Z. Wahrscheinlichkeitstheorie verw. Geb. 21, 329–338 (1972)

Normal Approximation and Large Deviations for the Robbins-Monro Process*

MICHAEL WOODROOFE

Section 1. Introduction

By a Robbins-Monro Process (RMP) we will understand a Markov Process, X_n, $n \geqq 1$, of the following form. Let $G(\cdot : x)$, $x \in R$, denote a one parameter family of distribution functions for which $G(y : \cdot)$ is measurable for every $y \in R$ and

$$\mu(x) = \int y\, G(dy : x)$$

is finite for every $x \in R$. Let X_1 be any square integrable random variable and let

$$X_{n+1} = X_n - a_n Y_n, \qquad n \geqq 1, \tag{1.1}$$

where $a_n = a/n$ with $a > 0$ and the conditional distribution function of Y_n given $X_1, \ldots, X_n$ is $G(\cdot : X_n)$ for every $n \geqq 1$. This process was introduced in [5] as a method of solving the equation

$$\mu(x) = 0 \tag{1.2}$$

when μ is unknown, and the Y_n's are observable. Thus, theorems which assert the convergence or non-convergence of X_n to a root of (1.2) are of interest. In particular, it is known that if (1.2) has a unique solution x_0, and if G is suitably well-behaved, then $\sqrt{n}(X_n - x_0)$ has an asymptotic normal distribution (see [3] and/or [6]). Here we will supplement this information by investigating the rate of convergence of the distribution function of $\sqrt{n}(X_n - x_0)$ to normality and the probability of large deviations by $X_n - x_0$ — that is, the probability that $X_n - x_0$ is either $\geqq \varepsilon$ or $\leqq -\varepsilon$ where ε is a small positive number which does not vary with n. In studying these questions, we may without loss of generality (and will) restrict our attention to the case that $x_0 = 0$.

Section 2. Preliminaries

We will have repeatedly to deal with products of the form

$$\beta_{nk} = \prod_{j=k+1}^{n} (1 - \beta a_j) \tag{2.1}$$

where $a_j = a/j$ with $a > 0$, $\beta > 0$, and by convention an empty product is one. The inequalities $1 - x \leqq \exp(-x)$, $x \in R$, and $\exp(-x - x^2) \leqq 1 - x$, $0 \leqq x \leqq \frac{2}{5}$, show

* This research was supported in part by the National Science Foundation under grant GP-8910.

immediately that

$$\beta_{nk} \leqq \left(\frac{k+1}{n+1}\right)^{a\beta}, \quad m \leqq k \leqq n, \tag{2.2}$$

$$\beta_{nk} \geqq \left(\frac{k}{n}\right)^{a\beta} \exp(-a^2 \beta^2/k), \quad m \leqq k \leqq n, \tag{2.3}$$

where m is the greatest integer less than $5\alpha\beta/2$. Some useful consequences of (2.2) and (2.3) are the following. (1) If $a\beta > \frac{1}{2}$, then

$$\tau_{nk}^2 = \sum_{j=k}^{n} n a_j^2 \beta_{nj}^2 = \frac{a^2}{2a\beta - 1}\left[1 - \left(\frac{k}{n}\right)^{2a\beta-1}\right] + o(1) \tag{2.4}$$

where (2.4) defines τ_{nk}^2 and $o(1)$ is uniform in $m \leqq k \leqq n$. In particular, $\tau_{nm}^2 \to a^2/(2a\beta - 1)$ as $n \to \infty$. (2) If $a\beta \geqq 1$, then

$$\max_{m \leqq k \leqq n} a_k \beta_{nk} \leqq 2a/n, \quad n \geqq m, \tag{2.5}$$

$$\min_{n \leqq 2k \leqq 2n} a_k \beta_{nk} \geqq d_1/n, \quad n \geqq 2m, \tag{2.6}$$

where d_1 is a positive constant depending only on a and β. Finally, (3)

$$\sum_{j=k}^{n} n^2 a_j^3 \beta_{nj}^3 \leqq 2a\tau_{nk}^2, \quad m \leqq k \leqq n, \tag{2.7}$$

follows trivially from (2.5). We will also need

Lemma 2.1. *If $a\beta \geqq 1$ and $0 < \delta \leqq 1$, then there are constants d_2 and d_3, depending on a, β, and δ, for which*

$$\sum_{k=m}^{n} n a_k^2 \beta_{nk}^2 \tau_{nk}^{-2} k^{-\delta} \leqq d_2 n^{-\delta}(1 + \log n), \tag{2.8}$$

$$\sum_{k=m}^{n} \sqrt{n}\, a_k \beta_{nk} \tau_{nk}^{-1} k^{-\delta} \leqq d_3 n^{\frac{1}{2}-\delta}(1 + \log n) \tag{2.9}$$

for all $n \geqq m$.

