag{5}$$

Now we choose an arbitrary continuous function $\psi(x)$ on X. Let us show that $\lim_{m\to\infty} c_m \int_X f^m(x)\psi(x)\mu(dx) = \psi(x^*)$. For any $\delta > 0$, there exists an $\varepsilon_3 > 0$ such that $|\psi(x) - \psi(x^*)| < \delta$ for $\rho(x, x^*) \leqq \varepsilon_3$ and $x \in X$. Put $\varepsilon_4 = \min\{\varepsilon_1, \varepsilon_3\}$. Then

$$\begin{aligned}
&\left| c_m \int_X \psi(x) f^m(x)\mu(dx) - \psi(x^*) \right| \\
&\leqq c_m \int_{B_4} f^m(x)|\psi(x) - \psi(x^*)|\mu(dx) + c_m \int_{D_4} f^m(x)|\psi(x) - \psi(x^*)|\mu(dx) \\
&\leqq \delta c_m \int_{B_4} f^m(x)\mu(dx) + 2 \max_{x \in X} |\psi(x)| c_m \int_{D_4} f^m(x)\mu(dx).
\end{aligned}$$

By relations (5) and the definition of weak convergence the lemma is proved.

The next assertion follows immediately from Lemma 1.

Corollary 1. *Let conditions* a)–c) *hold. Then for any* $s = 0, 1, \cdots$ *the distributions* $P(s+1, N_s; dx)$ *converge in variation as* $N_s \to \infty$ *to the limit* $P_s(dx)$, *where*

$$P_{s+1}(dx) = \left[\int_X P_s(dz) f(z)\right]^{-1} \int_X P_s(dy) f(y) Q_s(y, dx). \tag{6}$$

Let us introduce conditions that are sufficient for the distributions defined by (2) and (6) to converge weakly to $\varepsilon_{x^*}(dx)$ as $s \to \infty$.

Theorem 1. *Suppose that conditions* c)–f) *hold*; $Q_s(y, dx)$ (*or* $R(s, N_s, y; dx)$) *for any* $y \in X$ *converge weakly to* $\varepsilon_y(dx)$ *as* $s \to \infty$;

g) *for any* $\varepsilon > 0$ *there exist a* $\delta > 0$ *and a natural number* s_0 *such that* $P_s(B_\varepsilon(x^*)) \geqq \delta$ (*or* $P(s, N_{s-1}; B_\varepsilon(x^*)) \geqq \delta$) *for all* $s \geqq s_0$.

Then the sequence of distributions defined by (6) (*or, respectively, by* (2)) *converges weakly as* $s \to \infty$ *to* $\varepsilon_{x^*}(dx)$.

We carry out the proof for the sequence of distributions (6) (for (2) the proof is similar). From the sequence (6), choose a convergent subsequence $P_{s_k}(dx)$ (this is possible by Prokhorov's theorem [8], p. 37), and denote the limit by $Q_0(dx)$ (Q_0 is a probability measure on $\mathscr{B}$). From (6) it follows that the subsequence $P_{s_k+1}(dx)$ converges weakly to the distribution $Q_1(dx) = L_1 f(x) Q_0(dx)$ (L_1 is a normalization constant) and, similarly, $P_{s_k+m}(dx)$ converges weakly to the distribution $Q_m(dx) = L_m f^m(x) Q_0(dx)$. Let us show that there exists a subsequence P_{s_i} converging weakly to ε_{x^*}.

By Theorem 2.2 in [8], the set of all finite intersections of open balls with centers in a countable dense set in X and with rational radii is a countable set determining convergence. Extract from it the subset $\mathfrak{A}$ consisting of sets of Q_1-continuity. Enumerate the elements of $\mathfrak{A}$: $\mathfrak{A} = \{A_i\}_{i=1}^\infty$. Fix a sequence $\{\varepsilon_m\}_{m=0}^\infty$, $\varepsilon_m > 0$, $\varepsilon_m \to_{m\to\infty} 0$. Since $P_{s_k+m}(A) \to_{k\to\infty} Q_m(A)$ for any $A \in \mathfrak{A}$, there exists a subsequence $R_{1,m}(dx) = P_{s_{k_m}+m}(dx)$, for which $|R_{1,m}(A_1) - Q_m(A_1)| < \varepsilon_m$ for all $m = 0, 1, \cdots$. From the sequence $\{R_{k-1,m}(dx)\}_{m=0}^\infty$ ($k = 2, 3, \cdots$) we extract in similar fashion a subsequence $\{R_{k,m}(dx)\}_{m=0}^\infty$, for which $|R_{k,m}(A_k) - Q_m(A_k)| < \varepsilon_m$. The diagonal subsequence $R_{m,m}(dx)$ possesses the property $|R_{m,m}(A_i) - Q_m(A_i)| < \varepsilon_m$ for any $A_i \in \mathfrak{A}$, $i \leqq m$. This subsequence converges weakly to $\varepsilon_{x^*}(dx)$. Indeed, for any $A_i \in \mathfrak{A}$,

