REFERENCES

[1] D. R. Cox, "The design of an experiment in which certain treatment arrangements are inadmissible," *Biometrika*, Vol. 41 (1954), pp. 287–295.
[2] R. A. Fisher, *The Design of Experiments*, 5th ed., Oliver and Boyd, Ltd., 1949.
[3] P. M. Grundy and M. J. R. Healy, "Restricted randomization and quasi-Latin squares," *J. Roy. Stat. Soc.* (B), Vol. 12 (1950), pp. 286–291.
[4] J. O Irwin and M. G. Kendall, "Sampling moments of moments for a finite population," *Ann. Eugenics* Vol. 12 (1943–45), pp. 138–142.
[5] O. Kempthorne, *The Design and Analysis of Experiments*, J. Wiley and Sons, 1952.
[6] F. Yates, "Bases logiques de la planification des expériences, *Ann. Inst. H. Poincaré* Vol. 12 (1951), pp. 97–112.

ON STOCHASTIC APPROXIMATION METHODS[1]

By J. Wolfowitz

Cornell University

In [1] A. Dvoretzky proved the theorem quoted below, which implies all previous results on the convergence to a limit of stochastic approximation methods. (For a description of these results see [1].) In the present note we give a simple and, we think, perspicuous proof of this theorem which may be of help in further work. The present note is entirely self-contained and may be read without reference to [1].

Theorem. (Dvoretzky) *Let α_n, β_n and $\gamma_n (n = 1, 2, \cdots)$ be non-negative real numbers satisfying*

$$\lim_{n=\infty} \alpha_n = 0, \tag{1}$$

$$\sum_{n=1}^{\infty} \beta_n < \infty, \tag{2}$$

and

$$\sum_{n=1}^{\infty} \gamma_n = \infty. \tag{3}$$

Let θ be a real number and $T_n (n = 1, 2, \cdots)$ be measurable transformations satisfying

$$|T_n(r_1, \cdots, r_n) - \theta| \leqq \max[\alpha_n, (1 + \beta_n)|r_n - \theta| - \gamma_n] \tag{4}$$

for all real $r_1, \cdots, r_n$. Let X_1 and $Y_n (n = 1, 2, \cdots)$ be random variables and define[2]

Received December 16, 1955.

[1] This research was supported by the United States Air Force under Contract No. AF18-(600)-685 monitored by the Office of Scientific Research.

[2] In the proof of the theorem we will, for the sake of brevity, write $T_n(X_n)$ for

$$T_n(X_1, \cdots, X_n),$$

just as is done in [1]. No ambiguity will be caused by this.

$$X_{n+1}(\omega) = T_n(X_1(\omega), \cdots, X_n(\omega)) + Y_n(\tag{5}$$

for $n \geqq 1$.

Then the conditions $E\{X_1^2\} < \infty$,

$$\sum_{n=1}^{\infty} E\{Y_n^2\} < \infty \tag{6}$$

and

$$E\{Y_n \mid X_1, \cdots, X_n\} = 0 \tag{7}$$

with probability 1 for all n, *imply*

$$\lim_{n=\infty} E\{(X_n - \theta)^2\} = 0 \tag{8}$$

and

$$P\{\lim_{n=\infty} X_n = \theta\} = 1. \tag{9}$$

EXTENSION. *The theorem remains valid if* α_n *and* β_n *in* (4) *are replaced by non-negative functions* $\alpha_n(r_1, \cdots, r_n)$ *and* $\beta_n(r_1, \cdots, r_n)$ *respectively, provided: The functions* $\alpha_n(r_1, \cdots, r_n)$ *are uniformly bounded and*

$$\lim_{n=\infty} \alpha_n(r_1, \cdots, r_n) = 0 \tag{10}$$

uniformly for all sequences $r_1, \cdots, r_n, \cdots$; *the functions* $\beta_n(r_1, \cdots, r_n)$ *are measurable and*

$$\sum_{n=1}^{\infty} \beta_n(r_1, \cdots, r_n) \tag{11}$$

is uniformly bounded and uniformly convergent for all sequences $r_1, \cdots, r_n, \cdots$; *and for any* $L > 0$ *there exist non-negative functions* $\gamma_n(r_1, \cdots, r_n)$ *satisfying* (4), and

$$\sum_{n=1}^{\infty} \gamma_n(r_1, \cdots, r_n) = \infty \tag{12}$$

holds uniformly for all sequences $r_1, \cdots, r_n, \cdots$ *for which*

$$\sup_{n=1,2,\cdots} |r_n| < L. \tag{13}$$

PROOF: Without loss of generality we may take $\theta = 0$.