Proof. From (2.6) we have $\tau_{nk}^2 \geqq d_1^2(n-k+1)/n$ for $2k \geqq n \geqq 2m$, and therefore, $\tau_{nk}^2 \geqq d_1^2/2$ for $2k \leqq n$ and $n \geqq 2m$. It follows that

$$\sum_{2k \leqq n} n a_k^2 \beta_{nk}^2 \tau_{nk}^{-2} k^{-\delta} \leqq \frac{8a^2}{n d_1^2} \sum_{k=1}^{n} k^{-\delta}$$

$$\sum_{2k > n} n a_k^2 \beta_{nk}^2 \tau_{nk}^{-2} k^{-\delta} \leqq \frac{8a^2}{d_1^2 n^\delta} \sum_{k=1}^{n} k^{-1}$$

for $n \geqq 2m$. (2.8) follows easily, and (2.9) may be established by a similar argument.

Section 3. Normal Approximation

Our study of the rate of convergence of the distribution function of $\sqrt{n}\, X_{n+1}$ to normality will be conducted under several assumptions. First, we will assume the regression function μ to be approximately linear near 0 by requiring

$$\mu(0) = 0 \quad \text{and} \quad \mu'(0) = \beta > 0 \tag{3.1}$$

where $'$ denotes derivative. Our bound on the rate of convergence will then involve the sequence $\bar{g}_k = E(g(X_k))$ where $g(x) = |\mu(x) - \beta x|$, $x \in R$. The case $\mu'(0) < 0$ may, of course, be reduced to the case $\mu'(0) > 0$ by considering $-X_n$, $n \geqq 1$.

We will also place some conditions on the conditional distribution of $Z_n = Y_n - \mu(X_n)$ given X_n. Let $F(z:x) = G(z + \mu(x): x)$, $x, z \in R$, so that $F(\cdot : X_n)$ is a version of the conditional distribution function of Z_n given X_n. We will require the conditional variances

$$\sigma^2(x) = \int z^2 F(dz:x), \qquad x \in R,$$

to be bounded, say

$$\sigma^2(x) \leqq c_1, \qquad x \in R. \tag{3.2}$$

We will also require the conditional characteristic functions

$$\varphi(t:x) = \int e^{itz} F(dz:x), \qquad t, x \in R,$$

to be smooth near the point $(t, x) = (0, 0)$ in the following sense. Let $u(t:x) = \log \varphi(t:x)$, which exists and is bounded for $|t| \leqq 1/\sqrt{c_1}$ and $x \in R$ because $|\varphi(t:x) - 1| \leqq \sigma^2(x) t^2/2$ for $t, x \in R$. Then, we require the existence of positive constants c_2 and $h_1 \leqq 1/\sqrt{c_1}$ for which

$$|u(t:x) - u(t:0)| \leqq c_2 t^2 |x| \tag{3.3}$$

for $|t| \leqq h_1$ and all $x \in R$. Finally, we will require

$$c_3 = \int |z^3| F(dz:0) < \infty. \tag{3.4}$$

Condition (3.3) will be satisfied if the mixed partial derivative $\frac{\partial}{\partial x} \frac{\partial^2}{\partial t^2} \varphi(t:x)$ exists and is continuous in some neighborhood of $(0, 0)$. For then, observing that $\frac{\partial}{\partial x} \frac{\partial}{\partial t} u(t:x)\Big|_{t=0} = 0 = \frac{\partial}{\partial x} u(0:x)$ for small x, we may write

$$u(t:x) - u(t:0) = \int_0^x \int_0^t (t-s) \frac{\partial}{\partial y} \frac{\partial^2}{\partial s^2} u(s:y)\, ds\, dy$$

for small t and x, while $u(t:x)$ is bounded for small t and all x, as observed above.

Let $v_k^2 = E(X_k^2)$, $k \geqq 1$, let $\sigma_n^2 = \sigma^2 \tau_{nm}^2$ where $\sigma^2 = \sigma^2(0)$ and τ_{nm}^2 is defined by (2.4), and let Φ denote the standard normal distribution function.

Theorem 3.1. *Let X_n, $n \geqq 1$, be an RMP which satisfies conditions* (3.1), (3.2), (3.3), *and* (3.4). *Suppose also that $\sigma^2 = \sigma^2(0) > 0$ and that $a\beta \geqq 1$. Then, there is a constant c for which*

$$|Pr(\sqrt{n}\, X_{n+1} \leqq x) - \Phi(x \sigma_n^{-1})| \leqq \frac{c}{\sqrt{n}} + c \sum_{k=m}^{n} \left(n a_k^2 \beta_{nk}^2 v_k \tau_{nk}^{-2} + \sqrt{n}\, a_k \beta_{nk} \bar{g}_k \tau_{nk}^{-1}\right) \tag{3.5}$$

for all $x \in R$ and $n \geqq m$. The constant c depends on $a, \beta, c_1, c_2, c_3, h_1, v_m$, and σ^2.

