HAUSDORFF DIMENSION IN PROBABILITY THEORY[1]

BY
PATRICK BILLINGSLEY

1. Summary

Let $(\Omega, \mathcal{B}, \mu)$ be a probability measure space on which is defined a stochastic process $\{x_n\}$ with finite state space. In §§1–3 we define a notion of fractional dimension, in terms of μ and $\{x_n\}$, for any set $M \subset \Omega$. If Ω is the unit interval, if μ is Lebesgue measure, and if $\sum_{n=1}^{\infty} x_n(\omega)s^{-n}$ is, for each ω, the base s expansion of ω, the definition reduces to the classical one due to Hausdorff. In §§4 and 5 we obtain, under the assumption that $\{x_n\}$ is a Markov chain, the dimensions of certain sets defined in terms of the asymptotic relative frequencies of the various transitions $i \rightarrow j$. In §7 these theorems are specialized to the case in which $\{x_n\}$ is independent. In the classical case these results become extensions of theorems due to Eggleston [4, 5] and Volkmann [16, 17]. In §6 we use the preceding theorems to obtain a result on "generalized Lipschitz conditions" on certain measures, a result which reduces in the classical case to one of Kinney [11]. In §8 the dimensions obtained in the first part of the paper are shown to be related to entropy and certain allied concepts of information theory.

2. Introduction and definitions

Let $\{x_1, x_2, \cdots\}$ be a stochastic process defined on a probability measure space $(\Omega, \mathcal{B}, \mu)$. Suppose that the state space of the process is a finite set σ, the states of which for notational convenience we take to be the first s integers: $\sigma = \{1, 2, \cdots, s\}$. Thus $x_n \in \sigma$ with probability one for all n. A set of the form

$$\{\omega : x_k(\omega) = a_n, k = 1, \cdots, n\},$$

where $(a_1, a_2, \cdots, a_n)$ is a sequence of states, we call a *cylinder* or, more specifically, an *n-cylinder*. (While the sets $\{\omega : x_1(\omega) = a \text{ or } b\}$ and $\{\omega : x_n(\omega) = a\}$ are cylinders according to the usual definition, they are not according to the one given above, which will be adhered to throughout the paper.)

If M is a subset of Ω, if $\rho > 0$, and if $\mathcal{V}$ is an enumerable (possibly finite) collection of cylinders, then we say that $\mathcal{V}$ is a ρ-covering of M provided $\mu(v) < \rho$ for all $v \in \mathcal{V}$ and $M \subset V = \bigcup \{v : v \in \mathcal{V}\}$. (We will consistently denote a collection of cylinders by a script letter, with or without subscripts, and the union of the collection by the corresponding Latin letter.) If $\alpha \geqq 0$,

Received December 12, 1958.

[1] Presented to the Institute of Mathematical Statistics August 26, 1958. Most of the work on this paper was done while the author was a National Science Foundation Postdoctoral Fellow at Princeton University.