I. From (4) and (6) it follows readily that $EX_n^2 < \infty$ for any n.

II. Define $s(n)$ to be the sign of $[T_n(X_n)][X_n]$ if neither factor is zero, and $s(n) = 1$ if either factor is zero. Define $\pi(m, n) = \prod_{j=m}^{n} s(j)$, $Y_n' = \pi(1, n)Y_n$. The series $\sum_1^\infty Y_n'$ converges w.p.1, by Loève ([2], p. 387, D) and (6) and (7). Let

$$Z(m, n) = \sum_{j=m}^{n} Y_j'$$

For any δ and ϵ both >0, there exists $M'(\delta, \epsilon)$ such that

(14) $$P\left\{\sup_{\substack{m,n \\ M' \leqq m \leqq n}} |Z(m, n)| > \frac{\delta}{48}\right\} < \frac{\epsilon}{2}.$$

III. Let $d(m, m-1) = 1$ and, for $n \geqq m$,

$$d(m, n) = \prod_{j=m}^{n} (1 + \beta_j).$$

Consider the sum

$$S(m, n) = \sum_{j=m}^{n+1} d(j, n) Y'_{j-1},$$

which is equal to

(15) $$\sum_{j=m}^{n-1} Z((m-2), (j-1))[d(j, n) - d(j+1, n)] - Y'_{m-2}\, d(m, n) + Z((m-2), (n-1))\, d(n, n) + Y'_n.$$

Since $d(j, n) \geqq d(j+1, n)$ we have that the absolute value of (15) is not greater than

$$2\left[\sup_{\substack{j \\ m-1 \leqq j \leqq n}} |Z((m-2), (j-1))|\right](d(m, n)) + |Y_n|.$$

Hence, from (11) and (14) it follows that, for δ and ϵ both >0, there exists an $M''(\delta, \epsilon) \geqq M'(\delta, \epsilon)$ such that $d(m, \infty) < \frac{3}{2}$ for $m \geqq M''$ and

(16) $$P\left\{\sup_{\substack{m,n \\ M'' \leqq m \leqq n}} |Z(m, n)| < \frac{\delta}{48},\ \sup_{\substack{m,n \\ M'' \leqq m \leqq n}} |S(m, n)| < \frac{\delta}{8}\right\} > 1 - \frac{\epsilon}{2}.$$

Proof of (9) under the conditions of the extension. Let ϵ and δ be positive and arbitrary. It is sufficient to prove that

(17) $$P\{|X_n| < \delta \text{ for all } n \text{ sufficiently large}\} > 1 - \epsilon.$$

Let $M \geqq M''(\delta, \epsilon)$ be so large that, for $n \geqq M$, $\alpha_n < \delta/8$. Let L be so large that $L > \delta$ and

(18) $$\max_{1 \leqq j \leqq M} EX_j^2 < \frac{\epsilon L^2}{32M}.$$

We take this to be the L for which (12) holds. It also follows that

(19) $$P\left\{\max_{1 \leqq j \leqq M} |X_j| \leqq \frac{L}{4}\right\} > 1 - \frac{\epsilon}{2}.$$

Suppose that the following four conditions are fulfilled:

(20) The relations in curly brackets in (16);

$$|X_m| \leqq \frac{\delta}{4} \quad \text{for some} \quad m \geqq M; \tag{21}$$

$$|X_{m+j}| > \frac{\delta}{4}, \quad 1 \leqq j \leqq k; \tag{22}$$

$$|X_{m+k+1}| \leqq \frac{\delta}{4}. \tag{23}$$

Here $1 \leqq k \leqq \infty$. In case $k = \infty$, (22) is to hold for all $j \geqq 1$ and (23) is vacuous. (It will be clear by the time the proof is finished that k cannot $= \infty$.) Because $\alpha_n < \delta/8$ for $n \geqq M$ and because of (20), (21), and (22) it follows that

$$|T_{m+j}(X_{m+j})| > \alpha_{m+j}, \qquad 0 \leqq j \leqq k - 1, \tag{24}$$

$$\text{sign } X_{m+j+1} = \text{sign } T_{m+j}(X_{m+j}), \qquad 0 \leqq j \leqq k - 1. \tag{25}$$

Applying (4) (with the γ's zero) we obtain that X_{m+1} lies between zero and

$$s(m)(1 + \beta_m)X_m + Y_m. \tag{26}$$

Repeating this argument, we obtain that, for $1 \leqq j \leqq k$, X_{m+j} lies between 0 and

$$\begin{aligned} &s(m + j - 1)s(m + j - 2) \cdots s(m)\, d(m, m + j - 1)X_m \\ &+ s(m + j - 1) \cdots s(m + 1)\, d(m + 1, m + j - 1)Y_m + \cdots \\ &+ s(m + j - 1)\, d(m + j - 1, m + j - 1)Y_{m+j-2} + Y_{m+j-1}. \end{aligned} \tag{27}$$

The absolute value of (27) is not greater than

$$|X_m|\, d(m, m + j - 1) + |S(m + 1, m + j - 1)|. \tag{28}$$

Hence

$$|X_{m+j}| < \delta, \qquad 1 \leqq j \leqq k. \tag{29}$$

To prove (17) it remains only to show that the following conditions cannot both hold:

(30) the relations in curly brackets in (16) and (19);

$$|X_n| > \frac{\delta}{4} \quad \text{for all} \quad n \geqq M. \tag{31}$$

Applying the argument of the previous paragraph with δ replaced by L we obtain that

$$|X_n| < L \text{ for all } n \geqq 1. \tag{32}$$

Hence (12) holds. In view of (30) and (31) it follows that

$$|T(X_n)| > \alpha_n \text{ for all } n \geqq M - 1, \tag{33}$$

$$\text{sign } T_n(X_n) = \text{sign } X_{n+1} \text{ for all } n \geqq M - 1. \tag{34}$$

We may now, and do, apply the argument which led to (28), but with the γ's which satisfy (12). We conclude that, for all $n > M$, the absolute value of $|X_n|$ is not greater than

$$(35) \qquad | X_M | \, d(M, n-1) + | S(M+1, n-1) | - \sum_{j=M}^{n-1} \gamma_j$$

For n sufficiently large this becomes negative, contradicting (33) and hence (31). This completes the proof of (9).

