of using simply $\phi(x)$, x a single observation from the uniform distribution, one should use,

$$\{\phi(x) + \phi(1 - x) + \phi(y)\}/3$$

where

$$y = x \qquad 0 \leqq x \leqq \tfrac{1}{2}$$

$$= \tfrac{3}{2} - x \qquad \tfrac{1}{2} < x \leqq 1$$

for example. The reduced variance property of this estimate is a result of the above theorem. G consists of the identity transformation, the transformation $x \to 1 - x$, and the transformation $x \to y$. Each of these transformations then has weight $\frac{1}{3}$.

REFERENCES

[1] Hodges, J. L., Jr. and Lehmann, E. L. (1950). Some problems in minimax point estimation. *Ann. Math. Statist.* **21** 182–197.
[2] Hammersley, J. M. and Morton, K. W. (1956). A new Monte Carlo technique, antithetic variates. *Proc. Cambridge Philos. Soc.* **52** 449–475.

ON STOCHASTIC APPROXIMATIONS

By Shmuel Friedman

The Hebrew University, Jerusalem

0. Summary. The procedure of stochastic approximations suggested by Robbins-Monro [1], for reaching a zero point x_0 of a regression function, was shown by Dvoretzky [4], to be a convergent w. p. 1. and in mean square under certain conditions. In this paper we deal with two problems of modifying the process to acquire convergence under weaker conditions.

1. Introduction. Let $H(y/x)$ be a family of distribution functions, which correspond to the parameter x.

Let us write: $m(x) = \int y \, dH(y/x)$; $\sigma^2(x) = \int (y - m(x))^2 \, dH(y/x)$.

Let $\{a_n\}$ be a sequence of positive members, such that, $\sum a_n = \infty$, $\sum a_n^2 < \infty$.

Let x_1 be an arbitrary number. The Robbins-Monro process is defined recursively for all n by $x_{n+1} = x_n - a_n y_n$, where y_n is a chance variable with distribution function $H(y/x_n)$. The conditions for its convergence were shown to be:

$$|m(x)| \leqq L\,|x| + K. \tag{1}$$

$$\sigma^2(x) \leqq \sigma^2 < \infty. \tag{2}$$

$$\text{If} \quad x < x_0, \quad \text{then} \quad m(x) < 0, \quad \text{while if} \quad x > x_0, \quad \text{then} \quad m(x) > 0. \tag{3}$$

Received August 28, 1961; revised January 1, 1962.

(4) For every $\delta_2 > \delta_1 > 0, \quad \inf_{\delta_1 \leqq |x-x_0| \leqq \delta_2} |m(x)| \leqq 0.$

In this paper we shall use two modifications of the process. The first modification will remove the necessity for $m(x)$ and $\sigma^2(x)$ to be bounded by a linear function and a constant respectively. The second will lead to convergence when the point x_0 is not a unique zero point $m(x)$, and will be particularly useful when the function $m(x)$ stops being a constant.

2. Weakening of Conditions (1) and (2).

THEOREM. *Let $f(x)$ be a function which is positive and bounded in any finite interval. Let the following conditions be satisfied:*

$$|m(x)| \leqq (L\,|x| + K)f(x) \tag{5}$$

$$\sigma^2(x) \leqq \sigma^2 f^2(x) \tag{6}$$

and also Conditions (3) *and* (4). *Let us define $x_{n+1} = x_n - a_n y_n / f(x_n)$, for $n \geqq 1$, where x_1 is arbitrary. Then $x_1 \to x_0$* w.p.1. *and in mean square.*

PROOF. Let us define the new random variable $y^*(x) = y(x)/f(x)$, and let us denote its mean and variance by $m^*(x)$, $\sigma_*^2(x)$. Then:

$$|m^*(x)| = m(x)/f(x) \leqq L\,|x| + K. \tag{7}$$

$$\sigma_*^2(x) = \sigma^2(x)/f^2(x) \leqq \sigma^2. \tag{8}$$

(7) and (8) show that $y^*(x)$ satisfies Conditions (1) and (2). (3) and (4) are obvious, because $0 < f(x) < \infty$ in any finite interval, so that the sequence of random variables: $x_{n+1} = x_n - a_n y^*(x_n)$, tends to x_0 according to Dvoretzky's Theorem [4].

