one finds

$$e^{\frac{1}{2}x_0^2} \frac{\exp\left[-\frac{1}{2}\frac{(x_T - x_0)^2}{2T}\right]}{(2\pi T)^{\frac{1}{2}}} = \int_0^T d\theta(2-\theta)^{\frac{1}{2}} \exp\left[\frac{(x_0+a)^2}{4(2-\theta)}\right] \cdot Q_a(\theta \mid x_0) \frac{\exp\left[-\frac{1}{2}\frac{(x_T - a)^2}{2(T-\theta)}\right]}{[2\pi 2(T-\theta)]^{\frac{1}{2}}}.$$

Integrate on x_T from $-\infty$ to a to obtain

$$\pi^{-\frac{1}{2}} e^{\frac{1}{2}x_0^2} \int_{-\infty}^{(a-x_0)/(2T)^{\frac{1}{2}}} e^{-\frac{1}{2}u^2}\, du = \int_0^T d\theta(2-\theta)^{\frac{1}{2}} \exp\left[\frac{(x_0+a)^2}{4(2-\theta)}\right] Q_a(\theta \mid x_0)\tfrac{1}{2}.$$

Then $Q_a(T \mid x_0)$ can be obtained directly by differentiation with respect to T. A similar derivation can be carried out under the assumption $x_0 < a$. The combined result is

$$Q_a(T \mid x_0) = \frac{|x_0 - a| \exp\left\{-\frac{1}{2}\frac{[x_0(1-T) - a]^2}{T(2-T)}\right\}}{T[2\pi T(2-T)]^{\frac{1}{2}}}, x_0 \neq a, \quad 0 < T \leqq 1.$$

The author has been unable to obtain an expression for $Q_a(T \mid x_0)$ valid for $T > 1$.

REFERENCES

[1] Darling, D. A., and Siegert, A. J. F., "The first passage problem for a continuous Markov process," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 624–639.

[2] Rice, S. O., "Distribution of the duration of fades in radio transmission," *Bell System Tech. J.*, Vol. 37 (1958), pp. 620–623, 630–631.

[3] Siegert, Arnold J. F., "On the zeros of Markoffian random functions," Rand Corporation Memorandum RM-447, September 5, 1950.

[4] Siegert, A. J. F., "On the first passage time probability problem," *Phys. Rev.*, Vol. 81 (1951), pp. 617–623.

A NOTE ON THE ERGODIC THEOREM OF INFORMATION THEORY[1]

By K. L. Chung

Syracuse University

The purpose of this note is to extend the result of Breiman [1], [2] to an infinite alphabet, or equivalently, the result of Carleson [3] to convergence with probability one.

Let $\{\cdots, x_{-1}, x_0, x_1, \cdots\}$ be a stationary stochastic process taking values in a countable "alphabet" $\{a_i, i = 1, 2, \cdots\}$. Let

$$p(a_{i_1}, \cdots, a_{i_n}) = \mathcal{P}\{x_k = a_{i_k}, k = 1, \cdots, n\},$$

Received October 22, 1960.

[1] This research was supported in part by the Office of Scientific Research of the United States Air Force, under Contract No. AF 49 (638)-265.

and write $p_i = p(a_i)$ for short. Denoting by "lg" the logarithm to the base 2, we set

$$g_0 = \lg \frac{1}{p(x_0)}, \qquad g_k = \lg \frac{p(x_{-k}, \cdots, x_{-1})}{p(x_{-k}, \cdots, x_{-1}, x_0)},$$

$$g_0^{(i)} = \lg \frac{1}{p(a_i)}, \qquad g_k^{(i)} = \lg \frac{p(x_{-k}, \cdots, x_{-1})}{p(x_{-k}, \cdots, x_{-1}, a_i)}.$$

We have then

$$g_k \geqq \mathcal{E}\{g_{k+1} \mid x_0, \cdots, x_{-k}\}$$

and

$$\mathcal{E}\{g_k\} \leqq -\mathcal{E}\{\lg p(x_0)\}.$$

Hence $\{g_k, k = 0, 1, 2, \cdots\}$ is a nonnegative lower semimartingale provided that the "entropy" is finite:

$$H = -\mathcal{E}\{\lg p(x_0)\} = -\sum_{i=1}^{\infty} p_i \lg p_i < \infty. \tag{1}$$

Hence by the martingale convergence theorem, g_k converges with probability one as $k \to \infty$. To prove the ergodic theorem, namely that with probability one

$$\lim_{n\to\infty} n^{-1} \lg p(x_0, \cdots, x_{n-1}) = -H, \tag{2}$$

it is sufficient, following [1], to show that

$$\mathcal{E}\{\sup_{0\leqq k<\infty} g_k\} < \infty. \tag{3}$$

The inequality (3) implies also that the sequence $\{g_k, k = 0, 1, 2, \cdots\}$ is uniformly integrable, hence its convergence with probability one implies its convergence in mean (of order one). From this it follows (see [4]) that (2) holds also in mean. The last assertion has already been proved by Carleson [3]. We state our result as follows.