The fact that $EX_1^2 < \infty$ is used in the above proof only in order that $EX_n^2 < \infty$ for all n, and this latter fact is needed only for (8), and not for (9). For in the proof above we used the fact that $EX_n^2 < \infty$ only to obtain explicitly an L for which (19) holds. Such an L obviously exists whether or not $EX_n^2 < \infty$.

Proof of (8) under the conditions of the extension. Let $K = \max_{1 \leqq j < \infty} \alpha_j$. Let N be an integer to be chosen later. In view of (9) we have only to prove that $\lim_{n \to \infty} E\{(|X_n| - K)^+\}^2 = 0$. Let P denote probability measure and A be any set in the sample space which can be defined in terms of $X_1, \cdots, X_m$. We use the inequality

$$H_{m+1}(A) = \int_A ((| X_{m+1} | - K)^+)^2 \, dP = \int_A ((| T_m(X_m) + Y_m | - K)^+)^2 \, dP$$

$$(36) \qquad \leqq \int_A [Y_m^2 + ((| T_m(X_m) | - K)^+)^2] \, dP$$

$$\leqq \int_A [Y_m^2 + K\beta_m(1 + K\beta_m) + (1 + \beta_m)^2(1 + K\beta_m)((| X_m | - K)^+)^2] \, dP$$

which is in [1] and can be deduced from (4) and (7). Let $B(j)$ be the set $\{|X_{N+j}| \leqq K, |X_{N+i}| > K \text{ for } 0 \leqq i < j\}$, $D(j)$ the complement of

$$B(0) + B(1) + \cdots + B(j).$$

Iterate the inequality (36) to obtain an upper bound on $H_n(A)$, $n > N$, beginning the iteration at $m = N, N+1, \cdots, n-1$, respectively, and using as A the sets $B(0), B(1), \cdots, B(n-N-1)$, respectively. In each case the last term of the integrand of the right member of (36) vanishes. Adding, we obtain that $H_n(B(0) + \cdots + B(n-N))$ can be made arbitrarily small by making N sufficiently large.

It remains only to consider $H_n(D(n-N))$. For any point in $D(n-N)$ we have, as in (27), that ∞

$$(37) \qquad |X_n| \leqq |\pi(1, N-1) \, d(N, n-1) X_N + S(N+1, n-1)|$$

Hence, by Minkowski's inequality

$$(38) \qquad \left(\int_{D(n-N)} (X_n)^2 \, dP \right)^{\frac{1}{2}}$$
$$\leqq [d(1, \infty)] \left(\int_{D(n-N)} (X_N)^2 \, dP \right)^{\frac{1}{2}} + [d(1, \infty)] \left(\sum_{j=N}^{\infty} EY_j^2 \right)^{\frac{1}{2}}$$

The second term on the right of (38) can be made arbitrarily small by making N sufficiently large. The first term can be made arbitrarily small by making n sufficiently large, since $P\{D(n - N)\} \to 0$ as $n \to \infty$. This completes the proof of (8).

REFERENCES

[1] A. DVORETZKY, "On stochastic approximation," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1956.
[2] M. LOÈVE, *Probability Theory*, D. Van Nostrand Company, 1955.

ON THE DERIVATIVES OF A CHARACTERISTIC FUNCTION AT THE ORIGIN

BY E. J. G. PITMAN

University of Tasmania

1. Introduction. Let $F(x)$, $-\infty < x < \infty$, be a distribution function, and

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} \, dF(x)$$

its characteristic function, defined and continuous for all real t. Let k be a positive integer. If the kth moment of $F(x)$,

$$\mu_k = \int_{-\infty}^{\infty} x^k \, dF(x),$$

exists and is finite (integral absolutely convergent), $\phi(t)$ has a finite kth derivative for all real t given by

$$\phi^{(k)}(t) = i^k \int_{-\infty}^{\infty} x^k e^{itx} \, dF(x).$$

In particular,

$$\phi^{(k)}(0) = i^k \mu_k .$$

The existence and finiteness of μ_k is a sufficient condition for the existence and finiteness of $\phi^{(k)}(0)$. It can be shown (see [1]) that when k is even, this condition is also necessary; but when k is odd this is not so. Zygmund [2] has given a necessary and sufficient condition for the existence of $\phi'(0)$ and also one for the existence of a symmetric derivative of higher odd order at $t = 0$; but he imposes a certain condition (smoothness) on the characteristic function. In the following theorem the conditions are on the distribution function only.

Received October 5, 1955