Before we prove Theorem 3.1, let us investigate its consequences in the case of a *quasi-linear* RMP by which we mean an RMP for which

$$\mu(0)=0 \quad \text{and} \quad \gamma_1 \leqq \mu(x)/x \leqq \gamma_2, \qquad x \neq 0, \tag{3.6}$$

where γ_1 and γ_2 are positive constants. If X_n, $n \geqq 1$, is a quasi-linear RMP for which $a\gamma_1 > \frac{1}{2}$, then it can be shown (cf. [2]) that $v_k^2 \leqq c_4/k$, $k \geqq 1$, where c_4 is independent of k. If, in addition, μ is twice continuously differentiable near zero, then we will also have $g(x) \leqq c_5 x^2$, $x \in R$, where c_5 is independent of x, and therefore, $\bar{g}_k \leqq c_4 c_5/k$, $k \geqq 1$. Combining this remark with Lemma 2.1 now produces

Theorem 3.2. *If, in addition to the hypotheses of Theorem* 3.1, X_n, $n \geqq 1$, *is a quasi-linear RMP with* $a\gamma_1 > \frac{1}{2}$, *and* $g(x) \leqq c_5 x^2$, $x \in R$, *then there is a constant* c' *for which*

$$|Pr(\sqrt{n}\, X_{n+1} \leqq x) - \Phi(x\, \sigma_n^{-1})| \leqq \frac{c'}{\sqrt{n}}(1 + \log n)$$

for all $x \in R$ *and* $n \geqq m$.

Let φ_n denote the unconditional characteristic function of X_{n+1}, $n \geqq 0$. Then, to prove Theorem 3.1, it will suffice to exhibit a positive h, depending on $a, \beta, c_1, c_2, c_3, h_1, v_m$, and σ^2, for which

$$\int_{-h\sqrt{n}}^{h\sqrt{n}} |\varphi_n(t\sqrt{n}) - e^{-\frac{1}{2}\sigma_n^2 t^2}| \frac{dt}{|t|} \tag{3.7}$$

does not exceed the right side of (3.5). Theorem 3.1 will then follow from Theorem 2 of [4], pp. 196–200.

To derive the desired bound for (3.7), we begin with the remark that

$$E(e^{itX_{k+1}} | X_k = x) = \varphi(-t\, a_k : x) \exp(i t x - i t a_k \mu(x))$$

for $t, x \in R$ and $k \geqq 1$ by Eq. (1.1). For $|t a_k| \leqq h_1$, $x \in R$, and $k \geqq 1$, this may also be written

$$E(e^{itX_{k+1}} | X_k = x) = \varphi(-t\, a_k : 0)\, [e^{itx(1-\beta a_k)} + R_k(t, x)] \tag{3.8}$$

where

$$R_k(t, x) = \exp(u(-t\, a_k : x) - u(-t\, a_k : 0) + i t x - i t a_k \mu(x)) - \exp(i t x - i t x \beta a_k).$$

Moreover, by condition (3.3)

$$\begin{aligned} |R_k(t, x)| &\leqq \exp(|u(-t\, a_k : x) - u(-t\, a_k : 0)|) - 1 \\ &\qquad + |\exp(-i t a_k \mu(x)) - \exp(-i t x \beta a_k)| \\ &\leqq c_6\, a_k^2\, t^2\, |x| + |t|\, a_k\, g(x) \end{aligned} \tag{3.9}$$

for $|t a_k| \leqq h_1$, $x \in R$, and $k \geqq 1$ where $c_6 = c_2 e^2$. (Here we use the fact that $|u(t:x)| \leqq 1$ for $|t| \leqq 1/\sqrt{c_1}$.) Now replace t by $s = t\sqrt{n}\, \beta_{nk}$, where $|t| \leqq \sqrt{n}\, h_1/2a$, apply inequality

(2.5), take expectations in (3.8), and iterate for $k=n, \ldots, m$ to obtain

$$\begin{aligned} \varphi_n(t\sqrt{n}) = {} & \varphi_{m-1}(t\sqrt{n}\,\beta_{n\,m-1}) \prod_{k=m}^{n} \varphi(-t\sqrt{n}\,a_k\,\beta_{nk}:0) \\ & + \sum_{k=m}^{n} r_{nk}(t) \prod_{j=k}^{n} \varphi(-t\sqrt{n}\,a_j\,\beta_{nj}:0) \end{aligned} \tag{3.10}$$

where by (3.9)

$$\begin{aligned} |r_{nk}(t)| &= |E(R_k(t\sqrt{n}\,\beta_{nk}, X_k))| \\ &\leqq c_6\, n\, a_k^2\, \beta_{nk}^2\, v_k\, t^2 + \sqrt{n}\, a_k\, \beta_{nk}\, \bar{g}_k\, |t| \end{aligned} \tag{3.11}$$

for $k=m, \ldots, n$ and $n \geqq m$.