This theorem enables us to construct a convergence process when $|m(x)|$ and $\sigma^2(x)$ are bounded by known functions $f_1(x)$, $f_2(x)$. If we take as $f(x)$ $\max\{f_1(x), [f_2(x)]^{\frac{1}{2}}\}$.

This procedure is also applicable in the case when $f(x)$ is decreasing to zero for large values of x. In this case there is also convergence in the usual Robbins-Monro procedure, but the convergence is rather slow. By dividing by $f(x)$ we enlarge the step for big values of x.

Another transformation of the r.v.y. can sometimes be used if the variance does not exist for all values of x. In this case if we define $y^{**} = |y_n|^{\frac{1}{2}} \operatorname{sgn} y_n$; $x_{n+1} = x_n - a_n y_n^{**}$, then this process converges to x_0 under certain conditions.

3. Convergence to the point where $m(x)$ stops being a constant.

THEOREM. *Let Conditions* (1) *and* (2) *be fulfilled, and also the following conditions:*

(9) *If* $x < x_0$, *then* $m(x) = 0$; *while if* $x > x_0$, *then* $m(x) > 0$.

(10) *For every* $0 < \delta, \quad \inf_{\delta \leqq x - x_0 < \infty} |m(x)| > 0.$

If we choose a_i, δ_i such that: $a_i > 0$, $\sum a_i = \infty$, $\sum a_i^2 < \infty$; $\delta_i > 0$, $\sum a_i \delta_i =$

∞, $\delta_i \to 0$ *and if we define* $x_{n+1} = x_n + a_n(\delta_n - y_n)$ *then* $x_n \to x_0$ w.p.1. *and in mean square.*

PROOF. By definition, $x_{n+1} = x_n + a_n\delta_n - a_n m(x_n) - a_n(y_n - m(x_n))$. Let us put: $T_n(x_n) = x_n + a_n\delta_n - a_n m(x_n)$; $y'_n = a_n(y_n - m(x_n))$ then $E(y'_n/x_1 \cdots x_n) = 0$. By Condition (2): $E(y'_n) \leqq a_n^2\sigma^2$ so that $\sum Ey'^2_n < \infty$. Hence y' satisfies the conditions of Dvoretzky's Theorem [4].

Let us show that the conditions of Dvoretzky's Theorem on T_n are also satisfied.

Without loss of generality we assume: $x_0 = 0$. Let us put $\eta_n = a_n\delta_n$ ($\eta_n \to 0$), and let us define $\xi_n > 0$, such that $\inf_{\xi_n<x<1} m(x) > 2\delta_n$ and $\xi_n \to 0$. This is possible because of (10), and because $\delta_n \to 0$. Put

$$\Delta_n = \max(\xi_n, \eta_n). \tag{11}$$

It is clear that $\Delta_n > 0$, $\Delta_n \to 0$, $\sum \Delta_n = \infty$, and $\inf_{\Delta_n<x<1} m(x) > 2\delta_n$.

We shall show that the condition required by Dvoretzky's Theorem on $T_n(x_n)$ is fulfilled. For $x_n < -\Delta_n : |T_n(x_n)| = |x_n + a_n\delta_n| = |x_n + \eta_n| = |x_n| - \eta_n$, by (9). For $|x_n| \leqq \Delta_n : T_n(x_n) \leqq |x_n| + \eta_n + a_n(L|x_n| + K) \leqq \Delta_n(2 + a_nL) + a_nK$, by Condition (1). For $|x_n| > \Delta_n : T_n(x_n) = a_n\delta_n + x_n - a_n m(x_n)$. For n sufficiently large:

$$a_n\delta_n + x_n - a_n m(x_n) \geqq a_n\delta_n - Ka_n > -Ka_n. \tag{12}$$

Also: $a_n\delta_n + x_n - a_n m(x_n) < a_n\delta_n + x_n - a_n \cdot 2\delta_n = x_n - \eta_n$. This is true because, if $\Delta_n < x_n < 1$, then $m(x_n) > 2\delta_n$ by (11), and if $1 \leqq x_n < \infty$, then for $n > n_1$, $2\delta_n < \inf_{1 \leqq x<\infty} m(x)$.

