one finds
\begin{equation}
\begin{aligned}
e^{ib\theta} \frac{\exp\left[ -\frac{1}{2} \frac{(x_T - x_0)^2}{2T} \right]}{(2\pi T)^{1/2}} &= \int_0^T d\theta (2 - \theta)^{1/2} \exp\left[ \frac{(x_0 + a)^2}{4(2 - \theta)} \right] \\
&\quad \cdot Q_\omega(\theta \mid x_0)^{1/2}.
\end{aligned}
\end{equation}

Integrate on $x_T$ from $-\infty$ to $a$ to obtain
\begin{equation}
\pi^{-1} e^{ib\theta} \int_{-\infty}^{(a-x_0)/(2\theta)^{1/2}} e^{-u^2} du = \int_0^T d\theta (2 - \theta)^{1/2} \exp\left[ \frac{(x_0 + a)^2}{4(2 - \theta)} \right] Q_\omega(\theta \mid x_0)^{1/2}.
\end{equation}

Then $Q_\omega(T \mid x_0)$ 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
\begin{equation}
Q_\omega(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))^{1/2}}, x_0 \neq a, \quad 0 < T \leq 1.
\end{equation}

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

REFERENCES


---

**A NOTE ON THE ERGODIC THEOREM OF INFORMATION THEORY**

**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
\begin{equation}
p(a_{i_1}, \cdots, a_{i_n}) = \theta | x_k = a_{i_k}, \ k = 1, \cdots, n|,
\end{equation}

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

\[
g_0 = \lg \frac{1}{p(x_0)}, \quad g_k = \lg \frac{p(x_{-k}, \ldots, x_{-1})}{p(x_{-k}, \ldots, x_{-1}, x_0)},
\]

\[
g_0^{(i)} = \lg \frac{1}{p(a_i)}, \quad g_k^{(i)} = \lg \frac{p(x_{-k}, \ldots, x_{-1})}{p(x_{-k}, \ldots, x_{-1}, a_i)}.
\]

We have then

\[
g_k \geq \mathcal{E}\{g_{k+1} \mid x_0, \ldots, x_{-k}\}
\]

and

\[
\mathcal{E}\{g_k\} \leq -\mathcal{E}\{\lg p(x_0)\}.
\]

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

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

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

\[
\lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} p_i \lg p_i = -H,
\]

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

\[
\mathcal{E}\{\sup_{0 \leq k < \infty} g_k\} < \infty.
\]

The inequality (3) implies also that the \( \{g_k, k = 0, 1, 2, \cdots\} \) 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 \( \omega \) denote the sample point and define for each nonnegative integer \( m \)

\[
E_k(m) = \{\omega : \sup_{0 \leq i < k} g_j < m; g_k \geq m\},
\]

\[
E_k^{(i)}(m) = \{\omega : \sup_{0 \leq i < k} g_j^{(i)} < m; g_k^{(i)} \geq m\},
\]

\[
Z_i = \{\omega : x_0 = a_i\}.
\]

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 \( f(m) \geq 0 \) and write

\[
\phi[E_k(m)] = \sum_{i=1}^{\infty} \phi[E_k(m) \cap Z_i]
\]

\[
= \sum_{i \leq f(m)} \phi[E_k(m) \cap Z_i] + \sum_{i > f(m)} \phi[E_k(m) \cap Z_i].
\]
We have, since $g_k \geq m$ on $E_k(m)$,

$$\varphi\{E_k(m) \cap Z_i\} \leq 2^{-m} \varphi\{E_k^{(i)}(m)\};$$

(4) \[ \sum_{k=0}^{\infty} \sum_{i \geq f(m)} \varphi\{E_k(m) \cap Z_i\} \leq 2^{-m} \sum_{i \geq f(m)} \sum_{k=0}^{\infty} \varphi\{E_k^{(i)}(m)\} \leq 2^{-m} \sum_{i \geq f(m)} 1 \leq \frac{f(m)}{2^m}. \]

On the other hand, it is plain that

(5) \[ \sum_{k=0}^{\infty} \sum_{i \geq f(m)} \varphi\{E_k(m) \cap Z_i\} \leq \sum_{i \geq f(m)} \varphi\{Z_i\} = \sum_{i \geq f(m)} p_i. \]

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

(6) \[ \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \varphi\{E_k(m)\} \leq \sum_{m=0}^{\infty} \frac{f(m)}{2^m} + \sum_{i=1}^{\infty} f^{-1}(i)p_i. \]

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

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

Hence we have by (6),

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

Finally,

$$\mathbb{E}\{\sup_{0 \leq k < \infty} g_k\} \leq \sum_{m=0}^{\infty} \varphi\{\sup_{0 \leq k < \infty} g_k \geq m\} = \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \varphi\{E_k(m)\},$$

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

REFERENCES