We will obviously need an estimate of the products $\prod_{j=k}^{n} \varphi(-t\sqrt{n}\,a_j\,\beta_{nj}:0)$. An appropriate one may be obtained from Taylor's Theorem and the inequality $|u'''(t:0)| \leqq 7c_3$ for $|t| \leqq \sigma^3/5c_3$ ([4], p. 203). Together, they imply that for $|t| \leqq \sqrt{n}\,\sigma^3/10\,a\,c_3$ and $k \geqq m$

$$\sum_{j=k}^{n} u(-t\sqrt{n}\,a_j\,\beta_{nj}:0) = \frac{-t^2}{2}\sigma^2\,\tau_{nk}^2 + \frac{7}{6}\theta\, c_3\, t^3 \sqrt{n^3} \sum_{j=k}^{n} a_j^3\, \beta_{nj}^3 \tag{3.12}$$

where θ is a complex number of modulus at most one. Observe that by (2.7), the absolute value of (3.12) does not exceed $\frac{-1}{3} t^2\,\sigma^2\,\tau_{nk}^2$ for $|t| \leqq \sqrt{n}\,\sigma^2/14\,a\,c_3$ and $k \geqq m$.

We are now prepared to estimate (3.7). Let h be the minimum of $h_1/2a$, $\sigma^3/10\,a\,c_3$, and $\sigma^2/14\,a\,c_3$. Then, by (3.11) and (3.12)

$$\begin{aligned} & \sum_{k=m}^{n} \int_{-h\sqrt{n}}^{h\sqrt{n}} \left| r_{nk}(t) \prod_{j=k}^{n} \varphi(-t\sqrt{n}\,a_j\,\beta_{nj}:0) \right| \frac{dt}{|t|} \\ & \leqq \sum_{k=m}^{n} \int (c_6\, n\, a_k^2\, \beta_{nk}^2\, v_k\, |t| + \sqrt{n}\, a_k\, \beta_{nk}\, \bar{g}_k)\, e^{\frac{-1}{3}\sigma^2\,\tau_{nk}^2\, t^2}\, dt \\ & \leqq \sum_{k=m}^{n} (5\, c_6\, n\, a_k^2\, \beta_{nk}^2\, v_k\, \sigma^{-2}\, \tau_{nk}^{-2} + 5\sqrt{n}\, a_k\, \beta_{nk}\, \bar{g}_k\, \sigma^{-1}\, \tau_{nk}^{-1}) \end{aligned} \tag{3.13}$$

for $n \geqq m$. Moreover, it follows from (3.12), (2.7), and the inequality $|\varphi_{m-1}(s) - 1| \leqq v_m |s|$ that

$$\begin{aligned} & \int_{-h\sqrt{n}}^{h\sqrt{n}} \left| \varphi_{m-1}(t\sqrt{n}\,\beta_{nm-1}) \prod_{k=m}^{n} \varphi(-t\sqrt{n}\,a_k\,\beta_{nk}:0) - e^{\frac{-1}{2}\sigma_n^2 t^2} \right| \frac{dt}{|t|} \\ & \leqq \int \left(\sqrt{n}\, v_m\, \beta_{nm-1} + \frac{7}{3} a\, c_3\, n^{\frac{-1}{2}}\, \tau_{nm}^2\, t^2 \right) e^{\frac{-1}{3}\sigma_n^2 t^2}\, dt \\ & \leqq \sqrt{\pi}(2m\, v_m\, \sigma_n^{-1} + 7\, a\, c_3\, \tau_{nm}^2\, \sigma_n^{-3})/\sqrt{n} \end{aligned} \tag{3.14}$$

for $n \geqq m$. (3.10), (3.13), and (3.14) now combine to give the desired estimate for (3.7), thus completing the proof of Theorem 3.1.

