e.g. on $2 \mid \Sigma \mid^{1/2} / \operatorname{tr} \Sigma$. In the third place, it can be shown that the probability ratio is monotonic. This can be demonstrated either by starting from the Wishart distribution, or by using (2). However, in this example the latter way does not seem to be any simpler than the former. The moral seems to be that in some cases the utilization of the representation (1) or (2) leads to the results in a fast and elegant way, but in other cases the conventional approach may be more practical.

REFERENCES

[1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley and Sons, New York, 1958.

[2] M. S. Bartlett, "On the theory of statistical regression," *Proc. Roy. Soc. Edinburgh*, Vol. 53 (1933), pp. 260–283.

[3] G. Elfving, "A simple method of deducing certain distributions connected with multivariate sampling," *Skand. Aktuarietids.*, Vol. 30 (1947), pp. 56–74.

[4] Samuel Karlin, "Decision theory for Pólya type distributions. Case of two action, I." *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, University of California Press, 1956, pp. 115–128.

[5] E. L. Lehmann, "Ordered families of distributions," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 399–419.

[6] J. G. Mauldon, "Pivotal quantities for Wishart's and related distributions, and a paradox in fiducial theory," *J. Roy. Stat. Soc.*, Ser. B, Vol. 17 (1955), pp. 79–85.

[7] Daniel Teichroew and Rosedith Sitgreaves, "Computation of an empirical sampling distribution for the classification statistic *W*," *Probability and Statistics in Item Analysis and Classification Problems*, School of Aviation Medicine, USAF, Texas, No. 57–98 (1957).

[8] Robert A. Wijsman, "Random orthogonal transformations and their use in some classical distribution problems in multivariate analysis," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 415–423.

ON DVORETZKY'S STOCHASTIC APPROXIMATION THEOREM

C. Derman[1] and J. Sacks[2]

Columbia University

1. Introduction. A very general theorem was proved by Dvoretzky [2] on the convergence of transformations with superimposed random errors. This work followed that of Robbins-Monro [5] and others (see [6] for bibliography) and contains the most comprehensive results on convergence (with probability one and in mean square) of the stochastic approximation procedures of Robbins-

Received August 21, 1958.

[1] This research was sponsored by the Office of Naval Research under Contract Nonr-266(55), Project NR 042–099.

[2] This research was sponsored by the Office of Naval Research under Contract Nonr-266(51), Project NR 043–198.

Monro [5] and Kiefer-Wolfowitz [3]. Wolfowitz [7] provided another proof of Dvoretzky's theorem. In this note we provide a third proof of the probability one version which is of a simpler nature than the previous two. The method of proof also permits a direct extension to the multidimensional case. The multidimensional results obtained by Block [1] do not seem to include the result below. Mean square convergence does not seem to follow readily from our methods.

2. Fundamental lemmas. Here we prove two lemmas which are at the basis of our method of argument. The first lemma is used to prove Theorem 1 below—the one-dimensional version of Dvoretzky's theorem. Lemma 2 is used in the proof of Theorem 2—the multidimensional version of Theorem 1.

LEMMA 1. *Let* $\{a_n\}$, $\{b_n\}$, $\{c_n\}$, $\{\delta_n\}$, *and* $\{\xi_n\}$ *be sequences of real numbers satisfying*

(i) $\{a_n\}, \{b_n\}, \{c_n\}, \{\xi_n\}$ *are non-negative,*

(ii) $\lim_{n\to\infty} a_n = 0$, $\sum b_n < \infty$, $\sum c_n = \infty$, $\sum \delta_n$ *converges,*

and, for all n larger than some N_0,

(iii) $\xi_{n+1} \leqq \max(a_n, (1 + b_n)\xi_n + \delta_n - c_n)$.

Then, $\lim_{n\to\infty} \xi_n = 0$.