THEOREM. (1) *implies* (3) *and consequently* (2) *both in mean and with probability one.*

PROOF. Let ω denote the sample point and define for each nonnegative integer m

$$E_k(m) = \{\omega : \sup_{0\leqq j<k} g_j < m;\ g_k \geqq m\},$$

$$E_k^{(i)}(m) = \{\omega : \sup_{0\leqq j<k} g_j^{(i)} < m;\ g_k^{(i)} \geqq m\},$$

$$Z_i = \{\omega : x_0 = a_i\}.$$

We may suppose that the sequence $\{p_i, i = 1, 2, \cdots\}$ is nonincreasing since this can always be achieved by relabelling the alphabet. Let $f(m) \geqq 0$ and write

$$\mathcal{P}\{E_k(m)\} = \sum_{i=1}^{\infty} \mathcal{P}\{E_k(m) \cap Z_i\}$$

$$= \sum_{i\leqq f(m)} \mathcal{P}\{E_k(m) \cap Z_i\} + \sum_{i>f(m)} \mathcal{P}\{E_k(m) \cap Z_i\}.$$

We have, since $g_k \geqq m$ on $E_k(m)$,

$$\mathcal{P}\{E_k(m) \cap Z_i\} \leqq 2^{-m}\mathcal{P}\{E_k^{(i)}(m)\};$$

$$\sum_{k=0}^{\infty} \sum_{i \leqq f(m)} \mathcal{P}\{E_k(m) \cap Z_i\} \leqq 2^{-m} \sum_{i \leqq f(m)} \sum_{k=0}^{\infty} \mathcal{P}\{E_k^{(i)}(m)\} \tag{4}$$

$$\leqq 2^{-m} \sum_{i \leqq f(m)} 1 \leqq \frac{f(m)}{2^m}.$$

On the other hand, it is plain that

$$\sum_{k=0}^{\infty} \sum_{i > f(m)} \mathcal{P}\{E_k(m) \cap Z_i\} \leqq \sum_{i > f(m)} \mathcal{P}\{Z_i\} = \sum_{i > f(m)} p_i . \tag{5}$$

Let $f^{-1}(i)$ be the number of m such that $f(m) < i$, then $f^{-1}(i) \leqq 1 + \max\{m : f(m) < i\}$. Summing (4) and (5) over all m, we obtain

$$\sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \mathcal{P}\{E_k(m)\} \leqq \sum_{m=0}^{\infty} \frac{f(m)}{2^m} + \sum_{i=1}^{\infty} f^{-1}(i) p_i . \tag{6}$$

Now choose $f(m) = 2^m/(m + 1)^2$; a simple computation shows that there exist two positive constants A and B such that $f^{-1}(i) \leqq A \lg i + B$ for all $i \geqq 1$. Since $\{p_i\}$ is nonincreasing, we have $ip_i \leqq 1$ so that

$$\sum_{i=1}^{\infty} f^{-1}(i) p_i \leqq \sum_{i=1}^{\infty} \left(A \lg \frac{1}{p_i} + B\right) p_i = AH + B.$$

Hence we have by (6),

$$\sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \mathcal{P}\{E_k(m)\} \leqq \frac{\pi^2}{6} + AH + B.$$

Finally,

$$\mathcal{E}\{\sup_{0 \leqq k < \infty} g_k\} \leqq \sum_{m=0}^{\infty} \mathcal{P}\{\sup_{0 \leqq k < \infty} g_k \geqq m\} = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \mathcal{P}\{E_k(m)\},$$

which completes the proof that (1) implies (3).

REFERENCES

[1] Leo Breiman, "The individual ergodic theorem of information theory," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 809–811.

[2] Leo Breiman, "A correction to 'the individual ergodic theorem of information theory'," *Ann. Math. Stat.*, Vol. 31 (1960), pp. 809–810.

[3] L. Carleson, "Two remarks on the basic theorems of information theory," *Math. Scand.*, Vol. 6 (1958), pp. 175–180.

[4] A. Feinstein, *Foundations of Information Theory*, McGraw-Hill, New York, 1958.