It is interesting that the finiteness of

$$\int |z^3|\, F(dz:x) \tag{3.15}$$

for $x \neq 0$ is not explicitly required in Theorems 3.1 and 3.2. Nor is it implied by their other hypotheses. The RMP determined by $a=1$, $\mu(x)=x$, $x \in R$, $X_1=0$, and

$$\varphi(t:x)=\left[1+\frac{|t x|}{1+x^2}\right] \exp\left(\frac{-|t x|}{1+x^2}-\frac{t^2}{2}\right), \quad t, x \in R,$$

satisfies the hypotheses of Theorem 3.2. However, $\frac{\partial^3}{\partial t^3}\varphi(t:x)$ fails to exist at $t=0$ if $x \neq 0$, so that (3.15) must be infinite if $x \neq 0$ for this process.

Section 4. Large Deviations

In this section we will study the rate of convergence to zero of $Pr(X_{n+1} \leqq -\varepsilon)$, where X_n, $n \geqq 1$, is an RMP which satisfies the conditions listed below, and ε is a small positive number which does not vary with n. We will assume throughout this section that X_n, $n \geqq 1$, is an RMP which satisfies conditions (3.2) and (3.6) of the previous section and that, moreover, $a\gamma_1 \geqq 1$ in condition (3.6). These two assumptions will not be repeated in the statements of our lemmas and theorems. We will also require the existence of moment generating functions which we will denote by the symbol φ, thus changing our notation from that of the previous section. Explicitly, we require the existence of positive constants h_1 and c_2 (possibly different from the h_1 and c_2 of the previous section) for which

$$\varphi(t:x)=\int e^{tz}\, F(dz:x) \leqq c_2 \tag{4.1}$$

for $0 \leqq t \leqq 2h_1$. Here (4.1) defines φ, and F is as in the previous section. We will also require the existence of an integer $r \geqq 5a\gamma_2/2$, where γ_2 is as in condition (3.6), for which $E(e^{-tX_r})$ is finite for small positive values of t. In this case we will have

$$E(e^{-tX_r}) \leqq c_3, \quad 0 \leqq t \leqq h_2 \tag{4.2}$$

for appropriate values of $h_2>0$ and c_3. An easily checked condition which implies (4.2) will be given at the end of this section.

Lemma 4.1. *If* (4.1) *is satisfied, then there is a constant b, depending on h_1, c_1, and c_2, for which $\varphi(t:x) \leqq \exp(\frac{1}{2} b t^2)$ for $0 \leqq t \leqq h_1$ and all $x \in R$.*

Proof. By Taylor's Theorem and an obvious inequality, it will suffice to exhibit a b for which $\varphi''(t:x) \leqq b$ for $0 \leqq t \leqq h_1$ where $'$ denotes derivative with respect to t. This follows from

$$\begin{aligned}\varphi''(t:x) &= \int_{-\infty}^{0} z^2 e^{tz}\, F(dz:x)+\int_{0}^{\infty} z^2 e^{tz}\, F(dz:x)\\ &\leqq \sigma^2(x)+\left(\frac{2}{h_1}\right)^2 \varphi(t+h_1:x),\end{aligned}$$

condition (3.2), and condition (4.1) with $b=c_1+4c_2/h_1^2$.

Theorem 4.1. *Let X_n, $n \geqq 1$, be an RMP which satisfies conditions* (4.1) *and* (4.2). *Let φ_n denote the unconditional moment generating function of X_{n+1}, $n \geqq r-1$, let β_{nk} be defined by* (2.1) *with $\beta = \gamma_1$, and let $h = \min(h_1/2a, h_2/r)$. Then, $\varphi_n(-nt) \leqq (n+c_3)\exp(\frac{1}{2} n b \tau_{nr}^2 t^2)$ for $0 \leqq t \leqq h$ and $n \geqq r$ where b is as in Lemma* 4.1.

Proof. It follows easily from Eq. (1.1) and condition (3.6) that

$$E(e^{-tX_{k+1}} \mid X_k = x) \leqq \varphi(t a_k : x)\,[e^{-t(1-a_k\gamma_1)x} + 1]$$

for $0 \leqq t a_k \leqq h_1$ and $k \geqq r$. Now replace t by $n t \beta_{nk}$, take expectations, apply Lemma 4.1, (2.2), and (2.5), and iterate to obtain

$$\begin{aligned}\varphi_n(-nt) &\leqq \varphi_{r-1}(-nt\beta_{nr-1})\exp(\tfrac{1}{2} n b \tau_{nr}^2 t^2) + \sum_{k=r}^{n} \exp(\tfrac{1}{2} n b \tau_{nk}^2 t^2)\\ &\leqq (n+c_3)\exp(\tfrac{1}{2} n b \tau_{nr}^2 t^2)\end{aligned}$$

for $0 \leqq t \leqq h$ and $n \geqq r$.