PROOF. Let $N > N_0$ and write

$$B_n = \prod_{i=1}^{n} (1 + b_i). \tag{1}$$

Take $n > N$ and iterate (iii) back to N. This yields

$$\begin{aligned}\xi_{n+1} \leqq \max\Bigg(&\frac{B_n}{B_{N-1}}\,\xi_N + B_n \sum_{j=N}^{n} \frac{\delta_j - c_j}{B_j}, \\ &\max_{N \leqq k \leqq n} \left[\frac{B_n}{B_k}\, a_k + B_n \sum_{j=k+1}^{n} \frac{\delta_j - c_j}{B_j}\right]\Bigg).\end{aligned} \tag{2}$$

Now (i) and (ii) imply that B_n increases to B (say) which is finite. It can then be shown that $\sum_{j=1}^{\infty} \delta_j/B_j < \infty$ and $\sum_{j=1}^{\infty} c_j/B_j = \infty$. Since $(B_n/B_{N-1})\xi_N$ is finite we see that the first term in the right-hand side of (2) must be negative for large enough n and can therefore be ignored. Thus, for n large enough

$$\begin{aligned}\xi_{n+1} &\leqq \max_{N \leqq k \leqq n} \left(\frac{B_n}{B_k}\, a_k + B_n \sum_{j=k+1}^{n} \frac{\delta_j - c_j}{B_j}\right) \\ &\leqq B\left(\max_{k \geqq N} a_k + \max_{N \leqq k \leqq n} \left|\sum_{j=k+1}^{n} \delta_j/B_j\right|\right).\end{aligned} \tag{3}$$

Since $\sum \delta_j/B_j$ converges and $a_k \to 0$ the right member of (3) can be made arbitrarily small by choosing N large enough. This completes the proof of Lemma 1.

LEMMA 2. *Let* $\{a_n\}$, $\{c_n\}$, $\{\xi_n\}$ *be as in Lemma* 1. *Suppose*

(i) $\{\delta_n\}$ *are positive,* $\sum \delta_n < \infty$,

(ii) $\sum b_n$ *converges and* $\sum b_n^2 < \infty$.

Then, $\lim_{n\to\infty} \xi_n = 0$.

PROOF: Let N be large enough so that $|b_n| < 1$ for $n > N$. Since $\sum b_n$ converges we have, for $n, k > N$,

$$(4)\quad \begin{aligned} 0 < \frac{B_n}{B_k} = \prod_{i=k+1}^{n} (1 + b_i) &\leqq \exp\left[\sum_{i=k+1}^{n} b_i\right] \\ &\leqq \exp\left[\max_{n \geqq N} \max_{N \leqq k \leqq n} \left|\sum_{i=k+1}^{n} b_i\right|\right] \leqq A < \infty. \end{aligned}$$

Also, because of (ii) and the fact that

$$\frac{B_n}{B_N} \prod_{i=N+1}^{n} (1 - b_i) = \prod_{i=N+1}^{n} (1 - b_i^2),$$

we have

$$(5)\quad \lim_{n\to\infty} \frac{B_n}{B_N} > 0.$$

With (4) and (5) established the proof goes through as in Lemma 1.

We remark that, if the sequences $\{a_n\}, \cdots, \{\xi_n\}$ are random variables which satisfy the stated conditions with probability one, then the results of the lemmas hold with probability 1.

3. Stochastic approximation theorems.

THEOREM 1. (Dvoretzky). *Let* $\{X_n\}$, $\{T_n(X_1, \cdots, X_n)\}$, $\{Y_n(X_1, \cdots, X_n)\}$ *be sequences of real random variables with* X_1 *arbitrary and*

$$(6)\quad X_{n+1} = T_n(X_1, \cdots, X_n) + Y_n(X_1, \cdots, X_n).$$

Assume

$$(7)\quad E\{Y_n \mid X_1, \cdots, X_n\} = 0 \qquad \text{w.p.1},$$

$$(8)\quad \sum EY_n^2 < \infty,$$

and

$$(9)\quad |T_n| \leqq \max(\alpha_n, (1 + \beta_n)|X_n| - \gamma_n)$$

where α_n, β_n, γ_n *are positive numbers such that*

$$(10)\quad \alpha_n \to 0, \sum \beta_n < \infty, \qquad \sum \gamma_n = \infty.$$

Then $X_n \to 0$ w.p.1.

PROOF. We may assume that $\{\alpha_n\}$ is such that

$$(11)\quad \sum \frac{EY_n^2}{\alpha_n^2} < \infty.$$

For, if this is not the case, there is always a sequence α_n^* which satisfies (10) and (11). Taking $A_n = \max(\alpha_n, \alpha_n^*)$ we obtain a sequence which satisfies (9), (10), and (11).