Hence for $x_n > \Delta_n : |T_n(x_n)| \leqq \max(Ka_n, x_n - \eta_n)$. If we put $\alpha_n = Ka_n + \Delta_n(2 + \alpha_n L)$, we find that in all cases:

$$|T_n| \leqq \max(\alpha_n, |x_n| - \eta_n),$$

and

$$\alpha_n > 0, \qquad \alpha_n \to 0, \qquad \eta_n > 0, \qquad \sum \eta_n = \infty.$$

This shows that the conditions of Dvoretzky's Theorem [4] are satisfied, and $x_n \to x_0$ w.p.l. and in mean square.

Note that Condition (10) is stronger than (4), which is not sufficient in this case, as will be seen in the following example:

For $x \leqq 0$, $m(x) = 0$.

For $(1/2)\sum_1^n a_i\delta_i \leqq x < (1/2)\sum_1^{n+1} a_i\delta_i$, $m(x) = \delta_n/2$, $y(x) = m(x)$, $x_1 = 0$. Then $x_2 = a_1(\delta_1 - \delta_1/2) = a_1\delta_1/2$, and we can see by induction that, $x_{n+1} = (1/2)\sum_1^n a_n\delta_n$ so that $x_n \to \infty$.

This Theorem enables us to find the point where the regression function stops to be a constant, if the value of this constant is known.

We can replace (9) by:

(13) If $x > x_0$ then $m(x) > 0$, while if $x < x_0$, then $m(x) \leqq 0$,

and get the same result, so that we can also reach the point where $m(x)$ takes its last zero, which can be sometimes important. (By a similar process, we can reach the first zero, when $m(x) \geqq 0$, for every value of x exceeding that zero point.)

If we do not know the value of the constant, we can use the next Theorem, which imposes, however, sharper conditions on $m(x)$.

THEOREM. *Let the following Conditions be fulfilled.*

$$|m(x+1) - m(x)| < L|x| + K. \tag{14}$$

$$\sigma^2(x) \leqq \sigma^2 < \infty. \tag{15}$$

$$\text{If } x < x_0, \quad \text{then } \bar{D}m(x) = 0; \quad \text{while if } x > x_0, \quad \text{then } \underline{D}m(x) > 0. \tag{16}$$

$$\text{For every } \delta > 0, \quad \inf_{\delta < x - x_0 < \infty} \underline{D}m(x) > 0. \tag{17}$$

If we choose a_n, c_n, δ_n such that:

$$a_n > 0, \quad \sum a_n = \infty, \quad \sum a_n^2 < \infty, \quad \sum a_n^2/c_n^2 < \infty,$$

$$\delta_n > 0, \quad \delta_n \to 0, \quad \sum a_n \delta_n = \infty;$$

and if we define: $x_{n+1} = x_n - a_n\{[y(x_n + c_n) - y(x_n)]/c_n - \delta_n\}$, then $x_n \to x_0$ w.p.1. and in mean square.

The problem of finding the point where $m(x)$ stops being a constant, was suggested by Gutmann [3].

REFERENCES

[1] ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22** 400–407.

[2] BLUM, JULIUS (1954). Approximation methods which converge with probability one. *Ann. Math. Statist.* **25** 382–386.

[3] GUTMANN, LOUIS and RUTH (1959). An illustration of stochastic approximation. *Biometrics* **15** 551–560.

[4] DVORETZKY, ARYEH (1956). On stochastic approximation. *Proc. Third Berkeley Symp. Math. Statist. and Prob.* **1** 39–56. Univ. of California Press.

THE USE OF THE RANGE IN PLACE OF THE STANDARD DEVIATION IN STEIN'S TEST

BY WILLIAM KNIGHT

University of New Brunswick

A two sample procedure for obtaining a confidence interval of predetermined length for the mean, μ, of a normal distribution with unknown variance, σ^2, was devised by Stein [4] and generalized by Wormleighton [5]. In this procedure a first

Received May 16, 1962.