Corollary 4.1. *Let the hypotheses of Theorem* 4.1 *be satisfied and let d be an upper bound for $b\tau_{nr}^2$, $n \geqq r$. Then,*

$$Pr(X_{n+1} \leqq -x) \leqq (n+c_3)\begin{cases}\exp\left(\dfrac{-1}{2d}\, n x^2\right): & 0 \leqq x \leqq dh\\[2ex] \exp\left(\dfrac{-1}{2}\, n h x\right): & x \geqq dh.\end{cases}$$

Moreover, if $0 \leqq s \leqq h/2$, then $Pr(X_{n+1} \leqq -x) \leqq (n+c_3)\exp(-3nsx/2)$ for $x \geqq 4ds$.

Proof. The corollary follows easily from Theorem 4.1 and Bernstein's Inequality, $Pr(X_{n+1} \leqq -x) \leqq e^{-ntx}\varphi_n(-nt)$, $t, x \geqq 0$, on setting $t = x/d$, $t = h$, and $t = 2s$ in the three cases respectively.

We will now obtain a more precise estimate of $Pr(X_{n+1} \leqq -\varepsilon)$ than that provided by Corollary 4.1 under some additional assumptions. Let $u(t:x) = \log\varphi(t:x)$ for $0 \leqq t \leqq 2h_1$ and $x \in R$. Then, we will require the existence of $h_3 > 0$ and c_4 for which

$$|u(t:x) - u(t:0)| \leqq c_4 t^2 |x| \tag{4.3}$$

for $0 \leqq t \leqq h_3$ and $x \in R$. We will also require condition (3.1) to hold, and we will use the notation $g(x) = |\mu(x) - \beta x|$, $x \in R$, where $\beta = \mu'(0)$. Observe that $g(x) = o(x)$ as $x \to 0$ and that $a\beta \geqq 1$ since $a\gamma_1 \geqq 1$. In the remainder of this section, we will use $\beta = \mu'(0)$ in the definition of β_{nk} (see (2.1)).

Lemma 4.2. *If conditions* (3.1) *and* (4.1) *are satisfied, then*

$$q_0(t) = \lim (1/n) \log \prod_{k=r}^{n} \varphi(n t a_k \beta_{nk} : 0)$$

exists as $n \to \infty$ for $0 \leqq t \leqq h_1/a$. Moreover, letting $\sigma^2 = \sigma^2(0)$ and $\alpha = a^2\sigma^2/(2a\beta - 1)$, we have $q_0(t) = \frac{1}{2}\alpha t^2 + o(t^2)$ as $t \to 0$.

Proof. Let $0 \leqq t \leqq h_1/a$ and define f_n, $n \geqq 1$, on $(0, 1]$ by $f_n(s) = u(n\,t\,a_k\,\beta_{nk}:0)$ if $k-1 \leqq n\,s \leqq k$ where $r \leqq k \leqq n$ and $f_n(s) = 0$ for $s \leqq (r-1)/n$. Then, by (2.2), (2.3), (2.5), and (4.1), f_n converges boundedly to f as $n \to \infty$ where $f(s) = u(a\,t\,s^{\alpha\beta-1}:0)$, $0 < s \leqq 1$. Therefore,

$$q_0(t) = \lim \int_0^1 f_n(s)\,ds = \int_0^1 f(s)\,ds \tag{4.4}$$

exists as $n \to \infty$, as asserted. Moreover, the expansion $f(s) = \frac{1}{2}\sigma^2 a^2 t^2 s^{2\alpha\beta-2} + o(t^2)$ as $t \to 0$ may be integrated in (4.4) to yield $q_0(t) = \frac{1}{2}\alpha t^2 + o(t^2)$ as $t \to 0$.

The probablistic significance of the lemma is the following. If μ were linear, say $\mu(x) = \beta x$, and the random variables $Z_n = Y_n - \mu(X_n)$ were independent, then the moment generating function of $-n X_{n+1}$ would be

$$\varphi_{r-1}(-n\,t\,\beta_{nr-1}) \prod_{k=r}^{n} \varphi(n\,t\,a_k\,\beta_{nk}:0).$$

Moreover, an easy adaptation of the argument presented in [1], pp. 1017–1018, would show that for small positive $\varepsilon \lim(1/n) \log Pr(X_{n+1} \leqq -\varepsilon) = p_0(\varepsilon)$ as $n \to \infty$ where $p_0(\varepsilon) = \min_t (q_0(t) - \varepsilon t)$. That analogous behavior obtains without the assumptions of independence or linearity is the content of our final theorem.