Define $Z_n = Y_n \operatorname{sgn} T_n$. Then (7), (8), and (11) hold with Y_n replaced by

Z_n. Now (7) and (8) imply that $\sum Z_n$ converges w.p.1. From (11), the Chebyshev inequality, and the Borel-Cantelli lemma we conclude that

$$|Z_n| \leqq \alpha_n \tag{12}$$

w.p.1 for n large enough. Now, from (6) and (12) we can write w.p.1 that for n large enough

$$\begin{aligned} |X_{n+1}| &\leqq 2\,\alpha_n, && \text{if } |T_n| \leqq \alpha_n, \\ &= |T_n| + Z_n, && \text{if } |T_n| > \alpha_n. \end{aligned}$$

Hence

$$|X_{n+1}| \leqq \max(2\alpha_n, |T_n| + Z_n) \leqq \max(2\alpha_n, (1 + \beta_n)|X_n| + Z_n - \gamma_n)$$

for large enough n w.p.1. Lemma 1 with $\xi_n = |X_n|$, $a_n = 2\alpha_n$, $b_n = \beta_n$, $\delta_n = Z_n$, $c_n = \gamma_n$ yields the desired conclusion.

Remark: The above proof also goes through for the extended case considered by Dvoretzky where the α's, β's, γ's are allowed to be random with $\alpha_n \to 0$ uniformly w.p.1 and β_n, γ_n satisfying (9) and (10) w.p.1.

We now turn to the obvious multidimensional generalization of Theorem 1. The symbol $|t|$ will be used below to denote the length of a vector t.

THEOREM 2. *The conditions are the same as in Theorem* 1 *with these modifications*: $\{X_n\}$, $\{T_n\}$, $\{Y_n\}$ *are p-dimensional random vectors*; (8) *should be interpreted as* $\sum E|Y_n|^2 < \infty$; *the absolute value in* (9) *should be read as length. The conclusion is that* $|X_n| \to 0$ w.p.1.

PROOF: As in Theorem 1 we can assume

$$\sum \frac{E\,|Y_n|^2}{\alpha_n^2} < \infty. \tag{13}$$

For similar reasons we may also assume

$$\sum \alpha_n \gamma_n = \infty. \tag{14}$$

If we define, for random orthogonal transformations $P_n = P_{n, x_1, x_2, \cdots, x_n}$, $Z_n = P_n Y_n$, then $\{Z_n\}$ satisfies (7), (8), and (13) and, as a consequence of (13),

$$|Z_n| < \alpha_n \tag{15}$$

w.p.1 for n large enough. Choose P_n so that $P_n T_n = (|T_n|, 0, \cdots, 0)$ and notice that

$$|X_{n+1}|^2 = |P_n X_{n+1}|^2 = (|T_n| + Z_{n1})^2 + \sum_{r=2}^{p} Z_{nr}^2 \tag{16}$$

where Z_{nr} is the rth component of Z_n.

Fix ω (a point in the sample space) and choose $N(\omega)$ so that, for $n \geqq N(\omega)$ (15) holds. If $|T_n| > 2\alpha_n$ we have as consequences of (9) and (15)

$$0 < 2\alpha_n + Z_{n1} < |T_n| + Z_{n1} < (1 + \beta_n)|X_n| - \gamma_n + Z_{n1}, \tag{17}$$

$$(17') \qquad |X_n| > \frac{2\alpha_n + \gamma_n}{1 + \beta_n} = \rho_n \text{ (say)}.$$

Thus (16) and (17) yield

$$|X_{n+1}|^2 \leqq ((1 + \beta_n)|X_n| - \gamma_n + Z_{n1})^2 + \sum_{r=2}^{p} Z_{nr}^2$$
$$= (1 + \beta_n)^2|X_n|^2 + |Z_n|^2 + 2Z_{n1}(1 + \beta_n)|X_n|$$
$$- 2\gamma_n Z_{n1} - 2\gamma_n(1 + \beta_n)|X_n| + \gamma_n^2 .$$

Let $-c_n$ be the sum of the last three terms. Then (15) and (17′) show that $-c_n \leqq -2\alpha_n \gamma_n$ and, therefore, from (14), $\sum c_n = \infty$. Thus, whenever $|T_n| > 2\alpha_n$,

$$(18) \qquad |X_{n+1}|^2 \leqq (1 + \beta_n)^2 |X_n|^2 + \frac{2Z_{n1}(1 + \beta_n)}{\rho_n} |X_n|^2 + Z_n^2 - c_n .$$

If $|T_n| \leqq 2\alpha_n$, then $|X_{n+1}|^2 \leqq 9\alpha_n^2$. Thus, letting

$$1 + b_n = (1 + \beta_n)^2 + \frac{2Z_{n1}(1 + \beta_n)}{\rho_n},$$

we have

$$(19) \qquad |X_{n+1}|^2 \leqq \max\,(9\alpha_n^2 , (1 + b_n)|X_n|^2 + |Z_n|^2 - c_n).$$