$$|R_{m,m}(A_i)-\varepsilon_{x^*}(A_i)| \leqq |R_{m,m}(A_i)-Q_m(A_i)|+|Q_m(A_i)-\varepsilon_{x^*}(A_i)|.$$

For $m \geqq i$ the first term does not exceed ε_m and hence as $m \to \infty$ tends to zero. The second term tends to zero by conditions f) and g) of Lemma 2.

Thus there exists a subsequence $P_{s_i}(dx)$ converging to $\varepsilon_{x^*}(dx)$. From (6) it follows that $P_{s_i+1}(dx)$ converges to the same limit and hence any subsequence of the sequence $P_s(dx)$ converges to this limit. But this is true also for the sequence itself. The theorem is proved.

The next two assertions give sufficient conditions for the convergence of the sequence $P_s(dx)$ to $\varepsilon_{x^*}(dx)$ for two important special choices of the transition probabilities $Q_s(y, dx)$ (see [3]).

Corollary 2. *Suppose that the function f is calculated with no random error,*

$$Q_s(x, A) = \int_X 1_{[z\in A, f(x)\leqq f(z)]} P_s(x, dz) + 1_A(x) \int_X 1_{[f(z)<f(x)]} P_s(x, dz), \tag{7}$$

the transition probabilities $P_s(x, dz)$ converge weakly as $s \to \infty$ to $\varepsilon_x(dz)$ for all $x \in X$, conditions c)–f) *hold and* h) *the measure μ is absolutely continuous with respect to the measure P_0.*

Then the sequence of distributions defined by (6) *converges weakly to $\varepsilon_{x^*}(dx)$.*

PROOF. From conditions e) and h), it follows that $P_0(A(\delta))>0$ for any $\delta>0$, while from (7) it follows that $P_s(A(\delta)) \geqq \cdots \geqq P_0(A, \delta))$ for any $\delta > 0$ and $s = 0, 1, \cdots$. Using conditions d) and f) we see that g) is valid. All the conditions of Theorem 1 are fulfilled. The corollary is proved.

In order to model the random element η_s with distribution $Q_s(y, d\eta_s)$ defined by (7), it is first necessary to derive an independent realization ζ_s of the random element with distribution $P_s(y, d\zeta_s)$ and to set $\eta_s = \zeta_s$ if $f(\zeta_s) \geqq f(y)$ and $\eta_s = y$ if $f(\zeta_s) < f(y)$.

Corollary 3. *Suppose that $X \subset R^n$, $n \geqq 1$, $\mu = \mu_n|_{\mathscr{B}}$ (μ_n is Lebesgue measure), the conditions* c)–f) *and* h) *hold, the measures $Q_s(y, dx)$ are defined according to the formula*

$$Q_s(y, dx) = c_s(y)\beta_s^{-n}\varphi(\beta_s^{-1}(x-y))\mu_n(dx), \tag{8}$$

where φ is a continuous symmetric distribution density in R^n with bounded support, $\beta_s > 0$, $\sum_{s=0}^{\infty} \beta_s < \infty$, and $c_s(y)[\beta_s^{-n} \int_X \varphi(\beta_s^{-1}(x-y))\mu_n(dx)]^{-1}$.

Then the sequence of distributions $P_s(dx)$ defined by (6) *weakly converges to $\varepsilon_{x^*}(dx)$.*

PROOF. Under the above assumptions, the distributions $P_s(dx)$, $s \geqq 1$, have continuous densities with respect to Lebesgue measure. Denote them by $p_s(x)$. From (8) it follows that $p_s(x) > 0$ for any $s \geqq 0$ and for those $x \in X$ for which $f(x) \neq 0$. Let us show that g) holds. Take a fixed $\delta > 0$. From (6) and the fact that φ has compact support it follows for any s and ε that

$$P_{s+1}(A(\varepsilon+\varepsilon_s)) \geqq P_s(A(\varepsilon)), \tag{9}$$

where the quantity $\varepsilon_s \geqq 0$ is determined in terms of the extent of the support of the density φ, and $\sum_{s=0}^{\infty} \varepsilon_s = \text{const} \sum_{s=0}^{\infty} \beta_s < \infty$. Choose s_0 so that $\sum_{s=s_0}^{\infty} \varepsilon_s < \delta/2$ and let $\delta_1 = P_{s_0}(A(\delta/2))$. For any $s \geqq s_0$,

$$P_s(A(\delta)) \geqq P_{s_0}\left(A\left(\frac{\delta}{2} + \sum_{t=s_0}^{s} \varepsilon_t\right)\right) \geqq P_{s_0}\left(A\frac{\delta}{2}\right) = \delta_1 \tag{10}$$

and hence condition g) holds. The corollary is proved.