Theorem 4.2. *Let conditions* (3.1), (4.1), (4.2), *and* (4.3) *be satisfied. Then, as* $n \to \infty$, $q(t) = \limsup (1/n) \log \varphi_n(-n\,t)$ *is finite for* $0 \leqq t \leqq h$, *and*

$$q(t) \leqq q_0(t) + o(t^2) \tag{4.5}$$

as $t \to 0$. *Moreover, defining* $p(\varepsilon) = \min_t (q(t) - \varepsilon t)$ *for* $\varepsilon > 0$, *we have*

$$\limsup (1/n) \log Pr(X_{n+1} \leqq -\varepsilon) \leqq p(\varepsilon)$$

as $n \to \infty$ *for all* $\varepsilon > 0$. *Finally, if* $\sigma^2 > 0$, *then*

$$p(\varepsilon) \leqq \frac{-\varepsilon^2}{2\alpha} + o(\varepsilon^2) = p_0(\varepsilon) \tag{4.6}$$

as $\varepsilon \to 0$.

Proof. The first assertion of the theorem is a trivial consequence of Theorem 4.1, and the third is then a trivial consequence of Berstein's Inequality. Moreover, relation (4.6) follows easily from relation (4.5) and Lemma 4.2. Thus, it will suffice to demonstrate (4.5). Let H_k denote the unconditional distribution function of X_k, let $k \geqq r$, and let $0 \leqq s \leqq h/2$. Then, by condition (3.6), Lemma 4.1, and some familiar conditioning arguments

$$\begin{aligned}
\varphi_k(-k\,s) \leqq & \int_{-\infty}^{-\delta} \exp\left(\tfrac{1}{2} b\,a^2 s^2 - k\,s\,x(1 - a_k \gamma_1)\right) H_k(dx) \\
& + \int_{-\delta}^{0} \exp\left(-k\,s\,x + k\,s\,a_k\,\mu(x)\right) \varphi(k\,s\,a_k : x)\, H_k(dx) \\
& + \int_{0}^{\infty} \exp\left(\tfrac{1}{2} b\,a^2 s^2 - k\,s\,x(1 - a_k \gamma_2)\right) H_k(dx) \\
= & I_1 + I_2 + I_3 \qquad \text{say,}
\end{aligned}$$

where $\delta=4ds$ and d is as in Corollary 4.1. Integrating I_1 by parts, we find easily that $I_1 \leqq 8(k+c_3)\exp(\frac{1}{2}b\,a^2 s^2)$, while obviously $I_3 \leqq \exp(\frac{1}{2}b\,a^2 s^2)$. Therefore, $I_1+I_3 \leqq k\,c_5$ where c_5 is independent of $k \geqq r$ and $0<s\leqq h/2$. To estimate I_2 let $0\leqq s\leqq h_3/a$ and $k\geqq r$. Then,

$$\varphi(k\,s\,a_k:x)\exp\big(-k\,s\,x+k\,s\,a_k\,\mu(x)\big)=\varepsilon(x)\,\varphi(k\,s\,a_k:0)\exp\big(-k\,s\,x(1-\beta\,a_k)\big)$$

where for $-\delta\leqq x\leqq 0$, $\varepsilon(x)\leqq\exp\big(c_6\,s^3+a\,s\,g_1(s)\big)$ with $c_6=4d\,a^2\,c_4$ and $g_1(s)=\sup\{g(x)\colon -\delta\leqq x\leqq 0\}=o(s)$ as $s\to 0$. (The bounds on ε follow easily from (3.1) and (4.3).) It now follows that $I_2\leqq\exp\big(c_6\,s^3+a s g_1(s)\big)\,\varphi(k s a_k:0)\,\varphi_{k-1}\big(-k s(1-\beta\,a_k)\big)$ for $0\leqq s\leqq h_3/a$ and $k\geqq r$. Therefore,

$$\varphi_k(-k\,s)\leqq\exp\big(c_6\,s^3+a\,s\,g_1(s)\big)\,\varphi(k\,s\,a_k:0)\,\varphi_{k-1}\big(-k\,s(1-\beta\,a_k)\big)+k\,c_5$$

for $0\leqq s\leqq h_4=\min(h/2,\,h_3/a)$ and $k\geqq r$. Now replace $k\,s$ by $n\,t\,\beta_{nk}$, apply (2.5), and iterate to obtain

$$\varphi_n(-n\,t)\leqq(c_3+n^2\,c_5)\exp\big(8\,c_6\,n\,t^3+2\,a\,n\,t\,g_1(2\,t)\big)\prod_{k=r}^{n}\varphi(n\,t\,a_k\,\beta_{nk}:0)$$

for $0\leqq t\leqq h_4/2$ and $n\geqq r$. (4.5) follows easily.

