TWO APPROXIMATIONS TO THE ROBBINS-MONRO PROCESS

J. L. HODGES, JR. AND E. L. LEHMANN*
UNIVERSITY OF CALIFORNIA

1. Introduction

The following process was introduced by Robbins and Monro [1]. For each real number x let $Y(x)$ be a random variable such that $E[Y(x)] = M(x)$ exists. We assume that M is Borel measurable, that the regression equation $M(x) = \alpha$ has a single root θ, which we wish to estimate, and that $(x - \theta)[M(x) - \alpha] > 0$ for all $x \neq \theta$. An initial value x_1 and a sequence $\{a_n\}$ of positive numbers are selected. The $(n + 1)$st approximation to θ is defined inductively by the formula

$$x_{n+1} = x_n - a_n [Y(x_n) - \alpha] . \tag{1.1}$$

In [1], [2], conditions were investigated under which X_n tends to θ in mean square, and in [3], [4] for convergence with probability 1.

The statistician is naturally concerned with the speed of convergence, and with the choice of coefficients $\{a_n\}$ to maximize the speed. This problem was attacked by Chung [5] who studied the asymptotic behavior of the moments of X_n, and thereby was able to prove asymptotic normality under certain conditions.

Chung considers two cases, using different coefficients a_n and getting variances of different orders in the two:

(i) The "quasi-linear" case (theorem 9). Here, $a_n = c/n$, and $\sqrt{n}\,(X_n - \theta)$ tends in law to the normal distribution $N[0, \sigma^2c^2/(2\alpha_1 c - 1)]$ where $\alpha_1 = M'(\theta) > 0$ and σ^2 is the variance of $Y(\theta)$. The variance of X_n tends to 0 with the speed $1/n$ which a statistician would hope for. Chung proves optimum properties for these estimates. Among the assumptions of theorem 9 we mention particularly

$$\varliminf_{|x| \to \infty} \frac{M(x)}{x} > 0 , \tag{1.2}$$

which as Chung emphasizes is quite restrictive from the point of view of statistical applications, since it is not satisfied in any problem in which $M(x)$ is bounded. For example, the quantal response problem (in which up-and-down methods generally had their origin and to date their most important applications) is excluded.

(ii) The "bounded case" (theorem 6). Here, $M(x)$ is bounded, but unfortunately the coefficients a_n are taken to be $1/n^{1-\epsilon}$ where ϵ must exceed a positive number $1/2(1 + K_4)$ whose value depends on the problem. Chung now shows $n^{(1-\epsilon)/2}(x_n - \theta)$ to have a normal limit, so that the variance of x_n tends to 0 with the speed $1/n^{1-\epsilon}$. The statistician is naturally unhappy with estimates of such great variability.

* This paper was prepared with the partial support of the Office of Naval Research, and of the Office of Ordnance Research, U.S. Army, under Contract DA-04-200-ORD-355.