EPSILON ENTROPY OF PROBABILITY DISTRIBUTIONS

EDWARD C. POSNER and EUGENE R. RODEMICH
JET PROPULSION LABORATORY
CALIFORNIA INSTITUTE OF TECHNOLOGY

1. Introduction

This paper summarizes recent work on the theory of epsilon entropy for probability distributions on complete separable metric spaces. The theory was conceived [3] in order to have a framework for discussing the quality of data storage and transmission systems.

The concept of data source was defined in [4] as a probabilistic metric space: a complete separable metric space together with a probability distribution under which open sets are measurable, so that the Borel sets are measurable. An ε partition of such a space is a partition by measurable ε sets, which, depending on context, can be sets of diameter at most ε or sets of radius at most $\frac{1}{2}\varepsilon$, that is, sets contained in spheres of radius $\frac{1}{2}\varepsilon$. The entropy $H(U)$ of a partition U is the Shannon entropy of the probability of the distribution consisting of the measures of the sets of the partition. The (one shot) epsilon entropy of X with distribution μ, $H_{\varepsilon;\mu}(X)$, is defined by

$$H_{\varepsilon;\mu}(X) = \inf_U \{H(U);\, U \text{ an } \varepsilon \text{ partition}\} \tag{1.1}$$

and, except for roundoff in the entropy function, a term less than 1, $H_{\varepsilon;\mu}(X)$ is the minimum expected number of bits necessary to describe X to within ε when storage is not allowed. The inf in (1.1) was shown to be a min in [4].

For X a compact metric space, Kolmogorov's epsilon entropy $H_\varepsilon(X)$ is defined as

$$H_\varepsilon(X) = \min_U \{\log \text{card}\,(U);\, U \text{ an } \varepsilon \text{ partition}\} \tag{1.2}$$

and, except for roundoff in the logarithm, is the minimum number of bits necessary to describe X to within ε when words of fixed length are used.

Suppose one does experiments from X independently and then attempts storage or transmission. That is, take a cartesian product $X^{(n)}$ of X, with product measure $\mu^{(n)}$ and supremum metric. Thus, ε sets in the product are the subsets

This paper presents the results of one phase of research carried out under Contract NAS 7-100, sponsored by the National Aeronautics and Space Administration.