Two questions left unanswered by Theorem 4.2 are the following: (1) does $\lim(1/n)\log Pr(X_{n+1}\leqq-\varepsilon)$ necessarily exist as $n\to\infty$ under the hypotheses of Theorem 4.2 (or some minor variation theorem); and (2) if so, is it necessarily equal to $p_0(\varepsilon)$? We have been unable to answer the first of these questions, but we have found an example of an RMP for which the hypotheses of Theorem 4.2 are satisfied, and $q(t)<q_0(t)$ for $0\leqq t\leqq 1$. Since the latter inequality implies $p(\varepsilon)<p_0(\varepsilon)$ for sufficiently small $\varepsilon>0$, the answer to the second question must be "no".

The example is quite simple. Let X_n, $n\geqq 1$, be the RMP determined by $a=1$, $\mu(x)=x$, $x\in R$, $X_1=0$, and $u(t:x)=\frac{1}{2}\sigma^2(x)\,t^2$, $t, x\in R$, where

$$\sigma^2(x)=1+\frac{2x}{1+2x^2},\qquad x\in R.$$

For this process, an argument similar to that given in the proof of Theorem 4.2 will show that

$$\varphi_n(-n\,t)\leqq c\,n\exp\left(\frac{1}{2}\sum_{k=1}^{n}t_{nk}^2\right),\qquad 0\leqq t\leqq 1$$

where c is a constant independent of n, and the array t_{nk} is defined by $t_{nn}=t$ and

$$t_{nk}=\left[1-\frac{t_{nk+1}}{2k}\right]t_{nk+1},\qquad k\geqq n-1. \tag{4.7}$$

Since for $0\leqq t\leqq 1$, $t_{nk}\geqq t\left(1-\dfrac{1}{2k}\right)^{n-k}$, there is a $\delta>0$, independent of $n\geqq 1$ and $0\leqq t\leqq 1$, for which $t_{nk}\geqq t\,\delta$ for $0\leqq t\leqq 1$, $k\geqq n/2$, and $n\geqq 1$. Let $s_k=\log t_{nk}$. Then,

$s_k - s_{k+1} \leqq -t\delta/2k$, $k \geqq n/2$, by (4.7), so that $t_{nk} \leqq t(k/n)^{t\delta/2}$ for $0 \leqq t \leqq 1$, $k \geqq n/2$, and $n \geqq 1$. Therefore,

$$q(t) \leqq \frac{1}{2} t^2 \left[\frac{1}{2} + \int_{\frac{1}{2}}^{1} x^{t\delta} dx\right] < \frac{1}{2} t^2 = q_0(t)$$

for $0 \leqq t \leqq 1$, as asserted.

We conclude with two remarks concerning our assumptions. The first is the obvious comment that if (4.1) and (4.2) are changed to

$$\varphi(-t:x) \leqq c_2, \quad 0 \leqq t \leqq 2h_1, \quad x \in R, \tag{4.1'}$$

$$E(e^{tX_r}) \leqq c_3, \quad 0 \leqq t \leqq h_2, \tag{4.2'}$$

then we would obtain bounds for $\varphi_n(nt)$ which are exact analogues of the bounds derived for $\varphi_n(-nt)$ in Theorems 4.1 and 4.2. The second remark is that if (4.1) and (4.1') are both satisfied, and if $E(e^{tX_1})$ is finite for all t in some (two-sided) neighborhood of zero, then (4.2) and (4.2') are both satisfied for all $r \geqq 1$. This follows by induction from the inequality $\varphi_k(t) \leqq c_2[\varphi_{k-1}(t) + \varphi_{k-1}(-t\gamma_2 a_k)]$, which is valid for $|t a_k| \leqq 2h_1$.

Acknowledgement. I wish to thank a referee for bringing some inaccuracies and unclear exposition in an earlier version of this paper to my attention.

References

1. Bahadur, R.R., Rao, R.R.: On deviations of the sample mean. Ann. math. Statistics **31**, 1015–1027 (1960).
2. Chung, K.L.: On a stochastic approximation method. Ann. math. Statistics **25**, 463–483 (1954).
3. Fabian, V.: On asymptotic normality in stochastic approximation. Ann. math. Statistics **39**, 1327–1332 (1968).
4. Kolmogorov, A.N., Gnedenko, B.V.: Limit distributions for sums of independent random variables. Addison-Wesley 1954.
5. Robbins, H., Monro, S.: A stochastic approximation method. Ann. math. Statistics **22**, 400–407 (1951).
6. Schmetterer, L.: Stochastic approximation. Proc. Fourth Berkeley Sympos. math. Statist. Probab. **1**, 587–609 (1961).

Dr. Michael Woodroofe
The University of Michigan
Department of Mathematics
Ann Arbor, Michigan 48104
USA

(Received February 7, 1969/June 17, 1971)