In order to model the random vector η_s with distribution $Q_s(y, d\eta_s)$ defined by (8), it is necessary to determine an independent realization ζ_s of the random vector distributed with density φ, to verify that $y + \beta_s\zeta_s \in X$ (if not, to determine a new realization ζ_s) and to take $\eta_s = y + \beta_s\zeta_s$.

Just as Theorem 1, Corollaries 2 and 3 may be reformulated for the sequence (2). Let us do it for Corollary 3.

Corollary 4. *Let the conditions in Lemma* 1 *and Corollary* 3 *hold. Then there exists a sequence of natural numbers* N_s ($N_s \to \infty$ *as* $s \to \infty$) *such that the sequence of distributions* $P(s+1, N_s; dx)$ *defined by formula* (2) *converges weakly to* $\varepsilon_{x^*}(dx)$.

PROOF. We can repeat the proof of Corollary 3, changing only formulas (9) and (10). We require that N_s be so large that for any s, instead of (9) we have the inequality $P_{s+1}(A(\varepsilon+\varepsilon_s)) \geqq P_s(A(\varepsilon))(1-\delta_s)$, where $0<\delta_s<1$, and that $\sum_{s=0}^{\infty} \delta_s < \infty$ (this is possible by Lemma 1). Instead of (10) we have the inequalities

$$P_s(A(\delta)) \geqq P_{s_0}\left(A\left(\frac{\delta}{2}+\sum_{t=s_0}^{s} \delta_t\right)\right) \prod_{t=s_0}^{s} (1-\delta_t) \geqq \delta_1 \prod_{t=s_0}^{\infty} (1-\delta_t).$$

To complete the proof it remains to use the known fact: if $0<\delta_s<1$ for all $s=0, 1, \cdots$ and $\sum_{s=0}^{\infty} \delta_s < \infty$, then $\prod_{s=0}^{\infty} (1-\delta_s) > 0$.

In conclusion we point out that the efficiency of the algorithms in this paper is borne out by calculations carried out by the authors with the goal of finding the extremum of certain functions described in the literature as test functions.

The authors believe that the main results of the paper are the techniques that make it possible to construct and investigate algorithms for the global search for an extremum that are based on models of a sequence of measures converging to a δ-measure concentrated at the global extremum point.

Received by the editors
April 6, 1981

REFERENCES

[1] S. M. ERMAKOV, A. A. ZHIGLYAVSKII AND V. N. SOLNTSEV, *On a general scheme of random search for extrema of functions*, in: *Monte-Carlo Methods in Computational Mathematics and Mathematical Physics*, vol. 1, VTs SO AN SSSR, Novosibirsk, 1979, 17–24. (In Russian.)

[2] YU. A. SUSHKOV, *On a method of organizing a random search*, in: *The Investigation of Operations and Statistical Modeling*, no. 1, LGU, Leningrad, 1972, pp. 180–186. (In Russian.)

[3] S. M. ERMAKOV AND A. A. ZHIGLYAVSKII, *On a probabilistic search method in global optimization and the design of extremal experiments*, in: *Questions of Cybernetics*, No. 73, *Scientific Advice on a complex problem*, Cybernetics, Moscow, 1981, pp. 3–24. (In Russian.)

[4] S. M. ERMAKOV, *Die Monte-Carlo Methode und Verwandte Fragen*, R. Oldenbourg Verlag, Munich, Vienna, 1975.

[5] YU. V. PROKHOROV AND YU. A. ROZANOV, *Theory of Probability*, Nauka, Moscow, 1973. (In Russian.)

[6] M. LOÈVE, *Probability Theory*, Van Nostrand, Princeton, NJ, 3rd ed., 1963.

[7] J. R. BLUM, H. CHERNOFF, M. ROSENBLATT AND H. TEICHER, *Central limit theorems for interchangeable process*, Canadian J. Math., 10 (1958), 222–229.

[8] P. BILLINGSLEY, *Convergence of Probability Measures*, John Wiley, New York, 1968.