The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1249–1259

BOUNDS AND ASYMPTOTES FOR THE PERFORMANCE[1] OF MULTIVARIATE QUANTIZERS

BY PETER ELIAS

Massachusetts Institute of Technology

1. Introduction. Let $\mathbf{x} = (x_1, x_2, \cdots, x_N)$ be a vector-valued random variable with probability measure μ defined on the Lebesgue-measurable subsets of N-dimensional Euclidean space E^N. Let $\{R_i\}$, $1 \leqq i \leqq K$, be a set of K Lebesgue-measurable disjoint subsets of E^N, with

$$\sum_{i=1}^{K} \mu(R_i) = 1. \tag{1.1}$$

We define the K-region quantizer Q with quantization regions R_i as a function mapping the portion of E^N covered by the union of the R_i onto the integers 1 to K, given by

$$Q(\mathbf{x}) = i \qquad \text{for} \quad \mathbf{x} \in R_i. \tag{1.2}$$

Such a quantizer may be used as a model for the grouping of N-variate data (statistics), the quantization of signals (communications engineering) and analog-digital conversion (data-processing). It maps each $\mathbf{x}$ into the integer index i, $1 \leqq i \leqq K$, which labels the region R_i in which $\mathbf{x}$ falls, and saves only the value of i for further processing.

Quantization simplifies the handling of data, but introduces an error in the representation of $\mathbf{x}$, since $\mathbf{x}$ must be estimated by some function $\hat{\mathbf{x}}(i)$ of $Q(\mathbf{x}) = i$ alone. The first exploration of a quantization problem seems to be due to Sheppard [7], who analyzed the effect of quantization error on the estimate of the variance of the distribution μ of a scalar random variable $x(N = 1)$, assuming a smooth μ and equal intervals for the R_i.

Panter and Dite [6], also for $N = 1$, use the mean square value of the difference between x and its estimate $\hat{x}(Q(x))$

$$\overline{[x-\hat{x}(Q(x))]^2} = \int_{E^1} [x-\hat{x}(Q(x))]^2 \, d\mu \tag{1.3}$$

as a measure of performance. For fixed K they seek the minimum of this measure by moving the boundaries between the intervals R_i and by choosing the K values of $\hat{x}(i)$. For absolutely continuous μ with sufficiently smooth density $f = d\mu/dx$ they show that the minimum attainable value of their error measure is asymptotic in K to

$$\overline{[x-\hat{x}(Q(x))]^2} \sim (C_2/K^2)\{\int_{E^1} f^{\frac{1}{3}} \, d\lambda\}^3 \tag{1.4}$$

where λ is Lebesgue measure and C_2 is known. They credit this result to Pierre Aigrain.

Received August 21, 1969.

[1] Research sponsored by the Joint Services Electronics Program (Contract DA28-043-AMC-02536(E)).