BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 21, Number 2, October 1989

AN ERGODIC THEOREM FOR CONSTRAINED SEQUENCES OF FUNCTIONS

JOHN C. KIEFFER

I. Introduction. For each integer $n = 1, 2, \ldots$, let S_n be a finite nonempty set, and let $\mathfrak{F}_n$ be a nonempty family of real-valued functions on S_n. This announcement is concerned with the asymptotic behavior of sequences of functions $\{f_n\}$ which are constrained by $\{\mathfrak{F}_n\}$ in the sense that $f_n \in \mathfrak{F}_n$ for every n. Specifically, given an S_n-valued random variable Y_n $(n \geq 1)$, an examination is made of the almost sure asymptotic behavior of sequences of random variables of the form $\{f_n(Y_n)/n\}$, where $\{f_n\}$ is constrained by $\{\mathfrak{F}_n\}$ is the above sense. (See Theorem 1.) Of particular interest for applications is the context in which for some finite set A, and some stationary sequence $X_1, X_2, \ldots$ of A-valued random variables, we have $S_n = A^n$ and $Y_n = (X_1, \ldots, X_n)$ for every n. In this context, our main result (Theorem 2) is an ergodic theorem which gives sufficient conditions on the constraining sequence $\{\mathfrak{F}_n\}$ so that $\{f_n(X_1, \ldots, X_n)/n\}$ will converge almost surely when $\{f_n\}$ is a certain sequence of functions constrained by $\{\mathfrak{F}_n\}$. The subadditive ergodic theorem **[4]** for stationary, ergodic processes with finite state space and the Shannon-McMillan-Breiman theorem **[2]** are special cases of our main result.

At this point, we mention examples from information theory and statistics illustrating the utility of results of the type just described.

EXAMPLE 1.1 (INFORMATION THEORY). Let S_n be a set of messages, each of which has length n. Let $\mathfrak{F}_n$ be the family of all functions $f\colon S_n \to \{1, 2, \ldots\}$ for which there is a uniquely decipherable code **[1]** which assigns to each message $m \in S_n$ a binary codeword of length $f(m)$. Let Y_n be a random message from S_n. One may want to select $f_n \in \mathfrak{F}_n$ so that, with probability one, the codeword length per message length $f_n(Y_n)/n$ does not exceed a certain bound in the limit as the message length $n \to \infty$.

EXAMPLE 1.2 (STATISTICS). Let S_n be the set of all sequences of length n that can be formed from a finite set A. Let $\Theta(\mu)$ be a real parameter of an unknown probability distribution μ on A. Let Y_n be a random sample of size n drawn according to the distribution μ. A family $\mathfrak{F}_n$ of functions on S_n is specified, consisting of the statistics that are to be allowed as possible parameter estimators. It is desired to select a statistic $f_n \in \mathfrak{F}_n$ $(n \geq 1)$ so that $f_n(Y_n) \to \Theta(\mu)$ with probability one as the sample size $n \to \infty$, no matter what may be the distribution of μ.

Received by the editors November 14, 1988 and, in revised form, May 25, 1989.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 28D99; Secondary 60G10, 94A15.
Author supported by NSF Grant NCR-8702176.