We wish to apply Lemma 2 at this point and to do so we put $a_n = 9\alpha_n^2$, $\delta_n = |Z_n|^2$, $\xi_n = |X_n|^2$ and c_n, b_n as they are defined between (17′) and (19). Since $\sum Z_n^2 < \infty$ w.p.1 is a consequence of (8) we have only to verify that $\sum b_n$ and $\sum b_n^2$ converge. Since $\sum \beta_n < \infty$ and, consequently, $\sum \beta_n^2 < \infty$ we have to show that $\sum \frac{Z_{n1}(1 + \beta_n)}{\rho_n}$ converges. Since $E\{Z_{n1} \mid X_1, \cdots, X_n\} = 0$ w.p.1 and since, by (13),

$$\sum_{n=1}^{\infty} E \frac{(1 + \beta_n)^2 Z_{n1}^2}{\rho_n^2} \leqq \sum_{n=1}^{\infty} (1 + \beta_n)^4 \frac{E\,|Z_n|^2}{(2\alpha_n + \gamma_n)^2}$$
$$\leqq \max_n\,(1 + \beta_n)^4 \sum_{n=1}^{\infty} \frac{E\,|Z_n|^2}{\alpha_n^2} < \infty$$

we can apply Theorem D, p. 387, [4] to draw the desired conclusion. In similar fashion we can show $\sum b_n^2 < \infty$ and this completes the proof of the theorem.

Remark: Theorem 2 and its extension permitting random α_n, β_n, and γ_n enable us to prove convergence properties of the multidimensional analogues of the Kiefer-Wolfowitz procedure. In particular it fills a gap in [6], Section 5 where convergence w.p.1 is assumed. The author of [6] expresses his thanks to H. Kesten for bringing this to his attention.

REFERENCES

[1] H. D. Block, "On stochastic approximation," *ONR Report*, Department of Mathematics, Cornell University, Ithaca, New York.

[2] ARYEH DVORETZKY, "On stochastic approximation," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 39-56, University of California Press, 1956.
[3] J. KIEFER and J. WOLFOWITZ, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 462–466.
[4] MICHEL LOÈVE, *Probability Theory*, D. Van Nostrand Company, New York, 1955.
[5] HERBERT ROBBINS and SUTTON MONRO, "A stochastic approximation method," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 400–407.
[6] JEROME SACKS, "Asymptotic Distribution of Stochastic Approximation Procedures," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 373–405.
[7] J. WOLFOWITZ, "On Stochastic Approximation Methods," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1151–1155.

ON A PROBLEM OF ROBBINS

BY J. R. ISBELL

University of Washington

1. Introduction. This note concerns a sequential decision problem raised by Herbert Robbins [2]. The problem is not solved; in fact, it is not known if there is a uniformly best procedure. A procedure is given here which is uniformly better than the one proposed in [2] and is best at least in a special case.

The nature of the problem is this: given two coins with unknown probabilities p_1, p_2, of coming up heads, to prescribe a rule for making an infinite sequence of tosses, choosing the coin for the nth toss as a function of the history of the sequence since the $(n - r)$-th toss (inclusive). The memory length r is fixed. The aim is to maximize the frequency of heads.

The rule proposed here is best in case p_1 or p_2 is 0. We cannot say *the* best, since many rules have the same effects in this case. The rule may be briefly stated: "*Change coins when one coin shows tails r successive times, or when $r - 1$ tails with one coin are followed by a single toss with the other coin, which is tails*". Robbins' rule [2] calls for changing in these cases and further whenever the first toss with a new coin is tails. For $r \leqq 2$, the rules coincide. Otherwise the present rule is better except in two trivial cases, $p_1 = p_2$ and $\max(p_1, p_2) = 1$.

2. Formulation. The description of the memory requires some amplification for the case $n < r$. (None is given in [2].) Here we shall regard the sequence of tosses as a Markov process with 4^r states, namely the states of the memory. We consider that the process may begin in any state, and we propose to evaluate any procedure according to the results it yields starting from the worst possible state.

This is an artificial description which one might prefer to avoid. On the other hand, any decision procedure which might be optimal according to some other version of the problem but disqualified by our artificial start could be described

Received July 8, 1958; revised January 5, 1959.